On systems of continuity equations with nonlinear diffusion and nonlocal drifts

Guillaume Carlier, Maxime Laborde

On systems of continuity equations with nonlinear diffusion and nonlocal drifts

Guillaume Carlier [*], Maxime Laborde [†]

April 30, 2015

Abstract

This paper is devoted to existence and uniqueness results for classes of nonlinear diffusion equations (or systems) which may be viewed as *regular* perturbations of Wasserstein gradient flows. First, in the case where the drift is a gradient (in the physical space), we obtain existence by a semi-implicit Jordan-Kinderlehrer-Otto scheme. Then, in the nonpotential case, we derive existence from a regularization procedure and parabolic energy estimates. We also address the uniqueness issue by a displacement convexity argument.

Keywords: nonlinear diffusion, systems, interacting species, Wasserstein gradient flows, semi-implicit JKO scheme, nonlinear parabolic equations.

MS Classification: 35K15, 35K40, 49J40.

1 Introduction

The continuity equation with a density-dependent drift

$$\partial_t \rho = \mathrm{div}(\rho v), \quad \text{with } v = V[\rho]$$

is ubiquitous in modeling and arises in a variety of domains such as biology, particle physics, population dynamics, crowd modelling, opinion formation... It should actually come as no surprise since it captures the dynamics of a population of particles following the ODE $\dot{X} = -v(t, X)$ where $v = V[\rho]$

[*]CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE carlier@ceremade.dauphine.fr

[†]CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE laborde@ceremade.dauphine.fr.

depends itself on the density in a way (local, nonlocal, attractive, repulsive etc..) depending on which phenomena (aggregation, diffusion...) one aims to capture and the type of applications.

Of course, at this level of generality not much can be said on existence and uniqueness. There are however two cases which may be treated in a systematic way. The first one, is the *regular case* where $V[\rho]$ is a smooth vector field whatever the probability measure ρ is, with some uniform bounds on some of its derivatives and $\rho \mapsto V[\rho]$ is Lipschitz in the Wasserstein metric. In this regular case, existence and uniqueness can be proved as a simple exercise by the method of characteristics and a suitable fixed point argument. This regular case (a typical example being that of a convolution) is however rather restrictive and for instance rules out diffusion. The second case where there is a general theory is the *Wasserstein gradient flow case*. In this case, at least at a formal level, v may be written as $v = \nabla \frac{\delta \mathcal{E}}{\delta \rho}$ that is the gradient of the first variation of a functional $\mathcal{E}$ defined on measures. In their seminal paper [10], Jordan, Kinderlehrer and Otto discovered that the heat flow is the gradient flow of the entropy functional $\mathcal{E}(\rho) = \int \rho \log(\rho)$ which corresponds to the case $v = \frac{\nabla \rho}{\rho}$. The theory of Wasserstein gradient flows has been very succesful in addressing a variety of nonlinear evolution equations such as the porous medium equation [18], aggregation equations [3] or granular media equations [4]. This powerful theory is presented in a complete and detailed way in the reference book of Ambrosio, Gigli and Savaré [1].

The purpose of the present paper is a contribution to the following general question: can one hope for an existence/uniqueness theory in the case where V is the sum of a Wasserstein gradient flow term and a regular term (not necessarily a gradient). Our motivation for this question actually comes from systems. For instance, a simple but natural model, for the evolution of two (say) interacting species is:

$$\begin{cases} \partial_t \rho_1 = \nu_1 \Delta \rho_1 + \operatorname{div}(\rho_1 \nabla (F \star \rho_1 + G \star \rho_2)), \\ \partial_t \rho_2 = \nu_2 \Delta \rho_2 + \operatorname{div}(\rho_2 \nabla (H \star \rho_1 + K \star \rho_2)). \end{cases}$$

When $\nu_1 = \nu_2 = 0$ i.e. without diffusion, this is exactly the system studied by Di Francesco and Fagioli [8]. As emphasized in [8], if cross-interactions are symmetric i.e. $G = K$ (or more generally G and K are proportional), this system has a (product) Wasserstein gradient flow structure but this is certainly a restrictive and often unrealistic assumption in applications. This is why Di Francesco and Fagioli, still taking advantage of the similarity with Wasserstein gradient flows used a semi-implicit scheme à la Jordan-Kinderlehrer-Otto to obtain existence and uniqueness results. In [8], there

is no diffusion but clearly the structure of the system belongs to the *mixed* case where drifts can be decomposed as the sum of a Wasserstein gradient and a regular term. Of course, the semi-implicit scheme only makes sense when drifts are gradients.

Regarding systems with a gradient structure and in the presence of nonlinear diffusion, our first contribution is to establish strong enough convergence at the level of the semi-implicit scheme to recover a solution of the PDE at the limit. The delicate step is of course to pass to the limit in the nonlinear diffusion term, which can be done thanks to the powerful *flow interchange* argument of Matthes, McCann and Savaré [13] in a similar way as in the work of Di Francesco and Matthes [9]. We will then address the nonpotential case in which the drift may contain a nongradient (but regular) part. This case cannot be attacked by the semi-implicit minimization scheme and we will prove existence by suitably regularizing the diffusion and using standard parabolic energy estimates. Finally, we will derive an uniqueness result from displacement convexity of the energy.

The paper is organized as follows. In section 2, we consider the potential case, introduce a semi-implicit scheme à la Jordan-Kinderlehrer-Otto [10] and state a first existence result. Section 3 is devoted to the proof of this existence result. Section 4 extends the result to the case of systems (again in the case where all drifts are gradients). Section 5 proves existence for the non-potential case. The final section 6 shows uniqueness by a simple displacement convexity argument.

2 The potential case and the semi-implicit JKO scheme

Our aim is to solve the following nonlinear diffusion equation:

$$\partial_t \rho = \mathrm{div}\Big(\rho \nabla (E'(\rho) + U[\rho])\Big),\ \rho|_{t=0} = \rho_0, \tag{2.1}$$

on $(0, +\infty) \times \mathbb{T}^d$ where $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$ denotes the flat torus (we take periodic boundary conditions to simplify the exposition, we refer to the work of the second author [11] for extensions to $\mathbb{R}^d$ or a bounded domain) which we identify with the unit cube $[0,1]^d$ equipped with the quotient distance:

$$d^2(x,y) := \inf_{k \in \mathbb{Z}^d} |x - y + k|^2.$$

Denoting by $\mathcal{P}(\mathbb{T}^d)$ the set of Borel probability measures on $\mathbb{T}^d$, we assume the following assumption on the map $\rho \in \mathcal{P}(\mathbb{T}^d) \mapsto U(\rho)$:

$$\forall \rho \in \mathcal{P}(\mathbb{T}^d), U[\rho] \in W^{2,\infty}(\mathbb{T}^d), \text{ and } U[\rho] \geq 0, \tag{2.2}$$

$$\sup_{\rho\in\mathcal{P}(\mathbb{T}^d)} \{\|\nabla U[\rho]\|_{L^\infty} + \|(\Delta U[\rho])_+\|_{L^\infty}\} < +\infty \tag{2.3}$$

and there exists a constant C such that for all $(\rho,\nu)\in\mathcal{P}(\mathbb{T}^d)\times\mathcal{P}(\mathbb{T}^d)$

$$\|\nabla U[\rho] - \nabla U[\nu]\|_{L^\infty(\mathbb{T}^d)} \leq C W_2(\rho,\nu), \tag{2.4}$$

with $W_2(\rho,\nu)$ denoting the 2-Wasserstein distance between ρ and ν i.e.

$$W_2(\rho,\nu) := \inf_{\gamma\in\Pi(\rho,\nu)} \Big\{ \int_{\mathbb{T}^d\times\mathbb{T}^d} d^2(x,y)\mathrm{d}\gamma(x,y)\Big\}^{\frac{1}{2}}$$

where $\Pi(\mu,\nu)$ is the set of transport plans between ρ and ν i.e. the set of Borel probability measures on $\mathbb{T}^d\times\mathbb{T}^d$ having μ and ν as marginals. It is well known, see [23, 24, 20], that W_2 metrizes the weak star topology on $\mathcal{P}(\mathbb{T}^d)$ so that $(\mathcal{P}(\mathbb{T}^d), W_2)$ is a compact metric space.

As for the nonlinear diffusion term $\mathrm{div}(\rho\nabla(E'(\rho)))$ it is convenient to rewrite it as:

$$\mathrm{div}(\rho\nabla(E'(\rho))) = \Delta F'(\rho)$$

where $F'(t) := tE'(t) - E(t)$ so that $F''(t) = tE''(t)$. The typical energies E we have in mind are the following classical examples

- Entropy: $E(t) := t\log(t)$ so that $F'(t) = t$, $F''(t) = 1$ (which thus gives a linear diffusion driven by the laplacian),
- Porous media $E(t) = t^m$ with $m > 1$ so that $F'(t) = (m-1)t^m$, $F(t) = \frac{m-1}{m+1}t^{m+1}$.

We shall assume that E is a continuous convex function on $\mathbb{R}_+$ with $E(0) = 0$, E is of class C^2 on $(0,+\infty)$ and that there are constants $C > 0$ and $m \geq 1$ such that

$$E''(t) \geq \frac{t^{m-2}}{C},\ F'(t) = tE'(t) - E(t) \leq C(1+t^m),\ \forall t\in(0,+\infty). \tag{2.5}$$

Of course, these assumptions are satisfied in the examples above corresponding respectively to linear diffusion and the porous medium equation.

Finally, we assume that the initial condition $\rho_0\in\mathcal{P}(\mathbb{T}^d)$ satisfies

$$\int_{\mathbb{T}^d} E(\rho_0(x))\ \mathrm{d}x < +\infty \tag{2.6}$$

which with (2.5) in particular implies that $\rho_0\in L^m(\mathbb{T}^d)$ and $F'(\rho_0)\in L^1(\mathbb{T}^d)$.

A weak solution of (2.1) then is a curve $t \in (0,+\infty) \mapsto \rho(t,.) \in \mathcal{P}(\mathbb{T}^d)$ such that $F'(\rho) \in L^1_{\rm loc}((0,+\infty)\times\mathbb{T}^d)$ and

$$\int_0^{+\infty}\int_{\mathbb{T}^d}(\partial_t\phi\rho + \Delta\phi F'(\rho) - \nabla U[\rho]\cdot\nabla\phi\rho)\,{\rm d}x\,{\rm d}t = -\int_{\mathbb{T}^d}\phi(0,x)\rho_0(x)\,{\rm d}x \tag{2.7}$$

for every $\phi \in C^2_c([0,+\infty)\times\mathbb{T}^d)$.

Theorem 2.1. *Assume (2.2)-(2.3)-(2.4)-(2.5)-(2.6), then (2.1) admits at least one weak solution.*

The complete proof of this result is given in section 3. This proof is strongly based on a semi-implicit version of the Jordan-Kinderlehrer-Otto (JKO) scheme [10] as in Di Francesco and Fagioli [8]. More precisely given a time step $h>0$, we construct inductively a sequence $\rho_h^k \in \mathcal{P}(\mathbb{T}^d)$ by setting $\rho_h^0 = \rho_0$ and, given ρ_h^k we select ρ_h^{k+1} as a solution of

$$\inf_{\rho\in\mathcal{P}(\mathbb{T}^d)}\left\{\frac{1}{2h}W_2^2(\rho,\rho_h^k) + \mathcal{E}(\rho) + \mathcal{U}(\rho|\rho_h^k)\right\} \tag{2.8}$$

where

$$\mathcal{E}(\rho) := \begin{cases}\int_{\mathbb{T}^d}E(\rho(x))\,{\rm d}x, & \text{if } E(\rho)\in L^1,\\ +\infty, & \text{otherwise,}\end{cases} \qquad \mathcal{U}(\rho|\nu) := \int_{\mathbb{T}^d}U[\nu]{\rm d}\rho. \tag{2.9}$$

Note that assumption (2.5) ensures that $\mathcal{E}$ controls from above $\int_{\mathbb{T}^d}\rho^m{\rm d}x$ if $m>1$ and $\int_{\mathbb{T}^d}\rho\log(\rho){\rm d}x$ if $m=1$, so in any case sublevels of $\mathcal{E}$ are weakly relatively compact in $L^m(\mathbb{T}^d)$.

By standard lower semicontinuity and compactness arguments, it is clear that (2.8) possesses solutions so that one can indeed generate a sequence $(\rho_h^k)_{k\in\mathbb{N}}$ by the semi-implicit JKO scheme (2.8). It is even uniquely defined (but we won't really need it in the sequel) because each ρ_h^k remains absolutely continuous with respect to Lebesgue's measure, so that $\rho\mapsto W_2^2(\rho,\rho_h^k)$ is in fact strictly convex (see Proposition 7.19 of [20]) and the other terms $\mathcal{E}$ and $\mathcal{U}(.|\rho_h^k)$ are convex. We finally extend in a piecewise constant way the sequence $(\rho_h^k)_{k\in\mathbb{N}}$ i.e. set:

$$\rho_h(t,.) := \rho_h^k \text{ for } t\in((k-1)h,kh] \text{ and } k\in\mathbb{N}. \tag{2.10}$$

The proof detailed in the next section consists in showing that as $h\to 0$, one may recover a limit ρ which satisfies (2.1). This is the same strategy as in [8] but the tricky part consists in passing to the limit in the nonlinear diffusion term $F'(\rho_h)$. This will be done thanks to the powerful flow interchange argument of Matthes, McCann and Savaré [13] in a similar way as in the work of Di Francesco and Matthes [9].

3 Existence proof

The proof of Theorem 2.1 is divided into three parts. The first two parts concern a priori estimates on ρ_h and the last part consists in showing that passing to the limit in the Euler-Lagrange equation of (2.8) actually enables us to recover a solution of (2.1). The discussion on uniqueness is deferred to the final section 6. Of course, it is enough to work on a fixed finite time interval $(0, T)$, which we shall implicitly do below, we thus also set $N := [\frac{T}{h}] + 1$. In what follows C (respectively C_T) is a generic (resp. only depending on T) constant whose value may vary from one line to another

3.1 Basic a priori estimates

From the very definition of the JKO semi-implicit scheme (2.8) we have for every k:

$$\frac{1}{2h} W_2^2(\rho_h^{k+1}, \rho_h^k) \leq \mathcal{E}(\rho_h^k) - \mathcal{E}(\rho_h^{k+1}) + \int_{\mathbb{T}^d} U[\rho_h^k] \mathrm{d}(\rho_h^k - \rho_h^{k+1}). \tag{3.1}$$

Recall then that the 1-Wasserstein distance W_1 is defined by:

$$W_1(\rho, \nu) := \inf_{\gamma \in \Pi(\rho,\nu)} \Big\{ \int_{\mathbb{T}^d \times \mathbb{T}^d} d(x, y) \mathrm{d}\gamma(x, y) \Big\},$$

so that by Cauchy-Schwarz's inequality

$$W_1(\rho, \nu) \leq W_2(\rho, \nu).$$

The well-known Kantorovich duality (see [23, 24, 20]) states that W_1 can be also be expressed as

$$W_1(\rho, \nu) = \sup \Big\{ \int_{\mathbb{T}^d} \phi \mathrm{d}(\rho - \nu),\ \phi \text{ 1-Lipschitz on } \mathbb{T}^d \Big\}$$

so that $\int_{\mathbb{T}^d} \phi \mathrm{d}(\rho - \nu)$ is less than the Lipschitz constant of ϕ times $W_1(\rho, \nu)$. Thanks to these considerations, assumption (2.3) and Young's inequality, we get

$$\int_{\mathbb{T}^d} U[\rho_h^k] \mathrm{d}(\rho_h^k - \rho_h^{k+1}) \leq C W_2(\rho_h^{k+1}, \rho_h^k) \leq \frac{1}{4h} W_2^2(\rho_h^{k+1}, \rho_h^k) + C^2 h. \tag{3.2}$$

Together with (3.1), this gives

$$\frac{1}{4h} W_2^2(\rho_h^{k+1}, \rho_h^k) \leq \mathcal{E}(\rho_h^k) - \mathcal{E}(\rho_h^{k+1}) + Ch \tag{3.3}$$

summing between 0 and N and using the fact that E is bounded from below and (2.6) gives

$$\frac{1}{4h}\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1},\rho_h^k) \leq \mathcal{E}(\rho_0) - \mathcal{E}(\rho_h^N) + CNh \leq C(1+T), \tag{3.4}$$

as well as

$$\mathcal{E}(\rho_h^k) \leq \mathcal{E}(\rho_0) + Ckh, \tag{3.5}$$

which, thanks to (2.6) and (2.5), also gives

$$\sup_{t\in[0,T]} \|\rho_h(t,.)\|_{L^m}^m \leq C(1+T) \text{ if } m>1, \tag{3.6}$$

and

$$\sup_{t\in[0,T]} \int_{\mathbb{T}^d} \rho_h(t,x)\log(\rho_h(t,x))\mathrm{d}x \leq C(1+T), \text{ if } m=1. \tag{3.7}$$

With (3.4), we also have the Hölder like estimate:

$$W_2(\rho_h(t,.),\rho_h(s,.)) \leq C\sqrt{(1+T)}\sqrt{|t-s|+h},\ \forall (s,t)\in[0,T]^2. \tag{3.8}$$

Using (3.8) and refined versions of Ascoli-Arzelà Theorem (see [1]) and (3.6)-(3.7), one deduces the existence of a vanishing sequence $h_n \to 0$ and of a $\rho \in C^{0,1/2}([0,T],(\mathcal{P}(\mathbb{T}^d),W_2)) \cap L^\infty((0,T),L^m(\mathbb{T}^d))$ such that

$$\rho_{h_n} \rightharpoonup \rho \text{ in } L^m((0,T)\times\mathbb{T}^d), \lim_n \sup_{t\in[0,T]} W_2(\rho_{h_n}(t,.),\rho(t,.)) = 0. \tag{3.9}$$

Now, using (2.4), we deduce that $\nabla U[\rho_{h_n}]$ converges to $\nabla U[\rho]$ in $L^\infty((0,T)\times\mathbb{T}^d)$ and then

$$\rho_{h_n}\nabla U[\rho_{h_n}] \rightharpoonup \rho\nabla U[\rho] \text{ in } L^m((0,T)\times\mathbb{T}^d). \tag{3.10}$$

3.2 Refined a priori estimates by flow interchange

This is the key step in the proof which enables us to obtain strong convergence, in what follows we essentially follow similar arguments as in Di Francesco and Matthes [9]. For $\nu\in\mathcal{P}(\mathbb{T}^d)$, set

$$\mathrm{Ent}(\nu) := \begin{cases} \int_{\mathbb{T}^d} \nu(x)\log(\nu(x))\mathrm{d}x, \text{ if } \nu\log(\nu)\in L^1, \\ +\infty, \text{ otherwise.} \end{cases}$$

For $\nu \in \mathcal{P}(\mathbb{T}^d)$ with $\mathrm{Ent}(\nu) < +\infty$ let us denote by $e^{t\Delta}\nu := \eta(t,.)$, the solution at time t of the heat equation:

$$\partial_t \eta = \Delta\eta,\ \eta\ |_{t=0}= \nu. \tag{3.11}$$

It is well-known since the seminal work of [10] that the heat flow can be viewed as the gradient flow of Ent for W_2 (see [1] for the theory of gradient flows in metric spaces). Moreover the fact that Ent is *displacement convex*[1], gives (see [1] Theorem 11.1.4, [6], [16]), the following evolution variational equality:

$$\frac{1}{2}\frac{d^+}{dt}W_2^2(e^{t\Delta}\nu,\mu) \leq \mathrm{Ent}(\mu) - \mathrm{Ent}(e^{t\Delta}\nu),\ \forall t \geq 0,\ \mu \in \mathcal{P}(\mathbb{T}^d) \tag{3.12}$$

where we have used the notation:

$$\frac{d^+}{dt}f(t) = \limsup_{s\to 0^+} \frac{f(t+s)-f(t)}{s}.$$

Taking $e^{t\Delta}\rho_h^{k+1}$ as a competitor in the minimization (2.8) gives

$$\begin{aligned} 0 \leq \frac{1}{2h}\frac{d^+}{dt}\Big(W_2^2(e^{t\Delta}\rho_h^{k+1},\rho_h^k)\Big)_{|_{t=0}} + \frac{d^+}{dt}\Big(\mathcal{E}(e^{t\Delta}\rho_h^{k+1})\Big)_{|_{t=0}} \\ + \frac{d^+}{dt}\Big(\mathcal{U}(e^{t\Delta}\rho_h^{k+1}|\rho_h^k)\Big)_{|_{t=0}}. \end{aligned} \tag{3.13}$$

Since $e^{t\Delta}\rho_h^{k+1}$ is smooth for $t>0$, we can directly compute:

$$\frac{d}{dt}(\mathcal{E}(e^{t\Delta}\rho_h^{k+1})) = -\int_{\mathbb{T}^d} E''(e^{t\Delta}\rho_h^{k+1})|\nabla(e^{t\Delta}\rho_h^{k+1})|^2\ \mathrm{dx}$$

which, with (2.5) gives that for some positive constant $\lambda > 0$

$$\frac{d}{dt}(\mathcal{E}(e^{t\Delta}\rho_h^{k+1})) \leq -\lambda\int_{\mathbb{T}^d}|\nabla((e^{t\Delta}\rho_h^{k+1})^{\frac{m}{2}})|^2.$$

[1]See section 6 for a precise definition. Here, we are working on $\mathbb{T}^d$, but we should not worry about it, it is just if we were working on $\mathbb{R}^d$ with periodic functions only. The optimal transport map between absolutely continuous periodic measures is well-known, it is given by the gradient of a convex function F such that $F(x) - \frac{|x|^2}{2}$ is periodic (see Cordero-Erausquin [5]) and which is characterized by a Monge-Ampère equation. Displacement convexity of the entropy on the $\mathbb{T}^d$ can therefore be proved as in the euclidean case. Another way to see this is to remark that Bochner's formula on $\mathbb{T}^d$ is just the same as in $\mathbb{R}^d$, this does not change the Ricci curvature and thus, thanks to a celebrated result of Lott and Villani [12] and Sturm [22] (see in particular the proof of Theorem 4.9), this does not change the displacement convexity of Ent (with respect to Lebesgue's measure).

We then have

$$-\frac{d^+}{dt}\Big(\mathcal{E}(e^{t\Delta}\rho_h^{k+1})\Big) \geq \lambda \liminf_{s\to 0^+} \int_0^1 \int_{\mathbb{T}^d} |\nabla((e^{ts\Delta}\rho_h^{k+1})^{\frac{m}{2}})|^2 \mathrm{d}x\mathrm{d}t. \tag{3.14}$$

In a similar way, for $t > 0$, we have

$$\frac{d}{dt}\Big(\mathcal{U}(e^{t\Delta}\rho_h^{k+1}|\rho_h^k)\Big) = \int_{\mathbb{T}^d} \Delta(U[\rho_h^k])e^{t\Delta}\rho_h^{k+1}$$

and the right hand side is uniformly bounded from above thanks to (2.3). With (3.12)-(3.13)-(3.14) this gives:

$$\lambda h \liminf_{s\to 0^+} \int_0^1 \int_{\mathbb{T}^d} |\nabla((e^{ts\Delta}\rho_h^{k+1})^{\frac{m}{2}})|^2 \mathrm{d}x\mathrm{d}t \leq (\mathrm{Ent}(\rho_h^k) - \mathrm{Ent}(\rho_h^{k+1})) + Ch.$$

Since $e^{s\Delta}\rho_h^{k+1}$ converges strongly to ρ_h^{k+1} in L^m as $s \to 0^+$, $(e^{s\Delta}\rho_h^{k+1})^{\frac{m}{2}}$ converges strongly to $(\rho_h^{k+1})^{\frac{m}{2}}$ in L^2. By lower semicontinuity we deduce that $\nabla(\rho_h^{k+1})^{\frac{m}{2}} \in L^2$ and

$$h \int_{\mathbb{T}^d} |\nabla(\rho_h^{k+1})^{\frac{m}{2}})|^2 \mathrm{d}x \leq C(\mathrm{Ent}(\rho_h^k) - \mathrm{Ent}(\rho_h^{k+1}) + h). \tag{3.15}$$

Summing from $k = 0$ to $N - 1$ and using the fact that $\mathrm{Ent}(\rho_0)$ is finite gives

$$\int_0^T \int_{\mathbb{T}^d} |\nabla(\rho_h)^{\frac{m}{2}})|^2 \mathrm{d}x\mathrm{d}t \leq CNh + C(\mathrm{Ent}(\rho_0) - \mathrm{Ent}(\rho_h^N)) \leq C(1+T)$$

which, with (3.6), also gives

$$\int_0^T \|\rho_h(t,.)^{\frac{m}{2}}\|^2_{H^1(\mathbb{T}^d)} \mathrm{d}t \leq C_T. \tag{3.16}$$

We then observe that since the injection of $H^1(\mathbb{T}^d)$ in $L^2(\mathbb{T}^d)$ is compact and since $\eta \mapsto \eta^{\frac{2}{m}}$ maps continuously $L^2(\mathbb{T}^d)$ into $L^m(\mathbb{T}^d)$, sublevel sets of $\rho \mapsto \|\rho_h(t,.)^{\frac{m}{2}}\|_{H^1(\mathbb{T}^d)}$ are strongly relatively compact in $L^m(\mathbb{T}^d)$. Now arguing as in Di Francesco and Matthes [9] i.e. using the refined version of Aubin-Lions Lemma provided by Theorem 2 of Rossi and Savaré [19] gives that the family $(\rho_h)_h$ is relatively compact in $L^m((0,T)\times\mathbb{T}^d)$. The conclusion of this step is that (3.9) can be strenghtened to

$$\rho_{h_n} \to \rho \text{ strongly in } L^m((0,T)\times\mathbb{T}^d). \tag{3.17}$$

Now, thanks to the second part of (2.5) and Krasnoselskii's Theorem (see [7], chapter 2) $\rho \mapsto F'(\rho)$ is continuous from L^m to L^1 and then (3.17) implies that

$$F'(\rho_{h_n}) \to F'(\rho) \text{ strongly in } L^1((0,T)\times\mathbb{T}^d). \tag{3.18}$$

3.3 Passing to the limit in the Euler-Lagrange equation

Now, we write the Euler-Lagrange equation for (2.8) as in the seminal work of Jordan, Kinderlehrer and Otto [10] for the Fokker-Planck equation and Otto [17] for nonlinear diffusions. Let ξ be a C^∞ vector field on $\mathbb{T}^d$ and denote by X_t the flow of ξ:

$$\frac{d}{dt}X_t(x) = \xi(X_t(x)),\ X_0(x) = x.$$

Define then $\nu_t := X_{t\#}\rho_h^{k+1}$ so that the change of variables formula gives

$$\rho_h^{k+1} = \nu_t(X_t)\det(DX_t). \tag{3.19}$$

Since ρ_h^{k+1} solves (2.8), we have

$$0 \le \frac{1}{2h}\frac{d^+}{dt}\Big(W_2^2(\nu_t,\rho_h^k)\Big)_{|t=0} + \frac{d^+}{dt}\Big(\mathcal{E}(\nu_t)\Big)_{|t=0} + \frac{d^+}{dt}\Big(\mathcal{U}(\nu_t|\rho_h^k)\Big)_{|t=0}. \tag{3.20}$$

Let then $\gamma_h^k \in \Pi(\rho_h^k,\rho_h^{k+1})$ be such that

$$W_2^2(\rho_h^{k+1},\rho_h^k) = \int_{\mathbb{T}^d\times\mathbb{T}^d} d^2(x,y)\mathrm{d}\gamma_h^k(x,y).$$

Choosing $k(x,y)\in\mathbb{Z}^d$ such that $d^2(x,y) = |y+k(x,y)-x|^2$, by definition of ν_t, we then have

$$W_2^2(\nu_t,\rho_h^k) \le \int_{\mathbb{T}^d\times\mathbb{T}^d} |X_t(y)+k(x,y)-x|^2\mathrm{d}\gamma_h^k(x,y)$$

and since $X_t(y) = y + t\xi(y) + o(t)$, we get

$$\frac{1}{2h}\frac{d^+}{dt}\Big(W_2^2(\nu_t,\rho_h^k)\Big)_{|t=0} \le \int_{\mathbb{T}^d\times\mathbb{T}^d}\xi(y)\cdot\Big(\frac{y-x}{h}\Big)\mathrm{d}\gamma_h^k(x,y). \tag{3.21}$$

As for the differentiation of $\mathcal{E}(\nu_t)$, following [10]-[17], using (3.19) we write

$$\mathcal{E}(\nu_t) = \int_{\mathbb{T}^d} E\Big(\frac{\rho_h^{k+1}}{\det(DX_t)}\Big)\det(DX_t)$$

observing that for $\rho\ge 0$ and $\alpha>0$

$$\frac{d}{d\alpha}(E\Big(\frac{\rho}{\alpha}\Big)\alpha) = -F'\Big(\frac{\rho}{\alpha}\Big)$$

then thanks to (2.5), (3.6), the fact that $\det(X_t) = 1 + t\mathrm{div}(\xi) + o(t)$ (with a uniform $o(t)$) and Lebesgue's dominated convergence Theorem, one obtains:

$$\frac{d^+}{dt}\Big(\mathcal{E}(\nu_t)\Big)_{|_{t=0}} = -\int_{\mathbb{T}^d} F'(\rho_h^{k+1})\mathrm{div}(\xi). \tag{3.22}$$

In a similar way,

$$\frac{d^+}{dt}\Big(\mathcal{U}(\nu_t)\Big)_{|_{t=0}} = \int_{\mathbb{T}^d} \nabla U[\rho_h^k]\cdot\xi\rho_h^{k+1}. \tag{3.23}$$

Combining (3.21)-(3.22)-(3.23) and (3.20), and applying the previous to both ξ and $-\xi$ gives the Euler-Lagrange equation

$$\int_{\mathbb{T}^d\times\mathbb{T}^d} \xi(y)\cdot(x-y)\mathrm{d}\gamma_h^k(x,y) = -h\int_{\mathbb{T}^d} F'(\rho_h^{k+1})\mathrm{div}(\xi) + h\int_{\mathbb{T}^d}\nabla U[\rho_h^k]\cdot\xi\rho_h^{k+1}. \tag{3.24}$$

for every smooth vector field ξ. Now let $\phi \in C_c^\infty([0,T)\times\mathbb{T}^d)$ (which we extend by $\phi(0,.)$ on $(-h,0)$), we then have

$$\begin{aligned}\int_0^T\int_{\mathbb{T}^d}\partial_t\phi\rho_h &= \sum_{k=0}^N\int_{\mathbb{T}^d}(\phi(kh,.)-\phi((k-1)h,.))\rho_h^k\\ &= \sum_{k=0}^N\int_{\mathbb{T}^d}\phi(kh,.)(\rho_h^k-\rho_h^{k+1}) - \int_{\mathbb{T}^d}\phi(0,.)\rho_0\\ &= \sum_{k=0}^N\int_{\mathbb{T}^d\times\mathbb{T}^d}(\phi(kh,x)-\phi(kh,y))\mathrm{d}\gamma_h^k(x,y) - \int_{\mathbb{T}^d}\phi(0,.)\rho_0.\end{aligned}$$

Using a second order Taylor-Lagrange formula gives

$$\int_{\mathbb{T}^d\times\mathbb{T}^d}(\phi(kh,x)-\phi(kh,y))\mathrm{d}\gamma_h^k(x,y) = \int_{\mathbb{T}^d\times\mathbb{T}^d}\nabla\phi(kh,y)\cdot(x-y)\mathrm{d}\gamma_h^k(x,y)+R_h^k$$

with

$$|R_h^k| \le \|D^2\phi\|_{L^\infty}W_2^2(\rho_h^{k+1},\rho_h^k).$$

With (3.3) this gives that $\sum_{k=0}^N|R_h^k| \le C_T h$, so that applying (3.24) to $\xi = \nabla\phi(kh,.)$, and using the fact that, with (2.4), Cauchy-Schwarz inequality and (3.4), $\sum_{k=0}^N\|\nabla U[\rho_h^k]-\nabla U[\rho_h^{k+1}]\|_{L^\infty}\le C_T$, we finally get

$$\begin{aligned}\int_0^T\int_{\mathbb{T}^d}\partial_t\phi\rho_h = &-\int_0^T\int_{\mathbb{T}^d}F'(\rho_h)\Delta\phi\\ &+\int_0^T\int_{\mathbb{T}^d}\nabla U[\rho_h]\cdot\nabla\phi\rho_h - \int_{\mathbb{T}^d}\phi(0,.)\rho_0+\delta_h\end{aligned}$$

where δ_h goes to zero as $h\to 0$. Thanks to (3.10), (3.17) and (3.18) we may pass to the limit on the vanishing sequence h_n to find that the limit ρ is a solution of (2.1). This completes the proof of Theorem 2.1.

4 Extension to systems

Let us now consider the extension of (2.1) to systems for the evolution of l densities $\rho := (\rho_1, \dots, \rho_l)$ of interacting species:

$$\partial_t \rho_i = \operatorname{div}(\rho_i \nabla (E_i'(\rho_i) + U_i[\rho])) = 0,\ \rho_i|_{t=0} = \rho_{i,0},\ i = 1, \dots, l, \tag{4.1}$$

on $(0, +\infty) \times \mathbb{T}^d$. Let us assume that for every $i = 1, \dots, l$ the map $\rho \in \mathcal{P}(\mathbb{T}^d)^l \mapsto U_i(\rho)$ fullfills

$$\forall \rho \in \mathcal{P}(\mathbb{T}^d)^l, U_i[\rho] \in W^{2,\infty}(\mathbb{T}^d),\ \text{and } U_i[\rho] \geq 0, \tag{4.2}$$

$$\sup_{\rho \in \mathcal{P}(\mathbb{T}^d)^l} \{\|\nabla U_i[\rho]\|_{L^\infty} + \|(\Delta U_i[\rho])_+\|_{L^\infty}\} < +\infty \tag{4.3}$$

and there exists a constant C such that for all $(\rho, \nu) = ((\rho_1, \dots, \rho_l), (\nu_1, \dots, \nu_l)) \in \mathcal{P}(\mathbb{T}^d)^l \times \mathcal{P}(\mathbb{T}^d)^l$ and every i

$$\|\nabla U_i[\rho] - \nabla U_i[\nu]\|_{L^\infty(\mathbb{T}^d)} \leq C \sum_{j=1}^{l} W_2(\rho_j, \nu_j). \tag{4.4}$$

As in the previous section, we assume that for each i, E_i is a continuous convex function on $\mathbb{R}_+$ with $E_i(0) = 0$, E_i is of class C^2 on $(0, +\infty)$ and that there are constants $C > 0$ and $m_i \geq 1$ such that

$$E_i''(t) \geq \frac{t^{m_i - 2}}{C},\ tE_i'(t) - E_i(t) \leq C(1 + t^{m_i}),\ \forall t \in (0, +\infty). \tag{4.5}$$

Finally we assume that the initial condition $\rho_0 \in \mathcal{P}(\mathbb{T}^d)$ satisfies

$$\sum_{i=1}^{l} \int_{\mathbb{T}^d} E_i(\rho_{i,0}(x))\ \mathrm{d}x < +\infty \tag{4.6}$$

which with (4.5) in particular implies that $\rho_{i,0} \in L^{m_i}(\mathbb{T}^d)$. The semi-implicit JKO scheme then takes the following form: given a time step $h > 0$, we construct inductively a sequence $\rho_h^k \in \mathcal{P}(\mathbb{T}^d)^l$ by setting $\rho_h^0 = \rho_0$ and, given $\rho_h^k \in \mathcal{P}(\mathbb{T}^d)^l$ we select ρ_h^{k+1} as a solution of

$$\inf_{\rho \in \mathcal{P}(\mathbb{T}^d)^l} \left\{ \frac{1}{2h} \sum_{i=1}^{l} W_2^2(\rho_i, \rho_{i,h}^k) + \mathcal{E}(\rho) + \mathcal{U}(\rho | \rho_h^k) \right\} \tag{4.7}$$

where

$$\mathcal{E}(\rho) := \begin{cases} \sum_{i=1}^{l} \int_{\mathbb{T}^d} E_i(\rho_i(x))\ \mathrm{d}x, & \text{if } E_i(\rho_i) \in L^1, \\ +\infty, & \text{otherwise,} \end{cases}$$

and

$$\mathcal{U}(\rho|\nu) := \sum_{i=1}^{l} \int_{\mathbb{T}^d} U_i[\nu] \mathrm{d}\rho_i.$$

Extending in a piecewise constant way the sequence $(\rho_h^k)_{k\in\mathbb{N}}$ defines the $\mathcal{P}(\mathbb{T}^d)^l$-valued curve:

$$\rho_h(t,.) := \rho_h^k \text{ for } t \in ((k-1)h, kh] \text{ and } k \in \mathbb{N}. \tag{4.8}$$

Arguing exactly as in the proof detailed in section 3, there is strong convergence in $\Pi_{i=1}^l L^{m_i}((0,T)\times\mathbb{T}^d)$ of a sequence ρ_{h_n} to some limit curve ρ and passing to the limit in the Euler-Lagrange for (4.7) exactly gives:

Theorem 4.1. *Assume (4.2)-(4.3)-(4.4)-(4.5)-(4.6), then (4.1) admits at least one weak solution.*

5 The non potential case

We are now interested in the case where the drift may not be a gradient. More precisely, we consider the following nonlinear diffusion equation:

$$\partial_t \rho - \mathrm{div}(\rho \nabla(E'(\rho))) + \mathrm{div}(\rho V[\rho]) = 0,\ \rho|_{t=0} = \rho_0, \tag{5.1}$$

on $(0,T)\times\mathbb{T}^d$. Denoting by $H^{-1}(\mathbb{T}^d)$ the dual of $H^1(\mathbb{T}^d)$, we assume the following regularity on the drift term $V[\rho]$:

$$\forall \rho \in L^2 \cap \mathcal{P}(\mathbb{T}^d), V[\rho] \in W^{1,\infty}, \sup_{\rho\in L^2\cap\mathcal{P}(\mathbb{T}^d)} \{\|V[\rho]\|_{L^\infty} + \|\nabla V[\rho]\|_{L^\infty}\} < +\infty \tag{5.2}$$

and for every $R>0$, there exists a modulus ω_R such that, for every $(\rho,\nu) \in (L^2(\mathbb{T}^d)\cap\mathcal{P}(\mathbb{T}^d))^2$ such that $\|\rho\|_{H^{-1}(\mathbb{T}^d)} \le R$ and $\|\nu\|_{H^{-1}(\mathbb{T}^d)} \le R$, one has

$$\|V[\rho] - V[\nu]\|_{L^2(\mathbb{T}^d)} \le \omega_R(\|\rho-\nu\|_{H^{-1}(\mathbb{T}^d)}). \tag{5.3}$$

Typical examples of velocity fields $\rho \mapsto V[\rho]$ that satisfy the above assumptions are those of the form $V[\rho](x) = \int_{\mathbb{T}^d} B(x,y)\rho(y)dy$ with B smooth enough (but not necessarily a gradient with respect to x).

As before, E is convex on $\mathbb{R}_+$ and we define $F'(t) := tE'(t) - E(t)$ so that $F''(t) = tE''(t)$. We make the following assumptions on F (which are satisfied for instance when $E(t) = t^m$ with $m>1$ or $E(t) = t\log(t)$):

$$F \in C^2(\mathbb{R}_+,\mathbb{R}), F(0) = F'(0) = 0,\ F \text{ is convex}, \tag{5.4}$$

$$F'' \text{ is nondecreasing, and for every } \rho > 0,\ F''(\rho) > 0 \tag{5.5}$$

and there is a constant $C > 0$ such that

$$F'(\rho) \leq C(1 + \rho^2 + F(\rho)),\ \forall \rho \in \mathbb{R}_+. \tag{5.6}$$

As for the initial condition ρ_0 we assume that it is a probability density such that

$$\rho_0 \in H^1(\mathbb{T}^d),\ F(\rho_0) \in L^1(\mathbb{T}^d),\ F'(\rho_0) \in H^1(\mathbb{T}^d). \tag{5.7}$$

A nonnegative weak solution of the PDE

$$\partial_t \rho - \Delta(F'(\rho)) + \mathrm{div}(\rho V[\rho]) = 0,\ \rho|_{t=0} = \rho_0. \tag{5.8}$$

is by definition a function $\rho \in L^2((0,T) \times \mathbb{T}^d, \mathbb{R}_+)$ such that

$$F'(\rho) \in L^2((0,T), H^1(\mathbb{T}^d))$$

and

$$\int_0^T \int_{\mathbb{T}^d} (-\partial_t \phi \rho + \nabla F'(\rho) \cdot \nabla \phi - \rho V[\rho] \cdot \nabla \phi)\, \mathrm{d}x\mathrm{d}t = \int_{\mathbb{T}^d} \phi(0,x)\rho_0(x)\mathrm{d}x \tag{5.9}$$

for every $\phi \in C^1([0,T] \times \mathbb{T}^d)$ such that $\phi(T,.) = 0$.

Before we proceed to the existence proof, we need some preliminary results. Let us first study the continuity of the drift term $\rho = \rho(t,x) \mapsto V[\rho(t,.)](x)$. It is easy to see that when (5.2) and (5.3) are satisfied and ρ^n converges strongly in $L^2((0,T) \times \mathbb{T}^d)$ (hence in $L^2((0,T), H^{-1}(\mathbb{T}^d))$ to some ρ then $V[\rho^n]$ converges to $V[\rho]$ in $L^2((0,T) \times \mathbb{T}^d)$, but we wil need a variant in the sequel:

Lemma 5.1. *Assume that (5.2) and (5.3) are satisfied. Let ρ^n be a sequence in $L^2((0,T) \times \mathbb{T}^d)$ such that $\partial_t \rho^n \in L^2((0,T), H^{-1}(\mathbb{T}^d))$ with*

$$\sup_n \|\partial_t \rho^n\|_{L^2((0,T),H^{-1}(\mathbb{T}^d))} < +\infty, \tag{5.10}$$

and $\rho \in L^2((0,T) \times \mathbb{T}^d)$ such that $\rho^n \rightharpoonup \rho$ in $L^2((0,T) \times \mathbb{T}^d)$, then $V[\rho^n]$ converges to $V[\rho]$ strongly in $L^2((0,T) \times \mathbb{T}^d)$.

Proof. First observe that (5.10) implies that for some constant C one has

$$\|\rho^n(t,.) - \rho^n(s,.)\|_{H^{-1}} \leq C\sqrt{|t-s|},\ \forall n,\ \forall (t,s) \in (0,T)^2. \tag{5.11}$$

Let $t \in (0,T)$ and for $h \in (0,t)$ define

$$\overline{\rho}^n_{t,h}(x) := \frac{1}{h}\int_{t-h}^t \rho^n(s,x)ds,\ \overline{\rho}_{t,h} := \frac{1}{h}\int_{t-h}^t \rho(s,x)ds$$

thanks to (5.11), we obtain, for every n, t and h:

$$\|\rho^n(t,.) - \overline{\rho}^n_{t,h}\|_{H^{-1}} \leq C\sqrt{h},\ \|\rho(t,.) - \overline{\rho}_{t,h}\|_{H^{-1}} \leq C\sqrt{h}. \tag{5.12}$$

For fixed $h > 0$, $\overline{\rho}^n_{t,h} \rightharpoonup \overline{\rho}_{t,h}$ in $L^2(\mathbb{T}^d)$ as $n \to \infty$, and since the imbedding of $L^2(\mathbb{T}^d)$ into $H^{-1}(\mathbb{T}^d)$ is compact we also have $\|\overline{\rho}^n_{t,h} - \overline{\rho}_{t,h}\|_{H^{-1}(\mathbb{T}^d)} \to 0$ as $n \to \infty$. We then get

$$\|\rho^n(t,.) - \rho(t,.)\|_{H^{-1}} \leq 2C\sqrt{h} + \|\overline{\rho}^n_{t,h} - \overline{\rho}_{t,h}\|_{H^{-1}(\mathbb{T}^d)}$$

from which we deduce that $\|\rho^n(t,.) - \rho(t,.)\|_{H^{-1}}$ tends to 0. Thanks to (5.3), this implies that $\|V[\rho^n(t,.)] - V[\rho(t,.)]\|_{L^2(\mathbb{T}^d)}$ tends to 0. The claimed L^2 convergence then follows from (5.2) and Lebesgue's dominated convergence Theorem.

□

We now introduce a regularized nonlinearity to approximate (5.8) by a uniformly parabolic equation as follows. Let $\varepsilon \in (0,1)$, let δ_ε and M_ε be respectively the smallest ρ for which $F''(\rho) \geq \varepsilon$ and the largest ρ for which $F''(\rho) \leq \varepsilon^{-1}$. Let then F_ε be defined by

$$F_\varepsilon(\rho) := \begin{cases} F(\delta_\varepsilon) + F'(\delta_\varepsilon)(\rho - \delta_\varepsilon) + \frac{\varepsilon}{2}(\rho - \delta_\varepsilon)^2 \text{ if } \rho \in [0, \delta_\varepsilon]; \\ F(\rho) \text{ if } \rho \in [\delta_\varepsilon, M_\varepsilon], \\ F(M_\varepsilon) + F'(M_\varepsilon)(\rho - M_\varepsilon) + \frac{1}{2\varepsilon}(\rho - M_\varepsilon)^2 \text{ if } \rho \geq M_\varepsilon. \end{cases} \tag{5.13}$$

Clearly, by construction F_ε is convex and C^2 on $\mathbb{R}_+$ with

$$\varepsilon \leq F''_\varepsilon \leq \frac{1}{\varepsilon} \text{ on } \mathbb{R}_+ \tag{5.14}$$

and F_ε converges pointwise to F since δ_ε and M_ε converge respectively to 0 and $+\infty$. In fact, this approximation also has good Γ-convergence properties:

Lemma 5.2. *Let $\theta \in L^2((0,T) \times \mathbb{T}^d, \mathbb{R}_+)$, then*

$$\lim_{\varepsilon \to 0^+} \int_0^T \int_{\mathbb{T}^d} F_\varepsilon(\theta(t,x))\,dx\,dt = \int_0^T \int_{\mathbb{T}^d} F(\theta(t,x))\,dx\,dt \tag{5.15}$$

moreover if $\theta_\varepsilon \in L^2((0,T) \times \mathbb{T}^d), \mathbb{R}_+)$ weakly converges to θ in $\in L^2((0,T) \times \mathbb{T}^d)$, then

$$\liminf_{\varepsilon \to 0^+} \int_0^T \int_{\mathbb{T}^d} F_\varepsilon(\theta_\varepsilon(t,x))\,dx\,dt \geq \int_0^T \int_{\mathbb{T}^d} F(\theta(t,x))\,dx\,dt \tag{5.16}$$

Proof. Fatou's lemma first yields

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d}F_\varepsilon(\theta(t,x))\mathrm{d}x\mathrm{d}t\geq\int_0^T\int_{\mathbb{T}^d}F(\theta(t,x))\mathrm{d}x\mathrm{d}t$$

on the other hand

$$\int_0^T\int_{\mathbb{T}^d}F_\varepsilon(\theta(t,x))\mathrm{d}x\mathrm{d}t\leq\int_0^T\int_{\mathbb{T}^d}F(\theta(t,x))\mathrm{d}x\mathrm{d}t+\int\int_{\{\theta\leq\delta_\varepsilon\}}(F_\varepsilon(\theta)-F(\theta))\mathrm{d}x\mathrm{d}t$$

since the second term in the right hand side converges to 0, we easily deduce (5.15). Let us now assume that $\theta_\varepsilon\in L^2((0,T)\times\mathbb{T}^d,\mathbb{R}_+)$ weakly converges to θ in $\in L^2((0,T)\times\mathbb{T}^d)$. Let $\gamma>0$ (fixed for the moment) and denote by F^γ the function defined by

$$F^\gamma(\rho)=\begin{cases}F(\rho)\text{ if }\rho\in[0,\gamma],\\ F(\gamma)+F'(\gamma)(\rho-\gamma)\text{ if }\rho\geq\gamma\end{cases}$$

by construction F^γ is convex and below F. For $\varepsilon>0$ small enough so that $\gamma\in[\delta_\varepsilon,M_\varepsilon]$, we similarly define

$$F^\gamma_\varepsilon(\rho)=\begin{cases}F_\varepsilon(\rho)\text{ if }\rho\in[0,\gamma],\\ F(\gamma)+F'(\gamma)(\rho-\gamma)\text{ if }\rho\geq\gamma\end{cases}$$

so that F^γ_ε is convex and coincides with F^γ on $[\delta_\varepsilon,+\infty)$. We then have

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d}F_\varepsilon(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t\geq\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d}F^\gamma_\varepsilon(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t$$
$$\geq\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d}F^\gamma(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t+\liminf_{\varepsilon\to 0^+}\int\int_{\{\theta_\varepsilon\leq\delta_\varepsilon\}}(F_\varepsilon(\theta_\varepsilon)-F(\theta_\varepsilon))$$

the second term converges to 0 whereas by weak lower semi-continuity (thanks to the convexity of F^γ) we have

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d}F^\gamma(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t\geq\int_0^T\int_{\mathbb{T}^d}F^\gamma(\theta(t,x))\mathrm{d}x\mathrm{d}t,$$

hence

$$\liminf_{\varepsilon\to 0^+}\int_0^T\int_{\mathbb{T}^d}F_\varepsilon(\theta_\varepsilon(t,x))\mathrm{d}x\mathrm{d}t\geq\sup_{\gamma>0}\int_0^T\int_{\mathbb{T}^d}F^\gamma(\theta(t,x))\mathrm{d}x\mathrm{d}t$$

and then (5.16) easily follows from the previous inequality, the fact that F^γ converges monotonically to F and Beppo-Levi's monotone convergence Theorem. □

Theorem 5.3. *Assume (5.2)-(5.3)-(5.4)-(5.5)-(5.6)-(5.7), then (5.8) admits at least one weak nonnegative solution.*

Proof. The proof proceeds in three steps.

Step 1: Regularized equation. We first prove existence of a weak solution to the regularized equation:

$$\partial_t \rho^\varepsilon - \Delta(F_\varepsilon'(\rho^\varepsilon)) + \mathrm{div}(\rho^\varepsilon V[\rho^\varepsilon]) = 0,\ \rho^\varepsilon|_{t=0} = \rho_0. \tag{5.17}$$

For fixed $\nu \in L^2((0,T)\times\mathbb{T}^d, \mathbb{R}_+)$ the linear parabolic equation:

$$\partial_t \rho^\varepsilon - \mathrm{div}(F_\varepsilon''(\nu)\nabla\rho^\varepsilon) + \mathrm{div}(\rho^\varepsilon V[\nu]) = 0,\ \rho^\varepsilon|_{t=0} = \rho_0 \tag{5.18}$$

admits a unique weak solution which we denote $\rho^\varepsilon := T^\varepsilon(\nu)$, moreover ρ^ε is nonnegative by the maximum principle and $\rho^\varepsilon \in L^2((0,T), H^1(\mathbb{T}^d)) \cap C([0,T], L^2(\mathbb{T}^d))$ and $\partial_t\rho \in L^2((0,T), H^{-1}(\mathbb{T}^d))$ and more precisely, thanks to (5.2) and (5.14) there is a constant C_ε such that:

$$\int_0^T \int_{\mathbb{T}^d} (|\nabla\rho^\varepsilon|^2 + (\rho^\varepsilon)^2)\mathrm{d}x\mathrm{d}t + \int_0^T \|\partial_t\rho^\varepsilon\|_{H^{-1}}^2 \le C_\varepsilon. \tag{5.19}$$

Thanks to (5.3), it is easy to check that T^ε is a continuous map of $L^2((0,T)\times\mathbb{T}^d, \mathbb{R}_+)$. In addition, (5.19) and the Aubin-Lions lemma (see [2], [21]) imply that $T^\varepsilon(L^2((0,T)\times\mathbb{T}^d, \mathbb{R}_+))$ is relatively compact in $L^2((0,T)\times\mathbb{T}^d)$. Schauder's fixed point Theorem then ensures that T^ε admits at least one fixed point ρ^ε i.e. a solution of (5.17).

Step 2: A priori estimates. We aim now to derive estimates independent of ε on ρ^ε. Let $\delta > 0$ such that $\delta \in (\delta_\varepsilon, M_\varepsilon)$, we then take $(\rho^\varepsilon - \delta)_+$ as test-function in (5.17) (which is actually licit since this test-function belongs to $L^2((0,T), H^1(\mathbb{T}^d)))$ integrating between 0 and $t \in [0,T]$ this yields

$$\int_0^t \langle \partial_t\rho^\varepsilon, (\rho^\varepsilon - \delta)_+\rangle_{H^{-1},H^1} ds + \int_0^t \int_{\{\rho^\varepsilon \ge \delta\}} F_\varepsilon''(\rho^\varepsilon)|\nabla\rho^\varepsilon|^2 = \int_0^t \int_{\{\rho^\varepsilon\ge\delta\}} \rho^\varepsilon V[\rho^\varepsilon]\cdot\nabla\rho^\varepsilon$$

hence, using Young's inequality

$$\begin{aligned}\frac{1}{2}\|(\rho^\varepsilon(t,.)-\delta)_+\|_{L^2}^2 - \frac{1}{2}\|(\rho_0-\delta)_+\|_{L^2}^2 + \int_0^t\int_{\{\rho^\varepsilon\ge\delta\}} F_\varepsilon''(\rho^\varepsilon)|\nabla\rho^\varepsilon|^2 \\ \le C\Big(\frac{\nu}{2}\int_0^t\int_{\{\rho^\varepsilon\ge\delta\}} |\nabla\rho^\varepsilon|^2 + \frac{1}{2\nu}\int_0^t\int_{\{\rho^\varepsilon\ge\delta\}} (\rho^\varepsilon)^2\Big) \\ \le C\Big(\frac{\nu}{2}\int_0^t\int_{\{\rho^\varepsilon\ge\delta\}} |\nabla\rho^\varepsilon|^2 + \frac{1}{\nu}\int_0^t\int_{\{\rho^\varepsilon\ge\delta\}} [(\rho^\varepsilon-\delta)_+^2 + \delta^2]\Big)\end{aligned}$$

since $F_\varepsilon''(\delta) = F''(\delta) > 0$ and F'' nondecreasing, we can choose ν small enough so that the first term in the right hand side is absorbed by the left hand side of the inequality. Gronwall's lemma then gives

$$\sup_{t\in(0,T)} \|\rho^\varepsilon(t,.)\|_{L^2} \leq C \tag{5.20}$$

for a constant C that does not depend on ε. Next we take $F_\varepsilon'(\rho^\varepsilon)$ as test-function which similarly gives:

$$\begin{aligned}\int_{\mathbb{T}^d} F_\varepsilon(\rho^\varepsilon(t,.)) - \int_{\mathbb{T}^d} F_\varepsilon(\rho_0) + \int_0^t \int_{\mathbb{T}^d} |\nabla F_\varepsilon'(\rho^\varepsilon)|^2 &= \int_0^t \int_{\mathbb{T}^d} \rho^\varepsilon V[\rho^\varepsilon]\cdot \nabla F_\varepsilon'(\rho^\varepsilon) \\ &\leq C\Big(\frac{\nu}{2}\int_0^t \int_{\mathbb{T}^d} |\nabla F_\varepsilon'(\rho^\varepsilon)|^2 + \frac{1}{2\nu}\int_0^t \int_{\mathbb{T}^d} (\rho^\varepsilon)^2\Big)\end{aligned}$$

using (5.20) and chosing ν small enough we thus get

$$\sup_{t\in[0,T]} \int_{\mathbb{T}^d} F_\varepsilon(\rho_\varepsilon(t,.)) + \int_0^T \int_{\mathbb{T}^d} |\nabla F_\varepsilon'(\rho^\varepsilon)|^2 \leq C \tag{5.21}$$

for a constant C not depending on ε. Next we use (5.6) and (5.20)-(5.21) to deduce that

$$\sup_{t\in[0,T]} \int_{\mathbb{T}^d} F_\varepsilon'(\rho^\varepsilon) \leq C \tag{5.22}$$

together with Poincaré-Wirtinger inequality, using again (5.21), this gives

$$\|F_\varepsilon'(\rho^\varepsilon)\|_{L^2((0,T),H^1(\mathbb{T}^d))} \leq C. \tag{5.23}$$

Step 3: Passing to the limit. Let us set

$$u^\varepsilon := F_\varepsilon'(\rho^\varepsilon),\ \sigma^\varepsilon := \nabla u^\varepsilon - \rho^\varepsilon V[\rho^\varepsilon] \tag{5.24}$$

so that (5.17) can be rewritten as

$$\partial_t \rho^\varepsilon = \Delta u^\varepsilon - \operatorname{div}(\rho^\varepsilon V[\rho^\varepsilon]) = \operatorname{div}(\sigma^\varepsilon),\ \rho^\varepsilon|_{t=0} = \rho_0. \tag{5.25}$$

We know from the previous step that

$$\|\rho^\varepsilon\|_{L^\infty((0,T),L^2(\mathbb{T}^d))} + \|\sigma^\varepsilon\|_{L^2((0,T),L^2(\mathbb{T}^d))} + \|u^\varepsilon\|_{L^2((0,T),H^1(\mathbb{T}^d))} \leq C \tag{5.26}$$

as well as

$$\|\partial_t \rho^\varepsilon\|_{L^2((0,T),H^{-1}(\mathbb{T}^d))} \leq C. \tag{5.27}$$

Passing to subsequences if necessary, we may therefore assume that

$$\rho^\varepsilon \rightharpoonup \rho \text{ in } L^2((0,T)\times\mathbb{T}^d),\ u^\varepsilon \rightharpoonup u \text{ in } L^2((0,T),H^1(\mathbb{T}^d)) \tag{5.28}$$

and thanks to Lemma 5.1 and (5.27), we have

$$\sigma^\varepsilon \rightharpoonup \sigma := \nabla u - \rho V[\rho] \text{ in } L^2((0,T)\times\mathbb{T}^d). \tag{5.29}$$

Obviously one then has:

$$\partial_t\rho = \mathrm{div}(\sigma) = \Delta u - \mathrm{div}(\rho V[\rho]),\ \rho|_{t=0} = \rho_0. \tag{5.30}$$

So to establish that ρ is a weak nonnegative solution of (5.8), it is enough to prove that $u = F'(\rho)$. Thanks to the convexity of F this amounts to prove that

$$\int_0^T\int_{\mathbb{T}^d} F(\theta(t,x))\mathrm{d}x dt \geq \int_0^T\int_{\mathbb{T}^d} F(\rho(t,x))\mathrm{d}x dt + \int_0^T\int_{\mathbb{T}^d} u(\theta-\rho)\mathrm{d}x\mathrm{d}t \tag{5.31}$$

for every $\theta \in L^2((0,T)\times\mathbb{T}^d,\mathbb{R}_+)$. By definition of u^ε we know that

$$\int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\theta(t,x))\mathrm{d}x dt \geq \int_0^T\int_{\mathbb{T}^d} F_\varepsilon(\rho^\varepsilon(t,x))\mathrm{d}x dt + \int_0^T\int_{\mathbb{T}^d} u^\varepsilon(\theta-\rho^\varepsilon)\mathrm{d}x\mathrm{d}t. \tag{5.32}$$

Let us prove that

$$\lim_\varepsilon \int_0^T\int_{\mathbb{T}^d} u^\varepsilon\rho^\varepsilon = \int_0^T\int_{\mathbb{T}^d} u\rho. \tag{5.33}$$

For that purpose, let ψ^ε be the potential defined by

$$-\Delta\psi^\varepsilon = \rho^\varepsilon,\ \int_{\mathbb{T}^d}\psi^\varepsilon = 0,\ \psi^\varepsilon \in H^1(\mathbb{T}^d). \tag{5.34}$$

Thanks to (5.20), we have $\psi^\varepsilon \in L^\infty((0,T),H^2(\mathbb{T}^d))$ with a bound indepen-dendent of ε:

$$\|\nabla\psi^\varepsilon\|_{L^\infty((0,T),H^1(\mathbb{T}^d))} \leq C. \tag{5.35}$$

As for the time derivative of $\nabla\psi^\varepsilon$ we observe that

$$-\Delta(\partial_t\psi^\varepsilon) = \partial_t\rho^\varepsilon = \mathrm{div}(\sigma^\varepsilon)$$

so that, thanks to (5.26), we have $\partial_t\nabla\psi^\varepsilon \in L^2((0,T)\times\mathbb{T}^d)$ and more precisely

$$\|\partial_t\nabla\psi^\varepsilon\|_{L^2((0,T)\times\mathbb{T}^d)} \leq \|\sigma^\varepsilon\|_{L^2((0,T)\times\mathbb{T}^d)} \leq C$$

this proves that $\nabla\psi^\varepsilon$ is bounded in $H^1((0,T)\times\mathbb{T}^d)$, hence converges in $L^2((0,T)\times\mathbb{T}^d)$, up to an extraction if necessary, to ψ given by

$$-\Delta\psi=\rho,\ \int_{\mathbb{T}^d}\psi=0,\ \psi\in H^1(\mathbb{T}^d). \tag{5.36}$$

Weak convergence of ∇u^ε and strong convergence of $\nabla\psi^\varepsilon$ in L^2 then give

$$\begin{aligned}\lim_\varepsilon\int_0^T\int_{\mathbb{T}^d}u^\varepsilon\rho^\varepsilon&=\lim_\varepsilon\int_0^T\int_{\mathbb{T}^d}\nabla u^\varepsilon\nabla\psi^\varepsilon\\&=\int_0^T\int_{\mathbb{T}^d}\nabla u\nabla\psi=\int_0^T\int_{\mathbb{T}^d}u\rho\end{aligned}$$

which establishes (5.33). Next, we use Lemma 5.2, letting ε tend to 0^+, using (5.33) we obtain inequality (5.31) which proves that $u=F'(\rho)$ and so ρ is a weak solution of (5.8), concluding the proof.

□

The previous arguments again clearly adapt to systems. More precisely, let us consider the system for the evolution of l densities $\rho:=(\rho_1,\dots,\rho_l)$:

$$\partial_t\rho_i-\Delta(F_i'(\rho_i))+\operatorname{div}(\rho_iV_i[\rho])=0,\ \rho_i|_{t=0}=\rho_{i,0} \tag{5.37}$$

on $(0,+\infty)\times\mathbb{T}^d$. Assuming that each function F_i satisfies (5.4)-(5.5)-(5.6), that the initial conditions are probability densities which satisfy

$$\rho_{i,0}\in H^1(\mathbb{T}^d),\ F_i(\rho_{i,0})\in L^1(\mathbb{T}^d),\ F_i'(\rho_{i,0})\in H^1(\mathbb{T}^d),\ \forall i=1,\dots,l, \tag{5.38}$$

and for every $i=1,\dots,l$, the maps V_i satisfy

$$\forall\rho\in L^2(\mathbb{T}^d)^l, V_i[\rho]\in W^{1,\infty}(\mathbb{T}^d),\quad\sup_{\rho\in L^2(\mathbb{T}^d)^l}\{\|V_i[\rho]\|_{L^\infty}+\|\nabla V_i[\rho]\|_{L^\infty}\}<+\infty \tag{5.39}$$

and for every $R>0$, there exists a modulus ω_R such that, for every $(\rho,\nu)\in L^2(\mathbb{T}^d)^l\times L^2(\mathbb{T}^d)^l$ such that $\|\rho\|_{H^{-1}(\mathbb{T}^d)^l}\le R$ and $\|\nu\|_{H^{-1}(\mathbb{T}^d)^l}\le R$, one has

$$\|V_i[\rho]-V_i[\nu]\|_{L^2(\mathbb{T}^d)}\le\omega_R\Big(\sum_{j=1}^l\|\rho_j-\nu_j\|_{H^{-1}(\mathbb{T}^d)}\Big). \tag{5.40}$$

A direct adaptation of the proof of Theorem 5.3 gives

Theorem 5.4. *Assume that each function F_i satisfies (5.4)-(5.5)-(5.6), and that (5.38)-(5.39)-(5.40) are satisfied for $i=1,\dots,l$, then (5.37) admits at least one weak solution $(\rho_1,\dots,\rho_l)$ with each ρ_i nonnegative.*

6 Uniqueness

We end the paper by a general uniqueness argument based on geodesic convexity. For the purpose of our paper it is enough to consider an internal energy that is a functional defined on the subset $\mathcal{P}_{\rm ac}(\mathbb{T}^d)$ of $\mathcal{P}(\mathbb{T}^d)$ (consisting of absolutely continuous with respect to the Lebesgue measure on $\mathbb{T}^d$ elements of $\mathcal{P}(\mathbb{T}^d)$) and given by

$$\mathcal{E}(\rho) := \int_{\mathbb{T}^d} E(\rho(x))\ \mathrm{d}x,\ \rho \in \mathcal{P}_{\rm ac}(\mathbb{T}^d),$$

where $E : \mathbb{R}_+ \to \mathbb{R}$ is a convex function with $E(0) = 0$ and E smooth on $(0, +\infty)$. Given ρ and ν in $\mathcal{P}_{\rm ac}(\mathbb{T}^d)$, there is a unique optimal transport map T between ρ and ν (see Cordero-Erausquin [5] and McCann [14] for the case of a general Riemannian manifold) i.e. a map such that $T_{\#}\rho = \nu$ and

$$W_2^2(\rho, \nu) = \int_{\mathbb{T}^d} d^2(T(x), x)\rho(x)\mathrm{d}x.$$

Moreover, this map is given by the gradient of a convex function $T(x) = \nabla\phi(x)$ with ϕ convex on $\mathbb{R}^d$ and such that $x \mapsto \nabla\phi(x) - x$ is periodic on $\mathbb{R}^d$ (so that T indeed defines a map from $\mathbb{T}^d$ to itself). In fact, on $\mathbb{T}^d$, we should rather write

$$T(x) = \exp_x(-\nabla\psi(x)) \text{ where } \psi(x) = \frac{1}{2}|x|^2 - \phi(x).$$

Note in particular that $d(x, T(x)) = |\nabla\psi(x)|$ and then

$$W_2(\rho, \nu) = \int_{\mathbb{T}^d} |\nabla\psi(x)|^2\rho(x)\mathrm{d}x.$$

The Wasserstein geodesic between ρ and ν is then the curve $t \in [0, 1] \mapsto \nu_t$ given by the McCann's interpolation:

$$\nu_t := T_t\#\rho,\ T_t(x) = \exp_x(-t\nabla\psi(x)),$$

it is indeed a constant speed geodesic:

$$W_2(\nu_t, \nu_s) = |t - s|W_2(\rho, \nu)$$

and T_t is the optimal transport map from ρ to ν_t. Then the internal energy $\mathcal{E}$ is said to be *displacement convex* whenever for every ρ and ν in $\mathcal{P}_{\rm ac}(\mathbb{T}^d)$, defining ν_t as above one has

$$t \in [0, 1] \mapsto \mathcal{E}(\nu_t) \text{ is convex.}$$

If E satisfies McCann's condition:

$$r \in (0, +\infty) \mapsto r^d E(r^{-d}) \text{ is convex nonincreasing} \tag{6.1}$$

then it is well-known that $\mathcal{E}$ is displacement convex (see McCann [15]) and that for ρ, ψ and ν as above, one has (setting $F'(\rho) = \rho E'(\rho) - E(\rho)$):

$$\begin{aligned}
\mathcal{E}(\nu) - \mathcal{E}(\rho) &\geq \lim_{t\to 0^+} \int_{\mathbb{T}^d} \frac{E(\nu_t(x)) - E(\rho(x))}{t} \mathrm{d}x \\
&= \lim_{t\to 0^+} \int_{\mathbb{T}^d} \frac{1}{t} (E\Big(\frac{\rho(x)}{\det(DT_t(x))}\Big) \det(DT_t(x)) - E(\rho(x))) \mathrm{d}x \\
&= -\int_{\mathbb{T}^d} \nabla F'(\rho(x)) \cdot \nabla \psi \ \mathrm{d}x = -\int_{\mathbb{T}^d} \nabla E'(\rho(x)) \cdot \nabla \psi \rho(x) \ \mathrm{d}x
\end{aligned}$$

as soon $\nabla E'(\rho)$ in $L^2(\rho)$. Similarly, letting S ($S(y) = y - \nabla\theta(y)$) be the optimal map from ν to ρ and using the fact that $S(T(x)) = x$ i.e. $\nabla\theta(T(x)) = T(x) - x = -\nabla\psi(x)$, we get

$$\begin{aligned}
\mathcal{E}(\rho) - \mathcal{E}(\nu) &\geq -\int_{\mathbb{T}^d} \nabla E'(\nu(y)) \cdot \nabla\theta(y) \nu(y) \ \mathrm{d}y \\
&= -\int_{\mathbb{T}^d} \nabla E'(\nu(T(x)) \cdot \nabla\theta(T(x)) \rho(x) \ \mathrm{d}x \\
&= \int_{\mathbb{T}^d} \nabla E'(\nu(T(x)) \cdot \nabla\psi(x) \rho(x) \ \mathrm{d}x
\end{aligned}$$

so that summing the two inequalities above gives

$$0 \geq \int_{\mathbb{T}^d} (\nabla E'(\nu(T(x)) - \nabla E'(\rho(x)) \cdot \nabla\psi(x) \rho(x) \ \mathrm{d}x. \tag{6.2}$$

Now let us consider the system for the evolution of l densities $\rho = (\rho_1, \ldots, \rho_l)$:

$$\partial_t \rho_i = \mathrm{div}(\rho_i(\nabla E_i'(\rho_i) + V_i[\rho])), \tag{6.3}$$

on $(0, +\infty) \times \mathbb{T}^d$. Let us assume that each E_i is a convex function as above, that V_i maps $\mathcal{P}(\mathbb{T}^d)^l$ into $W^{1,\infty}(\mathbb{T}^d)$ and that for some constant C one has

$$|V_i[\rho](x) - V_i[\rho](y)| \leq C d(x, y), \ \forall \rho \in \mathcal{P}(\mathbb{T}^d)^l \text{ and } (x, y) \in \mathbb{T}^d \times T^d \tag{6.4}$$

and

$$\|\nabla V_i[\rho] - \nabla V_i[\nu]\|_{L^\infty(\mathbb{T}^d)} \leq C \sum_{j=1}^{l} W_2(\rho_j, \nu_j), \ \forall (\rho, \nu) \in \mathcal{P}(\mathbb{T}^d)^l \times \mathcal{P}(\mathbb{T}^d)^l. \tag{6.5}$$

Then the following uniqueness result holds:

Theorem 6.1. *Assume that the drifts V_i's satisfy (6.4)-(6.5) and that the E_i's satisfy McCann's condition (6.1). Let ρ and ν be two solutions on $(0,T)\times\mathbb{T}^d$ of (6.3) such that*

$$\int_0^T \sum_{i=1}^{l} \|v_{i,t}\|_{L^2(\rho_{i,t})} dt + \int_0^T \sum_{i=1}^{l} \|w_{i,t}\|_{L^2(\nu_{i,t})} dt < +\infty \tag{6.6}$$

with

$$v_{i,t} = \nabla E_i'(\rho_{i,t}) + V_i[\rho_t],\ w_{i,t} = \nabla E_i'(\nu_{i,t}) + V_i[\nu_t] \tag{6.7}$$

then for every $t\in[0,T]$ one has the following stability estimate

$$\sum_{i=1}^{l} W_2^2(\rho_{i,t},\nu_{i,t}) \le e^{4Ct} \sum_{i=1}^{l} W_2^2(\rho_{i,0},\nu_{i,0}), \tag{6.8}$$

which in particular gives uniqueness for the Cauchy problem for (6.3).

Proof. Using Theorem 5.24 and Corollary 5.25 from [20], assumption (6.6) guarantees that $t\mapsto W_2^2(\rho_{i,t},\nu_{i,t})$ is differentiable for a.e $t\in(0,T)$ with

$$\frac{d}{dt}\frac{1}{2}W_2^2(\rho_{i,t},\nu_{i,t}) = \int_{\mathbb{T}^d} \nabla\psi_{i,t}(x)\cdot(w_{i,t}(T_{i,t}(x)) - v_{i,t}(x))\rho_{i,t}(x)\mathrm{d}x$$

where $\nabla\psi_{i,t}$ is such that $T_{i,t} := \exp_x(-\nabla\psi_{i,t})$ is the optimal map between $\rho_{i,t}$ and $\nu_{i,t}$. Now thanks to McCann's condition and recalling (6.2), we have

$$\int_{\mathbb{T}^d} \nabla\psi_{i,t}(x)\cdot(\nabla E_i'(\nu_{i,t}(T_{i,t}(x))) - \nabla E_i'(\rho_{i,t})(x)))\rho_{i,t}(x)\mathrm{d}x \le 0,$$

so that, using (6.4)-(6.5) yields

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}W_2^2(\rho_{i,t},\nu_{i,t}) &\le \int_{\mathbb{T}^d} \nabla\psi_{i,t}(x)\cdot(V_i[\nu_t](T_{i,t}(x)) - V_i[\rho_t](x))\rho_{i,t}(x)\mathrm{d}x\\
&= \int_{\mathbb{T}^d} \nabla\psi_{i,t}(x)\cdot(V_i[\nu_t](T_{i,t}(x)) - V_i[\nu_t](x))\rho_{i,t}(x)\mathrm{d}x\\
&+ \int_{\mathbb{T}^d} \nabla\psi_{i,t}(x)\cdot(V_i[\nu_t](x) - V_i[\rho_t](x))\rho_{i,t}(x)\mathrm{d}x\\
&\le C\int_{\mathbb{T}^d} |\nabla\psi_{i,t}(x)| d(x,T_{i,t}(x))\rho_{i,t}(x)\mathrm{d}x\\
&+ C\Big(\int_{\mathbb{T}^d} |\nabla\psi_{i,t}(x)|^2\rho_{i,t}(x)\mathrm{d}x\Big)^{\frac{1}{2}} \sum_{j=1}^{l} W_2(\rho_{j,t},\nu_{j,t})\\
&= C\left(W_2^2(\rho_{i,t},\nu_{i,t}) + W_2(\rho_{i,t},\nu_{i,t})\sum_{j=1}^{l} W_2(\rho_{j,t},\nu_{j,t})\right).
\end{aligned}$$

Summing over i and using Gronwall's lemma gives the desired inequality (6.8).

□

Remark 6.2. In the uniqueness result stated in Theorem 6.1, the integrability condition (6.6) is made as an assumption since it ensures absolute continuity for the Wasserstein distance of the curves ρ and ν. In the potential case where $V_i = \nabla U_i$, under the assumptions of Theorem 4.1, it can be checked (see [11]) that solutions constructed by the semi-implicit JKO scheme actually satisfy (6.6).

Acknowledgements: G.C. gratefully acknowledges the hospitality of the Mathematics and Statistics Department at UVIC (Victoria, Canada), and the support from the CNRS, from the ANR, through the project ISOTACE (ANR-12- MONU-0013) and from INRIA through the *action exploratoire* MOKAPLAN.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probabiliy measures*, Lectures Math. ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[2] J.-P. Aubin, *Un théorème de compacité*, CRAS **256**, 5042–5044 (1963).

[3] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent, and D. Slepcev, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J. **156(2)**, 229–271 (2011).

[4] J. A. Carrillo, R. J. McCann, and C. Villani, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal. **179(2)**, 217-263 (2006).

[5] D. Cordero-Erausquin, *Sur le transport de mesures périodiques*, C.R. Acad. Sci. Paris Sér. I Math. **329**, 199–202 (1999).

[6] S. Daneri, G. Savaré, *Eulerian calculus for the displacement convexity in the Wasserstein distance*, SIAM J. Math. Anal. **40**, no. 3, 1104–1122 (2008).

[7] D.G. de Figueiredo, *Lectures on the Ekeland variational principle with applications and detours*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, 81. Published for the Tata Institute of Fundamental Research, Bombay; by Springer-Verlag, Berlin, (1989).

[8] M. Di Francesco, S. Fagioli, *Measure solutions for nonlocal interaction PDEs with two species*, Nonlinearity **26**, 2777—2808 (2013).

[9] M. Di Francesco, D. Matthes, *Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations*, Calc. Var. Partial Differential Equations **50**, no. 1-2, 199–230 (2014).

[10] R. Jordan, D. Kinderlehrer, F. Otto, *The Variational Formulation of the Fokker-Plank Equation*, SIAM J. of Math. Anal. **29**, 1-17 (1998).

[11] M. Laborde, *On some non linear evolution systems which are perturbations of Wasserstein gradient flows*, preprint 2015.

[12] J. Lott, C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Annals of Mathematics **169**, 903–991 (2009).

[13] D. Matthes, R.J. McCann, G. Savaré, *A family of nonlinear fourth order equations of gradient flow type*, Comm. Partial Differential Equations **34**, no. 10-12, 1352-1397 (2009).

[14] R.-J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. **11**, 589-608 (2001).

[15] R.-J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128**, 153-179 (1997).

[16] F. Otto, M. Westdickenberg, *Eulerian calculus for the contraction in the Wasserstein distance*, SIAM J. Math. Anal. **37**, no. 4, 1227–1255 (2005).

[17] F. Otto, *Doubly degenerate diffusion equations as steepest descent*, Manuscript (1996).

[18] F. Otto, *The geometry of dissipative evolution equations: the porous-medium equation*, Comm. Partial Differential Equations **26**, 101-174 (2001).

[19] R. Rossi, G. Savaré, *Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces*, Ann. Sc. Norm. Super. Pisa Cl. Sci. **5(2)**, 395–431 (2003).

[20] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Birkhäuser Verlag, Basel, 2015.

[21] J. Simon, *Compact sets in the space $L^p(0,T;B)$*, Ann. Mat. Pura Appl. **4**, 146, 65–96 (1987).

[22] K.-T. Sturm, *On the geometry of metric measure spaces*, Acta Math. **196**, 65–131 (2006).

[23] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, 2003.

[24] C. Villani, *Optimal transport: Old and New*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Heidelberg, 2009.