A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts

Guillaume Carlier *, Maxime Laborde †

November 3, 2016

Abstract

We prove an existence result for nonlinear diffusion equations in the presence of a nonlocal density-dependent drift which is not necessarily potential. The proof is constructive and based on the Helmholtz decomposition of the drift and a splitting scheme. The splitting scheme combines transport steps by the divergence-free part of the drift and semi-implicit minimization steps à la Jordan-Kinderlherer Otto to deal with the potential part.

Keywords: Wasserstein gradient flows, Jordan-Kinderlehrer-Otto scheme, splitting, nonlocal drift, nonlinear diffusions, Helmholtz decomposition.

MS Classification: 35K15, 35K40.

1 Introduction

In their seminal paper [7], Jordan-Kinderlehrer and Otto, showed that the Fokker-Planck equation

$$\partial_t \rho - \Delta\rho - \operatorname{div}(\rho \nabla V) = 0, \qquad \rho_{|t=0} = \rho_0$$

with a probability density ρ_0 as Cauchy datum can be viewed as the gradient flow for the Wasserstein metric of the relative entropy with respect to the equilibrium measure $\rho_\infty := e^{-V}$, $E(\rho|\rho_\infty) := \int \rho \log(\rho/\rho_\infty)$. They also introduced an implicit Euler-scheme (nowadays referred to as the JKO scheme): given a time step $h > 0$ and starting from $\rho_h^0 = \rho_0$, construct inductively a sequence ρ_h^k by:

$$\rho_h^{k+1} \in \operatorname{argmin}\left\{ \frac{1}{2h} W_2^2(\rho, \rho_h^k) + E(\rho|\rho_\infty) \right\}$$

where W_2 denotes the 2-Wasserstein distance (see section 2) and proved the convergence of its piecewise constant interpolation to the solution of the Fokker-Planck equation as $h \to 0$. The theory of Wasserstein gradient flows has developed rapidly in the last twenty years with many applications for instance to porous medium equations [14] or aggregation equations [6]. The reference textbook of Ambrosio, Gigli and Savaré [2] gives a very detailed account of this powerful theory, which enables to study general nonlinear diffusion equations of the form

$$\partial_t \rho - \Delta P(\rho) - \operatorname{div}(\rho \nabla V) = 0 \text{ where } P(\rho) := \rho F'(\rho) - F(\rho),$$

as Wasserstein gradient flows for the energy $\int (F(\rho) + V\rho)$.

The purpose of the present paper is to present a splitting transport-JKO scheme to study nonlinear diffusion equations (or more generally, systems) with a general density-dependent drift:

$$\partial_t \rho(t,x) - \Delta P(\rho(t,x)) - \operatorname{div}(\rho(t,x) U[\rho(t,.)](x)) = 0, \ t \geq 0, \ x \in \Omega,$$
$$(\nabla P(\rho) + \rho U[\rho]) \cdot \nu = 0 \text{ on } \partial\Omega, \ \rho_{|t=0} = \rho_0,$$

*Université Paris-Dauphine, PSL Research University, CNRS, CEREMADE, 75016, Paris, France and and MOKAPLAN, Inria-Paris, carlier@ceremade.dauphine.fr

†Université Paris-Dauphine, PSL Research University, CNRS, CEREMADE, 75016, Paris, France and MOKAPLAN, Inria-Paris laborde@ceremade.dauphine.fr.

where for every probability density ρ, $U[\rho]$ is a-not necessarily potential- vector field, for instance, in even dimensions, it can mix a gradient and Hamiltonian structures i.e. be of the form $\nabla V[\rho] + J\nabla H[\rho]$ (where J is the usual symplectic matrix). The potential case where $U[\rho] = \nabla V[\rho]$ can be studied by means of a semi-implicit JKO scheme introduced by Di Francesco and Fagioli [8] in the nondiffusive case and further developed by Laborde [11] for the case of a non linear diffusion. The idea of our splitting scheme is natural and consists in performing a Helmholtz decomposition of $U[\rho]$. We then treat the divergence-free part purely by (continuous in time) transport and the potential part by the semi-implicit JKO scheme. For the transport steps of the splitting scheme, we essentially need the divergence-free part to have some Sobolev regularity in x so as to be able to apply DiPerna-Lions theory, we will also need both the potential and divergence-free part of $\rho \mapsto U[\rho]$ to satisfy some Lipschitz continuity condition with respect to the Wasserstein distance. In our recent work [5], we studied the same type of equations or systems (in the periodic in space setting) by quite different arguments (approximation by uniformly parabolic equations). One advantage of the constructive splitting method presented here is that the transport steps by a divergence-free vector field preserve the internal energy, this is one way to overcome some difficulties discussed in [13] (section 5, variant 3).

The paper is organized as follows. Section 2 recalls some results from optimal transport and DiPerna Lions theory. Section 3 lists the various assumptions, explains the splitting scheme and gives the main result. Section 4 gives estimates on the discrete sequences of measures obtained by the splitting scheme. Convergence of the scheme as the time step goes to 0 to a solution of the PDE is proved in section 5. In the concluding section 6, we briefly discuss extension to systems and uniqueness issues.

2 Preliminaries

2.1 Wasserstein space

We recall some results from optimal transport theory that we will use in the sequel, we refer the reader to the textbooks of Villani [17, 18], Ambrosio, Gigli and Savaré [2] or Santambrogio [16] for a detailed exposition. Let Ω be an open subset of $\mathbb{R}^n$, we denote by $\mathcal{P}(\Omega)$ the set of Borel probability measures on Ω, $\mathcal{P}_2(\Omega) := \{\rho \in \mathcal{P}(\Omega) \ : \ M(\rho) := \int_\Omega |x|^2 d\rho(x) < +\infty\}$ and $\mathcal{P}_2^{\mathrm{ac}}(\Omega)$ the elements of $\mathcal{P}_2(\Omega)$ which in addition are absolutely continous with respect to the n-dimensional Lebesgue measure. Given ρ_0 and ρ_1 in $\mathcal{P}_2(\mathbb{R}^n)$ and denoting by $\Pi(\rho_0, \rho_1)$ the set of transport plans between ρ_0 and ρ_1 i.e. the set of Borel probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ having ρ_0 and ρ_1 as marginals, the 2-Wasserstein distance between ρ_0 and ρ_1, $W_2(\rho_0, \rho_1)$ is defined as the value of the optimal transport problem

$$W_2^2(\rho_0, \rho_1) = \inf_{\gamma \in \Pi(\rho_0,\rho_1)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 d\gamma(x, y).$$

Since this is a linear programming problem, it admits a dual formulation, which reads

$$\frac{1}{2} W_2^2(\rho_0, \rho_1) = \sup \Big\{ \int_{\mathbb{R}^n} \varphi d\rho_0 + \int_{\mathbb{R}^n} \psi d\rho_1 \ : \ \varphi(x) + \psi(y) \le \frac{1}{2}|x - y|^2 \Big\}.$$

Optimal potentials in the problem above are called Kantorovich potentials between ρ_0 and ρ_1, they can be taken semi-concave and their existence is well-known (see [17, 18, 2, 16]). If $\rho_0 \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ a celebrated result of Brenier [3] states that there is a unique optimal plan γ between ρ_0 and ρ_1 and it is induced by a transport map T i.e. is of the form $\gamma = (\mathrm{id}, T)_\# \rho_0$ and optimality is characterized by the fact that $T = \nabla u$ with u convex, moreover $T(x) = x - \nabla\varphi(x)$ where φ is the Kantorovich potential between ρ_0 and ρ_1. This in particular gives

$$W_2^2(\rho_0, \rho_1) = \int_{\mathbb{R}^n} |\nabla\varphi(x)|^2 \rho_0(x) dx.$$

Another important result is the Benamou-Brenier formula [4] which gives a dynamic formulation of $W_2^2(\rho_0, \rho_1)$ and expresses it as the infimum of the kinetic energy:

$$\int_0^1 \int_{\mathbb{R}^n} |v_t(x)|^2 d\rho_t(x) dt$$

among solutions of the continuity equation

$$\partial_t \rho + \mathrm{div}(\rho v) = 0, \qquad \rho_{|t=0} = \rho_0, \qquad \rho_{|t=1} = \rho_1.$$

2.2 Flows of weakly differentiable vector fields

We will also need to apply the DiPerna Lions theory [10] in the special case of divergence-free vector fields. Let $W \in W^{1,1}_{\rm loc}(\mathbb{R}^n, \mathbb{R}^n)$ be divergence-free $\mathrm{div}(W) = 0$ and with at most linear growth

$$|W(x)| \leqslant C(1 + |x|).$$

Then there exists a unique flow map $X : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$, $X \in C(\mathbb{R}_+, L^1_{\rm loc}(\mathbb{R}^n))$ such that

- $X(0, .) = \mathrm{id}$ and for a.e. x, $t \in \mathbb{R}_+ \mapsto X(., x)$ is a solution of the ODE $\dot{X} = W(X)$ i.e.:

$$X(t, x) = x + \int_0^t W(X(s, x))ds,\ t \geq 0,$$

- the map $t \mapsto X(t, x)$ satisfies the group property: $X(t, X(s, x)) = X(t + s, x)$ for a.e. x and every $t, s \geq 0$,
- for every $t \geqslant 0$, $X(t, .)$ preserves the n-dimensional Lebesgue measure.

Moreover given $\rho_0 \in \mathcal{P}^{\rm ac}_2(\mathbb{R}^n)$, $\rho(t, .) := X(t, .)_\# \rho_0 = \rho_0(X(t, .)^{-1})$ is the unique weak solution of the continuity equation

$$\partial_t \rho + \mathrm{div}(\rho W) = 0, \qquad \rho_{|t=0} = \rho_0, \tag{2.1}$$

which, since W is divergence-free, can also be rewritten as the transport equation $\partial_t \rho + \nabla\rho \cdot W = 0$. If we are given an open subset Ω of $\mathbb{R}^n$ with a smooth boundary and assume that the vector field W is tangential to $\partial\Omega$ (in the sense of traces), then the DiPerna-Lions flow X leaves Ω invariant so that if $\rho_0 \in \mathcal{P}^{\rm ac}_2(\Omega)$ (extended by 0 outside Ω, say), the solution $\rho_t = X(t, .)_\# \rho_0$ of (2.1) remains supported in Ω hence may be viewed as a curve with values in $\mathcal{P}^{\rm ac}_2(\Omega)$.

3 Assumptions and main result

Given a suitable convex nonlinearity F and its associated pressure $P(\rho) = \rho F'(\rho) - F(\rho)$ as well as a nonlocal drift $\rho \mapsto U[\rho]$, our goal is to solve

$$\partial_t \rho - \Delta P(\rho) - \mathrm{div}(\rho U[\rho]) = 0, \qquad \rho_{|t=0} = \rho_0, \tag{3.1}$$

on $(0, +\infty) \times \Omega$, where Ω is a smooth domain of $\mathbb{R}^n$ (not necessary bounded), in case Ω has a boundary, the previous equation is supplemented with the no-flux boundary condition (ν denotes the outer unit normal to $\partial\Omega$):

$$(\nabla P(\rho) + \rho U[\rho]) \cdot \nu = 0 \text{ on } \partial\Omega. \tag{3.2}$$

For every $\rho \in \mathcal{P}(\Omega)$, we assume that the Helmholtz decomposition of the vector field $U[\rho]$:

$$U[\rho] = -W[\rho] + \nabla V[\rho], \tag{3.3}$$

with

$$\nabla \cdot W[\rho] = 0,\ W[\rho] \cdot \nu = 0 \text{ on } \partial\Omega,$$

satisfies the following assumptions.

Assumptions on the potential part V:

- $\nabla V[\rho] \in L^\infty_{\rm loc}$ uniformly in ρ i.e for all $K \subset\subset \Omega$, there exists $C > 0$ such that for all $\rho \in \mathcal{P}(\Omega)$,

$$\|\nabla V[\rho]\|_{\infty, K} \leqslant C. \tag{3.4}$$

  Note that by Rademacher's condition, this condition implies that $V[\rho]$ is differentiable a.e.,

- $V[\rho]$ is semi-convex uniformly in ρ i.e there exists C such that for all $\rho \in \mathcal{P}(\Omega)$, for every $y \in \Omega$ and every $x \in \Omega$, point of differentiability of $V[\rho]$:

$$V[\rho](y) \geqslant V[\rho](x) + \langle \nabla V[\rho](x), y - x\rangle - \frac{C}{2}|y - x|^2. \tag{3.5}$$

- There exists $C \geqslant 0$ such that for all $\rho \in \mathcal{P}(\Omega)$, for all $x \in \Omega$:

$$V[\rho](x) \geqslant -C(1+|x|). \tag{3.6}$$

- $\nabla V[\rho] \in L^2(\rho)$ uniformly in ρ, i.e there exists $C > 0$ such that for all $\rho \in \mathcal{P}(\Omega)$,

$$\int_\Omega |\nabla V[\rho]|^2 \, d\rho \leqslant C. \tag{3.7}$$

- There exists $C > 0$ such that for all $\rho, \mu \in \mathcal{P}(\Omega)$,

$$\int_\Omega |\nabla V[\rho] - \nabla V[\mu]|^2 \, d\rho \leqslant C W_2^2(\rho, \mu). \tag{3.8}$$

Assumptions on the divergence-free part W:

- There exists $C > 0$ such that for all $\rho \in \mathcal{P}(\Omega)$ and

$$W[\rho] \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n) \qquad \text{and} \qquad |W[\rho](x)| \leqslant C(1+|x|), \text{ for all } x \in \mathbb{R}^n. \tag{3.9}$$

- There exists $C > 0$ such that for all $\rho, \mu \in \mathcal{P}(\Omega)$,

$$\int_\Omega |W[\rho] - W[\mu]|^2 \, d\rho \leqslant C W_2^2(\rho, \mu). \tag{3.10}$$

Before going further, let us briefly consider some examples of velocity fields $\rho \mapsto U[\rho]$ that satisfy the previous assumptions.

Example 1: Let us consider the case where Ω is a smooth bounded domain of $\mathbb{R}^n$ and $U[\rho]$ has the form

$$U[\rho](x) := \int_\Omega U(x,y) \mathrm{d}\rho(y)$$

where $U : \Omega \times \Omega \to \mathbb{R}^n$ satisfies the following regularity conditions

$$D_x U \in L^\infty_y(L^1_x), \ \mathrm{div}_x(U) \in L^\infty_y(C^\alpha_x), \ D_y U \in L^\infty_y(C^\alpha_x) \tag{3.11}$$

for some $\alpha \in (0,1)$. Then we perform an Helmholtz decomposition "y by y" by solving

$$\Delta_x V(.,y) = \mathrm{div}_x(U(.,y)) \text{ in } \Omega, \ \nabla_x V(.,y) \cdot \nu = U(.,y) \cdot \nu \text{ on } \partial\Omega,$$

defining then $V[\rho]$ by

$$V[\rho](x) := \int_\Omega V(x,y) \mathrm{d}\rho(y),$$

we have $U[\rho] = \nabla V[\rho] + W[\rho]$ where $W[\rho]$ is divergence-free and tangential to $\partial\Omega$. By elliptic regularity, (3.11) implies that $\nabla_x V \in L^\infty_y(C^{1+\alpha}_x)$ so that $V[\rho]$ is $C^{2+\alpha}$ uniformly in ρ which clearly implies (3.4), (3.5), (3.6) and (3.7). As for (3.8), we observe that since $\Delta_x \partial_{y_i} V = \mathrm{div}_x(\partial_{y_i} U)$ thanks to (3.11) and elliptic regularity we have $D^2_{xy} V \in L^\infty_y(C^\alpha_x) \subset L^\infty(\Omega \times \Omega)$, hence $\nabla_x V$ is Lipschitz in y, we then have

$$|\nabla V[\rho](x) - \nabla V[\mu](x)| \leq W_1(\rho,\mu) \|D^2_{xy} V\|_{L^\infty}$$

where W_1 denotes the 1-Wasserstein-distance, by Cauchy-Schwarz inequality $W_1^2 \leq W_2^2$ so that (3.8) holds with $C = \|D^2_{xy} V(x,.)\|^2_{L^\infty}$. Checking that $W[\rho] := U[\rho] - \nabla V[\rho]$ satisfies (3.9) and (3.10) is similar.

Example 2: The previous example can be straighforwardly generalized to higher-order interaction velocities i.e. to the case where $U[\rho]$ has the form

$$U[\rho](x) := \int_{\Omega^l} U(x, y_1, \cdots, y_l) \mathrm{d}\rho(y_1) \cdots \mathrm{d}\rho(y_l)$$

under a similar assumption as (3.11).

Example 3: Another important example and motivation comes from kinetic models where velocities may contain an Hamiltonian term. More precisely, consider the case where $\Omega = \mathbb{R}^{2d}$ and write generic elements of the phase-space as $x = (y, v)$ and $U[\rho] = \nabla V[\rho] + W[\rho]$ where the potential term $V[\rho]$ satisfies (3.4), (3.5), (3.6), (3.7) and (3.8) and

$$W[\rho](y, v) := \Big(-v, \int_{\mathbb{R}^d \times \mathbb{R}^d} \nabla H(y - z) \mathrm{d}\rho(z, w) \Big)$$

with $H \in C^{1,1}(\mathbb{R}^d)$. Then $W[\rho]$ is Lipschitz and one easily checks that (3.10) holds thanks to the fact that ∇H is Lipschitz.

Assumptions on the internal energy F and the associated pressure P:

The nonlinear diffusion term is given by a continuous strictly convex superlinear (i.e. $F(\rho)/\rho \to +\infty$ as $\rho \to +\infty$) function $F : \mathbb{R}^+ \to \mathbb{R}$ of class $\mathcal{C}^2((0, +\infty))$ which satisfies

$$F(0) = 0, \text{ and } P(\rho) \leqslant C(\rho + F(\rho)). \tag{3.12}$$

where $P(\rho) := \rho F'(\rho) - F(\rho)$ is the pressure associated to F. Moreover, we define $\mathcal{F} : \mathcal{P}(\Omega) \to \mathbb{R}$ by

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega F(\rho(x))\, dx & \text{if } \rho \ll \mathcal{L}^n, \\ +\infty & \text{otherwise.} \end{cases}$$

And we assume that

$$\mathcal{F}(\rho) \geqslant -C(1 + M(\rho))^\alpha, \qquad \text{for all } \rho \in \mathcal{P}(\Omega), \tag{3.13}$$

where $\alpha \in (0, 1)$ and $M(\rho) := \int_\Omega |x|^2\, d\rho(x)$ is the second moment of ρ.

The typical examples of energies we have in mind are $F(\rho) := \rho \log(\rho)$, which gives a linear diffusion driven by the laplacian, and $F(\rho) := \rho^m$ $(m > 1)$ which corresponds to the porous medium equation.

A weak solution of (3.1)-(3.2) is a curve $\rho : t \in (0, +\infty) \mapsto \rho(t, \cdot) \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ such that $\nabla P(\rho) \in \mathcal{M}^n([0, +\infty) \times \Omega)$ and

$$\int_0^\infty \left(\int_\Omega (\partial_t \phi \rho - \nabla \phi \cdot U[\rho]\rho) dx - \int_\Omega \nabla \phi \cdot d\nabla P(\rho) \right) dt = - \int_\Omega \phi(0, x) \rho_0(x)\, dx, \tag{3.14}$$

for every $\phi \in \mathcal{C}_c^\infty([0, +\infty) \times \mathbb{R}^n)$.

Our main result is the following:

Theorem 3.1. *Assume $\rho_0 \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ such that*

$$\mathcal{F}(\rho_0) < +\infty, \tag{3.15}$$

then (3.1) *admits at least one weak solution.*

The proof of this theorem is given in the next sections and is based on the following splitting scheme that combines pure transport steps by the divergence-free part of the drift U and Wasserstein gradient flow steps taking into account the potential V in a semi-implicit way. More precisely, given a time step $h > 0$, we construct by induction a sequence $\rho_h^k \in \mathcal{P}_2^{\mathrm{ac}}(\Omega)$ by setting $\rho_h^0 = \rho_0$ and given ρ_h^k we find ρ_h^{k+1} using the following scheme:

- **pure transport phase:** we introduce an intermediate measure, $\tilde{\rho}_h^{k+1}$ (with $\tilde{\rho}_h^0 = \rho_0$) defined by

$$\tilde{\rho}_h^{k+1} = X_h^k(h, \cdot)_\# \rho_h^k, \tag{3.16}$$

where X_h^k is solution of

$$\begin{cases} \partial_t X_h^k = W[\rho_h^k] \circ X_h^k, \\ X_h^k(0, \cdot) = \mathrm{id}\,. \end{cases} \tag{3.17}$$

Since $W[\rho_h^k]$ satisfies (3.9), as recalled in section 2, DiPerna-Lions theory [10] implies that X_h^k is well defined. Moreover, since $W[\rho_h^k]$ is divergence-free then X_h^k preserves the Lebesgue measure and leaves the domain Ω invariant thanks to the fact that $W[\rho]$ is tangential to $\partial\Omega$. Therefore $\tilde{\rho}_h^{k+1} = \rho_h^k(X_h^{k^{-1}})$ which implies the conservation of the internal energy:

$$\begin{aligned}
\mathcal{F}(\tilde{\rho}_h^{k+1}) &= \int_\Omega F(\rho_h^k(X_h^{k^{-1}}(x)))\,dx \\
&= \int_\Omega F(\rho_h^k(x))\,dx = \mathcal{F}(\rho_h^k).
\end{aligned} \tag{3.18}$$

In addition, we can see $\tilde{\rho}_h^{k+1}$ is the value at time h of the solution μ of the continuity equation

$$\begin{cases} \partial_t\mu + \operatorname{div}(\mu W[\rho_h^k]) = 0, \\ \mu_{|t=0} = \rho_h^k. \end{cases} \tag{3.19}$$

Thanks to these observations, we can easily control the W_2-distance between $\tilde{\rho}_h^{k+1}$ and ρ_h^k. Indeed, using Benamou-Brenier formula and (3.9), we obtain

$$\begin{aligned}
W_2^2(\tilde{\rho}_h^{k+1},\rho_h^k) &\leqslant h\int_0^h\int_\Omega |W[\rho_h^k]|^2\,d\mu_t dt \\
&\leqslant Ch\int_0^h\int_\Omega (1+|x|^2)\,d\mu_t dt \\
&\leqslant Ch\int_0^h (1+M(\mu_t))\;dt.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\frac{d}{dt}M(\mu_t) &= \int_\Omega |x|^2\partial_t\mu_t \\
&= -\int_\Omega |x|^2\operatorname{div}(W[\rho_h^k]\mu_t) \\
&= 2\int_\Omega x\cdot W[\rho_h^k]\mu_t \\
&\leqslant C(M(\mu_t)+1).
\end{aligned}$$

We obtain the last line using (3.9), Cauchy-Schwarz inequality and Young's inequality. Then,

$$M(\mu_t) \leqslant C(t+1)e^{Ct} \leqslant 2Ce^C \text{ for } t\leqslant 1,$$

which implies

$$W_2^2(\tilde{\rho}_h^{k+1},\rho_h^k) \leqslant Ch^2. \tag{3.20}$$

- **semi-implicit JKO scheme:** In the second step we use a semi-implict version of the Jordan-Kinderlehrer-Otto scheme [7], introduced by Di Francesco and Fagioli in [8] and used in [11], with $\tilde{\rho}_h^{k+1}$ being the measure defined in the previous step. More precisely, we select ρ_h^{k+1} as a solution of

$$\inf_{\rho\in\mathcal{P}_2^{ac}(\Omega)} \mathcal{E}_h(\rho|\tilde{\rho}_h^{k+1}) := W_2^2(\rho,\tilde{\rho}_h^{k+1}) + 2h\left(\mathcal{F}(\rho)+\mathcal{V}(\rho|\tilde{\rho}_h^{k+1})\right), \tag{3.21}$$

where

$$\mathcal{V}(\rho|\mu) := \int_\Omega V[\mu]\,d\rho.$$

By standard compactness and lower semicontinuity argument, (3.21) admits at least one solution (see for example [7, 11]) so the sequence ρ_h^k is well defined (it is even actually unique by strict convexity of $\mathcal{E}_h(.|\tilde{\rho}_h^{k+1})$).

To summarize, given a time step $h>0$, we construct by induction two sequences ρ_h^k and $\tilde{\rho}_h^k$ with the following splitting scheme: $\rho_h^0 = \tilde{\rho}_h^0 = \rho_0$ and for all $k\geqslant 0$,

$$\begin{cases} \tilde{\rho}_h^{k+1} = X_h^k(h,\cdot)_\#\rho_h^k, \\ \rho_h^{k+1} \in \operatorname{argmin}_{\rho\in\mathcal{P}_2^{ac}(\Omega)}\left\{W_2^2(\rho,\tilde{\rho}_h^{k+1}) + 2h\left(\mathcal{F}(\rho)+\mathcal{V}(\rho|\tilde{\rho}_h^{k+1})\right)\right\}. \end{cases} \tag{3.22}$$

We finally introduce three different interpolations:

- We denote ρ_h the usual piecewise constant interpolation of the sequence ρ_h^k

$$\rho_h(t,\cdot) := \rho_h^{k+1} \qquad \text{if } t \in (hk, h(k+1)], \tag{3.23}$$

- similarly, we interpolate in a piecewise constant way the sequence $\tilde{\rho}_h^k$:

$$\tilde{\rho}_h^1(t,\cdot) := \tilde{\rho}_h^{k+1} \qquad \text{if } t \in (hk, h(k+1)], \tag{3.24}$$

- finally, we denote by $\tilde{\rho}_h^2$ the *continuous* interpolation of $\tilde{\rho}_h^k$

$$\tilde{\rho}_h^2(t,\cdot) := X_h^k(t-hk,\cdot)_{\#}\rho_h^k \qquad \text{if } t \in (hk, h(k+1)]. \tag{3.25}$$

  We remark that on $(hk, h(k+1)]$, $\tilde{\rho}_h^2$ is the solution on $(0,h)$ of the continuity equation (3.19).

The next two sections are devoted to the proof of theorem 3.1. In section 4, we derive various estimates on the sequences generated by the splitting scheme above, in particular thanks to the Euler-Lagrange equation of the semi-implicit JKO steps. This enables us to pass to the limit as the time step goes to 0 (the difficult term being of course the nonlinear pressure term) and thus to conclude the existence proof, this is done in section 5.

4 Estimates

4.1 Basic a priori estimates

Using the semi-implicit JKO scheme we first obtain the following a priori estimates on ρ_h, $\tilde{\rho}_h^1$ and $\tilde{\rho}_h^2$.

Proposition 4.1. *There exists $h_0 > 0$, such that for $T > 0$, there exists $C > 0$ such that, for all h, k, with $h \in (0, h_0)$ and $hk < T$, $N = \lceil \frac{T}{h} \rceil$, we have*

$$M(\rho_h^k) \leqslant C, \tag{4.1}$$

$$\mathcal{F}(\rho_h^k) \leqslant C, \tag{4.2}$$

$$\textstyle\sum_{k=0}^{N-1} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) \leqslant Ch. \tag{4.3}$$

Proof. Using $\tilde{\rho}_h^{k+1}$ as a competitor of ρ_h^{k+1} in (3.21), we obtain

$$\frac{1}{2h} W_2^2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \leqslant \mathcal{F}(\tilde{\rho}_h^{k+1}) - \mathcal{F}(\rho_h^{k+1}) + \int_\Omega V[\tilde{\rho}_h^{k+1}]\left(\tilde{\rho}_h^{k+1} - \rho_h^{k+1}\right). \tag{4.4}$$

Let γ be the optimal transport plan between $\tilde{\rho}_h^{k+1}$ and ρ_h^{k+1}. Then we have

$$\begin{aligned}
\int_\Omega V[\tilde{\rho}_h^{k+1}]\left(\tilde{\rho}_h^{k+1} - \rho_h^{k+1}\right) &= \int_\Omega \left(V[\tilde{\rho}_h^{k+1}](x) - V[\tilde{\rho}_h^{k+1}](y)\right) d\gamma(x,y) \\
&= \int_\Omega \left(V[\tilde{\rho}_h^{k+1}](x) - V[\tilde{\rho}_h^{k+1}](y) + \nabla V[\tilde{\rho}_h^{k+1}](x)\cdot(y-x)\right) d\gamma(x,y) \\
&- \int_\Omega \nabla V[\tilde{\rho}_h^{k+1}](x)\cdot(y-x)\, d\gamma(x,y).
\end{aligned}$$

Using (3.5) for the first part and Cauchy-Schwarz inequality and (3.7) for the second part of the right hand side, we find

$$\begin{aligned}
\int_\Omega V[\tilde{\rho}_h^{k+1}]\left(\tilde{\rho}_h^{k+1} - \rho_h^{k+1}\right) &\leqslant \frac{C}{2}\int_\Omega |x-y|^2\,\gamma(x,y) + C\Big(\int_\Omega |x-y|^2\,\gamma(x,y)\Big)^{\frac{1}{2}} \\
&= \frac{C}{2} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) + C W_2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}).
\end{aligned}$$

Choosing $h \leq h_0 \leq \frac{1}{2C}$ and using Young's inequality

$$W_2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) \leq \frac{1}{8hC} W_2^2(\tilde{\rho}_h^{k+1}, \rho_h^{k+1}) + 2Ch,$$

(4.4) becomes

$$\frac{1}{8h}W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1}) \leqslant \mathcal{F}(\tilde\rho_h^{k+1}) - \mathcal{F}(\rho_h^{k+1}) + Ch.$$

Now using (3.18), to recover a telescopic sum, and summing over k, we obtain

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1}) \leqslant 8h\left(\mathcal{F}(\rho_0) - \mathcal{F}(\rho_h^N) + CT\right),$$

this inequality and (3.15) imply (4.2). In addition, since the lower bound of $\mathcal{F}$ is controlled by the second moment,

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1}) \leqslant 8h\left(\mathcal{F}(\rho_0) + C(1+M(\rho_h^N))^\alpha + CT\right). \tag{4.5}$$

But, with (4.5) and by standard arguments (see [7, 11]), we deduce that $M(\rho_h^k)$ satisfies (4.1) and then (4.5), (3.15) and (4.1) give (4.3). ☐

Remark 4.2. *Using estimate* (3.20) *between* $\tilde\rho_h^{k+1}$ *and* ρ_h^k*, and* (4.3)*, we also have*

$$\sum_{k=0}^{N-1} W_2^2(\rho_h^k,\rho_h^{k+1}) \leqslant Ch \text{ and } \sum_{k=0}^{N-1} W_2^2(\tilde\rho_h^k,\tilde\rho_h^{k+1}) \leqslant Ch.$$

Moreover, using (3.18)*, we have for all* $t\in[0,T]$*,*

$$\mathcal{F}(\tilde\rho_h^1(t)),\mathcal{F}(\tilde\rho_h^2(t)) \leqslant C \qquad \text{and} \qquad M(\tilde\rho_h^1(t)),M(\tilde\rho_h^2(t)) \leqslant C.$$

4.2 Discrete Euler-Lagrange equation and stronger estimates

Let us start with the Euler-Lagrange equation of (3.21).

Proposition 4.3. *For all* $k\geqslant 0$*, we have* $P(\rho_h^{k+1})\in W^{1,1}(\Omega)$ *and*

$$h\left(\nabla V[\tilde\rho_h^{k+1}]\rho_h^{k+1} + \nabla P(\rho_h^{k+1})\right) = -\nabla\varphi_h^{k+1}\rho_h^{k+1} \qquad a.e, \tag{4.6}$$

where φ_h^{k+1} *is a Kantorovich potential from* ρ_h^{k+1} *to* $\tilde\rho_h^{k+1}$ *(so that its gradient is unique* ρ_h^{k+1}*-a.e.) for* W_2*.*

Proof. The proof is the same as in [1, 11] for example. We start by taking the first variation in the semi-implicit JKO scheme along the flow of a smooth vector field. Let $\xi\in\mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$ be given and Φ_τ the corresponding flow defined by

$$\partial_\tau\Phi_\tau = \xi\circ\Phi_\tau,\ \Phi_0 = \mathrm{id}\,.$$

We define a pertubation of ρ_h^{k+1} by $\rho_\tau := \Phi_{\tau\#}\rho_h^{k+1}$. Then we get

$$\frac{1}{\tau}\left(\mathcal{E}_h(\rho_\tau|\tilde\rho_h^{k+1}) - \mathcal{E}_h(\rho_h^{k+1}|\tilde\rho_h^{k+1})\right) \geqslant 0. \tag{4.7}$$

By standard computations, we have

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(W_2^2(\rho_\tau,\tilde\rho_h^{k+1}) - W_2^2(\rho_h^{k+1},\tilde\rho_h^{k+1})) \leqslant \int_{\Omega\times\Omega}(x-y)\cdot\xi(x)\,d\gamma_h^{k+1}(x,y), \tag{4.8}$$

with γ_h^{k+1} is the W_2-optimal transport plan in $\Pi(\rho_h^{k+1},\tilde\rho_h^{k+1})$ and $\gamma_h^{k+1} = (\mathrm{id}\times T_h^{k+1})_\#\rho_h^{k+1}$ with $T_h^{k+1} = \mathrm{id}-\nabla\varphi_h^{k+1}$. Moreover, using (3.12), (4.2) and Lebesgue's dominated convergence theorem, we obtain

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{F}(\rho_\tau) - \mathcal{F}(\rho_h^{k+1})) \leqslant -\int_\Omega P(\rho_h^{k+1}(x))\,\mathrm{div}(\xi(x))\,dx. \tag{4.9}$$

Finally,

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}(\mathcal{V}(\rho_\tau|\tilde\rho_h^{k+1}) - \mathcal{V}(\rho_{i,h}^{k+1}|\tilde\rho_h^{k+1})) \leqslant \int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\xi\rho_h^{k+1}\,dx, \tag{4.10}$$

so that if we combine (4.7), (4.8), (4.9) and (4.10), and if we replace ξ by $-\xi$, we find that, for all $\xi \in \mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$,

$$\int_\Omega \nabla\varphi_h^{k+1}\cdot\xi\rho_h^{k+1} - h\int_\Omega P(\rho_h^{k+1})\operatorname{div}(\xi) + h\int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\xi\rho_h^{k+1} = 0. \tag{4.11}$$

Now we claim that $P(\rho_h^{k+1}) \in W^{1,1}(\Omega)$. Indeed, since P is controlled by F thanks to assumption (3.12), (4.2) gives $P(\rho_h^{k+1}) \in L^1(\Omega)$. Moreover, using (4.11), we obtain

$$\left|\int_\Omega P(\rho_h^{k+1})\operatorname{div}(\xi)\right| \leqslant \left[\int_\Omega \frac{|\nabla\varphi_h^k(y)|}{h}\rho_h^{k+1} + \int_\Omega |\nabla V[\tilde\rho_h^{k+1}]|\rho_h^{k+1}\right]\|\xi\|_{L^\infty(\Omega)}.$$

But using (3.8), (3.7), (4.3) and Cauchy-Schwarz inequality we get

$$\begin{aligned}
\int_\Omega |\nabla V[\tilde\rho_h^{k+1}]|\rho_h^{k+1} &\leqslant \left[\int_\Omega |\nabla V[\tilde\rho_h^{k+1}] - \nabla V[\rho_h^{k+1}]|\rho_h^{k+1} + \int_\Omega |\nabla V[\rho_h^{k+1}]|\rho_h^{k+1}\right]\\
&\leqslant \left[\Big(\int_\Omega |\nabla V[\tilde\rho_h^{k+1}] - \nabla V[\rho_h^{k+1}]|^2\rho_h^{k+1}\Big)^{1/2} + \Big(\int_\Omega |\nabla V[\rho_h^{k+1}]|^2\rho_h^{k+1}\Big)^{1/2}\right]\\
&\leqslant C\left[W_2(\tilde\rho_h^{k+1},\rho_h^{k+1}) + 1\right]\\
&\leqslant C.
\end{aligned}$$

We thus have

$$\left|\int_\Omega P(\rho_h^{k+1})\operatorname{div}(\xi)\right| \leqslant \left[\frac{W_2(\tilde\rho_h^{k+1},\rho_h^{k+1})}{h} + C\right]\|\xi\|_{L^\infty(\Omega)}.$$

This implies $P(\rho_h^{k+1}) \in BV(\Omega)$ and $\nabla P(\rho_h^{k+1}) = \left(-\nabla V[\tilde\rho_h^{k+1}]\rho_h^{k+1} - \frac{\nabla\varphi_h^{k+1}}{h}\rho_h^{k+1}\right)$ in $\mathcal{M}^n(\Omega)$. In fact, $P(\rho_h^{k+1})$ is in $W^{1,1}(\Omega)$ because $\nabla V[\tilde\rho_h^{k+1}]\rho_h^{k+1} + \frac{\nabla\varphi_h^k}{h}\rho_h^{k+1} \in L^1(\Omega)$ and we have proved (4.6).

□

We immediately deduce an $L^1((0,T), BV(\Omega))$ estimate for $P(\rho_h)$:

Corollary 4.4. *For all $T>0$, we have*

$$\|P(\rho_h)\|_{L^1((0,T);W^{1,1}(\Omega))} \leqslant CT. \tag{4.12}$$

Proof. If we integrate (4.6), we obtain

$$h\int_\Omega |\nabla P(\rho_h^{k+1})| \leqslant W_2(\rho_h^{k+1},\tilde\rho_h^{k+1}) + Ch,$$

Then we sum from $k=0$ to $N-1$ and thanks to (4.3), we have

$$\int_0^T\int_\Omega |\nabla P(\rho_h)| \leqslant CT.$$

We conclude thanks to (3.12) and (4.2). □

Proposition 4.5. *Let $h>0$, $N\in\mathbb{N}^*$, $T := Nh$, $t_k := hk$, for $k = 0,\cdots,N$, then, for every $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$*

$$\begin{aligned}
\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(\partial_t\phi(t,x) + W[\rho_h(t-h)]\cdot\nabla\phi(t,x))\,dxdt &= h\sum_{k=0}^{N-1}\int_\Omega \nabla P(\rho_h^{k+1}(x))\cdot\nabla\phi(t_k,x)\,dx\\
&+ h\sum_{k=0}^{N-1}\int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\nabla\phi(t_k,x)\rho_h^{k+1}\,dx\\
&+ \sum_{k=0}^{N-1}\int_{\Omega\times\Omega}\mathcal{R}[\phi(t_k,\cdot)](x,y)d\gamma_h^{k+1}(x,y)\\
&- \int_\Omega \rho_0(x)\phi(0,x)\,dx,
\end{aligned}$$

with, for all $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\Omega)}|x-y|^2,$$

and γ_h^{k+1} *is the optimal transport plan in* $\Pi(\rho_h^{k+1},\tilde\rho_h^{k+1})$.

Proof. Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, multiplying (4.6) by $\nabla\varphi$ and integrating on Ω, we obtain

$$-\int_\Omega \nabla\varphi_h^{k+1}\cdot\nabla\varphi\rho_h^{k+1} = h\left(\int_\Omega \nabla P(\rho_h^{k+1})\cdot\nabla\varphi + \int_\Omega \nabla V[\tilde\rho_h^{k+1}]\cdot\nabla\varphi\rho_h^{k+1}\right).$$

But, we can rewrite the left hand side by

$$-\int_\Omega \nabla\varphi_h^{k+1}\cdot\nabla\varphi\rho_h^{k+1} = \int_{\Omega\times\Omega}(y-x)\cdot\nabla\varphi(x)\,d\gamma_h^{k+1}(x,y).$$

A second-order Taylor-Lagrange formula then gives

$$\begin{aligned}\int_{\Omega\times\Omega}(y-x)\cdot\nabla\varphi(x)\,d\gamma_h^{k+1} &= \int_{\Omega\times\Omega}(\varphi(y)-\varphi(x))\;d\gamma_h^{k+1}(x,y) - \int_{\Omega\times\Omega}\mathcal{R}[\varphi](x,y)d\gamma_h^{k+1}(x,y)\\ &= \int_\Omega \varphi(\tilde\rho_h^{k+1}-\rho_h^{k+1}) - \int_{\Omega\times\Omega}\mathcal{R}[\varphi](x,y)d\gamma_h^{k+1}(x,y). \end{aligned} \tag{4.13}$$

Now let $\phi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$, we have

$$\begin{aligned}&\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(\partial_t\phi(t,x) + W[\rho_h(t-h)]\cdot\nabla\phi(t,x))\,dxdt\\ =\;& \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_\Omega \rho_h^k(x)(\partial_t\phi + W[\rho_h^k]\cdot\nabla\phi)(t,X_h^k(t-t_k,x))\,dxdt.\end{aligned}$$

We next observe that, on $[t_k,t_{k+1}]$,

$$\frac{d}{dt}[\phi(t,X_h^k(t-t_k,x))] = (\partial_t\phi + W[\rho_h^k]\cdot\nabla\phi)(t,X_h^k(t-t_k,x)).$$

Then,

$$\begin{aligned}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_\Omega \rho_h^k(x)(\partial_t\phi + W[\rho_h^k]\cdot\nabla\phi)(t,X_h^k(t-t_k,x))\,dxdt &= \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_\Omega \rho_h^k(x)\frac{d}{dt}[\phi(t,X_h^k(t-t_k,x))]\,dxdt\\ &= \sum_{k=0}^{N-1}\int_\Omega \rho_h^k(x)\left[\phi(t_{k+1},X_h^k(h,x)) - \phi(t_k,x)\right]\,dx\\ &= \sum_{k=0}^{N-1}\int_\Omega\left[\phi(t_{k+1},x)\tilde\rho_h^{k+1}(x) - \phi(t_k,x)\rho_h^k(x)\right]\,dx\\ &= \sum_{k=0}^{N-1}\int_\Omega \phi(t_{k+1},x)\left(\tilde\rho_h^{k+1}(x) - \rho_h^{k+1}(x)\right)\\ &\quad -\int_\Omega \phi(0,x)\rho_0(x)\,dx.\end{aligned}$$

Then the proof is complete by applying (4.13) with $\varphi = \phi(t_{k+1},\cdot)$.

□

5 Convergence and proof of Theorem 3.1

5.1 Weak and strong convergences

Using a refined version of Ascoli theorem (see [2]), estimate (4.3) and remark 4.2 and taking subsequences, if necessary, we have that, for every $T < +\infty$, ρ_h, $\tilde{\rho}_h^1$ and $\tilde{\rho}_h^2$ converge in $L^\infty((0,T), W_2)$ to some respective limits ρ, $\tilde{\rho}^1$ and $\tilde{\rho}^2$:

$$\sup_{t\in[0,T]} \max(W_2(\rho_h(t,.),\rho(t,.)), W_2(\tilde{\rho}_h^1(t,.),\tilde{\rho}^1(t,.)), W_2(\tilde{\rho}_h^2(t,.),\tilde{\rho}^2(t,.))) \to 0 \text{ as } h \to 0.$$

In fact, these three sequences have to converge to the same limit, ρ. Indeed, for all $\varphi \in \mathcal{C}_c^\infty((0,T)\times\Omega)$,

$$\begin{aligned}
\int_0^T \int_\Omega \varphi(\rho_h - \tilde{\rho}_h^1) &= \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \int_\Omega \varphi(\rho_h^{k+1} - \tilde{\rho}_h^{k+1}) \\
&\leqslant Ch \sum_{k=0}^{N-1} W_2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \\
&\leqslant ChN^{1/2} \left(\sum_{k=0}^{N-1} W_2^2(\rho_h^{k+1}, \tilde{\rho}_h^{k+1}) \right)^{1/2} \\
&\leqslant CT^{1/2}h,
\end{aligned}$$

because of (4.3). With a similar computation, we find that ρ_h and $\tilde{\rho}_h^2$ converge to the same limit. We thus have

$$\sup_{t\in[0,T]} \max(W_2(\rho_h(t,.),\rho(t,.)), W_2(\tilde{\rho}_h^1(t,.),\rho(t,.)), W_2(\tilde{\rho}_h^2(t,.),\rho(t,.))) \to 0 \text{ as } h \to 0. \tag{5.1}$$

Moreover it is classical to deduce from (4.3) and remark 4.2 an Hölder-like estimate of the form $W_2(\rho_h(t,.),\rho_h(s,.)) \leq C\sqrt{|t-s|+h}$ from which one deduces that the limit curve ρ actually belongs to $\mathcal{C}^{1/2}((0,T),W_2)$. This kind of convergence will be enough to pass to the limit in $\nabla V[\tilde{\rho}_h^1]\rho_h$ and $W[\rho_h]\tilde{\rho}_h^2$, because of assumptions (3.8) and (3.10), but we will need a stronger convergence to deal with the nonlinear diffusion term $P(\rho_h)$. Fo this purpose, we will use an extension of the Aubin-Lions Lemma due to Rossi and Savaré in [15]:

Theorem 5.1 (th. 2 in [15])**.** *On a Banach space B, let be given*

- *a normal coercive integrand $\mathcal{G} : B \to \mathbb{R}^+$, i.e, $\mathcal{G}$ is l.s.c and its sublevels are relatively compact in B,*
- *a pseudo-distance $g : B\times B \to [0,+\infty]$, i.e, g is l.s.c and $[g(\rho,\mu)=0,\ \rho,\mu\in B$ with $\mathcal{G}(\rho),\mathcal{G}(\mu)<\infty] \Rightarrow \rho=\mu$.*

Let $T>0$ and U be a set of measurable functions $u : (0,T)\to B$. Under the hypotheses that

$$\sup_{u\in U} \int_0^T \mathcal{G}(u(t))\,dt < +\infty \qquad \text{and} \qquad \lim_{h\searrow 0} \sup_{u\in U} \int_0^{T-h} g(u(t+h),u(t))\,dt = 0, \tag{5.2}$$

U contains a subsequence $(u_n)_{n\in\mathbb{N}}$ which converges (strongly in B) in measure with respect to $t\in(0,T)$ to a limit $u_\star : (0,T)\to B$.

We now apply this theorem to $B = L^1(\Omega)$, $U = \{\rho_h\}_h$, g defined by

$$g(\rho,\mu) := \begin{cases} W_2(\rho,\mu) & \text{if } \rho,\mu\in\mathcal{P}_2(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

and $\mathcal{G}$ by

$$\mathcal{G}(\rho) := \begin{cases} \mathcal{F}(\rho) + \|P(\rho)\|_{BV(\Omega)} + M(\rho) & \text{if } \rho\in\mathcal{P}_2^{ac}(\Omega), P(\rho)\in BV(\Omega) \text{ and } F(\rho)\in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

Lemma 5.2. *$\mathcal{G}$ is l.s.c on $L^1(\Omega)$ and its sublevels are relatively compact in $L^1(\Omega)$.*

Proof. Let us start by proving that sublevels of $\mathcal{G}$ are relatively compact in $L^1(\Omega)$. Let

$$A_c := \left\{\rho \in L^1(\Omega)\ :\ \mathcal{G}(\rho) \leqslant c\right\}$$

and (ρ_k) be a sequence in A_c then $P(\rho_k)$ is bounded in $BV(\Omega)$ thus, up to a subsequence $P(\rho_k)$ converges to some Φ in $L^1_{\text{loc}}(\Omega)$ and a.e.. Since P is continuous, one to one and its inverse is continuous, ρ_k converges to $\rho := P^{-1}(\Phi)$ a.e.; and, since $\mathcal{G}(\rho_k) \leqslant c$ and F is superlinear, ρ_k is uniformly integrable, using Vitali's convergence theorem, we obtain that ρ_k converges to ρ in $L^1(K \cap \Omega)$ for every compact K. To conclude that there is convergence in $L^1(\Omega)$, we use the fact that the second momentum of ρ_k is uniformly bounded:

$$\begin{aligned}\int_\Omega |\rho_k - \rho| &\leqslant \int_{\Omega\setminus B_R} \frac{|x|^2}{R^2}|\rho_k - \rho| + \int_{B_R\cap\Omega} |\rho_k - \rho| \\ &\leqslant \frac{2c}{R^2} + \int_{B_R\cap\Omega} |\rho_k - \rho|.\end{aligned}$$

The first term in the right hand can be made arbitrary small by choosing R large enough and the second term converges to zero by $L^1(B_R \cap \Omega)$-convergence.

Now we have to show the lower semi-continuity of $\mathcal{G}$ on $L^1(\Omega)$. Let (ρ_k) be a sequence which converges strongly to ρ in $L^1(\Omega)$ with (without loss of generality) $\sup_k \mathcal{G}(\rho_k) \leqslant C < +\infty$. Without loss of generality, we can assume that ρ_k converges to ρ a.e. Since $\sup_k \mathcal{G}(\rho_k) \leqslant C$, $P(\rho_k)$ is uniformly bounded in $BV(\Omega)$ so $P(\rho_k)$ converges weakly to μ in $BV(\Omega)$. Moreover, $P(\rho_k)$ converges strongly to μ in $L^1_{\text{loc}}(\Omega)$. We can conclude that $\mu = P(\rho)$ and by lower semi-continuity of $\mathcal{F}$, M and the BV-norm we have

$$\mathcal{G}(\rho) \leqslant \liminf_{k\nearrow+\infty} \mathcal{G}(\rho_k).$$

$\square$

Thanks to lemma 5.2, to apply theorem 5.1, it remains to verify (5.2). The first condition of (5.2) is satisfied because of the estimate on the momentum, (4.1), on the internal energy $\mathcal{F}$, (4.2) and on the gradient of $P(\rho_h)$ (4.12). The second condition of (5.2) comes from the estimate on the distance (4.3) and remark 4.2 (see for example [9] for a detailed proof). Then theorem 5.1 implies that ρ_h converges in measure with respect to t in $L^1(\Omega)$ to ρ. Since convergence in measure implies a.e convergence, up to a subsequence, we may also assume that $\rho_h(t,.)$ converges strongly in $L^1(\Omega)$ to $\rho(t,.)$ for a.e. t. Then Lebesgue's dominated convergence theorem implies that ρ_h converges strongly in $L^1((0,T)\times\Omega)$ to ρ.

Thanks to (3.12) and (4.2) $P(\rho_h)$ is uniformly bounded in $L^\infty((0,T),L^1(\Omega))$. In addition, by corollary 4.4, $P(\rho_h)$ is uniformly bounded in $L^1((0,T),W^{1,1}(\Omega))$. Thanks to the Sobolev embedding, we deduce that $P(\rho_h)$ is uniformly bounded in $L^\infty((0,T),L^1(\Omega)) \cap L^1((0,T),L^{n/n-1}(\Omega))$. To have uniform integrability of $P(\rho_h)$ both in the time and space variables, the following will be useful:

Lemma 5.3. *Let $p > 1$, $q := \frac{2p-1}{p}$ and $f \in L^\infty((0,T),L^1(\Omega)) \cap L^1((0,T),L^p(\Omega))$ then $f \in L^q((0,T)\times\Omega)$ and we have*

$$\|f\|^q_{L^q((0,T)\times\Omega)} \leqslant \|f\|^{q-1}_{L^\infty_t(L^1_x)} \|f\|_{L^1_t(L^p_x)}.$$

Proof. Writing

$$\frac{1}{q} = \frac{p}{2p-1} = \theta + \frac{1-\theta}{p}, \quad \text{with } \theta = \frac{p-1}{2p-1} \in (0,1),$$

and observing that $(1-\theta)q = 1$ and $\theta q = q - 1$, the interpolation inequality yields that

$$\|f\|^q_{L^q_x} \leqslant \|f\|^{q-1}_{L^1_x} \|f\|_{L^p_x},$$

since $f \in L^\infty((0,T),L^1(\Omega)) \cap L^1((0,T),L^p(\Omega))$ this implies that $f \in L^q((0,T)\times\Omega)$ and

$$\begin{aligned}\|f\|^q_{L^q((0,T)\times\Omega)} = \int_0^T \|f\|^q_{L^q_x} &\leqslant \int_0^T \|f\|^{q-1}_{L^1_x}\|f\|_{L^p_x} \\ &\leqslant \|f\|^{q-1}_{L^\infty_t(L^1_x)}\|f\|_{L^1_t(L^p_x)}.\end{aligned}$$

$\square$

Applying lemma 5.3 we deduce that $P(\rho_h)$ is uniformly bounded in $L^{(n+1)/n}((0,T)\times\Omega)$. This implies that $P(\rho_h)$ is uniformly integrable and since we know that it converges a.e. to $P(\rho)$, we can deduce from Vitali's convergence theorem that $P(\rho_h)$ converges strongly to $P(\rho)$ in $L^1_{\mathrm{loc}}((0,T)\times\Omega)$.

Thanks to Corollary 4.4 we deduce that $\nabla P(\rho_h)$ converges vaguely to $\nabla P(\rho)$ in $\mathcal{M}^n_{\mathrm{loc}}((0,T)\times\Omega)$. In fact, we have $\nabla P(\rho)$ in $\mathcal{M}^n((0,T)\times\Omega)$ and narrow convergence of $\nabla P(\rho_h)$ to $\nabla P(\rho)$, thanks to Prokhorov Theorem and the following tightness estimate:

Lemma 5.4. *The family $\nabla P(\rho_h)$, viewed as vector-valued measures on $[0,T]\times\Omega$, is tight, more precisely, for every h and every A measurable, $A\subset\Omega$*

$$\int_0^T\int_A|\nabla P(\rho_h)|\le C\Big(1+\sqrt{h}\Big)\Big(\int_0^T\int_A\rho_h(t,x)\,dx\,dt\Big)^{1/2}. \tag{5.3}$$

Proof. Integrating (4.6) on $(0,T)\times A$ together with Cauchy Schwarz inequality and (4.3), we get (taking $N=\lceil\frac{T}{h}\rceil+1$, say)

$$\begin{aligned}
\int_0^T\int_A|\nabla P(\rho_h)| &\leqslant \sum_{k=0}^N\int_A|\nabla\varphi_h^{k+1}|\rho_h^{k+1}+\int_0^T\int_A|\nabla V[\tilde\rho_h^1]|\rho_h\\
&\leqslant \sum_{k=0}^N\Big(\int_\Omega|\nabla\varphi_h^{k+1}|^2\rho_h^{k+1}\Big)^{1/2}\Big(\int_A\rho_h^{k+1}\Big)^{1/2}+\int_0^T\int_A|\nabla V[\tilde\rho_h^1]|\rho_h\\
&\leqslant \Big(\sum_{k=0}^N W_2^2(\tilde\rho_h^{k+1},\rho_h^{k+1})\Big)^{1/2}\Big(\int_0^T\int_A\rho_h(t,x)\mathrm{d}x\mathrm{d}t\Big)^{1/2}+\int_0^T\int_A|\nabla V[\tilde\rho_h^1]|\rho_h\\
&\leqslant C\sqrt{h}\Big(\int_0^T\int_A\rho_h(t,x)\mathrm{d}x\mathrm{d}t\Big)^{1/2}+\int_0^T\int_A|\nabla V[\tilde\rho_h^1]|\rho_h.
\end{aligned}$$

Moreover, with Cauchy Schwarz inequality, (3.7) and (3.8), we also have

$$\begin{aligned}
\int_0^T\int_A|\nabla V[\tilde\rho_h^1]|\rho_h &\leqslant \int_0^T\int_A|\nabla V[\rho_h]|\rho_h+\int_0^T\int_A|\nabla V[\tilde\rho_h^1]-\nabla V[\rho_h]|\rho_h\\
&\leqslant C\Big(1+\sqrt{h}\Big)\Big(\int_0^T\int_A\rho_h(t,x)\mathrm{d}x\mathrm{d}t\Big)^{1/2},
\end{aligned}$$

which proves (5.3). The tightness of $\nabla P(\rho_h)$ therefore immediately follows from that of ρ_h and (5.3). □

We can summarize all of this in the next result:

Theorem 5.5. *Up to a subsequence ρ_h converges strongly in $L^1((0,T)\times\Omega)$, $P(\rho_h)$ converges strongly to $P(\rho)$ in $L^1_{\mathrm{loc}}((0,T)\times\Omega)$ and $\nabla P(\rho_h)$ converges to $\nabla P(\rho)$ narrowly in $\mathcal{M}((0,T)\times\Omega)$.*

5.2 End of the proof of theorem 3.1

In this section we finish the proof of theorem 3.1. We have to pass to the limit in all terms in proposition 4.5. The linear term (with time derivative) and the diffusion term converge to the desired result because $\tilde\rho_h^2$ converges to ρ in $L^\infty([0,T],W_2)$ and $\nabla P(\rho_h)$ converges to $\nabla P(\rho)$ narrowly in $\mathcal{M}^n([0,T]\times\Omega)$. Remainder term goes to 0 when h goes to 0 because of (4.3).

So we just have to check the convergence in the transport terms, in what follows the test-function ϕ belongs again to $\mathcal{C}_c^\infty([0,T)\times\mathbb{R}^n)$.

- *term in W:* We have to show:

$$\int_0^T\int_\Omega\tilde\rho_h^2(t,x)W[\rho_h(t-h)](x)\cdot\nabla\phi(t,x)\,dx\,dt\to\int_0^T\int_\Omega\rho(t,x)W[\rho(t,\cdot)](x)\cdot\nabla\phi(t,x)\,dx\,dt.$$

We first have

$$\int_0^T\int_\Omega \tilde\rho_h^2(t,x)W[\rho_h(t-h)](x)\cdot\nabla\phi(t,x)\,dxdt = \int_0^T\int_\Omega \tilde\rho_h^2(t,x)W[\tilde\rho_h^2(t)](x)\cdot\nabla\phi(t,x)\,dxdt + \int_0^T\int_\Omega \tilde\rho_h^2(t,x)(W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)](x))\cdot\nabla\phi(t,x)\,dxdt.$$

The second term in the right hand side goes to zero when h goes to 0. Indeed,

$$\Big|\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)](x))\cdot\nabla\phi(t,x)\,dxdt\Big| \leqslant C\int_0^T\left(\int_\Omega \tilde\rho_h^2(t,x)|W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)(x)]|^2\,dx\right)^{1/2}dt,$$

then using (3.10),

$$\begin{aligned}\Big|\int_0^T\int_\Omega \tilde\rho_h^2(t,x)(W[\rho_h(t-h)](x)-W[\tilde\rho_h^2(t)](x))\cdot\nabla\phi(t,x)\,dxdt\Big| &\leqslant C\int_0^T W_2(\rho_h(t-h),\tilde\rho_h^2(t))\,dt\\ &\leqslant C\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} W_2(\rho_h^k, X_h^k(t-t_k)_\#\rho_h^k)\,dt\\ &\leqslant CTh,\end{aligned}$$

because of (3.20). Moreover, using (3.10), we get

$$\begin{aligned}&\Big|\int_0^T\int_\Omega(\tilde\rho_h^2(t,x)W[\tilde\rho_h^2(t)](x)\cdot\nabla\phi(t,x)-\rho(t,x)W[\rho(t)](x)\cdot\nabla\phi(t,x))\,dxdt\Big|\\ &\qquad\leqslant C\int_0^T\int_\Omega \tilde\rho_h^2(t,x)|W[\tilde\rho_h^2(t)](x)-W[\rho(t)](x))|\,dxdt\\ &\qquad+\Big|\int_0^T\int_\Omega(\rho(t,x)-\tilde\rho_h^2(t,x))W[\rho(t)](x)\cdot\nabla\phi(t,x)\,dxdt\Big|\\ &\qquad\leqslant CT\sup_{t\in[0,T]}W_2(\tilde\rho_h^2(t),\rho(t))+\Big|\int_0^T\int_\Omega(\rho(t,x)-\tilde\rho_h^2(t,x))W[\rho(t)](x)\cdot\nabla\phi(t,x)\,dxdt\Big|\end{aligned}$$

the first term in the right hand-side converges to 0 because of (5.1). As for the second one, it also converges to 0, because $W[\rho]\cdot\nabla\phi$ belongs to $L^\infty((0,T)\times\Omega)$ and $\tilde\rho_h^2$ is uniformly integrable by remark 4.2 and the superlinearity of F, hence, up to a subsequence it converges to ρ weakly in $L^1((0,T)\times\Omega)$.

- *term in ∇V:* We claim that

$$h\sum_{k=0}^{N-1}\int_\Omega \nabla V[\tilde\rho_h^{k+1}](x)\cdot\nabla\phi(t_k,x)\rho_h^{k+1}\,dx \to \int_0^T\int_\Omega \nabla V[\rho(t,\cdot)](x)\cdot\nabla\phi(t,x)\rho(t,x)\,dxdt.$$

The proof is the same as the previous one for W, using (3.8), (3.4) and the convergence of ρ_h to ρ in $L^1((0,T)\times\Omega))\cap L^\infty((0,T),W_2)$.

6 On extension to systems and uniqueness

The splitting transport-JKO scheme described above, can easily be adapted, under suitable assumptions to the case of systems for the evolution of N species coupled by nonlocal drifts:

$$\partial_t\rho_i-\Delta P_i(\rho_i)-\mathrm{div}(\rho_iU_i[\rho_1,\cdots,\rho_N])=0,\ \rho_i(0,.)=\rho_{i,0},\ i=1,\cdots,N, \tag{6.1}$$

where $P_i(s)=sF_i'(s)-F_i(s)$ is the pressure associated to a strictly convex superlinear function F_i with corresponding internal energy $\mathcal{F}_i(\rho_i):=\int F_i(\rho_i(x))dx$. Decomposing each drift $U_i[\rho_1,\cdots,\rho_N]=\nabla V_i[\rho_1,\cdots,\rho_N]-W_i[\rho_1,\cdots,\rho_N]$ with $\mathrm{div}(W_i[\rho_1,\cdots,\rho_N])=0$ and under similar assumptions as in paragraph 3, one can show, by similar arguments as above, convergence as $h\to 0$ to a solution of (6.1) of the following splitting scheme.

Starting form $\rho_{i,h}^0=\rho_{i,0}$ and given $\rho_h^k=(\rho_{1,h}^k,\cdots\rho_{N,h}^k)$ we find $\rho_h^{k+1}=(\rho_{1,h}^{k+1},\cdots\rho_{N,h}^{k+1})$ by:

- setting $\tilde{\rho}_{i,h}^{k+1} = X_{i,h}^k(h,.)_{\#}\rho_{i,h}^k$ where

$$\partial_t X_{i,h}^k = W_i[\rho_h^k] \circ X_{i,h}^k,\ X_{i,h}^k(0,\cdot) = \mathrm{id},$$

- defining $\tilde{\rho}_h^{k+1} = (\tilde{\rho}_{1,h}^{k+1}, \cdots, \tilde{\rho}_{N,h}^{k+1})$, $\rho_h^{k+1} = (\rho_{1,h}^{k+1}, \cdots \rho_{N,h}^{k+1})$ is obtained by the semi-implicit JKO scheme:

$$\rho_{i,h}^{k+1} = \underset{\rho_i \in \mathcal{P}_2^{ac}(\Omega)}{\operatorname{argmin}} \left\{ W_2^2(\rho_i, \tilde{\rho}_{i,h}^{k+1}) + 2h \left(\mathcal{F}_i(\rho_i) + \int_\Omega V_i[\tilde{\rho}_h^{k+1}]\rho_i \right) \right\}.$$

Finally, let us say a few words on uniqueness which we have not addressed here, but which can be obtained at least in two ways: either by assuming some displacement semiconvexity of the internal energy and proving some exponential in time contraction estimate on the W_2 distance between two solutions (see [8, 11]), or by assuming some nondegeneracy of the diffusion and establishing some H^{-1} contraction estimate (see [5]).

References

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

[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Math., ETH Zürich, 2005.

[3] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (4), 375–417 (1991).

[4] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84**(3), 375–393 (2000).

[5] G. Carlier and M. Laborde, *Remarks on continuity equations with nonlinear diffusion and nonlocal drifts*, J. Math. Anal. Appl. **444** (2), 1690Ð1702 (2016).

[6] 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).

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

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

[9] M. Di Francesco and D. Matthes, *curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations*, Calc. Var. and PDEs **50**(1-2), 199-230 (2012).

[10] R.J DiPerna and P.L Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math. **98**, 511-547, (1989).

[11] M. Laborde, *On some non linear evolution systems which are perturbations of Wasserstein gradient flows*, to appear in Radon Series on Computational and Applied Mathematics.

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

[13] A.R. Mészáros and F. Santambrogio, *Advection-diffusion equations with density constraints*, to appear in Analysis and PDEs (2016).

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

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

[16] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Birkäuser Verlag, Basel, 2015.

[17] C. Villani, *Topics in optimal transportation*, volume 58 of graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2003.

[18] C. Villani, *Optimal transport: Old and New*, Springer Verlag (Grundlehren der mathematischen Wissenschaften), 2008.