



**HAL**  
open science

# A two-phase two-fluxes degenerate Cahn-Hilliard model as constrained Wasserstein gradient flow

Clément Cancès, Daniel Matthes, Flore Nabet

► **To cite this version:**

Clément Cancès, Daniel Matthes, Flore Nabet. A two-phase two-fluxes degenerate Cahn-Hilliard model as constrained Wasserstein gradient flow. 2017. hal-01665338v1

**HAL Id: hal-01665338**

**<https://hal.science/hal-01665338v1>**

Preprint submitted on 15 Dec 2017 (v1), last revised 6 Mar 2019 (v3)

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

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

# A TWO-PHASE TWO-FLUXES DEGENERATE CAHN-HILLIARD MODEL AS CONSTRAINED WASSERSTEIN GRADIENT FLOW

CLÉMENT CANCÈS, DANIEL MATTHES, AND FLORE NABET

ABSTRACT. We study a so-called non-local Cahn-Hilliard model obtained as a constrained Wasserstein gradient flow of some Ginzburg-Landau energy. When compared to the more classical local degenerate Cahn-Hilliard model studied in [C. M. Elliott and H. Garcke, *SIAM J. Math. Anal.*, 27(2):404–423, 1996], the non-local model appears to take advantage of a larger flexibility on the phase fluxes to dissipate faster the energy, as confirmed by numerical simulations. We prove the existence of a solution to non-local problem by proving the convergence of the JKO minimizing movement scheme.

**Keywords.** Multiphase flow, Cahn-Hilliard type system, constrained Wasserstein gradient flow

**AMS subjects classification.** 35K65, 35K41, 49J40, 76T99

## 1. MOTIVATION AND PRESENTATION OF THE MODEL

1.1. **Motivation.** The celebrated Cahn-Hilliard model was introduced by Cahn and Hilliard [12] in order to model the evolution of metallic alloys in which each phase aims at being pure and to have a minimal perimeter. A similar approach was proposed by de Gennes [19] to model the segregation of polymer blends. In the two-phase case, only one of the two phases is described with its concentration  $c_1 \in [0, 1]$  and its flux  $\mathbf{J}_1$ . The local conservation of mass

$$\partial_t c_1 + \nabla \cdot \mathbf{J}_1 = 0$$

holds, and the other phase whose concentration  $c_2$  is equal to  $1 - c_1$  is assumed to have a flux equal to  $\mathbf{J}_2 = -\mathbf{J}_1$  (the total flux is therefore equal to  $\mathbf{0}$ ). The flux  $\mathbf{J}_1$  is equal to the opposite of some degenerate mobility times the gradient of the generalized chemical potential:

$$\mathbf{J}_1 = -\eta(c_1)\nabla\mu,$$

with  $\eta(0) = \eta(1) = 0$  (typically,  $\eta(c_1) = mc_1(1 - c_1)$ ). In the above relation,  $\mu$  is the functional derivative of some Ginzburg-Landau energy

$$\mathcal{E}(c) = \int \left( \frac{\alpha}{2} |\nabla c|^2 + F(c) \right) d\mathbf{x},$$

where, in absence of thermal agitation,  $F : [0, 1] \rightarrow \mathbb{R}$  is assumed to be concave with minima in 0 and 1, typically  $F(c) = \chi c(1 - c)$  (see for instance [7]). This modeling has been justified by Gurtin [27] by assuming the presence of micro-forces in the material.

The Cahn-Hilliard model has been used later on in the context of multiphase fluid flows because of its capacity to separate different phases (see for instance [9, 10]). In this setting, rather than the micro-force balance, it is natural to assume that the total flux  $\mathbf{J}_{\text{tot}} = \mathbf{J}_1 + \mathbf{J}_2$  is divergence free but not equal to  $\mathbf{0}$  in general. In this work, we propose a model of Cahn-Hilliard type where the total flux  $\mathbf{J}_{\text{tot}}$  is not equal to  $\mathbf{0}$ . This nonlocal model is closely related to the model derived

by E and Palffy-Muhoray in [22]. This model is obtained as a Wasserstein gradient flow of the Ginzburg-Landau energy but in the situation where both phases are allowed to move independently under the constraint that  $c_1 + c_2 = 1$ . This property was depicted by Otto and E in [40] but without thorough justification.

The nonlocal model is derived formally in Section 1.2, then compared to the classical (or local) degenerate Cahn-Hilliard model in Section 1.3. Numerical illustrations of its behavior are given in Section 1.4. In Section 2, we introduce the necessary material to prove our main result, that is the existence of a weak solution to our model. This existence result is obtained by showing the convergence of a minimizing movement scheme *à la* Jordan, Kinderlehrer, and Otto [29]. Finally, Section 3 is devoted to the proof of the convergence of the minimizing movement scheme.

**1.2. Derivation of the model.** We consider an incompressible mixture composed of two phases flowing within an open convex subset  $\Omega$  of  $\mathbb{R}^d$  with  $d \leq 3$ . The fluid is incompressible, so its composition at time  $t \geq 0$  is fully described by the saturations  $c_i(\mathbf{x}, t) \in [0, 1]$ ,  $i \in \{1, 2\}$ , i.e., the volume ratio of the phase  $i$  in the fluid. This leads to the constraint

$$(1) \quad c_1 + c_2 = 1 \quad \text{in } \Omega \times (0, \infty).$$

The motions of each phase is governed by a linear transport equation

$$(2) \quad \partial_t c_i + \nabla \cdot (c_i \mathbf{v}_i) = 0 \quad \text{in } \Omega \times (0, \infty),$$

where  $\mathbf{v}_i : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}^d$  denotes the speed of the phase  $i$ , and  $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2)$ . At the initial time  $t = 0$ , the composition of the mixture is given by  $\mathbf{c}^0 = (c_1^0, c_2^0) : \Omega \rightarrow [0, 1]^2$ , and we assume that

$$(3) \quad c_1^0 + c_2^0 = 1 \quad \text{in } \Omega.$$

We assume that

$$0 < \int_{\Omega} c_1^0 d\mathbf{x} < |\Omega|,$$

excluding the trivial situation where  $c_1 \equiv 1$  or  $c_1 \equiv 0$ . The boundary  $\partial\Omega$  of  $\Omega$  is impervious, hence the speed  $\mathbf{v}_i$  belongs to

$$\mathcal{Z}(\mathbf{c}) = \left\{ \mathbf{V} : \Omega \rightarrow (\mathbb{R}^d)^2 \mid c_i \mathbf{v}_i \cdot \mathbf{n} = 0 \right\},$$

where  $\mathbf{n}$  denotes the outward normal to  $\partial\Omega$ . As a consequence, the volume of the phase  $i$  is conserved along time:

$$(4) \quad \int_{\Omega} c_i(\mathbf{x}, t) d\mathbf{x} = \int_{\Omega} c_i^0(\mathbf{x}) d\mathbf{x}, \quad t \geq 0, \quad i \in \{1, 2\}.$$

At each time  $t \geq 0$ , the saturations  $\mathbf{c}(t)$  belongs to the set

$$\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2,$$

of the admissible saturation states, where

$$\mathcal{A}_i = \left\{ c_i \in L^1(\Omega; \mathbb{R}_+) \mid \int_{\Omega} c_i(\mathbf{x}) d\mathbf{x} = \int_{\Omega} c_i^0(\mathbf{x}) d\mathbf{x} \right\}, \quad i \in \{1, 2\}.$$

To any configuration  $\mathbf{c} \in \mathcal{A}$ , we can associate an energy

$$\mathcal{E}(\mathbf{c}) = \mathcal{E}_{\text{Dir}}(\mathbf{c}) + \mathcal{E}_{\text{chem}}(\mathbf{c}) + \mathcal{E}_{\text{therm}}(\mathbf{c}) + \mathcal{E}_{\text{cons}}(\mathbf{c}) + \mathcal{E}_{\text{ext}}(\mathbf{c}).$$

The Dirichlet energy

$$\mathcal{E}_{\text{Dir}}(\mathbf{c}) = \begin{cases} \sum_{i \in \{1,2\}} \frac{\alpha_i}{2} \int_{\Omega} |\nabla c_i|^2 d\mathbf{x} & \text{if } \mathbf{c} \in H^1(\Omega)^2, \\ +\infty & \text{otherwise;} \end{cases} \quad \alpha_i \geq 0,$$

penalizes the variations of the saturation profiles. The chemical energy is assumed to be proportional to the saturation of both phases:

$$\mathcal{E}_{\text{chem}}(\mathbf{c}) = \chi \int_{\Omega} c_1 c_2 d\mathbf{x}, \quad \text{with } \chi > 0.$$

The term  $\mathcal{E}_{\text{therm}}$  originates from the thermal agitation and is given by

$$\mathcal{E}_{\text{therm}}(\mathbf{c}) = \sum_{i \in \{1,2\}} \theta_i \int_{\Omega} H(c_i) d\mathbf{x}, \quad \theta_i \geq 0$$

where  $H$  is the Boltzmann entropy

$$H(c) = c \log(c) - c + 1 \geq 0, \quad \forall c \geq 0.$$

The case  $\theta_1 = \theta_2 = 0$  is called the deep quench limit [11]. The constraint (1) is incorporated directly in the energy by setting

$$\mathcal{E}_{\text{cons}}(\mathbf{c}) = \int_{\Omega} E_{\text{cons}}(\mathbf{c}) d\mathbf{x}, \quad \text{where } E_{\text{cons}}(\mathbf{c}) = \begin{cases} 0 & \text{if } c_1 + c_2 = 1, \\ +\infty & \text{otherwise.} \end{cases}$$

The component  $\mathcal{E}_{\text{ext}}$  comes from an exterior potential that acts on the volume of the fluid, like for instance gravity or electrostatic forces. We assume that there exist external potentials  $\Psi_i \in W^{1,1}(\Omega)$  with  $\Psi_1 - \Psi_2 \in L^2(\Omega)$  such that

$$\mathcal{E}_{\text{ext}}(\mathbf{c}) = \sum_{i \in \{1,2\}} \int_{\Omega} c_i(\mathbf{x}) \Psi_i(\mathbf{x}) d\mathbf{x}.$$

The chemical part  $\mathcal{E}_{\text{chem}}$  of the energy functional is not convex but is smooth, whereas the other components  $\mathcal{E}_{\text{Dir}}$ ,  $\mathcal{E}_{\text{cons}}$ , and  $\mathcal{E}_{\text{therm}}$  are all convex (in the usual linear sense:  $\mathcal{E}_{\text{Dir}}$  and  $\mathcal{E}_{\text{cons}}$  are not displacement convex in McCann's sense [37]). As a consequence, the energy functional  $\mathcal{E}$  admits a (local) subdifferential  $\partial\mathcal{E}(\mathbf{c})$  at each admissible  $\mathbf{c} \in \mathcal{A}$  defined by

$$\partial\mathcal{E}(\mathbf{c}) = \left\{ \mathbf{w} = (w_1, w_2) : \Omega \rightarrow \mathbb{R}^2 \mid \mathcal{E}(\hat{\mathbf{c}}) - \mathcal{E}(\mathbf{c}) - \int_{\Omega} \mathbf{w} \cdot (\hat{\mathbf{c}} - \mathbf{c}) d\mathbf{x} \geq o(\|\hat{\mathbf{c}} - \mathbf{c}\|_{L^1(\Omega)}) \right\}.$$

Denote by  $H^*(\Omega)$  the dual set of  $\{u \in H^1(\Omega) \mid \int_{\Omega} u d\mathbf{x} = 0\}$ , and by

$$\mathcal{X} = \{\mathbf{c} \in H^1(\Omega; \mathbb{R}_+)^2 \mid c_1 + c_2 = 1 \text{ a.e. in } \Omega\}$$

the domain of  $\mathcal{E}$ , i.e.,

$$(5) \quad \mathcal{E}(\mathbf{c}) < \infty \quad \Leftrightarrow \quad \mathbf{c} \in \mathcal{X}.$$

It follows from elementary calculations (similar to those of [14] in the context of porous media flows) that the subdifferential  $\partial\mathcal{E}(\mathbf{c})$  of  $\mathcal{E}$  at  $\mathbf{c} \in \mathcal{A} \cap \mathcal{X}$  is made of the elements  $\mathbf{w} = (w_1, w_2)$  of  $(H^*(\Omega))^2$  such that

$$(6) \quad w_1 - w_2 = -\alpha \Delta c_1 + \chi(1 - 2c_1) + f(c_1) + (\Psi_1 - \Psi_2),$$

where we have set  $\alpha = \alpha_1 + \alpha_2$  and

$$(7) \quad f(c_1) = \log \left( \frac{c_1^{\theta_1}}{(1-c_1)^{\theta_2}} \right) = \theta_1 \log(c_1) - \theta_2 \log(c_2).$$

We define then the chemical potential  $\mu_i$  of the phase  $i$  by

$$(8) \quad \mu_i = w_i - \theta_i \log(c_i) - \Psi_i, \quad i \in \{1, 2\},$$

so that (6) turns to

$$(9) \quad \mu_1 - \mu_2 = -\alpha \Delta c_1 + \chi(1 - 2c_1).$$

Until now,  $\boldsymbol{\mu}$  is only defined up to an additive constant. This degree of freedom is eliminated by imposing for almost all  $t > 0$  that

$$(10) \quad \int_{\Omega} \bar{\mu}(\mathbf{x}, t) d\mathbf{x} = 0, \quad \text{where } \bar{\mu} = c_1 \mu_1 + c_2 \mu_2.$$

The relation (9) is complemented by homogeneous Neumann boundary conditions

$$(11) \quad -\nabla c_1 \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega \times \mathbb{R}_+.$$

We are interested in the rigorous derivation of the PDEs that govern the gradient flow of the energy  $\mathcal{E}(\mathbf{c})$  in  $\mathcal{A}$  endowed with a tensorized Wasserstein distance to be introduced in Section 2.1. To this end, we prove the convergence of the minimizing movement (or JKO) scheme [29]. But let us first provide formal calculations based on the framework of generalized gradient flows of [38, 41, 14] in order to identify the underlying PDEs.

The motion of the phases induces a viscous dissipation given by

$$(12) \quad \mathfrak{D}(\mathbf{c}, \mathbf{V}) = \sum_{i \in \{1, 2\}} \frac{1}{2m_i} \int_{\Omega} c_i |\mathbf{v}_i|^2 d\mathbf{x}, \quad \forall \mathbf{c} \in \mathcal{A}, \forall \mathbf{v} \in \mathcal{Z}^2,$$

where  $m_i$  is the mobility coefficient of the phase  $i$ . We suppose as in [14] that at each time  $t \geq 0$ , the phase speeds  $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2)$  is selected by the following *steepest descent condition*:

$$\mathbf{V} \in \operatorname{argmin}_{\mathbf{V} \in \mathcal{Z}(\mathbf{c})} \left( \mathfrak{D}(\mathbf{c}, \mathbf{V}) + \max_{\mathbf{w} \in \partial \mathcal{E}(\mathbf{c})} \sum_{i \in \{1, 2\}} \int_{\Omega} c_i \mathbf{v}_i \cdot \nabla w_i d\mathbf{x} \right).$$

Assume that the min and the max can be swapped in the above formula, then optimizing over  $\mathbf{V}$  first yields

$$(13) \quad \mathbf{v}_i = -m_i \nabla w_i \quad \text{on } \{c_i > 0\} \text{ for } i \in \{1, 2\}.$$

Maximizing the result over the elements  $\mathbf{w} \in \partial \mathcal{E}(\mathbf{c})$  leads to the elliptic equation

$$-\nabla \cdot \left( \sum_{i \in \{1, 2\}} m_i c_i \nabla w_i \right) = 0.$$

The above equation combined with (2) and (13) implies that  $\partial_t(c_1 + c_2) \equiv 0$ , hence that the algebraic constraint (1) on the saturation remains satisfied along time. To sum up, the system of

partial differential equations corresponding to the variational modeling presented above is obtained by combining the equations (1), (2), (6), (11), and (13). It leads to the system

$$(14) \quad \begin{cases} \partial_t c_i - \nabla \cdot (m_i c_i \nabla w_i) = 0 & \text{for } i \in \{1, 2\}, \\ c_1 + c_2 = 1, \\ w_1 - w_2 = -\alpha \Delta c_1 + \chi(1 - 2c_1) + f(c_1) + (\Psi_1 - \Psi_2), \end{cases}$$

to be satisfied in some appropriate sense in  $\Omega \times (0, \infty)$ . Introducing the chemical potential  $\mu_i$  as in (8), the problem rewrites

$$\begin{cases} \partial_t c_i - \nabla \cdot (m_i c_i \nabla (\mu_i + \Psi_i)) = m_i \theta_i \Delta c_i & \text{for } i \in \{1, 2\}, \\ c_1 + c_2 = 1, \\ \mu_1 - \mu_2 = -\alpha \Delta c_1 + \chi(1 - 2c_1). \end{cases}$$

The system is complemented with initial conditions

$$(15) \quad (c_i)_{|t=0} = c_i^0 \in H^1(\Omega) \text{ with } c_i^0 \geq 0 \text{ and } c_1^0 + c_2^0 = 1 \text{ in } \Omega,$$

and boundary conditions

$$(16) \quad c_i \nabla (\mu_i + \Psi_i) \cdot \mathbf{n} = \nabla c_i \cdot \mathbf{n} = 0 \text{ on } \partial\Omega \times (0, \infty).$$

In the next section, we highlight some differences between the non-local model (14) and the local degenerate Cahn-Hilliard model that has been studied for instance in [23].

**1.3. Comparison with the classical degenerate Cahn-Hilliard model.** Even in the simple situation where the external potentials  $\Psi_i$  are equal to 0 and where  $\theta_1 = \theta_2 = 0$ , the system (14) differs as soon as  $d \geq 2$  from the local degenerate Cahn-Hilliard model that can be written as

$$(17) \quad \partial_t c - \nabla \cdot (\eta(c) \nabla \mu) = 0, \quad \mu = -\alpha \Delta c + \chi(1 - 2c), \quad \eta(c) = \frac{m_1 m_2 c(1 - c)}{m_1 c + m_2(1 - c)}.$$

We refer to [23] for the existence of weak solutions to (17) (complemented with suitable boundary conditions) and to [4] for the extension of the model to the case of  $N$  phases ( $N \geq 3$ ). Here,  $\mu$  is the generalized chemical potential that is defined as the difference of the phase chemical potentials.

The energy  $\tilde{\mathcal{E}}$  associated to (17) is similar to the one of our problem, i.e.,

$$\tilde{\mathcal{E}}(c) = \frac{\alpha}{2} \int_{\Omega} |\nabla c|^2 d\mathbf{x} + \chi \int_{\Omega} c(1 - c) d\mathbf{x}.$$

But both the equation governing the motion (2) and the dissipation (12) have to be modified. More precisely, the continuity equation (2) must be replaced by its nonlinear counterpart

$$\partial_t c + \nabla \cdot (\eta(c) \mathbf{v}) = 0,$$

while the dissipation is now given by

$$\tilde{\mathcal{D}}(c, \mathbf{v}) = \int_{\Omega} \eta(c) |\mathbf{v}|^2 d\mathbf{x}.$$

Therefore, the PDEs (17) can still be interpreted as the gradient flow of the energy  $\tilde{\mathcal{E}}$ , but the geometry is different: rather than considering some classical quadratic Wasserstein distance for each phase and to constrain the sum of the concentrations to be equal to 1 (as it will be the case for our approach), the set

$$\mathcal{A} = \left\{ c \in L^1(\Omega; \mathbb{R}_+) \mid \int_{\Omega} c(\mathbf{x}, t) d\mathbf{x} = \int_{\Omega} c_1^0(\mathbf{x}) d\mathbf{x} \right\}$$

has to be equipped with the weighted Wasserstein metric corresponding to the concave mobility  $\eta$ . We refer to [21] for the description of the corresponding metric and to [33] for the rigorous recovery of (17) by a gradient flow approach.

The difference between the non-local model (14) and the local one (17) can also be seen as follows. Summing the first equation of (14) for  $i \in \{1, 2\}$  yields

$$(18) \quad \nabla \cdot \mathbf{J}_{\text{tot}} = 0, \quad \text{where} \quad \mathbf{J}_{\text{tot}} = \sum_{i \in \{1,2\}} c_i \mathbf{v}_i = - \sum_{i \in \{1,2\}} m_i c_i \nabla \mu_i.$$

The equation for  $c_1$  can then be rewritten

$$(19) \quad \partial_t c_1 + \nabla \cdot (\rho(c_1) \mathbf{J}_{\text{tot}} - \eta(c_1) \nabla (\mu_1 - \mu_2)) = 0, \quad \text{with} \quad \rho(c) = \frac{m_1 c}{m_1 c + m_2 (1 - c)}.$$

Thus our model (14) boils down to the local Cahn-Hilliard equation as soon as  $\mathbf{J}_{\text{tot}} \equiv \mathbf{0}$ . This is the case when  $d = 1$  because of (18), but no longer if  $d \geq 2$ . Since our non-local model does not impose that  $\mathbf{J}_1 = -\mathbf{J}_2$ , it allows for additional motions. These motions—corresponding to the transport term  $\nabla \cdot (\rho(c_1) \mathbf{J}_{\text{tot}})$  in (19)—contribute to the dissipation as shows the formula

$$\mathfrak{D}(\mathbf{c}, -\nabla \mu) = \int_{\Omega} \frac{|\mathbf{J}_{\text{tot}}|^2}{m_1 c_1 + m_2 c_2} d\mathbf{x} + \tilde{\mathfrak{D}}(c, -\nabla (\mu_1 - \mu_2)).$$

Therefore, and as already noticed by Otto and E in [40], the instantaneous dissipation corresponding to a phase configuration  $\mathbf{c}$  is greater for the non-local model (14) than for the local model (17) and the energy decreases faster.

**1.4. Numerical illustration.** The goal of this section is to illustrate the behavior of the model (14) and to compare it with the classical degenerate Cahn-Hilliard problem (17). In order to solve numerically (14) we use an implicit in time finite volume scheme with upstream mobility described in [17] and inspired from the oil engineering context [25]. The mesh is triangular and assumed to fulfill the so-called orthogonality condition [28, 24] (this amounts to requiring the mesh to be Delaunay) so that the diffusive fluxes can be approximated thanks to a two-point flux approximation in a consistent way. As it is exposed in [17], the scheme is positivity preserving (i.e.,  $0 \leq c_{i,h} \leq 1$ ), it is energy diminishing (the discrete counterpart of the energy is decreasing) and entropy stable. It leads to a nonlinear system of algebraic equations to be solved at each time step. It is shown in [17] that this system admits (at least) one solution that is computed thanks to the Newton-Raphson method.

Concerning the problem (17), we use a similar approach, but since the mobility function  $\eta$  is no longer monotone, we have to use an implicit Godunov scheme to discretize it as a generalization of the upstream mobility (see for instance [16]). Here again, the discrete solution remains bounded between 0 and 1, the energy is decreasing and the entropy remains bounded. Here again, the resulting nonlinear system is solved at each time step by the mean of the Newton-Raphson method.

**Remark 1.1.** *Alternative numerical methods have been proposed in order to solve degenerate Cahn-Hilliard problems. We won't perform here an exhaustive list, but let us mention the contributions of Barrett et al. based on conformal finite elements [3, 4]. Even though very efficient, these methods have the drawback of requiring a small stabilization to be tuned following the mesh size. Unless one considers non-smooth energies as in [8], the scheme does not preserve the bounds  $0 \leq c_{i,h} \leq 1$ . These difficulties are overpassed in our approach by using some entropy stable hyperbolic fluxes to discretize the mobilities.*

Since our model has a Wasserstein gradient flow structure, it would be natural to use a Lagrangian method as for instance in [6, 35, 30, 18]. The main problem with this approach is that both phase move with their own speed, therefore such an approach would impose to move two meshes simultaneously. It is then rather unclear how to manage the constraint (1) in this case. For this reason, it seems more suitable to stick to an Eulerian description. An alternative approach to solve numerically our problem would therefore be to adapt the ALG2-JKO algorithm of Benamou et al. [5] to our setting.

We propose two different test cases that will allow to illustrate the difference between the local model (17) and the non-local model (14). For both of them, we do not consider any exterior potential, i.e.,  $\Psi_i = 0$ , and we neglect the thermal diffusion, i.e.,  $\theta_i = 0$ . Both phase mobilities  $m_i$  are assumed to be equal to 1, and we choose  $\alpha = 3.10^{-4}$  and  $\chi = 0.96$ .

1.4.1. *Test case 1: from a cross to a circle.* We start from an initial data that is the characteristic function of a cross. Since  $\alpha \ll \chi$ , it follows from the Modica and Mortola' result [39] that the free energy is close to the perimeter of a characteristic set (up to a multiplicative constant). This means that (up to a small regularization) both the local and the non-local Cahn-Hilliard models aim at minimizing the perimeter of the sets  $\{c_1 = 0\}$  and  $\{c_1 = 1\}$  corresponding to pure phases. Since the non-local model allows for more movements (cf. Section 1.3), the energy (thus the perimeter) should decay faster for the nonlocal model. This is indeed what we observe on Figures 1 and 2.

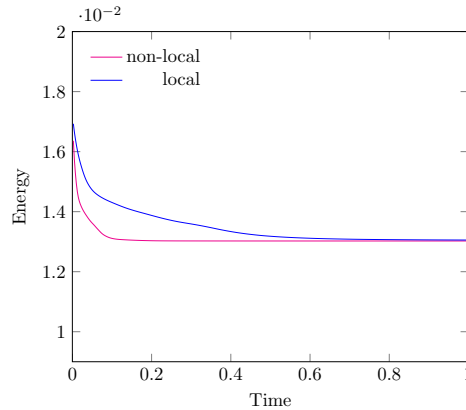


FIGURE 1. Evolution of the energy  $\mathcal{E}(\mathbf{c})$  along time for the non-local model (14) and the local one (17). The decay of the energy is faster for the non-local model, as shown in Section 1.3.

1.4.2. *Test case 2: Spinodal decomposition.* Similarly to the local model, the non-local model is able to reproduce the spinodal decomposition for mixtures. In order to illustrate this fact, we start from an initial data which consists in a constant concentration plus a small random perturbation:

$$c_1^0(\mathbf{x}) = 0.5 + r(\mathbf{x}), \quad r \ll 1.$$

Since  $c_1 = 0.5$  is very unfavorable from an energetic point of view, both phase will separate very rapidly, letting areas with pure phase appear. Then these area will cluster in order to minimize their perimeter. We plot on Figure 3 some snapshots to illustrate the spinodal decomposition



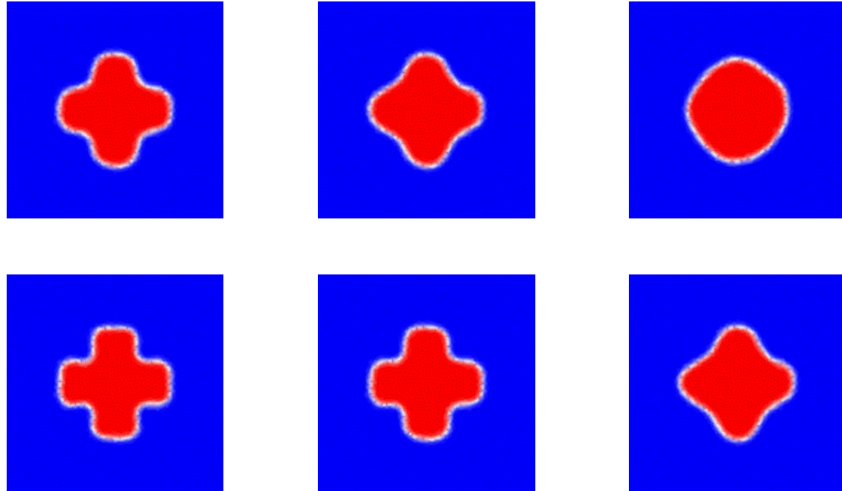


FIGURE 2. Evolution along time of the numerical solution to the non-local problem (14) (top) and to the local problem (17) (bottom). Snapshots at time  $t = 10^{-2}$  (left),  $t = 2.10^{-2}$  (middle), and  $t = 10^{-1}$  (right).

corresponding to models (14) and (17). On Figure 4, we compare the evolution of the energy along time for spinodal decomposition corresponding to both models. As expected, the energy decay is faster for the non-local model than for the local one. But contrarily to Test case 1, the solutions seem to converge towards different steady states.

## 2. WASSERSTEIN GRADIENT FLOW, JKO SCHEME AND MAIN RESULT

**2.1. Wasserstein distance.** As a preliminary to the introduction of the minimizing movement scheme, we introduce some necessary material related to Wasserstein (or Monge-Kantorovich) distances between nonnegative measures of prescribed mass that are absolutely continuous w.r.t. to the Lebesgue measure. We refer to Santambrogio's monograph [42] for an introduction to optimal transportation and to the Wasserstein distances, and to Villani's big book [43] for a more complete presentation.

Given two elements  $c_i$  and  $\check{c}_i$  of  $\mathcal{A}_i$  ( $i \in \{1, 2\}$ ), a map  $\mathbf{t} : \Omega \rightarrow \Omega$  is said to send  $c_i$  on  $\check{c}_i$  (we write  $\check{c}_i = \mathbf{t}\#c_i$ ) if

$$\int_A \check{c}_i(\mathbf{x})d\mathbf{x} = \int_{\mathbf{t}^{-1}(A)} c_i(\mathbf{x})d\mathbf{x}, \quad \text{for all Borel subset } A \text{ of } \Omega.$$

The Wasserstein distance  $W_i(c_i, \check{c}_i)$  with quadratic cost function between  $c_i$  and  $\check{c}_i$  is then defined by

$$(20) \quad W_i(c_i, \check{c}_i) = \inf_{\mathbf{t} \text{ s.t. } \check{c}_i = \mathbf{t}\#c_i} \left( \frac{1}{m_i} \int_{\Omega} |\mathbf{x} - \mathbf{t}(\mathbf{x})|^2 c_i(\mathbf{x})d\mathbf{x} \right)^{1/2}, \quad i \in \{1, 2\}.$$

In (20), the infimum is in fact a minimum, and  $\mathbf{t}$  is the gradient of a convex function. In our context of fluid flows, the cost for moving the mass of the phase  $i$  from a configuration  $c_i$  to another

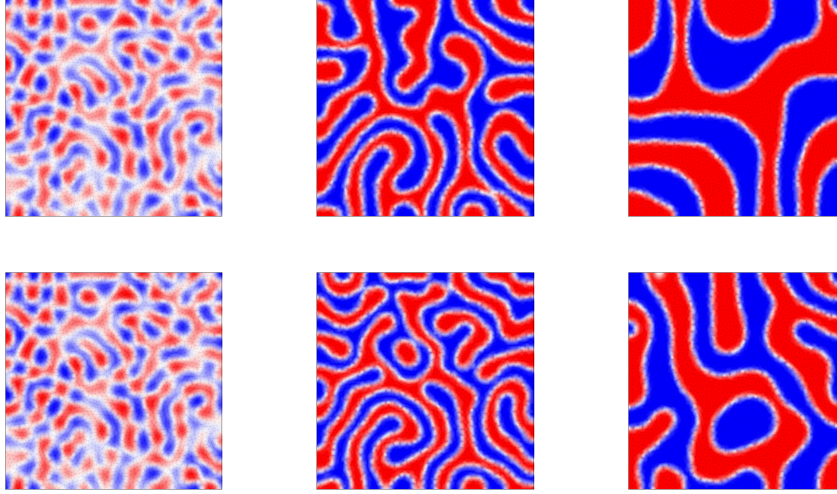


FIGURE 3. Snapshots at times  $t = 6 \cdot 10^{-3}$  (left),  $t = 5 \cdot 10^{-2}$  (middle), and  $t = 1$  (right) illustrating the spinodal decomposition governed by model (14) (top) and model (17) (bottom).

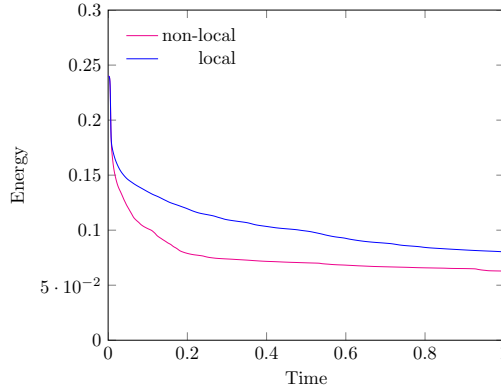


FIGURE 4. Evolution of the energy  $\mathcal{E}(\mathbf{c})$  along time for the non-local model (14) and the local one (17) for the spinodal decomposition test case.

configuration  $\check{c}_i$  regardless to the other phase is equal to  $W_i^2(c_i, \check{c}_i)$ . The multiplying factor  $1/m_i$  is natural since the more mobile is the phase, the less expensive are its displacements. We can then define the tensorized Wasserstein distance  $\mathbf{W}$  on  $\mathcal{A}$  by

$$\mathbf{W}^2(\mathbf{c}, \check{\mathbf{c}}) = W_1^2(c_1, \check{c}_1) + W_2^2(c_2, \check{c}_2), \quad \forall \mathbf{c}, \check{\mathbf{c}} \in \mathcal{A}.$$

In the core of the proof, we will make an extensive use of the Kantorovich dual problem. More precisely, we will use the fact that

$$(21) \quad \frac{1}{2}W_i^2(c_i, \check{c}_i) = \sup_{\substack{\varphi_i \in L^1(c_i), \psi_i \in L^1(\check{c}_i) \\ \varphi_i(\mathbf{x}) + \psi_i(\mathbf{y}) \leq \frac{|\mathbf{x} - \mathbf{y}|^2}{2m_i}}} \int_{\Omega} c_i(\mathbf{x})\varphi_i(\mathbf{x})d\mathbf{x} + \int_{\Omega} \check{c}_i(\mathbf{y})\psi_i(\mathbf{y})d\mathbf{y}.$$

Here,  $L^1(\rho)$  denotes the sets of integrable functions for the measure with density  $\rho$ . Here again, the supremum is in fact a maximum, and the Kantorovich potentials  $(\varphi_i, \psi_i)$  achieving the sup in (21) are  $dc_i \otimes d\check{c}_i$  unique up to an additive constant. The optimal transportation  $\mathbf{t}_i$  sending  $c_i$  on  $\check{c}_i$  achieving the inf in (20) is related to the Kantorovich potential by

$$\mathbf{t}_i(\mathbf{x}) = \mathbf{x} - m_i \nabla \varphi_i(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega$$

with  $\varphi_i$  achieving the sup in (21). As a consequence, the formula (20) provides

$$(22) \quad W_i^2(c_i, \check{c}_i) = m_i \int_{\Omega} c_i |\nabla \varphi_i|^2 d\mathbf{x}, \quad i \in \{1, 2\}$$

to be used in the sequel.

**2.2. The JKO scheme and the approximate solution.** We have now all the necessary material at hand to define the minimizing movement scheme. Let  $\tau > 0$  and  $\mathbf{c}^{n-1} \in \mathcal{A} \cap \mathcal{X}$ , then define the functional  $\mathcal{F}_{\tau}^n : \mathcal{A} \rightarrow \mathbb{R} \cup \{+\infty\}$  by setting

$$\mathcal{F}_{\tau}^n(\mathbf{c}) = \mathcal{E}(\mathbf{c}) + \frac{1}{2\tau} \mathbf{W}^2(\mathbf{c}, \mathbf{c}^{n-1}), \quad \forall \mathbf{c} \in \mathcal{A}.$$

The functional  $\mathcal{F}_{\tau}^n$  is bounded from below since all its components are (recall that  $\Psi_i \in L^1(\Omega)$ ). Then we define

$$(23) \quad \mathbf{c}^n \in \operatorname{argmin}_{\mathbf{c} \in \mathcal{A}} \mathcal{F}_{\tau}^n(\mathbf{c}).$$

The existence of such a minimizer is the purpose of the following statement.

**Proposition 2.1.** *For any  $\mathbf{c}^{n-1} \in \mathcal{A} \cap \mathcal{X}$ , there exists (at least) one solution  $\mathbf{c}^n \in \mathcal{A} \cap \mathcal{X}$  to the minimization scheme (23).*

*Proof.* Let  $(\mathbf{c}^{n,k})_{k \geq 0}$  be a minimizing sequence with  $\mathbf{c}^{n,0} = \mathbf{c}^{n-1}$ , then

$$(24) \quad \mathcal{E}(\mathbf{c}^{n,k}) \leq \mathcal{E}(\mathbf{c}^{n-1}) < \infty \quad \forall k \geq 0.$$

We infer from (5) that  $\mathbf{c}^{n,k} \in \mathcal{X}$ , hence  $0 \leq c_i^{n,k} \leq 1$  for all  $k \geq 0$  and  $i \in \{1, 2\}$ . Moreover, it follows from the definition of the energy  $\mathcal{E}$  that  $\|\nabla c_i^{n,k}\|_{L^2(\Omega)} \leq C$  for all  $k \geq 0$ . Hence, using the Poincaré-Wirtinger inequality, we get that  $\|c_i^{n,k}\|_{H^1(\Omega)} \leq C$ . In particular, we can assume that the sequence  $(\mathbf{c}^{n,k})_{k \geq 0}$  converges weakly in  $H^1(\Omega)^2$ , in the  $L^\infty(\Omega)^2$  weak- $\star$  sense and almost everywhere (hence strongly in  $L^p(\Omega)^2$  for all  $p \in (1, +\infty)$ ) towards  $\mathbf{c}^n \in \mathcal{X}$ . It follows from the lower-semi continuity of  $\mathcal{E}$  w.r.t. the weak- $\star$  topology of  $L^\infty \cap H^1(\Omega)^2$  (all the components of  $\mathcal{E}$  are continuous, excepted  $\mathcal{E}_{\text{Dir}}$  that is merely l.s.c.) that

$$\mathcal{E}(\mathbf{c}^n) \leq \liminf_{k \rightarrow \infty} \mathcal{E}(\mathbf{c}^{n,k}).$$

Besides, since  $\mathbf{c}^{n,k}$  tends to  $\mathbf{c}^n$  in  $L^1(\Omega)^2$ , and thanks to the lower semi-continuity of  $\mathbf{W}$  w.r.t. the topology induced on  $\mathcal{A}$  by  $L^1(\Omega)^2$ , then

$$\mathbf{W}(\mathbf{c}^n, \mathbf{c}^{n-1}) \leq \liminf_{k \rightarrow \infty} \mathbf{W}(\mathbf{c}^{n,k}, \mathbf{c}^{n-1}).$$

As a consequence,  $\mathbf{c}^n$  is a minimizer of  $\mathcal{F}_\tau^n$ .  $\square$

From a sequence  $(\mathbf{c}^n)_{n \geq 1}$  of iterated solutions to the scheme (23), we deduce a discrete solution  $\mathbf{c}_\tau : \mathbb{R}_+ \rightarrow \mathcal{A} \cap \mathcal{X}$  defined by

$$(25) \quad \mathbf{c}_\tau(0) = \mathbf{c}^0, \quad \mathbf{c}_\tau(t) = \mathbf{c}^n \text{ if } t \in ((n-1)\tau, n\tau].$$

We can also build approximate phase potentials  $(\boldsymbol{\mu}^n)_{n \geq 1}$ . More precisely, for all  $n \geq 1$ , let  $\boldsymbol{\varphi}^n = (\varphi_1^n, \varphi_2^n)$  and  $\boldsymbol{\psi}^n = (\psi_1^n, \psi_2^n)$  be Kantorovich potentials sending  $\mathbf{c}^n$  on  $\mathbf{c}^{n-1}$ , i.e.,

$$(26) \quad \frac{1}{2} \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) = \sum_{i \in \{1,2\}} \left( \int_{\Omega} c_i^n \varphi_i^n d\mathbf{x} + \int_{\Omega} c_i^{n-1} \psi_i^n d\mathbf{y} \right),$$

then  $\varphi_i^n$  and  $\psi_i^n$  can be shifted vertically by some arbitrary  $\alpha_i^n \in \mathbb{R}$  without changing the value of the right-hand side in the above expression. As it will be established in the proof of Lemma 3.5, there are suitable vertical shifts and functions  $F_i^n$  such that  $F_i^n = 0$  a.e. in  $\{c_i > 0\}$  such that

$$(27) \quad \mu_i^n := -\frac{\varphi_i^n}{\tau} - \Psi_i - \theta_i \log(c_i^n) + F_i^n$$

satisfies

$$\mu_1^n - \mu_2^n = -\alpha \Delta c_1^n + \chi c_1^n c_2^n \quad \text{and} \quad \int_{\Omega} (c_1^n \mu_1^n + c_2^n \mu_2^n) d\mathbf{x} = 0.$$

We can then define  $\boldsymbol{\mu}_\tau(t) = (\mu_{1,\tau}(t), \mu_{2,\tau}(t))$  for  $t > 0$  by setting  $\mu_{i,\tau}(t) = \mu_i^n$  for  $t \in ((n-1)\tau, n\tau]$ .

**2.3. Weak solutions.** The goal of this section is to state our main result, that is the convergence of the JKO scheme. It requires the introduction of the notion of weak solution that will be obtained at the limit when the approximation parameter  $\tau$  tends to 0.

**Definition 2.2.**  $(\mathbf{c}, \boldsymbol{\mu})$  is said to be a weak solution to the problem (14), (15), and (16) if

- the phase concentrations  $c_i$  belong to  $L^\infty(\mathbb{R}_+; H^1(\Omega)) \cap L_{\text{loc}}^2((0, T); H^2(\Omega)) \cap C(\mathbb{R}_+; L^2(\Omega))$  with  $c_i(0, \cdot) = c_i^0$ ,  $\nabla c_i \cdot \mathbf{n} = 0$  on  $\partial\Omega \times \mathbb{R}_+$  and  $c_1 + c_2 = 1$  a.e. in  $\Omega \times \mathbb{R}_+$ ;
- the phase potentials  $\mu_i$  belong to  $L_{\text{loc}}^2(\mathbb{R}_+; L^{d/d-1}(\Omega))$  and are such that  $\sqrt{c_i} \nabla \mu_i \in L^2(\mathbb{R}_+; L^2(\Omega))^d$  and

$$\int_{\Omega} (c_1(\mathbf{x}, t) \mu_1(\mathbf{x}, t) + c_2(\mathbf{x}, t) \mu_2(\mathbf{x}, t)) d\mathbf{x} = 0, \quad \text{for a.e. } t \geq 0;$$

- the relation (9) on the difference of the phase potentials holds almost everywhere in  $\Omega \times \mathbb{R}_+$ ;
- for all  $\xi \in C^2(\bar{\Omega})$  and all  $t_1, t_2 \in \mathbb{R}_+$  with  $t_2 \geq t_1$ , there holds

$$(28) \quad \int_{\Omega} (c_i(\mathbf{x}, t_2) - c_i(\mathbf{x}, t_1)) \xi(\mathbf{x}) d\mathbf{x} + m_i \int_{t_1}^{t_2} \int_{\Omega} (c_i \nabla(\mu_i + \Psi_i) + \theta_i \nabla c_i) \cdot \nabla \xi d\mathbf{x} dt = 0.$$

Here is the convergence theorem for the minimization scheme. The existence of a weak solution is a by-product.

**Theorem 2.3** (Convergence of the minimizing movement scheme). *Let  $(\tau_n)_{n \geq 1} \subset (0, 1)$  be a sequence tending to 0 as  $n$  tends to  $\infty$ , and let  $(\mathbf{c}_{\tau_n}, \boldsymbol{\mu}_{\tau_n})_{n \geq 1}$  be a corresponding sequence of approximate solution, then, up to the extraction of a subsequence,*

$$\begin{aligned} c_{i, \tau_n} &\xrightarrow[n \rightarrow \infty]{} c_i \quad \text{in the } L^\infty((0, T); H^1(\Omega))\text{-weak-}\star \text{ sense,} \\ \|c_{i, \tau_n}(\cdot, t) - c_i(\cdot, t)\|_{L^2(\Omega)} &\xrightarrow[n \rightarrow \infty]{} 0 \quad \text{for all } t \in [0, T], \\ c_{i, \tau_n} &\xrightarrow[n \rightarrow \infty]{} c_i \quad \text{in } L^2((0, T); W^{1, d}(\Omega)), \\ c_{i, \tau_n} &\xrightarrow[n \rightarrow \infty]{} c_i \quad \text{weakly in } L^2((0, T); H^2(\Omega)), \\ \mu_{i, \tau_n} &\xrightarrow[n \rightarrow \infty]{} \mu_i \quad \text{for the weak topology of } L^2((0, T); L^{d/d-1}(\Omega)), \\ c_{i, \tau_n} \nabla \mu_{i, \tau_n} &\xrightarrow[n \rightarrow \infty]{} c_i \nabla \mu_i \quad \text{weakly in } L^2(Q_T)^d, \end{aligned}$$

where  $(\mathbf{c}, \boldsymbol{\mu})$  is a weak solution in the sense of Definition 2.2.

The approximation of the weak solution thanks to a minimizing movement scheme provides a rigorous foundation to the fact that our model can be reinterpreted as a gradient flow in the metric space  $\mathcal{A}$  endowed with the Wasserstein metric  $\mathbf{W}$ . The remaining of the paper is devoted to proof of the Theorem 2.3. It is based on compactness arguments applied to the sequence  $(\mathbf{c}_\tau, \boldsymbol{\mu}_\tau)_{\tau > 0}$ .

### 3. PROOF OF THEOREM 2.3

We first establish some estimates on the approximate solution  $\mathbf{c}_\tau$ . The very classical energy estimate and some straightforward consequences are stated in Sections 3.1. In Section 3.2, we show that the approximate solution remains bounded away from 0 and 1 if the thermal diffusion coefficients  $\theta_i$  are positive. Section 3.3 where the flow interchange technique initially introduced in [34] is applied. The Euler-Lagrange equation are then obtained in Section 3.4 thanks to a linearization technic inspired from the work of Maury *et al.* [36]. The convergence of the approximate solution is finally established in Section 3.5.

**3.1. Energy and distance estimate.** Testing by  $\mathbf{c}^{n-1}$  in (23) yields

$$(29) \quad \mathcal{E}(\mathbf{c}^n) + \frac{1}{2\tau} \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) \leq \mathcal{E}(\mathbf{c}^{n-1}), \quad \forall n \geq 1.$$

Summing (29) over  $n$ , and using that  $\mathcal{E}(\mathbf{c}) \geq \mathcal{E}_\star > -\infty$  for all  $\mathbf{c} \in \mathcal{A}$ , we obtain the square distance estimate

$$(30) \quad \frac{1}{\tau} \sum_{n \geq 1} \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) \leq 2(\mathcal{E}(\mathbf{c}^0) - \mathcal{E}_\star) < +\infty.$$

This readily gives the approximate 1/2-Hölder estimate

$$(31) \quad \mathbf{W}(\mathbf{c}^{n_2}, \mathbf{c}^{n_1}) \leq C \sqrt{|n_2 - n_1| \tau}, \quad \forall n_1, n_2 \geq 0.$$

Bearing in mind the definition (25) of the approximate solution  $\mathbf{c}_\tau$ , we get that

$$(32) \quad \mathbf{W}(\mathbf{c}_\tau(t), \mathbf{c}_\tau(s)) \leq C \sqrt{|t - s| + \tau}, \quad \forall s, t \geq 0.$$

We also deduce from the energy estimate (29) that

$$(33) \quad \mathcal{E}(\mathbf{c}^n) \leq \mathcal{E}(\mathbf{c}^0) < \infty.$$

We deduce that

$$(34) \quad \int_{\Omega} |\nabla c_1^n|^2 d\mathbf{x} \leq \frac{2}{\alpha} \left( \mathcal{E}(c^0) + \sum_{i \in \{1,2\}} \|\Psi_i\|_{L^1} \right) < \infty, \quad \forall n \geq 0.$$

**3.2. Positivity of the discrete solution in presence of thermal agitation.** The formula (27) suggests to give a proper sense to the quantity  $\theta_i \log(c_i^n)$ . This is the purpose of the following lemma, which is an adaptation to our framework of [42, Lemma 8.6].

**Lemma 3.1.** *Let  $\mathbf{c}^n$  be a minimizer of  $\mathcal{F}_\tau^n$  as in (23). Assume that  $\theta_i > 0$ , then  $c_i^n > 0$  a.e. in  $\Omega$ . Moreover,  $\theta_i \log(c_i^n) \in L^1(\Omega)$ .*

*Proof.* We define  $\bar{\mathbf{c}} = (\bar{c}_1, \bar{c}_2)$  the constant element of  $\mathcal{X} \cap \mathcal{A}$  given by

$$\bar{c}_i = \frac{1}{|\Omega|} \int_{\Omega} c_i^0 d\mathbf{x} > 0.$$

Then given  $\epsilon \in (0, 1)$ , we introduce  $\mathbf{c}^\epsilon \in \mathcal{X} \cap \mathcal{A}$  which is defined by

$$(35) \quad c_i^\epsilon = \epsilon \bar{c}_i + (1 - \epsilon) c_i^n = c_i^n + \epsilon(\bar{c}_i - c_i^n).$$

Note that  $c_i^\epsilon > 0$  everywhere in  $\Omega$ . By optimality of  $\mathbf{c}^n$ , the inequality

$$\begin{aligned} \mathcal{E}_{\text{therm}}(\mathbf{c}^n) - \mathcal{E}_{\text{therm}}(\mathbf{c}^\epsilon) &\leq \mathcal{E}_{\text{Dir}}(\mathbf{c}^\epsilon) - \mathcal{E}_{\text{Dir}}(\mathbf{c}^n) + \mathcal{E}_{\text{chem}}(\mathbf{c}^\epsilon) - \mathcal{E}_{\text{chem}}(\mathbf{c}^n) \\ &\quad + \mathcal{E}_{\text{ext}}(\mathbf{c}^\epsilon) - \mathcal{E}_{\text{ext}}(\mathbf{c}^n) + \frac{1}{2\tau} (\mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) - \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1})) \end{aligned}$$

holds for all  $\epsilon \in (0, 1)$ . It follows from (35) that

$$\mathcal{E}_{\text{Dir}}(\mathbf{c}^\epsilon) \leq \mathcal{E}_{\text{Dir}}(\mathbf{c}^n),$$

that

$$\mathcal{E}_{\text{chem}}(\mathbf{c}^\epsilon) \leq \mathcal{E}_{\text{chem}}(\mathbf{c}^n) + \epsilon \int_{\Omega} (\bar{c}_1 - c_1^n)(1 - 2c_1^n) d\mathbf{x} \leq \mathcal{E}_{\text{chem}}(\mathbf{c}^n) + C\epsilon,$$

and that

$$\mathcal{E}_{\text{ext}}(\mathbf{c}^\epsilon) = \mathcal{E}_{\text{ext}}(\mathbf{c}^n) + \epsilon \sum_{i \in \{1,2\}} \langle c_i^n - \bar{c}_i, \Psi_i \rangle_{H^1, (H^1)'} \leq \mathcal{E}_{\text{ext}}(\mathbf{c}^n) + C\epsilon.$$

Moreover, the convexity of  $\mathbf{W}^2$  yields

$$\mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) \leq \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) + \epsilon (\mathbf{W}^2(\bar{\mathbf{c}}, \mathbf{c}^{n-1}) - \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1})) \leq \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) + C\epsilon.$$

Here and all along this proof,  $C$  may depend on  $n$  but not  $\epsilon$ . Combining the above inequalities, we obtain that

$$(36) \quad \mathcal{E}_{\text{therm}}(\mathbf{c}^n) - \mathcal{E}_{\text{therm}}(\mathbf{c}^\epsilon) = \sum_{i \in \{1,2\}} \theta_i \int_{\Omega} (H(c_i^n) - H(c_i^\epsilon)) d\mathbf{x} \leq C\epsilon.$$

Let  $i \in \{1, 2\}$  be such that  $\theta_i > 0$ , then denote by  $A = \{\mathbf{x} \in \Omega \mid c_i^n(\mathbf{x}) > 0\}$ . Then convexity of  $H$  implies that

$$(37) \quad H(c_i^n) - H(c_i^\epsilon) \geq (c_i^n - c_i^\epsilon) \log(c_i^\epsilon) = \epsilon(c_i^n - \bar{c}_i) \log(c_i^\epsilon) \quad \text{a.e. in } A.$$

On the other hand, we remark that

$$(38) \quad H(c_i^n(\mathbf{x})) - H(c_i^\epsilon(\mathbf{x})) = 1 - H(\epsilon \bar{c}_i) = -\epsilon \bar{c}_i \log(\epsilon \bar{c}_i) + \epsilon \bar{c}_i, \quad \forall \mathbf{x} \in A^c.$$

Integrating (37) over  $A$  and (38) over  $A^c$  and using (36), we obtain that

$$(39) \quad -\bar{c}_i \log(\epsilon \bar{c}_i) |A^c| + \int_A (c_i^n - \bar{c}_i) \log(c_i^\epsilon) \leq C.$$

Since  $(c_i^n - \bar{c}_i) \log(c_i^\epsilon) \geq (c_i^n - \bar{c}_i) \log(\bar{c}_i)$ , one gets that

$$-\bar{c}_i \log(\epsilon \bar{c}_i) |A^c| + \int_A (c_i^n - \bar{c}_i) \log(\bar{c}_i) \leq C.$$

We let  $\epsilon$  tend to 0 and get a contradiction unless  $|A^c| = 0$ . Thus we have proved that  $A = \Omega$  (up to a negligible set). The function  $(c_i^n - \bar{c}_i) \log(c_i^\epsilon)$  is bounded from below by  $(c_i^n - \bar{c}_i) \log(\bar{c}_i)$  that belongs to  $L^1(\Omega)$ , then we can apply Fatou's Lemma when  $\epsilon \rightarrow 0$  and claim that

$$\int_\Omega (c_i^n - \bar{c}_i) \log(c_i^n) d\mathbf{x} \leq \liminf_{\epsilon \rightarrow 0} \int_\Omega (c_i^n - \bar{c}_i) \log(c_i^\epsilon) d\mathbf{x} \leq C.$$

This latter inequality imposes that  $\log(c_i^n)$  belongs to  $L^1(\Omega)$ .  $\square$

**3.3. Flow interchange and entropy estimate.** In the next lemma, our goal is to get an improved regularity estimate on  $\mathbf{c}$  by the mean of the flow interchange technique.

**Lemma 3.2.** *There exists  $C$  depending only on  $\alpha, \chi, \Psi_i, \mathbf{c}^0, m_i, T$ , such that*

$$(40) \quad \int_0^T \int_\Omega |\Delta \mathbf{c}_\tau|^2 d\mathbf{x} dt \leq C, \quad \forall T > 0.$$

Since  $\Omega$  is convex, it implies

$$(41) \quad \|\mathbf{c}_\tau\|_{L^2((0,T);H^2(\Omega))} \leq C.$$

*Proof.* Let  $\check{c}_i$  ( $i \in \{1, 2\}$ ) be the unique solution to

$$(42) \quad \begin{cases} \partial_t \check{c}_i - \Delta \check{c}_i = 0 & \text{in } \Omega \times (0, \infty), \\ \nabla \check{c}_i \cdot \mathbf{n} = 0 & \text{on } \partial\Omega \times (0, \infty), \\ (\check{c}_i)|_{t=0} = c_i^n & \text{in } \Omega. \end{cases}$$

Then it is easy to check that  $\check{c}_1(\cdot, t) + \check{c}_2(\cdot, t) = 1$  and that  $\int_\Omega \check{c}_i(\mathbf{x}, t) dt = \int_\Omega c_i^0(\mathbf{x}) d\mathbf{x}$ , whereas  $c_i(\cdot, t) \geq 0$  owing to the maximum principle. Therefore,  $\check{\mathbf{c}} = (\check{c}_1, \check{c}_2)$  belongs to  $\mathcal{A} \cap \mathcal{X}$  and is an admissible competitor in (23).

The heat equation (42) can be reinterpreted as the gradient flow of the Boltzmann entropy  $\mathcal{H}(c) = \int_\Omega c \log(c) d\mathbf{x}$ , which is displacement convex. Therefore, applying the *Evolution Variational Inequality* [1, Definition 4.5] centered at  $c_i^{n-1}$  to the flow yields

$$\frac{1}{2} \frac{d}{dt} W^2(\check{c}_i(t), c_i^{n-1}) \leq \mathcal{H}(c_i^{n-1}) - \mathcal{H}(\check{c}_i(t)).$$

Dividing the previous inequality by  $\tau m_i$  and summing over  $i$  leads to

$$(43) \quad \frac{d}{dt} \left( \frac{1}{2\tau} \mathbf{W}^2(\check{\mathbf{c}}(t), \mathbf{c}^{n-1}) \right) \leq \sum_{i \in \{1,2\}} \frac{\mathcal{H}(c_i^{n-1}) - \mathcal{H}(\check{c}_i(t))}{m_i \tau}, \quad \forall t > 0.$$

The solution  $\check{c}_i$  of (42) belongs to  $C^\infty((0, \infty); H^2(\Omega))$  with  $\check{c}_i(\mathbf{x}, t) > 0$  if  $t > 0$ , so the solution is regular enough to justify the chain rules in the calculations below that give

$$(44) \quad \frac{d}{dt} \mathcal{E}(\check{\mathbf{c}}(t)) = A_{\text{Dir}}(t) + A_{\text{chem}}(t) + A_{\text{therm}}(t) + A_{\text{ext}}(t).$$

In the above relation, the term  $A_{\text{Dir}}(t)$  comes from the Dirichlet energy and writes

$$(45) \quad A_{\text{Dir}}(t) = \frac{d}{dt} \frac{\alpha}{2} \int_{\Omega} |\nabla \check{c}_1(t)|^2 d\mathbf{x} = -\alpha \int_{\Omega} \Delta \check{c}_1(t) \partial_t \check{c}_1(t) d\mathbf{x} = -\alpha \int_{\Omega} |\Delta \check{c}_1(t)|^2 d\mathbf{x}.$$

The term  $A_{\text{chem}}$  comes from the chemical energy and writes

$$\begin{aligned} A_{\text{chem}}(t) &= \frac{d}{dt} \chi \int_{\Omega} \check{c}_1(t)(1 - \check{c}_1(t)) d\mathbf{x} \\ &= \chi \int_{\Omega} (1 - 2\check{c}_1(t)) \partial_t \check{c}_1(t) d\mathbf{x} = 2\chi \int_{\Omega} |\nabla \check{c}_1(t)|^2 d\mathbf{x}. \end{aligned}$$

Since the Dirichlet energy is decreasing along the trajectories of the heat equation (42), we get that

$$(46) \quad A_{\text{chem}}(t) \leq 2\chi \int_{\Omega} |\nabla c_1^n|^2 d\mathbf{x} \leq C$$

thanks to (34). The term  $A_{\text{therm}}$  coming from the thermal diffusion is non-positive. Indeed,

$$\begin{aligned} A_{\text{therm}}(t) &= \sum_{i \in \{1,2\}} \theta_i \frac{d}{dt} \int_{\Omega} H(\check{c}_i(t)) d\mathbf{x} = \sum_{i \in \{1,2\}} \theta_i \int_{\Omega} \log(\check{c}_i(t)) \partial_t \check{c}_i(t) d\mathbf{x} \\ &= \sum_{i \in \{1,2\}} \theta_i \int_{\Omega} \log(\check{c}_i(t)) \Delta \check{c}_i(t) d\mathbf{x} = - \sum_{i \in \{1,2\}} \theta_i \int_{\Omega} \nabla \log(\check{c}_i(t)) \cdot \nabla \check{c}_i(t) d\mathbf{x}. \end{aligned}$$

Since  $\check{c}_i(t)$  is smooth on  $\bar{\partial}\Omega$  and bounded away from 0, we can write that

$$(47) \quad A_{\text{therm}}(t) = -4 \sum_{i \in \{1,2\}} \theta_i \int_{\Omega} |\nabla \sqrt{\check{c}_i(t)}|^2 d\mathbf{x} \leq 0.$$

The contributions  $A_{\text{ext}}$  is related to the potential energy  $\mathcal{E}_{\text{ext}}$  and writes

$$\begin{aligned} A_{\text{ext}}(t) &= \frac{d}{dt} \sum_{i \in \{1,2\}} \int_{\Omega} \Psi_i \check{c}_i(t) d\mathbf{x} = \int_{\Omega} (\Psi_1 - \Psi_2) \partial_t \check{c}_1(t) d\mathbf{x} \\ &= \int_{\Omega} (\Psi_1 - \Psi_2) \Delta \check{c}_1(t) d\mathbf{x} \leq \frac{\alpha}{2} \|\Delta \check{c}_1(t)\|_{L^2}^2 + \frac{1}{2\alpha} \|\Psi_1 - \Psi_2\|_{L^2}^2. \end{aligned}$$

It follows from (45) that

$$(48) \quad A_{\text{Dir}}(t) + A_{\text{ext}}(t) \leq -\frac{\alpha}{2} \|\Delta \check{c}_1(t)\|_{L^2}^2 + C, \quad \forall t > 0.$$

Since  $\check{\mathbf{c}}$  remains in  $\mathcal{X}$ , there is no contribution coming from the constraint  $\mathcal{E}_{\text{cons}}$  in (44). The combination of (43)–(48) ensures that

$$\frac{d}{dt} \mathcal{F}_{\tau}^n(\check{\mathbf{c}}(t)) + \frac{\alpha}{2} \int_{\Omega} |\Delta \check{c}_1(t)|^2 d\mathbf{x} \leq C \left( 1 + \sum_{i \in \{1,2\}} \frac{\mathcal{H}(c_i^{n-1}) - \mathcal{H}(\check{c}_i(t))}{\tau} \right), \quad \forall t > 0.$$

The solution  $\check{c}_i$  to (42) belongs to  $C([0, \infty); L^p(\Omega))$  (see for instance [13]), hence  $\mathcal{H}(\check{c}_i(t))$  tends to  $\mathcal{H}(c_i^n)$  as  $t$  tends to 0. Since  $c^n$  is a minimizer in (23), one has necessarily that

$$\limsup_{t \rightarrow 0^+} \frac{d}{dt} \mathcal{F}_{\tau}^n(\check{\mathbf{c}}(t)) \geq 0,$$



otherwise  $\check{c}(t)$  for some small  $t > 0$  would be a better competitor than  $\mathbf{c}^n$  in (23). Therefore, we obtain that

$$\liminf_{t \rightarrow 0^+} \int_{\Omega} |\Delta \check{c}_1(t)|^2 d\mathbf{x} \leq C \left( 1 + \sum_{i \in \{1,2\}} \frac{\mathcal{H}(c_i^{n-1}) - \mathcal{H}(\check{c}_i(t))}{\tau} \right).$$

Let  $(t_\ell)_{\ell \geq 1} \subset \mathbb{R}_+$  be a sequence tending to 0 achieving the liminf in the previous inequality, then the sequence  $(\Delta \check{c}_1(t_\ell))_{\ell \geq 1}$  is bounded in  $L^2(\Omega)$ . Since  $\check{c}_1(t_\ell)$  tends to  $c_1^n$  in  $L^2(\Omega)$ , we can identify the limit as  $\Delta c_1^n$ . The lower semi-continuity of the norm for the weak convergence yields

$$\int_{\Omega} |\Delta c_1^n|^2 d\mathbf{x} \leq \liminf_{t \rightarrow 0^+} \int_{\Omega} |\Delta \check{c}_1(t)|^2 d\mathbf{x}.$$

As a consequence, we obtain that

$$\int_{\Omega} |\Delta c_1^n|^2 d\mathbf{x} \leq C \left( 1 + \sum_{i \in \{1,2\}} \frac{\mathcal{H}(c_i^{n-1}) - \mathcal{H}(c_i^n)}{\tau} \right), \quad \forall n \geq 1.$$

Multiplying by  $\tau$  and summing over  $n \in \{1, \dots, \lceil \frac{T}{\tau} \rceil\}$  for a finite time horizon  $T \geq \tau$  leads to

$$(49) \quad \int_0^T \int_{\Omega} |\Delta \mathbf{c}_\tau|^2 d\mathbf{x} dt \leq C,$$

where we have used  $0 \leq \mathcal{H}(\mathbf{c}) \leq C$  for all  $\mathbf{c} \in \mathcal{X} \cap \mathcal{A}$ , and that  $|\Delta c_2^n| = |\Delta c_1^n|$ . Thus  $\mathbf{c}^n$  solves a Poisson equation with homogeneous Neumann boundary condition and  $L^2$  right-hand side, hence (41) holds (see for instance [31] for the case of a smooth  $\partial\Omega$  and to [26] in the non-smooth case).  $\square$

**3.4. Euler Lagrange equations.** The goal of this section is to characterize the minimizer  $\mathbf{c}^n$  of (23). The first step consists in recovering the homogeneous Neumann boundary condition (11) for the approximate solution  $\mathbf{c}_\tau$ .

**Lemma 3.3.** *Let  $\mathbf{c}^n$  be a solution to (23), then*

$$(50) \quad \nabla c_i^n \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega.$$

*Proof.* Let us first remark that the above property is trivial if  $c_i^0 \equiv 0$  or  $c_i^0 \equiv 1$ , since in these cases, one has  $c_1^n \equiv 0$  or  $c_1^n \equiv 1$  respectively. In what follows, we will assume that

$$(51) \quad 0 < \frac{1}{|\Omega|} \int_{\Omega} c_1^0 d\mathbf{x} < 1.$$

Since  $d < 4$ ,  $H^2(\Omega)$  is embedded in  $C(\bar{\Omega})$ . Hence there exists  $\mathbf{x}_m \in \Omega$ ,  $\eta > 0$ , and  $\rho > 0$  such that  $\eta < c_1^n(\mathbf{x}) < 1 - \eta$  for all  $\mathbf{x} \in \Omega$  such that  $|\mathbf{x} - \mathbf{x}_m| \leq \rho$ . We assume without loss of generality that  $\overline{B_{\mathbf{x}_m, \rho}} \subset \Omega$ .

Let  $\epsilon \in (0, \rho)$ , then define the cut-off function

$$\nu^\epsilon(\mathbf{x}) = \left( 1 - \frac{\text{dist}(\mathbf{x}, \partial\Omega)}{\epsilon} \right)^+, \quad \forall \mathbf{x} \in \bar{\Omega}.$$

Let  $h \in H^1(\Omega)$  be such that  $0 \leq c_1^n + h \leq 1$  a.e. in  $\Omega$ , then, for  $\epsilon$  smaller than some  $\epsilon_0 \in (0, 1)$ , we set

$$c_1^\epsilon = c_1^n + h\nu^\epsilon + r^\epsilon$$

where  $r^\epsilon \in C_c^\infty(B_{\mathbf{x}_m, \rho-\epsilon})$  is such that

$$\int_{\Omega} c_1^\epsilon d\mathbf{x} = \int_{\Omega} c_1^n d\mathbf{x} = \int_{\Omega} c_1^0 d\mathbf{x}, \quad 0 \leq c^\epsilon \leq 1, \quad \forall \epsilon \in (0, \epsilon_0).$$

Since  $r^\epsilon$  is just there to fulfill the preservation of mass, we can assume that it has a constant sign, so that

$$(52) \quad \|h\nu^\epsilon\|_{L^1} \leq |\partial\Omega|\epsilon, \quad \|r^\epsilon\|_{L^1} \leq |\partial\Omega|\epsilon, \quad \|r^\epsilon\|_{L^\infty} \leq C\epsilon, \quad \|\nabla r^\epsilon\|_{L^2} \leq C\epsilon, \quad \|c_1^n - c_1^\epsilon\|_{L^1} \leq 2|\partial\Omega|\epsilon.$$

Denoting by  $\mathbf{c}^\epsilon = (c_1^\epsilon, 1 - c_1^\epsilon) \in \mathcal{A} \cap \mathcal{X}$ , the optimality (23) of  $\mathbf{c}^n$  implies that

$$(53) \quad \mathcal{E}(\mathbf{c}^\epsilon) + \frac{1}{2\tau} \mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) \geq \mathcal{E}(\mathbf{c}^n) + \frac{1}{2\tau} \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}), \quad \forall \epsilon \in (0, \epsilon_0).$$

Let  $(\varphi^n, \psi^n)$  be some Kantorovich potentials corresponding to the optimal transport between  $\mathbf{c}^n$  and  $\mathbf{c}^{n-1}$  as in (26), then because of (21), one has

$$\mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) \leq 2 \sum_{i \in \{1,2\}} \left( \int_{\Omega} c_i^\epsilon \varphi_i^n d\mathbf{x} + \int_{\Omega} c_i^{n-1} \psi_i^n d\mathbf{y} \right).$$

Hence, subtracting (26) to the previous relation, one gets that

$$\mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) \leq \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) + 2 \sum_{i \in \{1,2\}} \|c_i^\epsilon - c_i^n\|_{L^1(\Omega)} \|\varphi_i^n\|_{L^\infty(\Omega)}.$$

Since  $\Omega$  is bounded,  $\varphi_i^n$  is Lipschitz continuous and we can assume that  $\int_{\Omega} \varphi_i^n d\mathbf{x} = 0$ , so that  $\|\varphi_i\|_\infty \leq C$ . Therefore, using (52), we obtain that

$$(54) \quad \mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) \leq \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1}) + C\epsilon, \quad \forall \epsilon \in (0, \epsilon_0).$$

Thanks to (52), one gets that

$$(55) \quad \mathcal{E}_{\text{ext}}(\mathbf{c}^\epsilon) + \mathcal{E}_{\text{chem}}(\mathbf{c}^\epsilon) + \mathcal{E}_{\text{therm}}(\mathbf{c}^\epsilon) \leq \mathcal{E}_{\text{ext}}(\mathbf{c}^n) + \mathcal{E}_{\text{chem}}(\mathbf{c}^n) + \mathcal{E}_{\text{therm}}(\mathbf{c}^\epsilon) + \zeta_\epsilon,$$

with  $\zeta_\epsilon \rightarrow 0$  as  $\epsilon \rightarrow 0$ . Concerning the Dirichlet energy, we infer from (52) that

$$(56) \quad \mathcal{E}_{\text{Dir}}(\mathbf{c}^\epsilon) \leq \mathcal{E}_{\text{Dir}}(\mathbf{c}^n) + \alpha \int_{\Omega} \nabla c_1^n \cdot \nabla h\nu^\epsilon d\mathbf{x} + \alpha \int_{\Omega} \nabla c_1^n \cdot \nabla \nu^\epsilon h d\mathbf{x} + C\epsilon.$$

Since  $c_1^n$  belongs to  $H^2(\Omega) \subset W^{1,6}(\Omega)$  (recall that  $d \leq 3$ ), it follows from Hölder inequality that

$$(57) \quad \left| \int_{\Omega} \nabla c_1^n \cdot \nabla h\nu^\epsilon d\mathbf{x} \right| \leq \|\nabla c_1^n\|_{L^6(\Omega)} \|\nabla h\|_{L^2(\Omega)} \|\nu^\epsilon\|_{L^3(\Omega)} \leq C\epsilon^{1/3}.$$

Combining (54)–(57) in (53) and letting  $\epsilon$  tend to 0 leads to

$$(58) \quad \int_{\partial\Omega} \nabla c_1^n \cdot \mathbf{n} h d\mathbf{x} = \lim_{\epsilon \rightarrow 0} \int_{\Omega} \nabla c_1^n \cdot \nabla \nu^\epsilon h d\mathbf{x} \geq 0, \quad \forall h \in H^{1/2}(\partial\Omega) \text{ s.t. } 0 \leq c_1^n + h \leq 1.$$

We have used above that  $\nabla c_1^n$  admits a strong trace in  $H^{1/2}(\Omega)^d$  on  $\partial\Omega$  since  $c_1^n$  belongs to  $H^2(\Omega)$ . Its normal component is characterized by

$$\nabla c_1^n(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \lim_{\epsilon \rightarrow 0} \frac{c_1^n(\mathbf{x}) - c_1^n(\mathbf{x} - \epsilon \mathbf{n}(\mathbf{x}))}{\epsilon} \quad \text{for a.e. } \mathbf{x} \in \partial\Omega.$$

This last relation implies that

$$(59a) \quad \nabla c_1^n \cdot \mathbf{n} \geq 0 \quad \text{on } \Gamma^{(1)} := \{\mathbf{x} \in \partial\Omega \mid c_1^n(\mathbf{x}) = 1\},$$

$$(59b) \quad \nabla c_1^n \cdot \mathbf{n} \leq 0 \quad \text{on } \Gamma^{(2)} := \{\mathbf{x} \in \partial\Omega \mid c_1^n(\mathbf{x}) = 0\}.$$

The combination of (58) with (59) yields (50).  $\square$

The next step consists in getting a linearized version of the optimality condition on  $\mathbf{c}^n$ . Recall that  $f(c) = \log\left(\frac{c^{\theta_1}}{(1-c)^{\theta_2}}\right)$ .

**Lemma 3.4.** *There exists a solution  $\mathbf{c}^n \in \mathcal{A} \cap \mathcal{X}$  to (23) such that, denoting by  $\varphi^n = (\varphi_1^n, \varphi_2^n)$  the corresponding (backward) Kantorovich potentials, and by*

$$(60) \quad F^n = \frac{\varphi_1^n}{\tau} - \frac{\varphi_2^n}{\tau} - \alpha \Delta c_1^n + f(c_1^n) + \chi(1 - 2c_1^n) + \Psi_1 - \Psi_2, \quad \forall n \geq 1,$$

then the following linearized optimality condition is fulfilled:

$$(61) \quad \int_{\Omega} F^n c_1^n d\mathbf{x} \leq \int_{\Omega} F^n c_1 d\mathbf{x}, \quad \forall \mathbf{c} = (c_1, c_2) \in \mathcal{A}.$$

*Proof.* The proof is directly inspired from [36, Lemma 3.1], see also [15, Lemma 3.2]. Assume first that  $c_i^{n-1} > 0$  almost everywhere in  $\Omega$ , so that the Kantorovich potentials  $\varphi_1^n, \psi_1^n$  from  $c_1^n$  to  $c_1^{n-1}$  are uniquely defined after normalizing  $\varphi_i(\mathbf{x}_{\text{ref}}) = 0$  for some arbitrary  $\mathbf{x}_{\text{ref}} \in \Omega$ . Let  $\mathbf{c}$  be an arbitrary element of  $\mathcal{A} \cap \mathcal{X}$ , then we perturb  $\mathbf{c}^n$  into

$$(62) \quad \mathbf{c}^\epsilon = (1 - \epsilon)\mathbf{c}^n + \epsilon\mathbf{c}, \quad \forall \epsilon \in (0, 1).$$

Clearly,  $\mathbf{c}^\epsilon$  belongs to  $\mathcal{A} \cap \mathcal{X}$ . Defining  $(\varphi_i^\epsilon, \psi_i^\epsilon)$  the Kantorovich potential from  $c_i^\epsilon$  to  $c_i^{n-1}$ , we infer from (21) that

$$\begin{cases} \frac{1}{2} W_i^2(c_i^\epsilon, c_i^{n-1}) = \int_{\Omega} \varphi_i^\epsilon(\mathbf{x}) c_i^\epsilon(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \psi_i^\epsilon(\mathbf{y}) c_i^{n-1}(\mathbf{y}) d\mathbf{y}, \\ \frac{1}{2} W_i^2(c_i^n, c_i^{n-1}) \geq \int_{\Omega} \varphi_i^\epsilon(\mathbf{x}) c_i^n(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \psi_i^\epsilon(\mathbf{y}) c_i^{n-1}(\mathbf{y}) d\mathbf{y}. \end{cases}$$

Subtracting the two above relations and using the definition (62) of  $\mathbf{c}^\epsilon$ , one gets

$$\frac{1}{2\tau} (W_i^2(c_i^\epsilon, c_i^{n-1}) - W_i^2(c_i^n, c_i^{n-1})) \leq \epsilon \int_{\Omega} \frac{\varphi_i^\epsilon}{\tau} (c_i - c_i^n) d\mathbf{x}.$$

Hence, using  $\mathbf{c}^n, \mathbf{c} \in \mathcal{X}$ , one gets that

$$(63) \quad \frac{1}{2\tau} (\mathbf{W}^2(\mathbf{c}^\epsilon, \mathbf{c}^{n-1}) - \mathbf{W}^2(\mathbf{c}^n, \mathbf{c}^{n-1})) \leq \epsilon \int_{\Omega} \left( \frac{\varphi_1^\epsilon}{\tau} - \frac{\varphi_2^\epsilon}{\tau} \right) (c_1 - c_1^n) d\mathbf{x}.$$

On the other hand, the convexity of  $\mathcal{E}_{\text{Dir}}$  and  $\mathcal{E}_{\text{therm}}$ , the linearity of  $\mathcal{E}_{\text{ext}}$ , and the concavity of  $\mathcal{E}_{\text{chem}}$  yield

$$(64) \quad \begin{aligned} \int_{\Omega} (\mathcal{E}(\mathbf{c}^\epsilon) - \mathcal{E}(\mathbf{c}^n)) &\leq \epsilon \int_{\Omega} \alpha \nabla c_1^\epsilon \cdot \nabla (c_1 - c_1^n) d\mathbf{x} + \epsilon \int_{\Omega} \chi(1 - 2c_1^n)(c_1 - c_1^n) d\mathbf{x} \\ &\quad + \epsilon \int_{\Omega} f(c_1^\epsilon)(c_1 - c_1^n) d\mathbf{x} + \epsilon \int_{\Omega} (\Psi_1 - \Psi_2)(c_1 - c_1^n) d\mathbf{x}. \end{aligned}$$

Bearing in mind that  $c_1^n$  is a minimizer, the combination of (63) with (64) leads to

$$\begin{aligned} 0 &\leq \frac{\mathcal{F}_\tau^n(\mathbf{c}^\epsilon) - \mathcal{F}_\tau^n(\mathbf{c}^n)}{\epsilon} \leq \int_{\Omega} \left( \frac{\varphi_1^\epsilon}{\tau} - \frac{\varphi_2^\epsilon}{\tau} \right) (c_1 - c_1^n) d\mathbf{x} + \int_{\Omega} \alpha \nabla c_1^\epsilon \cdot \nabla (c_1 - c_1^n) d\mathbf{x} \\ &\quad + \int_{\Omega} \chi(1 - 2c_1^n)(c_1 - c_1^n) d\mathbf{x} + \int_{\Omega} f(c_1^\epsilon)(c_1 - c_1^n) d\mathbf{x} + \int_{\Omega} (\Psi_1 - \Psi_2)(c_1 - c_1^n) d\mathbf{x}, \quad \forall \epsilon > 0. \end{aligned}$$

We can consider the limit  $\epsilon \rightarrow 0$  in the right-hand side of the above expression. Bearing in mind the definition (62) of  $\mathbf{c}^\epsilon$ , it is clear that  $c_1^\epsilon$  converges in  $H^1(\Omega)$  towards  $c_1^n$  and that  $f(c_1^\epsilon)$  converges in  $L^1(\Omega)$  towards  $f(c_1^n)$ , while  $\varphi_i^\epsilon$  converges uniformly towards  $\varphi_i^n$  (see for instance [42, Theorem 1.52]), so that (61) holds thanks to Lemma 3.3.

Assume now that  $c_i^{n-1} = 0$  on some part of  $\Omega$ , then we first evolve  $\mathbf{c}^{n-1}$  on a small time  $\delta > 0$  thanks to the flow (42), yields  $\mathbf{c}^{n-1,\delta} = (c_1^{n-1,\delta}, c_2^{n-1,\delta}) \in \mathcal{A} \cap \mathcal{X}$  with  $c_i^{n-1,\delta} = \check{c}_i(\delta) > 0$  in  $\Omega$ . This allows to define

$$\mathcal{F}_\tau^{n,\delta}(\mathbf{c}) = \frac{\mathbf{W}^2(\mathbf{c}, \mathbf{c}^{n-1,\delta})}{2\tau} + \mathcal{E}(\mathbf{c}),$$

as well as

$$\mathbf{c}^{n,\delta} \in \operatorname{argmin}_{\mathbf{c} \in \mathcal{A} \cap \mathcal{X}} \mathcal{F}_\tau^{n,\delta}$$

one of its minimizers. Define  $F^{n,\delta}$  by replacing in (60)  $c_i^n$  by  $c_i^{n,\delta}$  and  $\varphi_i^n$  by the Kantorovich potential  $\varphi_i^{n,\delta}$  sending  $c_i^{n,\delta}$  to  $c_i^{n-1,\delta}$ , then we get from the reasoning above that

$$(65) \quad \int_{\Omega} F^{n,\delta}(c_1 - c_1^{n,\delta}) \geq 0, \quad \forall \mathbf{c} = (c_1, c_2) \in \mathcal{A} \cap \mathcal{X}.$$

It results from Lemma 3.2 that  $-\Delta c_1^{n,\delta}$  is uniformly bounded in  $L^2(\Omega)$  w.r.t.  $\delta$ , hence  $(c_1^{n,\delta})_{\delta>0}$  is sequentially relatively compact in  $H^1(\Omega)$ . Up to the extraction of a subsequence,  $\mathbf{c}^{n,\delta}$  converges strongly in  $H^1(\Omega)^2$  towards  $\mathbf{c}^n$ , while  $f(c_1^{n,\delta})$  converges in  $L^1(\Omega)$  towards  $f(c_1^n)$  and  $-\Delta c_1^{n,\delta}$  converges weakly to  $-\Delta c_1^n$ . The Kantorovich potentials  $\varphi_i^{n,\delta}$  converge uniformly towards the Kantorovich potential  $\varphi_i^n$  sending  $c_i^n$  to  $c_i^{n-1}$  (cf. [42, Theorem 1.52]). Since the solutions to (42) belong to  $C(\mathbb{R}_+; L^1(\Omega)^2)$ ,  $c_i^{n-1,\delta}$  tends to  $c_i^{n-1}$  in  $L^1(\Omega)$  as  $\delta$  tends to 0. Therefore, the functional  $\mathcal{F}_\tau^{n,\delta}$   $\Gamma$ -converges towards  $\mathcal{F}_\tau^n$  as  $\delta$  tends to 0, thus the limit  $\mathbf{c}^n$  of  $\mathbf{c}^{n,\delta}$  is a minimizer of  $\mathcal{F}_\tau^n$ . Passing to the limit  $\delta \rightarrow 0$  in (65) provides that (61) holds even though  $c_i^n$  vanishes on some parts of  $\Omega$ .  $\square$

**Lemma 3.5.** *Let  $\mathbf{c}^n$  be a minimizer of (23) such that (61) holds, then there exists  $\boldsymbol{\mu}^n = (\mu_1^n, \mu_2^n) : \Omega \rightarrow \mathbb{R}^2$  such that*

$$(66) \quad \mu_1^n - \mu_2^n = -\alpha \Delta c_1^n + \chi(1 - 2c_1^n),$$

$$(67) \quad \mu_i^n = -\frac{\varphi_i^n}{\tau} - \Psi_i - \theta_i \log(c_i^n) \quad \text{d}c_i^n\text{-a.e. in } \Omega,$$

$$(68) \quad \int_{\Omega} \bar{\mu}^n \, \text{d}\mathbf{x} = 0, \quad \text{where } \bar{\mu}^n = c_1^n \mu_1^n + c_2^n \mu_2^n.$$

Moreover, we have the following uniform estimates w.r.t.  $\tau$ :

$$(69) \quad \|\mu_{i,\tau}\|_{L^2((0,T); L^{d/(d-1)}(\Omega))} \leq C, \quad \|\mu_{1,\tau} - \mu_{2,\tau}\|_{L^2(Q_T)} \leq C, \quad \|\sqrt{c_{i,\tau}} \nabla \mu_{i,\tau}\|_{L^2(Q_T)} \leq C.$$

*Proof.* Since (61) holds and since  $0 \leq c_1^n \leq 1$ , the bathtub principle [32, Theorem 1.14] implies the existence of some  $\ell \in \mathbb{R}$  such that

$$(70) \quad c_1^n(\mathbf{x}) = \begin{cases} 0 & \text{if } F^n(\mathbf{x}) > \ell, \\ 1 & \text{if } F^n(\mathbf{x}) < \ell, \end{cases} \quad c_2^n(\mathbf{x}) = \begin{cases} 1 & \text{if } F^n(\mathbf{x}) > \ell, \\ 0 & \text{if } F^n(\mathbf{x}) < \ell. \end{cases}$$

Up to a vertical shift of the Kantorovich potential  $\varphi_1^n$ , we can assume without loss of generality that  $\ell = 0$ . Define

$$\mu_1^n = -\frac{\varphi_1^n}{\tau} - \Psi_1 - \theta_1 \log(c_1^n) + (F^n)_+, \quad \text{and} \quad \mu_2^n = -\frac{\varphi_2^n}{\tau} - \Psi_2 - \theta_2 \log(c_2^n) + (F^n)_-,$$

then (66) holds in view of (60). It follows from (49) that

$$(71) \quad \iint_{Q_T} (\mu_{1,\tau} - \mu_{2,\tau})^2 \, d\mathbf{x}dt \leq C.$$

Moreover, it results from (70) that

$$(F^n)_+ = 0 \quad \text{d}c_1^n\text{-a.e. in } \Omega, \quad (F^n)_- = 0 \quad \text{d}c_2^n\text{-a.e. in } \Omega,$$

so that (67) is fulfilled. Now, shifting vertically both Kantorovich potentials  $\varphi_i^n$  by the same constant, we can assume without loss of generality that (68) holds. The definition of  $\bar{\mu}^n$  implies that

$$\begin{aligned} \nabla \bar{\mu}^n &= \sum_{i \in \{1,2\}} c_i^n \nabla \mu_i^n + \nabla c_1^n (\mu_1^n - \mu_2^n) \\ &= - \sum_{i \in \{1,2\}} \left( c_i^n \frac{\nabla \varphi_i^n}{\tau} - c_i^n \nabla \Psi_i - \theta_i \nabla c_i^n \right) + \nabla c_1^n (\mu_1^n - \mu_2^n). \end{aligned}$$

Using the triangle inequality, Cauchy-Schwarz inequality, and  $0 \leq c_i^n \leq 1$ , we get

$$\begin{aligned} \|\nabla \bar{\mu}^n\|_{L^1(\Omega)} &\leq \frac{1}{\tau} \sum_{i \in \{1,2\}} \left( \int_{\Omega} c_i^n \right)^{1/2} \left( \int_{\Omega} c_i^n |\nabla \varphi_i^n|^2 \, d\mathbf{x} \right)^{1/2} \\ &\quad + \sum_{i \in \{1,2\}} \|\nabla \Psi_i\|_{L^1(\Omega)} + \|\nabla c_1^n\|_{L^2(\Omega)^d} \left( |\Omega|^{1/2} (\theta_1 + \theta_2) + \|\mu_1^n - \mu_2^n\|_{L^2(\Omega)} \right). \end{aligned}$$

Since  $c^n \in \mathcal{A}$ , one has  $\int_{\Omega} c_i^n \, d\mathbf{x} = \int_{\Omega} c_i^0 \, d\mathbf{x}$ . Thus it follows from (22), (30), (34) and (71) that

$$(72) \quad \sum_{n=0}^{\lceil \frac{T}{\tau} \rceil} \tau \|\nabla \bar{\mu}^n\|_{L^1(\Omega)}^2 \leq C.$$

Bearing (68) in mind, we can use the Poincaré-Sobolev inequality and get that

$$(73) \quad \|\bar{\mu}_\tau\|_{L^2((0,T); L^{d/(d-1)}(\Omega))} \leq C.$$

Since

$$\mu_1^n = \bar{\mu}^n + c_2^n (\mu_1^n - \mu_2^n), \quad \mu_2^n = \bar{\mu}^n + c_1^n (\mu_2^n - \mu_1^n),$$

we deduce from (71) the desired  $L^2((0,T); L^{d/(d-1)}(\Omega))$  estimates on the phase potentials  $\mu_i^n$ . Finally, the combination of the relations (67), (22) and (30) yield

$$\iint_{Q_T} c_{i,\tau} |\nabla \mu_{i,\tau}|^2 \, d\mathbf{x}dt \leq C.$$

□

The following lemma is a first step towards the recovery of the weak formulation (28).

**Lemma 3.6.** *For any  $\xi \in C^2(\bar{\Omega})$ , there holds*

$$(74) \quad \left| \int_{\Omega} (c_i^n - c_i^{n-1}) \xi d\mathbf{x} + \tau m_i \int_{\Omega} (c_i^n \nabla (\mu_i^n + \Psi_i) + \theta_i \nabla c_i^n) \cdot \nabla \xi d\mathbf{x} \right| \leq \frac{1}{2} W_i^2(c_i^n, c_i^{n-1}) \|D^2 \xi\|_{\infty}.$$

*Proof.* The optimal transport map

$$\mathbf{t}_i^n(\mathbf{x}) = \mathbf{x} - m_i \nabla \varphi_i^n(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega$$

sending  $c_i^n$  to  $c_i^{n-1}$  maps  $\Omega$  into itself because  $\Omega$  is convex. Therefore, since  $c_i^{n-1} = \mathbf{t}_i^n \# c_i^n$  and thanks to (67), one gets that

$$\begin{aligned} \int_{\Omega} (c_i^n - c_i^{n-1}) \xi d\mathbf{x} + \tau m_i \int_{\Omega} (c_i^n \nabla (\mu_i^n + \Psi_i) + \theta_i \nabla c_i^n) \cdot \nabla \xi d\mathbf{x} \\ = \int_{\Omega} (\xi(\mathbf{x}) - \xi(\mathbf{t}_i^n(\mathbf{x})) - m_i \nabla \xi(\mathbf{x}) \cdot \nabla \varphi_i^n(\mathbf{x})) c_i^n(\mathbf{x}) d\mathbf{x} \end{aligned}$$

for all  $\xi \in C^2(\bar{\Omega})$ . The Taylor expansion of  $\xi$  at point  $\mathbf{x}$  provides

$$|\xi(\mathbf{t}_i^n(\mathbf{x})) - \xi(\mathbf{x}) + m_i \nabla \xi(\mathbf{x}) \cdot \nabla \varphi_i^n(\mathbf{x})| \leq \frac{1}{2} \|D^2 \xi\|_{\infty} |\mathbf{t}_i^n(\mathbf{x}) - \mathbf{x}|^2, \quad \forall \mathbf{x} \in \Omega,$$

so that

$$\begin{aligned} \left| \int_{\Omega} (c_i^n - c_i^{n-1}) \xi d\mathbf{x} + \tau m_i \int_{\Omega} (c_i^n \nabla (\mu_i^n + \Psi_i) + \theta_i \nabla c_i^n) \cdot \nabla \xi d\mathbf{x} \right| \\ \leq \frac{1}{2} \|D^2 \xi\|_{\infty} \int_{\Omega} |\mathbf{t}_i^n(\mathbf{x}) - \mathbf{x}|^2 c_i^n(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

which is exactly the desired result.  $\square$

**3.5. Convergence towards a weak solution.** The goal of this section is to consider the limit  $\tau \rightarrow 0$ . This requires some compactness on the approximate phase field  $\mathbf{c}_{\tau}$  and on the approximate potential  $\mu_{\tau}$ . In what follows,  $\mathcal{A}$  is equipped with the topology corresponding to the distance  $\mathbf{W}$ .

**Proposition 3.7.** *There exist  $\mathbf{c} \in C([0, T]; L^2(\Omega)) \cap L^2((0, T); H^2(\Omega)) \cap L^{\infty}((0, T); H^1(\Omega))$  with  $\mathbf{c}(t) \in \mathcal{A} \cap \mathcal{X}$  for a.e.  $t \in [0, T]$ , and  $\mu \in L^2((0, T); L^{d/(d-1)}(\Omega))$  such that, up to the extraction of a subsequence, the following convergence properties hold:*

$$(75a) \quad c_{i,\tau} \xrightarrow{\tau \rightarrow 0} c_i \quad \text{in the } L^{\infty}((0, T); H^1(\Omega))\text{-weak-}\star \text{ sense,}$$

$$(75b) \quad \|c_{i,\tau}(\cdot, t) - c_i(\cdot, t)\|_{L^2(\Omega)} \xrightarrow{\tau \rightarrow 0} 0 \quad \text{for all } t \in [0, T],$$

$$(75c) \quad c_{i,\tau} \xrightarrow{\tau \rightarrow 0} c_i \quad \text{in } L^2((0, T); W^{1,d}(\Omega)),$$

$$(75d) \quad c_{i,\tau} \xrightarrow{\tau \rightarrow 0} c_i \quad \text{weakly in } L^2((0, T); H^2(\Omega)),$$

$$(75e) \quad \mu_{i,\tau} \xrightarrow{\tau \rightarrow 0} \mu_i \quad \text{for the weak topology of } L^2((0, T); L^{d/d-1}(\Omega)),$$

$$(75f) \quad c_{i,\tau} \nabla \mu_{i,\tau} \xrightarrow{\tau \rightarrow 0} c_i \nabla \mu_i \quad \text{weakly in } L^2(Q_T)^d.$$

*Proof.* All the convergence properties stated below occur up to the extraction of a subsequence when  $\tau$  tends to 0. We deduce from Estimate (34) that the family  $(\mathbf{c}_{\tau})_{\tau > 0}$  is bounded in  $L^{\infty}((0, T); H^1(\Omega))$ . Hence we can assume that  $\mathbf{c}_{\tau}$  tends to some  $\mathbf{c}$  in the  $L^{\infty}((0, T); H^1(\Omega))$ -weak- $\star$  sense. Moreover, since  $0 \leq c_{i,\tau} \leq 1$ , we also have that  $0 \leq c_i \leq 1$  a.e. in  $Q_T$ . We also infer from Estimate (41) that  $\mathbf{c}_{\tau}$  converges weakly in  $L^2((0, T); H^2(\Omega))$  towards  $\mathbf{c}$ .

As a consequence of the  $L^\infty$  bound on  $c_{i,\tau}$  and of the Benamou-Brenier formula, we get that

$$\|c_i^{(1)} - c_i^{(2)}\|_{H^1(\Omega)'} \leq \frac{1}{m_i} W_i(c_i^{(1)}, c_i^{(2)}), \quad \forall c_i^{(1)}, c_i^{(2)} \in \mathcal{A}_i,$$

(see more precisely [42, Lemma 3.4]). Therefore, we infer from (32) that

$$\|c_{i,\tau}(t) - c_{i,\tau}(s)\|_{H^1(\Omega)'} \leq C\sqrt{|t-s| + \tau}, \quad \forall s, t \in [0, T].$$

Let  $\Delta t > 0$  and let  $t \in [0, T - \Delta t]$ , then

$$\|c_{i,\tau}(t + \Delta t) - c_{i,\tau}(t)\|_{L^2(\Omega)}^2 = \|c_{i,\tau}(t + \Delta t) - c_{i,\tau}(t)\|_{H^1} \|c_{i,\tau}(t + \Delta t) - c_{i,\tau}(t)\|_{(H^1)' } \leq C\sqrt{\Delta t + \tau}.$$

Bearing in mind the  $L^\infty((0, T); H^1(\Omega))$  estimate on  $c_{i,\tau}$ , we can apply a refined version of the Arzelà-Ascoli theorem [2, Prop. 3.3.1] to obtain that  $c_i \in C([0, T]; L^2(\Omega))$  and that

$$c_{i,\tau}(t) \xrightarrow{\tau \rightarrow 0} c_i(t) \quad \text{in } L^2(\Omega) \text{ for all } t \in [0, T].$$

Together with the estimate  $0 \leq c_{i,\tau} \leq 1$ , we deduce that  $\mathbf{c} \in C([0, T]; L^p(\Omega))^2$  and that

$$\mathbf{c}_\tau \xrightarrow{\tau \rightarrow 0} \mathbf{c} \quad \text{in } L^p(Q_T)^2, \quad \forall p \in [1, +\infty).$$

This implies in particular some strong convergence in  $L^2(Q_T)$ , which can be combined with the weak convergence in  $L^2((0, T); H^2(\Omega))$  thanks to interpolation arguments to derive some strong convergence in  $L^2((0, T); H^s(\Omega))$  for any  $s < 2$ . The continuous embedding of  $H^s(\Omega)$  into  $W^{1,d}(\Omega)$  when  $s \geq 1 + \max(0, \frac{d-2}{2})$  (see for instance [20, Theorem 6.7]) ensures that

$$(76) \quad \mathbf{c}_\tau \xrightarrow{\tau \rightarrow 0} \mathbf{c} \quad \text{in } L^2((0, T); W^{1,d}(\Omega))^2.$$

Let us switch to the phase potentials  $\mu$ . Thanks to Lemma 3.5, we have (uniform w.r.t.  $\tau$ )  $L^2((0, T); L^{d/(d-1)}(\Omega))$  estimates on  $\mu_{i,\tau}$ . Hence there exists  $\mu$  in  $L^2((0, T); L^{d/(d-1)}(\Omega))^2$  such that

$$(77) \quad \mu_{i,\tau} \xrightarrow{\tau \rightarrow 0} \mu_i \quad \text{weakly in } L^2((0, T); L^{d/(d-1)}(\Omega)).$$

In Lemma 3.5, we also established a (uniform w.r.t.  $\tau$ )  $L^2(Q_T)^d$  estimate on  $(\sqrt{c_{i,\tau}} \nabla \mu_{i,\tau})_{\tau > 0}$ . Since  $0 \leq c_{i,\tau} \leq 1$ , it implies a uniform  $L^2(Q_T)^d$  estimate on  $(c_{i,\tau} \nabla \mu_{i,\tau})_{\tau > 0}$ . Therefore, there exists  $\vartheta_i \in L^2(Q_T)^d$  such that  $c_{i,\tau} \nabla \mu_{i,\tau}$  converges weakly in  $L^2(Q_T)^d$  to  $\vartheta_i$  as  $\tau$  tends to 0. It remains to show that  $\vartheta_i = c_i \nabla \mu_i$ . First, the distributions  $c_i$  and  $\nabla \mu_i$  can be multiplied since  $c_i$  belongs to  $L^2((0, T); W^{1,d}(\Omega))$  and  $\nabla \mu_i$  belongs to  $L^2((0, T); W^{-1,d/(d-1)}(\Omega))$ . Moreover, for all  $\phi \in C_c^\infty(Q_T)^d$ , one has

$$\iint_{Q_T} c_{i,\tau} \nabla \mu_{i,\tau} \cdot \phi \, d\mathbf{x} dt = - \iint_{Q_T} \mu_{i,\tau} (\nabla c_{i,\tau} \cdot \phi + c_{i,\tau} \nabla \cdot \phi) \, d\mathbf{x} dt.$$

Thanks to (76) and (77), we can pass in the limit in the right-hand side of the above expression. This leads to

$$\iint_{Q_T} c_{i,\tau} \nabla \mu_{i,\tau} \cdot \phi \, d\mathbf{x} dt \xrightarrow{\tau \rightarrow 0} - \iint_{Q_T} \mu_i (\nabla c_i \cdot \phi + c_i \nabla \cdot \phi) \, d\mathbf{x} dt = \langle c_i \nabla \mu_i, \phi \rangle_{\mathcal{D}', \mathcal{D}}.$$

As a consequence,  $\vartheta_i = c_i \nabla \mu_i$  in the distributional sense, thus also in  $L^2(Q_T)$ .

Since  $\mathbf{c}_\tau(t)$  converges in  $L^2(\Omega)^2$  towards  $\mathbf{c}(t)$  for all  $t \in [0, 1]$ , there holds

$$\int_{\Omega} c_i(t) \, d\mathbf{x} = \int_{\Omega} c_i^0 \, d\mathbf{x}, \quad c_1(\mathbf{x}, t) + c_2(\mathbf{x}, t) = 1 \quad \text{for all } t \in [0, T].$$

Moreover, since  $\mathbf{c}(t)$  belongs to  $H^1(\Omega)$  for a.e.  $t \in (0, T)$ ,  $\mathbf{c}(t)$  belongs to  $\mathcal{A} \cap \mathcal{X}$  for a.e.  $t \in (0, T)$ . This concludes the proof of Proposition 3.7.  $\square$

We have all the necessary convergence properties to pass to the limit  $\tau \rightarrow 0$  and to identify the limit  $(\mathbf{c}, \boldsymbol{\mu})$  exhibited in Proposition 3.7 as a weak solution in the sense of Definition 2.2.

**Proposition 3.8.** *Let  $(\mathbf{c}, \boldsymbol{\mu})$  be as in Proposition 3.7, then  $(\mathbf{c}, \boldsymbol{\mu})$  is a weak solution to the problem (14), (15)–(16) in the sense of Definition 2.2.*

*Proof.* Since  $c_{i,\tau}$  and  $\mu_{i,\tau}$  tend weakly in  $L^2((0, T); H^2(\Omega))$  and  $L^2((0, T); L^{d/(d-1)}(\Omega))$  towards  $c_i$  and  $\mu_i$  respectively, we can pass to the limit in (66). Moreover, one can also pass to the limit in the relation  $0 = \nabla c_{1,\tau} \cdot \mathbf{n}$  established in Lemma 3.3, leading to  $\nabla c_1 \cdot \mathbf{n} = 0$  on  $\partial\Omega$ .

It only remains to recover the weak formulation (28). Let  $t_1, t_2 \in [0, T]$  with  $t_2 \geq t_1$ , then summing (74) over  $n \in \{\lceil \frac{t_1}{\tau} \rceil + 1, \dots, \lceil \frac{t_2}{\tau} \rceil\}$  yields

$$\left| \int_{\Omega} (c_{i,\tau}(t_2) - c_{i,\tau}(t_1)) \xi \, d\mathbf{x} + m_i \int_{\lceil \frac{t_1}{\tau} \rceil \tau}^{\lceil \frac{t_2}{\tau} \rceil \tau} \int_{\Omega} (c_{i,\tau} \nabla (\mu_{i,\tau} + \Psi_i) + \theta_i \nabla c_{i,\tau}) \cdot \nabla \xi \, d\mathbf{x} dt \right| \leq \frac{1}{2} \|D^2 \xi\|_{\infty} \sum_{n=\lceil \frac{t_1}{\tau} \rceil}^{\lceil \frac{t_2}{\tau} \rceil} W_i^2(c_i^n, c_i^{n-1}) \leq C\tau,$$

the last inequality being the consequence of the squared distance estimate (30). We can pass to the limit  $\tau \rightarrow 0$  in the above relation thanks to Proposition 3.7.  $\square$

We have finally proved Theorem 2.3 that is a combination of Propositions 3.7 and 3.8.

**Acknowledgements.** C. Cancès and F. Nabet acknowledge the support of the French National Research Agency (ANR) through grant ANR-13-JS01-0007-01 (project GEOPOR).

#### REFERENCES

- [1] L. Ambrosio and N. Gigli. A user’s guide to optimal transport. In *Modelling and optimisation of flows on networks*, volume 2062 of *Lecture Notes in Math.*, pages 1–155. Springer, Heidelberg, 2013.
- [2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [3] J. W. Barrett, J. F. Blowey, and H. Garcke. Finite element approximation of the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Numer. Anal.*, 37(1):286–318, 1999.
- [4] J. W. Barrett, J. F. Blowey, and H. Garcke. On fully practical finite element approximations of degenerate Cahn–Hilliard systems. *M2AN Math. Model. Numer. Anal.*, 35(4):713–748, 2001.
- [5] J.-D. Benamou, G. Carlier, and M. Laborde. An augmented Lagrangian approach to Wasserstein gradient flows and applications. In *Gradient flows: from theory to application*, volume 54 of *ESAIM Proc. Surveys*, pages 1–17. EDP Sci., Les Ulis, 2016.
- [6] A. Blanchet, V. Calvez, and J. A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.
- [7] J. F. Blowey and C. M. Elliott. The Cahn–Hilliard gradient theory for phase separation with nonsmooth free energy. I. Mathematical analysis. *European J. Appl. Math.*, 2(3):233–280, 1991.
- [8] J. F. Blowey and C. M. Elliott. The Cahn–Hilliard gradient theory for phase separation with nonsmooth free energy. II. Numerical analysis. *European J. Appl. Math.*, 3(2):147–179, 1992.
- [9] F. Boyer, C. Lapuerta, S. Minjeaud, B. Piar, and M. Quintard. Cahn–Hilliard / Navier–Stokes model for the simulation of three-phase flows. *Transp. Porous Media*, 82:463–483, 2010.
- [10] F. Boyer and F. Nabet. A DDFV method for a Cahn–Hilliard/Stokes phase field model with dynamic boundary conditions. *ESAIM Math. Model. Numer. Anal.*, 51(5):1691–1731, 2017.



- [11] J. W. Cahn, C. M. Elliott, and A. Novick-Cohen. The Cahn-Hilliard equation with a concentration dependent mobility: motion by minus the Laplacian of the mean curvature. *European J. Appl. Math.*, 7(3):287–301, 1996.
- [12] J. W. Cahn and J. E. Hilliard. Free energy of a nonuniform system. i: Interfacial free energy. *J. Chem. Phys.*, 28:258–267, 1958.
- [13] C. Cancès and T. Gallouët. On the time continuity of entropy solutions. *J. Evol. Equ.*, 11(1):43–55, 2011.
- [14] C. Cancès, T. O. Gallouët, and L. Monsaingeon. The gradient flow structure of immiscible incompressible two-phase flows in porous media. *C. R. Acad. Sci. Paris Sér. I Math.*, 353:985–989, 2015.
- [15] C. Cancès, T. O. Gallouët, and L. Monsaingeon. Incompressible immiscible multiphase flows in porous media: a variational approach. *Anal. PDE*, 10(8):1845–1876, 2017.
- [16] C. Cancès and C. Guichard. Convergence of a nonlinear entropy diminishing Control Volume Finite Element scheme for solving anisotropic degenerate parabolic equations. *Math. Comp.*, 85(298):549–580, 2016.
- [17] C. Cancès and F. Nabet. Finite volume approximation of a degenerate immiscible two-phase flow model of cahn–hilliard type. In C. Cancès and P. Omnes, editors, *Finite Volumes for Complex Applications VIII - Methods and Theoretical Aspects*, volume 199 of *Springer Proceedings in Mathematics and Statistics*, pages 431–438, Cham, 2017. Springer International Publishing.
- [18] J. A. Carrillo, B. Düring, D. Matthes, and M. S. McCormick. A lagrangian scheme for the solution of nonlinear diffusion equations using moving simplex meshes. arXiv:1702.01707, 2017.
- [19] P. G. de Gennes. Dynamics of fluctuations and spinodal decomposition in polymer blends. *J. Chem. Phys.*, 72:4756–4763, 1980.
- [20] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhikers guide to the fractional Sobolev spaces. *Bull. Sci. math.*, 136:521573, 2012.
- [21] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.
- [22] W. E and P. Palffy-Muhoray. Phase separation in incompressible systems. *Phys. Rev. E*, 55:R3844–R3846, Apr 1997.
- [23] C. M. Elliott and H. Garcke. On the Cahn-Hilliard equation with degenerate mobility. *SIAM J. Math. Anal.*, 27(2):404–423, 1996.
- [24] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. Ciarlet, P. G. (ed.) et al., in *Handbook of numerical analysis*. North-Holland, Amsterdam, pp. 713–1020, 2000.
- [25] R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *M2AN Math. Model. Numer. Anal.*, 37(6):937–972, 2003.
- [26] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.
- [27] M. E. Gurtin. Generalized Ginzburg-Landau and Cahn-Hilliard equations based on a microforce balance. *Physica D: Nonlinear Phenomena*, 92(3):178 – 192, 1996.
- [28] R. Herbin. An error estimate for a finite volume scheme for a diffusionconvection problem on a triangular mesh. *Numer. Methods Partial Differential Equations*, 11(2):165–173, 1995.
- [29] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [30] O. Junge, D. Matthes, and H. Osberger. A fully discrete variational scheme for solving nonlinear Fokker-Planck equations in multiple space dimensions. to appear in *SIAM J. Numer. Anal.*
- [31] O. A. Ladyzhenskaya and N. N. Ural'tseva. *Linear and quasilinear elliptic equations*, volume 46 of *Mathematics in Science and Engineering*. Academic Press, 1968.
- [32] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.
- [33] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.
- [34] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34:1352–1397, 2009.
- [35] D. Matthes and H. Osberger. A convergent Lagrangian discretization for a nonlinear fourth-order equation. *Found. Comput. Math.*, 17(1):73–126, 2017.
- [36] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.
- [37] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

- [38] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.
- [39] L. Modica and S. Mortola. Un esempio di  $\gamma$ -convergenza. *Boll. Un. Mat. Ital. B*, 14(1):285299, 1980.
- [40] F. Otto and W. E. Thermodynamically driven incompressible fluid mixtures. *J. Chem. Phys.*, 107(23):10177–10184, 1997.
- [41] M. A. Peletier. Variational modelling: Energies, gradient flows, and large deviations. Lecture Notes, Würzburg. Available at <http://www.win.tue.nl/~mpeletie>, Feb. 2014.
- [42] F. Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Progress in Nonlinear Differential Equations and Their Applications 87. Birkhäuser Basel, 1 edition, 2015.
- [43] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

INRIA, UNIV. LILLE, CNRS, UMR 8524 - LABORATOIRE PAUL PAINLEVÉ, F-59000 LILLE ([clement.cances@inria.fr](mailto:clement.cances@inria.fr))

ZENTRUM FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN, 85747 GARCHING, GERMANY ([matthes@ma.tum.de](mailto:matthes@ma.tum.de))

CMAP, CENTRE DE MATHÉMATIQUES APPLIQUÉES, ÉCOLE POLYTECHNIQUE, ROUTE DE SACLAY, 91128 PALAISEAU CEDEX, FRANCE ([flore.nabet@polytechnique.edu](mailto:flore.nabet@polytechnique.edu))