

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. Archive for Rational Mechanics and Analysis, 2019, 233 (2), pp.837-866. 10.1007/s00205-019-01369-6 . hal-01665338v3

HAL Id: hal-01665338
https://hal.science/hal-01665338v3

Submitted on 6 Mar 2019

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

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

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 non-local version of the Cahn-Hilliard dynamics for phase separation in a two-component incompressible and immiscible mixture with linear mobilities. In difference to the celebrated local model with nonlinear mobility, it is only assumed that the divergences of the two fluxes — but not necessarily the fluxes themselves — annihilate each other. Our main result is a rigorous proof of existence of weak solutions. The starting point is the formal representation of the dynamics as a constrained gradient flow in the Wasserstein metric. We then show that time-discrete approximations by means of the incremental minimizing movement scheme converge to a weak solution in the limit. Further, we compare the non-local model to the classical Cahn-Hilliard model in numerical experiments. Our results illustrate the significant speed-up in the decay of the free energy due to the higher degree of freedom for the velocity fields.

1. Motivation and presentation of the model

1.1. **Introduction.** We are interested in a non-local Cahn-Hilliard system

$$\partial_t c_i - \boldsymbol{\nabla} \cdot (m_i c_i \boldsymbol{\nabla} (\mu_i + \Psi_i)) = m_i \theta_i \Delta c_i \quad \text{for } i \in \{1, 2\}, \tag{1a}$$

$$c_1 + c_2 = 1, \tag{1b}$$

$$\mu_1 - \mu_2 = -\alpha \Delta c_1 + \chi(1 - 2c_1), \tag{1c}$$

modelling the flow of an incompressible and immiscible mixture in a bounded convex domain Ω. Above, m_1 and m_2 are positive mobility constants, $\chi > 0$ is a parameter for the chemical activity, and the non-negative numbers θ_1 and θ_2 control the degree of thermal agitation. The system is complemented with initial conditions

$$(c_i)_{|_{t=0}} = c_i^0 \in H^1(\Omega) \text{ with } c_i^0 \geq 0 \text{ and } c_1^0 + c_2^0 = 1 \text{ in } \Omega, \tag{2}$$

and boundary conditions

$$c_i \boldsymbol{\nabla}(\mu_i + \Psi_i) \cdot \boldsymbol{n} = \boldsymbol{\nabla} c_i \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega \times (0, \infty). \tag{3}$$

Equation (1a) can be rewritten as a conservation law

$$\partial_t c_i + \boldsymbol{\nabla} \cdot \boldsymbol{J}_i = 0, \tag{4}$$

where

$$\boldsymbol{J}_i = -m_i c_i \boldsymbol{\nabla} (\mu_i + \Psi_i + \theta_i \log(c_i)).$$

Summing (4) over $i \in \{1, 2\}$ and using (1b) yields

$$\boldsymbol{\nabla} \cdot \boldsymbol{J}_{\text{tot}} = 0 \quad \text{where} \quad \boldsymbol{J}_{\text{tot}} = \boldsymbol{J}_1 + \boldsymbol{J}_2$$

1991 *Mathematics Subject Classification.* 35K65, 35K41, 49J40, 76T99.

Key words and phrases. Multiphase flow, Cahn-Hilliard type system, constrained Wasserstein gradient flow.

denotes the total flux. In particular, we do not require that $\boldsymbol{J}_{\rm tot} = \boldsymbol{0}$ as for the classical (or local) degenerate Cahn-Hilliard model [16, 24, 20]. The relaxation of the constraint from vanishing total flux to divergence free total flux was initially proposed by E and Palffy-Muhoray in [19] and studied formally by Otto and E in [35]. It was in particular noticed in [35] that the system (1) can be interpreted as the constrained Wasserstein gradient flow of some Ginzburg-Landau energy where the velocity field $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2)$ transporting the concentrations $\boldsymbol{c} = (c_1, c_2)$ has to preserve the constraint (1b).

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 global existence of a weak solution to the model (1)–(3). This existence result is obtained by showing the convergence of a minimizing movement scheme *à la* Jordan, Kinderlehrer, and Otto [26]. 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 a bounded open convex subset Ω 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 $c_i(\boldsymbol{x}, t) \in [0, 1]$, $i \in \{1, 2\}$, i.e., the volume ratio of the phase i in the fluid. This leads to the constraint

$$c_1 + c_2 = 1 \qquad \text{in } \Omega \times (0, \infty). \tag{5}$$

We assume that this constraint is already satisfied at the initial time $t = 0$, where the composition of the mixture is given by $\boldsymbol{c}^0 = (c_1^0, c_2^0) : \Omega \to [0, 1]^2$, i.e.

$$c_1^0 + c_2^0 = 1 \qquad \text{in } \Omega. \tag{6}$$

We further assume that both phases have positive mass.

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

$$\partial_t c_i + \boldsymbol{\nabla} \cdot (c_i \boldsymbol{v}_i) = 0 \quad \text{in } \Omega \times (0, \infty), \tag{7}$$

where $\boldsymbol{v}_i : \Omega \times \mathbb{R}_+ \to \mathbb{R}^d$ denotes the speed of the phase i, and $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2)$. To enforce mass conservation, the boundary $\partial\Omega$ of Ω is assumed to be impervious, hence

$$c_1 \boldsymbol{v}_1 \cdot \boldsymbol{n} = c_2 \boldsymbol{v}_2 \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega, \tag{8}$$

where $\boldsymbol{n}$ denotes the outward normal to $\partial\Omega$, so that Gauss' theorem implies

$$\int_\Omega c_i(\boldsymbol{x}, t) \mathrm{d}\boldsymbol{x} = \int_\Omega c_i^0(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = |\Omega| \overline{c}_i, \qquad t \geq 0,\ i \in \{1, 2\}. \tag{9}$$

Consequently, at each time $t \geq 0$, the saturations $\boldsymbol{c}(t)$ belong to the set $\boldsymbol{\mathcal{A}} = \mathcal{A}_1 \times \mathcal{A}_2$, where

$$\mathcal{A}_i = \left\{ c_i : \Omega \to [0, 1] \text{ measurable} \;\middle|\; \int_\Omega c_i(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = |\Omega| \overline{c}_i \right\}, \qquad i \in \{1, 2\}.$$

To each configuration $\boldsymbol{c} \in \boldsymbol{\mathcal{A}}$, we associate the energy

$$\mathcal{E}(\boldsymbol{c}) = \mathcal{E}_{\rm Dir}(\boldsymbol{c}) + \mathcal{E}_{\rm chem}(\boldsymbol{c}) + \mathcal{E}_{\rm therm}(\boldsymbol{c}) + \mathcal{E}_{\rm cons}(\boldsymbol{c}) + \mathcal{E}_{\rm ext}(\boldsymbol{c}),$$

whose components are as follows.

- **Dirichlet energy:** for given $\alpha_1, \alpha_2 > 0$,

$$\mathcal{E}_{\rm Dir}(\boldsymbol{c}) := \begin{cases} \displaystyle\sum_{i \in \{1,2\}} \frac{\alpha_i}{2} \int_\Omega |\boldsymbol{\nabla} c_i|^2 \mathrm{d}\boldsymbol{x} & \text{if } \boldsymbol{c} \in H^1(\Omega)^2, \\ +\infty & \text{otherwise,} \end{cases}$$

penalizes the spatial variations of the saturation profiles.

- **Chemical energy:** for a given $\chi > 0$,
$$\mathcal{E}_{\rm chem}(\boldsymbol{c}) = \chi \int_\Omega c_1 c_2 \mathrm{d}\boldsymbol{x}$$
measures the impurity of the mixture.
- **Thermal energy:** for given $\theta_1, \theta_2 \geq 0$,
$$\mathcal{E}_{\rm therm}(\boldsymbol{c}) = \sum_{i\in\{1,2\}} \theta_i \int_\Omega H(c_i)\mathrm{d}\boldsymbol{x}, \quad \text{with} \quad H(c) = c\log(c) - c + 1 \geq 0,$$
is the free thermal energy. The case $\theta_1 = \theta_2 = 0$ is called the deep quench limit [9].
- **Constraint:**
$$\mathcal{E}_{\rm cons}(\boldsymbol{c}) = \int_\Omega E_{\rm cons}(\boldsymbol{c})\mathrm{d}\boldsymbol{x}, \quad \text{where} \quad E_{\rm cons}(\boldsymbol{c}) = \begin{cases} 0 & \text{if } c_1 + c_2 = 1, \\ +\infty & \text{otherwise,} \end{cases}$$
realizes the constraint (5).
- **Exterior potential:** for given external potentials $\Psi_1, \Psi_2 \in H^1(\Omega)$,
$$\mathcal{E}_{\rm ext}(\boldsymbol{c}) = \sum_{i\in\{1,2\}} \int_\Omega c_i(\boldsymbol{x})\Psi_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$$
is the potential energy related to exterior forces like gravity or electrostatic forces.

All of these components are convex in $\boldsymbol{c}$ *except* for the chemical energy $\mathcal{E}_{\rm chem}$, which, however, is smooth. Denote by
$$\boldsymbol{\mathcal{X}} = \left\{ \boldsymbol{c} \in H^1(\Omega;[0,1])^2 \,\middle|\, c_1 + c_2 = 1 \text{ a.e. in } \Omega \right\}$$
the domain of $\mathcal{E}$, i.e.,
$$\mathcal{E}(\boldsymbol{c}) < \infty \quad \Leftrightarrow \quad \boldsymbol{c} \in \boldsymbol{\mathcal{X}}. \tag{10}$$

Before entering into the rigorous derivation of the PDEs that govern the gradient flow of the energy $\mathcal{E}(\boldsymbol{c})$ with respect to a tensorized Wasserstein distance, we provide formal calculations based on the framework of generalized gradient flows of [33, 36, 11] in order to identify the underlying PDEs.

The motion of the phases induces the viscous dissipation
$$\mathfrak{D}(\boldsymbol{c}, \boldsymbol{V}) = \sum_{i\in\{1,2\}} \frac{1}{2m_i} \int_\Omega c_i |\boldsymbol{v}_i|^2 \mathrm{d}\boldsymbol{x}, \qquad \forall \boldsymbol{c} \in \boldsymbol{\mathcal{A}},\ \forall \boldsymbol{V} \in \boldsymbol{\mathcal{Z}}(\boldsymbol{c}), \tag{11}$$
where m_i is the mobility coefficient of the ith phase, and where
$$\boldsymbol{\mathcal{Z}}(\boldsymbol{c}) = \left\{ \boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2) : \Omega \to \left(\mathbb{R}^d\right)^2 \,\middle|\, c_i \boldsymbol{v}_i \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega \right\}$$
denotes the space of the admissible (but unconstrained) vector fields. This quantity is closely related to the tensorized Wasserstein distance through Benamou-Brenier formula [5], see (27) later on. We suppose as in [11] that at each time $t \geq 0$, the phase speeds $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2)$ is selected by the following *steepest descent condition*:
$$\boldsymbol{V} \in \underset{\widetilde{\boldsymbol{V}} = (\widetilde{\boldsymbol{v}}_1, \widetilde{\boldsymbol{v}}_2) \in \boldsymbol{\mathcal{Z}}(\boldsymbol{c})}{\operatorname{argmin}} \left(\mathfrak{D}(\boldsymbol{c}, \widetilde{\boldsymbol{V}}) + \sup_{\boldsymbol{w} \in \partial\mathcal{E}(\boldsymbol{c})} \sum_{i\in\{1,2\}} \int_\Omega c_i \widetilde{\boldsymbol{v}}_i \cdot \boldsymbol{\nabla} w_i \mathrm{d}\boldsymbol{x} \right), \tag{12}$$

where $\mathcal{E}$'s subdifferential

$$\partial\mathcal{E}(\boldsymbol{c}) = \left\{ \boldsymbol{w} \in L^2(\Omega;\mathbb{R}^d)^2 \middle| \mathcal{E}(\widehat{\boldsymbol{c}}) - \mathcal{E}(\boldsymbol{c}) - \int_\Omega \boldsymbol{w}\cdot(\widehat{\boldsymbol{c}}-\boldsymbol{c})\mathrm{d}\boldsymbol{x} \geq o\left(\|\widehat{\boldsymbol{c}}-\boldsymbol{c}\|_{L^2(\Omega)}\right) \right\}$$

is non-empty at each $\boldsymbol{c} \in \boldsymbol{\mathcal{X}}$ with $c_1 \in H^2(\Omega)$ and $f(c_1) \in L^2(\Omega)$, and there, its elements $\boldsymbol{w} = (w_1, w_2)$ are characterized by (c.f. [11])

$$w_1 - w_2 = -\alpha\Delta c_1 + \chi(1-2c_1) + f(c_1) + (\Psi_1 - \Psi_2). \tag{13}$$

In formula (13), we have set $\alpha = \alpha_1 + \alpha_2$ and

$$f(c_1) = \log\left(\frac{c_1^{\theta_1}}{(1-c_1)^{\theta_2}}\right) = \theta_1\log(c_1) - \theta_2\log(c_2). \tag{14}$$

The integral in (12) has to be understood as the evaluation of the element $\boldsymbol{w} = (w_1, w_2)$ in $\mathcal{E}$'s subdifferential at $\boldsymbol{c}$ on the infinitesimal changes $\partial_t c_i = -\boldsymbol{\nabla}\cdot(c_i\boldsymbol{v}_i)$ induced by $\boldsymbol{V}$. Since condition (13) only determines the difference $w_1 - w_2$, but not the w_i individually, is easily seen that the maximum above is $+\infty$ unless $\nabla\cdot(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = 0$ holds. That implies that any minimizer $\boldsymbol{V}$ is such that the constraint (5) is preserved.

In order to identify $\boldsymbol{V}$, we swap minimization and maximization in the variational problem (12). The inner minimization is a quadratic problem in $\boldsymbol{V}$, with solution

$$\boldsymbol{v}_i = -m_i\boldsymbol{\nabla}w_i \quad \text{on } \{c_i > 0\} \text{ for } i \in \{1,2\}. \tag{15}$$

The outer maximization in then a quadratic problem for $\boldsymbol{w} \in \partial\mathcal{E}(\boldsymbol{c})$, which amounts to

$$-\boldsymbol{\nabla}\cdot\left(\sum_{i\in\{1,2\}} m_i c_i \boldsymbol{\nabla}w_i\right) = 0. \tag{16}$$

Together with the condition (13), this elliptic equation determines $\boldsymbol{w}$ uniquely.

To sum up, the system of partial differential equations implied on the solution to the variational problem (12) is

$$\partial_t c_i - \boldsymbol{\nabla}\cdot(m_i c_i \boldsymbol{\nabla}w_i) = 0 \quad \text{for } i \in \{1,2\}, \tag{17a}$$

$$c_1 + c_2 = 1, \tag{17b}$$

$$w_1 - w_2 = -\alpha\Delta c_1 + \chi(1-2c_1) + f(c_1) + (\Psi_1 - \Psi_2), \tag{17c}$$

to be satisfied in $\Omega\times(0,\infty)$. The system relation (17) is complemented by homogeneous Neumann boundary conditions

$$-\boldsymbol{\nabla}c_1\cdot\boldsymbol{n} = 0 \quad \text{on } \partial\Omega\times\mathbb{R}_+ \tag{18}$$

and the no-flux conditions (8). Notice that with these boundary conditions, the set (17) of equations is (formally) closed in the sense that if $\boldsymbol{c}$ is known at some instance of time, then the w_i are uniquely determined (up to irrelevant global additive constants) by (17c) and by the elliptic equation (16) that follows from adding (17a) for $i = 1, 2$ in combination with the conservation $\partial_t(c_1 + c_2) = 0$ implied by (17b). Hence $\boldsymbol{c}$'s time derivatives $\partial_t c_i$ are determined as well.

At this point, it is natural to define the chemical potential μ_i of the phase i by

$$\mu_i = w_i - \theta_i\log(c_i) - \Psi_i, \qquad i \in \{1,2\}, \tag{19}$$

so that (17c) turns to

$$\mu_1 - \mu_2 = -\alpha\Delta c_1 + \chi(1-2c_1). \tag{20}$$

So far, $\boldsymbol{\mu}=(\mu_1,\mu_2)$ is only defined up to an additive constant. This degree of freedom is eliminated by imposing for almost all $t>0$ that

$$\int_\Omega \overline{\mu}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}=0, \quad \text{where } \overline{\mu}=c_1\mu_1+c_2\mu_2. \tag{21}$$

Whilst the phase chemical potential μ_i have a clear physical sense only on $\{c_i>0\}$, the global chemical potential $\overline{\mu}$ remains meaningful in the whole Ω. In particular, its spatial variations can be controlled, and thus $\overline{\mu}$ itself too thanks to a Poincaré-Wirtinger inequality.

With the help of the chemical potentials, the system (17) turns into (1), to be complemented with boundary conditions (3) and initial conditions (2).

In the next section, we highlight some differences between the non-local model (1) and the local degenerate Cahn-Hilliard model that has been studied for instance in [20].

1.3. **Comparison with the classical degenerate Cahn-Hilliard model.** Even in the simple situation where the external potentials Ψ_i are equal to 0 and where $\theta_1=\theta_2=0$, the system (17) differs as soon as $d\geq 2$ from the local degenerate Cahn-Hilliard model that can be written as

$$\partial_t c-\boldsymbol{\nabla}\cdot(\eta(c)\boldsymbol{\nabla}\mu)=0, \qquad \mu=-\alpha\Delta c+\chi(1-2c), \qquad \eta(c)=\frac{m_1m_2c(1-c)}{m_1c+m_2(1-c)}. \tag{22}$$

We refer to [20] for the existence of weak solutions to (22) (complemented with suitable boundary conditions) and to [4] for the extension of the model to the case of N phases ($N\geq 3$). Here, μ is the generalized chemical potential that is defined as the difference of the phase chemical potentials.

The energy $\widetilde{\mathcal{E}}$ associated to (22) is similar to the one of our problem, i.e.,

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

But both the equation governing the motion (7) and the dissipation (11) have to be modified. More precisely, the continuity equation (7) must be replaced by its nonlinear counterpart

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

while the dissipation is now given by

$$\widetilde{\mathfrak{D}}(c,\boldsymbol{v})=\int_\Omega\eta(c)|\boldsymbol{v}|^2\mathrm{d}\boldsymbol{x}.$$

Therefore, the PDEs (22) can still be interpreted as the gradient flow of the energy $\widetilde{\mathcal{E}}$, 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

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

has to be equipped with the weighted Wasserstein metric corresponding to the concave mobility η. We refer to [18] for the description of the corresponding metric and to [29] for the rigorous recovery of (22) by a gradient flow approach.

The difference between the non-local model (17) and the local one (22) can also be seen as follows. Summing the first equation of (17) for $i\in\{1,2\}$ yields

$$\boldsymbol{\nabla}\cdot\boldsymbol{J}_{\mathrm{tot}}=0, \quad \text{where} \quad \boldsymbol{J}_{\mathrm{tot}}=\sum_{i\in\{1,2\}}c_i\boldsymbol{v}_i=-\sum_{i\in\{1,2\}}m_ic_i\boldsymbol{\nabla}\mu_i. \tag{23}$$

The equation for c_1 can then be rewritten

$$\partial_t c_1 + \boldsymbol{\nabla} \cdot (\rho(c_1)\boldsymbol{J}_{\mathrm{tot}} - \eta(c_1)\boldsymbol{\nabla}(\mu_1 - \mu_2)) = 0, \quad \text{with} \quad \rho(c) = \frac{m_1 c}{m_1 c + m_2(1-c)}. \tag{24}$$

Thus our model (17) boils down to the local Cahn-Hilliard equation as soon as $\boldsymbol{J}_{\mathrm{tot}} \equiv \boldsymbol{0}$. This is the case when $d = 1$ because of (23), but no longer if $d \geq 2$. Since our non-local model does not impose that $\boldsymbol{J}_1 = -\boldsymbol{J}_2$, it allows for additional motions. These motions —corresponding to the transport term $\boldsymbol{\nabla} \cdot (\rho(c_1)\boldsymbol{J}_{\mathrm{tot}})$ in (24)— contribute to the dissipation as shows the formula

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

Therefore, and as already noticed by Otto and E in [35], the instantaneous dissipation corresponding to a phase configuration $\boldsymbol{c}$ is greater for the non-local model (17) than for the local model (22) and the energy decreases faster.

1.4. **Numerical illustration.** The goal of this section is to illustrate the behavior of the model (17) and to compare it with the classical degenerate Cahn-Hilliard problem (22). In order to solve numerically (17) we use an implicit in time finite volume scheme with upstream mobility described in [14] and inspired from the oil engineering context [22]. The mesh is triangular and assumed to fulfill the so-called orthogonality condition [25, 21] (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 [14], 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 [14] that this system admits (at least) one solution that is computed thanks to the Newton-Raphson method.

Concerning the problem (22), we use a similar approach, but since the mobility function η 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 [13]). 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 [7, 31, 27, 15]*. 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* (5) *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. [6] *to our setting , see* [10].

We propose two different test cases that will allow to illustrate the difference between the local model (22) and the non-local model (17). 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.

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 and we choose $\alpha = (3.6).10^{-4}$ and $\chi = 0.8$. Since $\alpha \ll \chi$, it follows from the Modica and Mortola' result [34] 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.

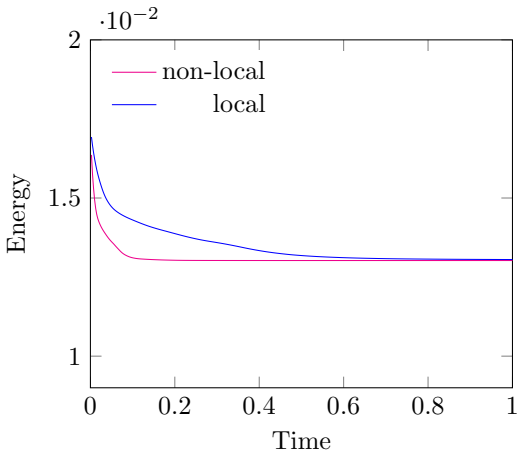


FIGURE 1. Evolution of the energy $\mathcal{E}(\boldsymbol{c})$ along time for the non-local model (17) and the local one (22). The decay of the energy is faster for the non-local model, as shown in Section 1.3.

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:

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

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 choose $\alpha = 3.10^{-4}$ and $\chi = 0.96$ and we plot on Figure 3 some snapshots to illustrate the spinodal decomposition corresponding to models (17) and (22). 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.

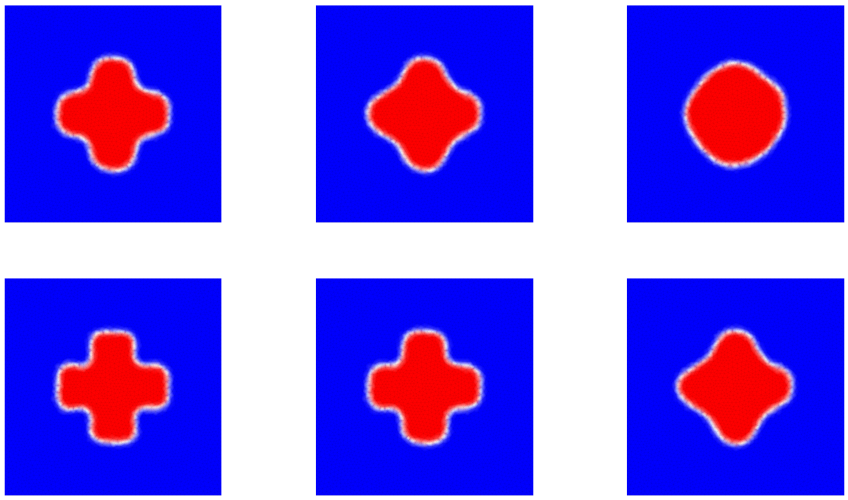


FIGURE 2. Evolution along time of the numerical solution to the non-local problem (17) (top) and to the local problem (22) (bottom). Snapshots at time $t = 10^{-2}$ (left), $t = 2.10^{-2}$ (middle), and $t = 10^{-1}$ (right).

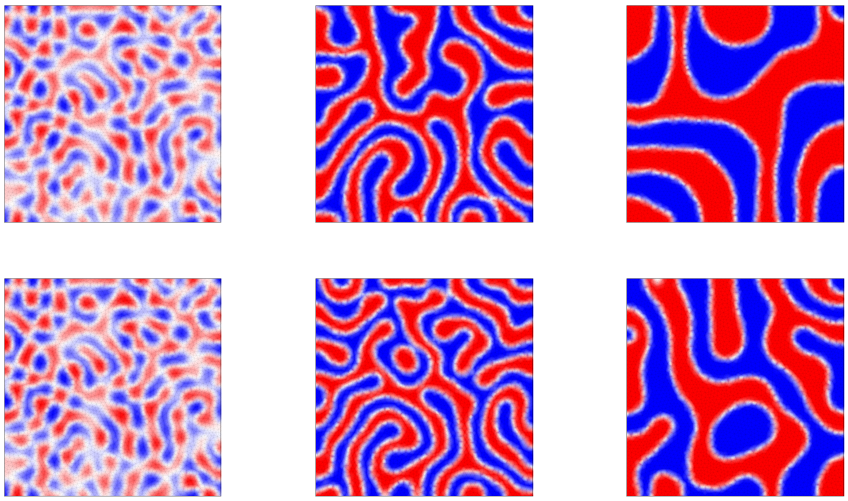


FIGURE 3. Snapshots at times $t = 6.10^{-3}$ (left), $t = 5.10^{-2}$ (middle), and $t = 1$ (right) illustrating the spinodal decomposition governed by model (17) (top) and model (22) (bottom).

2. Wasserstein gradient flow, JKO scheme and main result

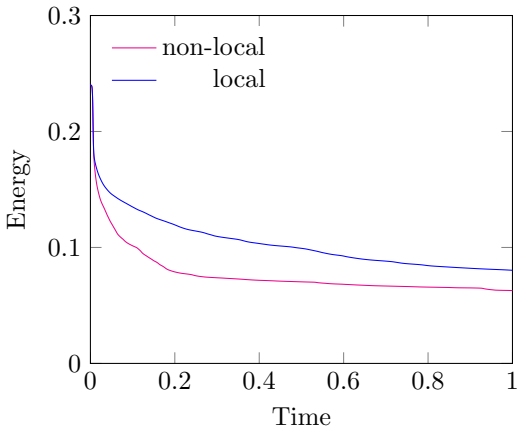


FIGURE 4. Evolution of the energy $\mathcal{E}(\boldsymbol{c})$ along time for the non-local model (17) and the local one (22) for the spinodal decomposition test case.

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 [37] for an introduction to optimal transportation and to the Wasserstein distances, and to Villani's big book [38] for a more complete presentation.

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

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

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

$$W_i(c_i,\check{c}_i) = \inf_{\boldsymbol{t} \text{ s.t. } \check{c}_i = \boldsymbol{t}\#c_i} \left(\frac{1}{m_i} \int_\Omega |\boldsymbol{x} - \boldsymbol{t}(\boldsymbol{x})|^2 c_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \right)^{1/2}, \qquad i \in \{1,2\}. \tag{25}$$

In (25), the infimum is in fact a minimum, and $\boldsymbol{t}$ 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 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 $\boldsymbol{W}$ on $\boldsymbol{\mathcal{A}}$ by

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

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

$$\frac{1}{2}W_i^2(c_i,\check{c}_i) = \sup_{\substack{\varphi_i \in L^1(c_i), \psi_i \in L^1(\check{c}_i) \\ \varphi_i(\boldsymbol{x}) + \psi_i(\boldsymbol{y}) \le \frac{|\boldsymbol{x}-\boldsymbol{y}|^2}{2m_i}}} \int_\Omega c_i(\boldsymbol{x})\varphi_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega \check{c}_i(\boldsymbol{y})\psi_i(\boldsymbol{y})\mathrm{d}\boldsymbol{y}. \tag{26}$$

Here, $L^1(\rho)$ denotes the sets of integrable functions for the measure with density ρ. Here again, the supremum is in fact a maximum, and the Kantorovich potentials (φ_i, ψ_i) achieving the sup in (26) are $\mathrm{d}c_i \otimes \mathrm{d}\check{c}_i$ unique up to an additive constant. The optimal transportation $\boldsymbol{t}_i$ sending c_i on $\check{c}_i$ achieving the inf in (25) is related to the Kantorovich potential by

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

with φ_i achieving the sup in (26). As a consequence, the formula (25) provides

$$W_i^2(c_i, \check{c}_i) = m_i \int_\Omega c_i \left|\boldsymbol{\nabla}\varphi_i\right|^2 \mathrm{d}\boldsymbol{x}, \qquad i \in \{1,2\} \tag{27}$$

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 $\boldsymbol{c}^{n-1} \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$, then define the functional $\mathcal{F}_\tau^n : \boldsymbol{\mathcal{A}} \to \mathbb{R} \cup \{+\infty\}$ by setting

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

The functional $\mathcal{F}_\tau^n$ is bounded from below since all its components are. Then we define

$$\boldsymbol{c}^n \in \operatorname*{argmin}_{\boldsymbol{c}\in\boldsymbol{\mathcal{A}}} \mathcal{F}_\tau^n(\boldsymbol{c}). \tag{28}$$

In case that $\mathcal{F}_\tau^n$ possesses several mimimizers, a selection must be made: if c_1^{n-1} and c_2^{n-1} are both positive a.e. in Ω, then any minimizer is fine. However, if one of the two densities vanishes on a set of positive measure, then we need to select $\boldsymbol{c}^n$ such that it is an accumulation point in the weak $H^2(\Omega)^2$-topology for $\delta \to 0$ of the set of minimizers of the functionals

$$\mathcal{F}_\tau^{n,\delta}(\boldsymbol{c}) = \mathcal{E}(\boldsymbol{c}) + \frac{1}{2\tau}\boldsymbol{W}^2(\boldsymbol{c}, \boldsymbol{c}^{n-1,\delta}), \quad \text{with} \quad c_i^{n-1,\delta} = (1-\delta)c_i^{n-1} + \delta\overline{c}_i \tag{29}$$

where $\overline{c}_i$ is defined by (9). Under certain conditions, for instance in the thermally agitated situation where $\theta_1, \theta_2 > 0$, it is known a priori that both c_1^n and c_2^n are strictly positive, and hence no selection is necessary, see e.g. Lemma 3.1. The proof of the existence of $\boldsymbol{c}^n$ solution to (28) is given below. The proof that one can indeed select the minimizer in the way described above is postponed to Corollary 3.3 in Section 3.4, where we show compactness properties of the set of minimizers.

Proposition 2.1. *For any $\boldsymbol{c}^{n-1} \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$, there exists (at least) one solution $\boldsymbol{c}^n \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$ to the minimization scheme* (28).

Proof. Let $\left(\boldsymbol{c}^{n,k}\right)_{k\geq 0}$ be a minimizing sequence with $\boldsymbol{c}^{n,0} = \boldsymbol{c}^{n-1}$ and such that

$$\mathcal{E}(\boldsymbol{c}^{n,k}) \leq \mathcal{E}(\boldsymbol{c}^{n-1}) < \infty \qquad \text{for all } k \geq 0. \tag{30}$$

We infer from (10) that $\boldsymbol{c}^{n,k} \in \boldsymbol{\mathcal{X}}$, hence $0 \leq c_i^{n,k} \leq 1$ for all $k \geq 0$ and $i \in \{1,2\}$. Moreover, in view of the presence of the Dirichlet energy in $\mathcal{E}$, it follows by means of the Poincaré-Wirtinger inequality that the $c_i^{n,k}$ are k-uniformly bounded in $H^1(\Omega)$. Passing to a subsequence if necessary, the $c_i^{n,k}$ converge to limits c_i^n, weakly in $H^1(\Omega)$ thanks to Alaoglu's theorem, and also strongly in $L^1(\Omega)$ by Rellich's compactness lemma. All the components of $\mathcal{E}$ are sequentielly lower semi-continuous with respect to this convergence, and so

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

Finally, also the Wasserstein distance is lower semi-continuous with respect to strong convergence in $L^1(\Omega)$,

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

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

Each sequence $(\boldsymbol{c}^n)_{n\geq 1}$ of iterated solutions to the scheme (28) is accompanied by approximate phase potentials $(\boldsymbol{\mu}^n)_{n\geq 1}$: these are introduced in such a way that in the time-continuous limit, the continuity equations $\partial_t c_i = \nabla\cdot[m_i c_i \nabla(\mu_i + \Psi_i)]$ hold. The suitable quantities μ_i^n are identified (somewhat *a posteriori*) by comparing the optimality conditions for (28) with the persued limit PDE system (1).

In order to justify the formal definition of $\boldsymbol{\mu}^n$ that we give below in (33), we anticipate an auxiliary result from Section 3.4, that can be understood as a formulation of the time-discrete Euler-Lagrange equations. It involves the (backward) Kantorovich potentials $\boldsymbol{\varphi}^n = (\varphi_1^n, \varphi_2^n)$ sending $\boldsymbol{c}^n$ on $\boldsymbol{c}^{n-1}$. A subtle point is to overcome the inherent non-uniqueness of φ_1^n and φ_2^n — particularly if one of the c_i^n's vanishes on a set of positive measure — by making a suitable selection, as will be explained in Section 3.4. We note that the intricate selection procedure for the minimizer $\boldsymbol{c}^n$ in (28) enters precisely at this point.

Lemma 2.2. *At each step $n \geq 1$, there exist Kantorovich potentials $\boldsymbol{\varphi}^n = (\varphi_1^n, \varphi_2^n)$ for sending $\boldsymbol{c}^n$ to $\boldsymbol{c}^{n-1}$ such that $F^n : \Omega \to \mathbb{R}$, given by*

$$F^n = \frac{\varphi_1^n}{\tau} - \frac{\varphi_2^n}{\tau} - \alpha\Delta c_1^n + f(c_1^n) + \chi(1-2c_1^n) + \Psi_1 - \Psi_2, \tag{31}$$

satisfies

$$F^n \leq 0 \quad \text{a.e. in } \{c_1^n > 0\}, \qquad F^n \geq 0 \quad \text{a.e. in } \{c_2^n > 0\}. \tag{32}$$

With the particular choice of $\boldsymbol{\varphi}^n$ from the lemma, we define $\boldsymbol{\mu}^n = (\mu_1^n, \mu_2^n) : \Omega \to \mathbb{R}^2$ by

$$\mu_1^n := -\frac{\varphi_1^n}{\tau} - \Psi_1 - \theta_1 \log(c_1^n) + (F^n)_+, \quad \text{and} \quad \mu_2^n := -\frac{\varphi_2^n}{\tau} - \Psi_2 - \theta_2 \log(c_2^n) + (F^n)_-. \tag{33}$$

Actually, since F^n is invariant under simultaneous addition of a global constant to both φ_1^n and φ_2^n, we can assume that their normalization is chosen such that

$$\int_\Omega \overline{\mu}^n \mathrm{d}\boldsymbol{x} = 0, \qquad \text{where} \quad \overline{\mu}^n = c_1^n\mu_1^n + c_2^n\mu_2^n. \tag{34}$$

Note that it follows directly from the definition of $\boldsymbol{\mu}^n$ that

$$\mu_1^n - \mu_2^n = -\alpha\Delta c_1^n + \chi(1-2c_1^n) \qquad \text{a.e. in } \Omega. \tag{35}$$

Moreover, thanks to (32), we have that

$$\mu_i^n = -\frac{\varphi_i^n}{\tau} - \Psi_i - \theta_i \log(c_i^n) \qquad \text{a.e. in } \{c_i^n > 0\}. \tag{36}$$

With these quantities at hand, we define further the piecewise constant interpolants $\boldsymbol{c}_\tau : \mathbb{R}_+ \to \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mu}_\tau : \mathbb{R}_+ \to \mathbb{R}^2$ by

$$\boldsymbol{c}_\tau(t) = (c_1^n, c_2^n), \quad \boldsymbol{\mu}_\tau(t) = (\mu_1^n, \mu_2^n) \qquad \text{for } t \in ((n-1)\tau, n\tau]. \tag{37}$$

Any pair $(\boldsymbol{c}_\tau, \boldsymbol{\mu}_\tau)$ that has been obtained in this way from an iterated minimizer $(\boldsymbol{c}^n)_{n\geq 1}$ will be referred to as *τ-approximate solution* below.

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 τ tends to 0.

Definition 2.3. *A pair $(\boldsymbol{c}, \boldsymbol{\mu})$ of phase concentrations $\boldsymbol{c} = (c_1, c_2) : \Omega \to [0,1]$ and phase potentials $\boldsymbol{\mu} = (\mu_1, \mu_2) : \Omega \to \mathbb{R}$ is said to be a* weak solution *to the problem* (17), (2), *and* (3) *if*

- ***Regularity of concentrations:*** $c_i \in L^\infty(\mathbb{R}_+; H^1(\Omega)) \cap L^2_{\mathrm{loc}}(\mathbb{R}_+; H^2(\Omega)) \cap C(\mathbb{R}_+; L^2(\Omega))$*;*
- ***Regularity of potentials:*** $\mu_i \in L^2_{\mathrm{loc}}(\mathbb{R}_+; L^{q_d}(\Omega))$ *with* $q_d = 2$ *if* $d \in \{1,2\}$, *and* $q_d = \frac{3}{2}$ *if* $d = 3$, *as well as* $c_i \boldsymbol{\nabla} \mu_i \in L^2(\mathbb{R}_+, L^2(\Omega))^d$, *and (normalization)*

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

- ***Initial and boundary conditions:*** $c_i(\cdot,0) = c_i^0$ *on* Ω, *and* $\boldsymbol{\nabla} c_i \cdot \boldsymbol{n} = 0$ *on* $\partial\Omega \times \mathbb{R}_+$*;*
- ***Volume constraint:*** $c_1 + c_2 = 1$ *a.e. in* $\Omega \times \mathbb{R}_+$*;*
- ***Continuity equations:*** *for all* $\xi \in C^2(\overline{\Omega})$ *and all* $t_1, t_2 \in \mathbb{R}_+$ *with* $t_2 \geq t_1$,

$$\int_\Omega (c_i(\boldsymbol{x},t_2) - c_i(\boldsymbol{x},t_1))\xi(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + m_i \int_{t_1}^{t_2} \int_\Omega (c_i \boldsymbol{\nabla}(\mu_i + \Psi_i) + \theta_i \boldsymbol{\nabla} c_i) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t = 0; \tag{38}$$

- ***Steepest descent:*** (20) *holds almost everywhere in* $\Omega \times \mathbb{R}_+$, *i.e.*

$$\mu_1 - \mu_2 = -\alpha\Delta c_1 + \chi(1 - 2c_1). \tag{39}$$

Here is the convergence theorem for the minimization scheme. The global existence of a weak solution is a by-product. All along the paper, T denotes an arbitrary finite time horizon, and we make use of the shorten notation Q_T for the space-time cylinder $\Omega \times (0,T)$.

Theorem 2.4 (Convergence of the minimizing movement scheme)**.** *Assume that parameters $\alpha_1, \alpha_2, \chi > 0$ and $\theta_1, \theta_2 \geq 0$ as well as external potentials $\Psi_1, \Psi_2 \in H^1(\Omega)$ are prescribed. Let $\boldsymbol{c}^0 = (c_1^0, c_2^0)$ be an initial condition satisfying* (6).

Then, for any sequence $(\tau_n)_{n\geq 1} \subset (0,1)$ with $\tau_n \to 0$ and a corresponding sequence of τ_n-approximate solutions $\left(\boldsymbol{c}_{\tau_n}, \boldsymbol{\mu}_{\tau_n}\right)_{n\geq 1}$, one can select a sub-sequence (not relabeled) such that,

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

and the limit $(\boldsymbol{c}, \boldsymbol{\mu})$ is a weak solution in the sense of Definition 2.3.

We recall that the definition of τ-approximate solutions involves the selection of particular minimizers in (28) unless all densities have full support.

3. Proof of Theorem 2.4

We first establish some estimates on the approximate solution $\boldsymbol{c}_\tau$. The very classical energy estimate and some straightforward consequences are exposed in Section 3.1. In Section 3.2, we show that the approximate solution remains bounded away from 0 and 1 if the thermal diffusion coefficients θ_i are positive. Then in Section 3.3, we make use of the flow interchange technique initially introduced in [30] to get enhanced regularity estimates on the approximate solutions. The Euler-Lagrange equation are then obtained in Section 3.4 thanks to a linearization technique inspired from the work of Maury *et al.* [32]. The convergence of the approximate solution towards a weak solution is finally established in Section 3.5.

3.1. **Energy and distance estimate.** By definition of $\boldsymbol{c}^n$ in (28), one has $\mathcal{F}_\tau^n(\boldsymbol{c}^n) \leq \mathcal{F}_\tau^n(\boldsymbol{c}^{n-1})$, i.e.,

$$\mathcal{E}(\boldsymbol{c}^n) + \frac{1}{2\tau}\boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1}) \leq \mathcal{E}(\boldsymbol{c}^{n-1}), \qquad \forall n \geq 1. \tag{40}$$

Summing (40) over n, and using that $\mathcal{E}(\boldsymbol{c}) \geq \mathcal{E}_\star > -\infty$ for all $\boldsymbol{c} \in \boldsymbol{\mathcal{A}}$, we obtain the square distance estimate

$$\frac{1}{\tau}\sum_{n\geq 1} \boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1}) \leq 2\left(\mathcal{E}(\boldsymbol{c}^0) - \mathcal{E}_\star\right) < +\infty. \tag{41}$$

This readily gives the approximate 1/2-Hölder estimate

$$\boldsymbol{W}(\boldsymbol{c}^{n_2}, \boldsymbol{c}^{n_1}) \leq C\sqrt{|n_2 - n_1|\tau}, \qquad \forall n_1, n_2 \geq 0. \tag{42}$$

Bearing in mind the definition (37) of the approximate solution $\boldsymbol{c}_\tau$, we get that

$$\boldsymbol{W}(\boldsymbol{c}_\tau(t), \boldsymbol{c}_\tau(s)) \leq C\sqrt{|t-s| + \tau}, \qquad \forall s, t \geq 0. \tag{43}$$

We also deduce from the energy estimate (40) that

$$\mathcal{E}(\boldsymbol{c}^n) \leq \mathcal{E}(\boldsymbol{c}^0) < \infty. \tag{44}$$

We deduce that

$$\int_\Omega |\boldsymbol{\nabla} c_1^n|^2 \mathrm{d}\boldsymbol{x} \leq \frac{2}{\alpha}\left(\mathcal{E}(\boldsymbol{c}^0) + \sum_{i\in\{1,2\}} \|\Psi_i\|_{L^1}\right) < \infty, \qquad \forall n \geq 0. \tag{45}$$

3.2. **Positivity of the discrete solution in presence of thermal agitation.** The formula (36) 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 [37, Lemma 8.6].

Lemma 3.1. *Let $\boldsymbol{c}^n$ be a minimizer of $\mathcal{F}_\tau^n$ as in (28). Assume that $\theta_i > 0$, then $c_i^n > 0$ a.e. in Ω. Moreover, $\theta_i \log(c_i^n) \in L^1(\Omega)$.*

Proof. With a slight abuse of notation, we denote by $\overline{\boldsymbol{c}} = (\overline{c}_1, \overline{c}_2)$ the constant element of $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ given by (9). Fix n. Given $\epsilon \in (0,1)$, we introduce $\boldsymbol{c}^\epsilon \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ by

$$c_i^\epsilon = \epsilon\overline{c}_i + (1-\epsilon)c_i^n = c_i^n + \epsilon(\overline{c}_i - c_i^n). \tag{46}$$

Note that $c_i^\epsilon > 0$ everywhere in Ω, and that

$$c_i^\epsilon - \overline{c}_i = (1-\epsilon)(c_i^n - \overline{c}_i) \tag{47}$$

By optimality of $\boldsymbol{c}^n$, the inequality

$$\mathcal{E}_{\mathrm{therm}}(\boldsymbol{c}^n) - \mathcal{E}_{\mathrm{therm}}(\boldsymbol{c}^\epsilon) \leq \mathcal{E}_{\mathrm{Dir}}(\boldsymbol{c}^\epsilon) - \mathcal{E}_{\mathrm{Dir}}(\boldsymbol{c}^n) + \mathcal{E}_{\mathrm{chem}}(\boldsymbol{c}^\epsilon) - \mathcal{E}_{\mathrm{chem}}(\boldsymbol{c}^n) + \mathcal{E}_{\mathrm{ext}}(\boldsymbol{c}^\epsilon) - \mathcal{E}_{\mathrm{ext}}(\boldsymbol{c}^n) + \frac{1}{2\tau}\left(\boldsymbol{W}^2(\boldsymbol{c}^\epsilon, \boldsymbol{c}^{n-1}) - \boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1})\right)$$

holds for all $\epsilon \in (0,1)$. Then — with generic constants C that may change from line to line, but are independent of ϵ — it follows directly from the definition of $\boldsymbol{c}^\epsilon$ in (46) that

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

that

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

and that

$$\mathcal{E}_{\mathrm{ext}}(\boldsymbol{c}^\epsilon) = \mathcal{E}_{\mathrm{ext}}(\boldsymbol{c}^n) + \epsilon \sum_{i\in\{1,2\}} \langle c_i^n - \overline{c}_i \,,\, \Psi_i \rangle_{L^\infty(\Omega), L^1(\Omega)} \leq \mathcal{E}_{\mathrm{ext}}(\boldsymbol{c}^n) + C\epsilon.$$

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

$$\boldsymbol{W}^2(\boldsymbol{c}^\epsilon, \boldsymbol{c}^{n-1}) \leq \boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1}) + \epsilon\left(\boldsymbol{W}^2(\overline{\boldsymbol{c}}, \boldsymbol{c}^{n-1}) - \boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1})\right) \leq \boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1}) + C\epsilon.$$

Combining the above inequalities, we obtain that

$$\mathcal{E}_{\mathrm{therm}}(\boldsymbol{c}^n) - \mathcal{E}_{\mathrm{therm}}(\boldsymbol{c}^\epsilon) = \sum_{i\in\{1,2\}} \theta_i \int_\Omega (H(c_i^n) - H(c_i^\epsilon))\mathrm{d}\boldsymbol{x} \leq \frac{C\epsilon}{\tau}. \tag{48}$$

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

$$H(c_i^n) - H(c_i^\epsilon) \geq (c_i^n - c_i^\epsilon)\log(c_i^\epsilon) = \epsilon(c_i^n - \overline{c}_i)\log(c_i^\epsilon) \quad \text{a.e. in } A. \tag{49}$$

On the other hand, since $c_i^\epsilon = \epsilon\overline{c}_i$ a.e. in A^c,

$$H(c_i^n) - H(c_i^\epsilon) = 1 - H(\epsilon\overline{c}_i) = -\epsilon\overline{c}_i\log(\epsilon\overline{c}_i) + \epsilon\overline{c}_i \quad \text{a.e. in } A^c. \tag{50}$$

Integrating (49) over A and (50) over A^c and using (48), we obtain after division by $\epsilon > 0$ that

$$-\overline{c}_i\log(\epsilon\overline{c}_i)|A^c| + \int_A (c_i^n - \overline{c}_i)\log(c_i^\epsilon) \leq \frac{C}{\tau}. \tag{51}$$

In view of (47), we have

$$(c_i^n - \overline{c}_i)\log(c_i^\epsilon) \geq (c_i^n - \overline{c}_i)\log(\overline{c}_i), \tag{52}$$

and therefore

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

Letting ϵ tend to 0 produces a contradiction unless $|A^c| = 0$. Thus we have proved that $A = \Omega$ (up to a negligible set). Moreover, in view of (52), and since $(c_i^n - \overline{c}_i)\log(\overline{c}_i) \in L^1(\Omega)$, Fatou's Lemma applies and leads to

$$\int_\Omega (c_i^n - \overline{c}_i)\log(c_i^n)\mathrm{d}\boldsymbol{x} \leq \liminf_{\epsilon\to 0} \int_\Omega (c_i^n - \overline{c}_i)\log(c_i^\epsilon)\mathrm{d}\boldsymbol{x} \leq \frac{C}{\tau}.$$

This latter inequality imposes that $\log(c_i^n)$ belongs to $L^1(\Omega)$ for fixed $\tau > 0$. □

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

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

$$\sum_{i\in\{1,2\}} \theta_i \int_0^T \int_\Omega |\boldsymbol{\nabla}\sqrt{c_{i,\tau}}|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}t + \int_0^T \int_\Omega |\Delta \boldsymbol{c}_\tau|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}t \le C, \qquad \forall T > 0. \tag{53}$$

Since Ω is convex, it implies

$$\|\boldsymbol{c}_\tau\|_{L^2((0,T);H^2(\Omega))} \le C. \tag{54}$$

Moreover, there holds

$$\boldsymbol{\nabla} c_i^n \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega. \tag{55}$$

Proof. Below, $s \ge 0$ denotes an auxiliary time variable. Let $\check{c}_i$ ($i \in \{1,2\}$) be the unique solution to

$$\begin{cases} \partial_s \check{c}_i = \Delta \check{c}_i & \text{in } \Omega \times (0,\infty), \\ \boldsymbol{\nabla}\check{c}_i \cdot \boldsymbol{n} = 0 & \text{on } \partial\Omega \times (0,\infty), \\ (\check{c}_i)_{|_{s=0}} = c_i^n & \text{in } \Omega. \end{cases} \tag{56}$$

It is easy to check that $\int_\Omega \check{c}_i(\boldsymbol{x},s)\mathrm{d}\boldsymbol{x} = \int_\Omega c_i^0(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$, hence each $\check{\boldsymbol{c}}(s) = (\check{c}_1(s), \check{c}_2(s))$ is an admissible competitor in (28), i.e., $\mathcal{F}_\tau^n(\boldsymbol{c}^n) \le \mathcal{F}_\tau^n(\check{\boldsymbol{c}}(s))$. After rearraging terms in $\mathcal{F}_\tau^n$ and dividing by $s > 0$, the passage to the limit $s \to 0$ in this inequality produces

$$-\limsup_{s\to 0} \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}(\check{\boldsymbol{c}}(s)) \le \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \left(\frac{1}{2\tau}\boldsymbol{W}^2(\check{\boldsymbol{c}}(s), \boldsymbol{c}^{n-1})\right). \tag{57}$$

To estimate the derivative on the right hand side above, we use that the heat equation (56) is the gradient flow of the Boltzmann entropy functional $\mathcal{H}(c) = \int_\Omega H(c)\mathrm{d}\boldsymbol{x}$, which is displacement convex. Therefore, solutions $\check{\boldsymbol{c}}$ to (56) satisfy the *Evolution Variational Inequality* [1, Definition 4.5] centered at c_i^{n-1}, that is

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s} W_i^2(\check{c}_i(s), c_i^{n-1}) \le \frac{\mathcal{H}(c_i^{n-1}) - \mathcal{H}(\check{c}_i(s))}{m_i}.$$

Division by τ and summation over i leads to

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{2\tau}\boldsymbol{W}^2(\check{\boldsymbol{c}}(s), \boldsymbol{c}^{n-1})\right) \le \sum_{i\in\{1,2\}} \frac{\mathcal{H}(c_i^{n-1}) - \mathcal{H}(\check{c}_i(s))}{m_i\tau}, \qquad \forall s > 0. \tag{58}$$

The estimate for the left-hand side of (57) is more combersome. We estimate the s-derivate of the various parts of $\mathcal{E}$ individually. To begin with, notice that there is no contribution from $\mathcal{E}_{\mathrm{cons}}$, since the constraint $\check{c}_1(s) + \check{c}_2(s) = 1$ is preserved by (56). For working on the other parts of $\mathcal{E}$, we use that the solution $\check{c}_i$ of (56) belongs to $C^\infty((0,\infty); H^2(\Omega))$, is positive on Ω for $s > 0$, and satisfies homogeneous Neumann boundary conditions. That makes $\check{c}_i$ regular enough to justify the following calculations, and in particular the integration by parts. For the Dirichlet energy, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathrm{Dir}}(\check{\boldsymbol{c}}(s)) = \frac{\mathrm{d}}{\mathrm{d}s}\frac{\alpha}{2}\int_\Omega |\boldsymbol{\nabla}\check{c}_1(s)|^2\mathrm{d}\boldsymbol{x} = -\alpha\int_\Omega \Delta\check{c}_1(s)\partial_s\check{c}_1(s)\mathrm{d}\boldsymbol{x} = -\alpha\int_\Omega |\Delta\check{c}_1(s)|^2\,\mathrm{d}\boldsymbol{x}. \tag{59}$$

The derivative of the chemical energy amounts to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathrm{chem}}(\check{\boldsymbol{c}}(s)) &= \frac{\mathrm{d}}{\mathrm{d}s}\chi\int_\Omega \check{c}_1(s)(1-\check{c}_1(s))\mathrm{d}\boldsymbol{x}\\ &= \chi\int_\Omega (1-2\check{c}_1(s))\partial_s\check{c}_1(s)\mathrm{d}\boldsymbol{x} = 2\chi\int_\Omega |\boldsymbol{\nabla}\check{c}_1(s)|^2\mathrm{d}\boldsymbol{x},\end{aligned}$$

and since $s\mapsto \mathcal{E}_{\mathrm{Dir}}(\check{\boldsymbol{c}}(s))$ is non-increasing by the previous calculation, we can further conclude that

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathrm{chem}}(\check{\boldsymbol{c}}(s)) \le 2\chi\int_\Omega |\boldsymbol{\nabla}c_1^n|^2\mathrm{d}\boldsymbol{x} \le C, \tag{60}$$

where the second inequality follows from (45). Next, for the thermal energy,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathrm{therm}}(\check{\boldsymbol{c}}(s)) &= \sum_{i\in\{1,2\}}\theta_i\frac{\mathrm{d}}{\mathrm{d}s}\int_\Omega H(\check{c}_i(s))\mathrm{d}\boldsymbol{x} = \sum_{i\in\{1,2\}}\theta_i\int_\Omega \log(\check{c}_i(s))\partial_s\check{c}_i(s)\mathrm{d}\boldsymbol{x}\\ &= \sum_{i\in\{1,2\}}\theta_i\int_\Omega \log(\check{c}_i(s))\Delta\check{c}_i(s)\mathrm{d}\boldsymbol{x} = -\sum_{i\in\{1,2\}}\theta_i\int_\Omega \boldsymbol{\nabla}\log(\check{c}_i(s))\cdot\boldsymbol{\nabla}\check{c}_i(s)\mathrm{d}\boldsymbol{x},\end{aligned}$$

and since $\check{c}_i(t)$ is smooth on $\overline{\Omega}$ and bounded away from 0, we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathrm{therm}}(\check{\boldsymbol{c}}(s)) = -4\sum_{i\in\{1,2\}}\theta_i\int_\Omega |\boldsymbol{\nabla}\sqrt{\check{c}_i(s)}|^2\mathrm{d}\boldsymbol{x}. \tag{61}$$

Finally, we obtain for the derivative of the potential

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{E}_{\mathrm{ext}}(\check{\boldsymbol{c}}(s)) &= \frac{\mathrm{d}}{\mathrm{d}s}\sum_{i\in\{1,2\}}\int_\Omega \Psi_i\check{c}_i(s)\mathrm{d}\boldsymbol{x} = \int_\Omega(\Psi_1-\Psi_2)\partial_s\check{c}_1(s)\mathrm{d}\boldsymbol{x}\\ &= \int_\Omega(\Psi_1-\Psi_2)\Delta\check{c}_1(s)\mathrm{d}\boldsymbol{x} \le \frac{\alpha}{2}\|\Delta\check{c}_1(s)\|_{L^2}^2 + \frac{1}{2\alpha}\|\Psi_1-\Psi_2\|_{L^2}^2.\end{aligned}$$

In combination with (59), this implies

$$\frac{\mathrm{d}}{\mathrm{d}s}\big(\mathcal{E}_{\mathrm{Dir}}(\check{\boldsymbol{c}}(s)) + \mathcal{E}_{\mathrm{ext}}(\check{\boldsymbol{c}}(s))\big) \le -\frac{\alpha}{2}\|\Delta\check{c}_1(s)\|_{L^2}^2 + C, \qquad \forall s>0. \tag{62}$$

Since the initial condition c_i^n in (56) belongs to $H^1(\Omega)$, we have that $\check{c}_i(s)\to c_i^n$ in $H^1(\Omega)$ as $s\to 0$. In particular, $\mathcal{H}(\check{c}_i(s))\to\mathcal{H}(c_i^n)$ as $s\to 0$. Now we substitute (58) and (60), (61), (62) into (57), obtaining

(63)

$$\limsup_{s\to 0}\left\{\int_\Omega |\Delta\check{c}_1(s)|^2\mathrm{d}\boldsymbol{x} + \sum_{i\in\{1,2\}}\theta_i\int_\Omega\left|\boldsymbol{\nabla}\sqrt{\check{c}_i(t)}\right|^2\mathrm{d}\boldsymbol{x}\right\} \le C\left(1+\sum_{i\in\{1,2\}}\frac{\mathcal{H}(c_i^{n-1})-\mathcal{H}(c_i^n)}{m_i\tau}\right).$$

Since the initial condition c_i^n in (56) belongs to $H^1(\Omega)$, we have that $\check{c}_i(s)\to c_i^n$ in $H^1(\Omega)$ as $s\to 0$. Convexity of the the functionals $c\mapsto\int|\Delta c|^2\mathrm{d}\boldsymbol{x}$ and $c\mapsto\int|\nabla\sqrt{c}|^2\mathrm{d}\boldsymbol{x}$ implies lower semi-continuity with respect to convergence in $H^1(\Omega)$, and therefore

$$\int_\Omega |\Delta c_1^n|^2\mathrm{d}\boldsymbol{x} + \sum_{i\in\{1,2\}}\theta_i\int_\Omega\left|\boldsymbol{\nabla}\sqrt{c_i^n}\right|^2\mathrm{d}\boldsymbol{x} \le C\left(1+\sum_{i\in\{1,2\}}\frac{\mathcal{H}(c_i^{n-1})-\mathcal{H}(c_i^n)}{m_i\tau}\right), \qquad \forall n\ge 1.$$

Multiplying by τ and summing over $n \in \left\{1, \dots, \left\lceil \frac{T}{\tau} \right\rceil\right\}$ leads to (53); here we use that $0 \leq \sum_{i\in\{1,2\}} \mathcal{H}(c_i) \leq 2|\Omega|$ for all measurable $c_1, c_2 : \Omega \to [0,1]$, and that $|\Delta c_2^n| = |\Delta c_1^n|$.

Concerning the boundary condition (55): observe that the solution $\check{c}_i$ to (56) satisfies in particular $\nabla \check{c}_i(s) \cdot \boldsymbol{n} = 0$ at each $s > 0$. Since Ω is convex, one can show (see [23, Chapter 3]) that

$$\int_\Omega \|\nabla^2 \check{c}_i(s)\|^2 \mathrm{d}\boldsymbol{x} \leq \int_\Omega |\Delta \check{c}_i(s)|^2 \mathrm{d}\boldsymbol{x}.$$

In combination with the fact that $\check{c}_i(s)$ has values in $[0,1]$ only, it follows that the L^2-norm of $\Delta \check{c}_i(s)$ controls the full H^2-norm of $\check{c}_i(s)$, i.e.,

$$\|\check{c}_i(s)\|^2_{H^2(\Omega)} \leq C\left(1 + \int_\Omega |\Delta \check{c}_i(s)|^2 \mathrm{d}\boldsymbol{x}\right). \tag{64}$$

On the other hand, (63) implies in particular that $\Delta \check{c}_i(s)$ remains bounded in L^2 as $s \to 0$. By Alaoglu, and since we already know that $\check{c}_i(s)$ tends to c_i^n in $H^1(\Omega)$, it follows that $\check{c}_i(s)$ converges to c_i^n weakly in $H^2(\Omega)$ as $s \to 0$. Now, since the trace operator mapping $c \in H^2(\Omega)$ to $\nabla c \cdot \boldsymbol{n} \in L^2(\partial\Omega)$ is weakly continuous, we conclude that $\nabla c_i^n \cdot \boldsymbol{n} = 0$ as well.

Finally, we obtain (54): since we have just shown that c_i^n satisfies homogeneous Neumann boundary conditions, the same argument as above shows that (64) holds also with c_i^n in place of $\check{c}_i(s)$. Summation of (53) over n directly produces (54). □

Corollary 3.3. *At each step $n \geq 1$, there exists among the minimizers in* (28) *at least one which is a weak $H^2(\Omega)^2$-limit point for $\delta \to 0$ of the set of minimizers of the functionals $\mathcal{F}_\tau^{n,\delta}$ defined in* (29).

Proof. Since $\boldsymbol{c}^{n-1,\delta} \to \boldsymbol{c}^{n-1}$ in $L^1(\Omega)^2$ as $\delta \to 0$, the $\mathcal{F}_\tau^{n,\delta}$ Γ-converge to $\mathcal{F}_\tau^n$, for instance in the topology induced by $\boldsymbol{W}$ on $\boldsymbol{\mathcal{A}}$. In complete analogy to the previous proof, we obtain for each minimizer $\boldsymbol{c}^{n,\delta}$ of $\mathcal{F}_\tau^{n,\delta}$ that

$$\|c^{n,\delta}\|^2_{H^2} \leq C\left(1 + \sum_{i\in\{1,2\}} \frac{\mathcal{H}(c_i^{n-1,\delta}) - \mathcal{H}(c_i^{n,\delta})}{m_i \tau}\right) \leq \frac{C}{\tau}.$$

Therefore, there is a sequence $\boldsymbol{c}^{n,\delta_k}$ that converges weakly in $H^2(\Omega)^2$ to a limit $\boldsymbol{c}^n$ that minimizes $\mathcal{F}_\tau^n$. □

3.4. **Euler Lagrange equations.** The goal of this section is to characterize the minimizer $\boldsymbol{c}^n$ of (28). We have anticipated the main result already in Lemma 2.2. It is a consequence of the following linearized optimality condition for $\boldsymbol{c}^n$.

Lemma 3.4. *In each step $n \geq 1$, there exist Kantorovich potentials $\boldsymbol{\varphi}^n = (\varphi_1^n, \varphi_2^n)$ such that the quantity F^n given in* (31) *satisfies the following linearized optimality condition:*

$$\int_\Omega F^n c_1^n \mathrm{d}\boldsymbol{x} \leq \int_\Omega F^n c_1 \mathrm{d}\boldsymbol{x}, \qquad \forall \boldsymbol{c} = (c_1, c_2) \in \boldsymbol{\mathcal{A}}. \tag{65}$$

Proof. The proof is inspired from [32, Lemma 3.1], see also [12, Lemma 3.2]. Assume first that $c_i^{n-1} > 0$ almost everywhere in Ω, so that the Kantorovich potentials φ_1^n, ψ_1^n from c_1^n to c_1^{n-1} are unique up to addition of a global constant, that we fix at an arbitrary value for this proof. For any given $\boldsymbol{c} \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$, we perturb $\boldsymbol{c}^n$ into

$$\boldsymbol{c}^\epsilon = (1-\epsilon)\boldsymbol{c}^n + \epsilon \boldsymbol{c}, \qquad \forall \epsilon \in (0,1). \tag{66}$$

Clearly, $\boldsymbol{c}^\epsilon$ belongs to $\boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$. Defining $(\varphi_i^\epsilon, \psi_i^\epsilon)$ the Kantorovich potential from c_i^ϵ to c_i^{n-1}, we infer from (26) that

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

Subtracting the two above relations and using the definition (66) of $\boldsymbol{c}^\epsilon$, one gets

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

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

$$\frac{1}{2\tau} \left(\boldsymbol{W}^2(\boldsymbol{c}^\epsilon, \boldsymbol{c}^{n-1}) - \boldsymbol{W}^2(\boldsymbol{c}^n, \boldsymbol{c}^{n-1}) \right) \leq \epsilon \int_\Omega \left(\frac{\varphi_1^\epsilon}{\tau} - \frac{\varphi_2^\epsilon}{\tau} \right) (c_1 - c_1^n) \, \mathrm{d}\boldsymbol{x}. \tag{67}$$

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

$$\begin{aligned} \mathcal{E}(\boldsymbol{c}^\epsilon) - \mathcal{E}(\boldsymbol{c}^n) \leq \epsilon \int_\Omega \alpha \boldsymbol{\nabla} c_1^\epsilon \cdot \boldsymbol{\nabla}(c_1 - c_1^n) \mathrm{d}\boldsymbol{x} + \epsilon \int_\Omega \chi(1 - 2c_1^n)(c_1 - c_1^n) \mathrm{d}\boldsymbol{x} \\ + \epsilon \int_\Omega f(c_1^\epsilon)(c_1 - c_1^n) \mathrm{d}\boldsymbol{x} + \epsilon \int_\Omega (\Psi_1 - \Psi_2)(c_1 - c_1^n) \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{68}$$

Bearing in mind that c_1^n is a minimizer, the combination of (67) with (68) leads to

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

We can consider the limit $\epsilon \to 0$ in the right-hand side of the above expression. From the definition (66) of $\boldsymbol{c}^\epsilon$, it is clear that c_1^ϵ 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 φ_i^ϵ converges uniformly towards φ_i^n (see for instance [37, Theorem 1.52]), so that (65) holds thanks to Lemma 3.2.

Assume now that $c_i^{n-1} = 0$ on some part of Ω. By definition of $\boldsymbol{c}^n$ in and after (28), there exists a sequence of minimizers $\boldsymbol{c}^{n,\delta}$ of the respective functionals $\mathcal{F}_\tau^{n,\delta}$ given in (29) that converge weakly in $H^2(\Omega)^2$ to $\boldsymbol{c}^n$. Defining $F^{n,\delta}$ as F^n in (31) upon replacing c_i^n by $c_i^{n,\delta}$ and φ_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

$$\int_\Omega F^{n,\delta}(c_1 - c_1^{n,\delta}) \geq 0, \qquad \forall \boldsymbol{c} = (c_1, c_2) \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}. \tag{69}$$

The weak convergence of the $c_i^{n,\delta}$ to c_i^n in $H^2(\Omega)$, and the induced uniform convergence of the Kantorovich potentials $\varphi_i^{n,\delta}$ to Kantorovich potentials φ_i^n sending c_i^n to c_i^{n-1} (cf. [37, Theorem 1.52]) is sufficient to deduce (65) from (69) in the limit $\delta \to 0$. □

We are now in the position to prove Lemma 2.2 which has been essential for the definition of the potentials $\boldsymbol{\mu}$.

Proof of Lemma 2.2. Since $0 \le c_1^n \le 1$, the minimization problem (65) can be solved thanks to the bathtub principle [28, Theorem 1.14]. It amounts to saturate the sublevel sets of F^n with $c_1^n = 1$ until all the mass $\int_\Omega c_i^0$ has been allocated. This implies the existence of some $\ell \in \mathbb{R}$ such that

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

Given F^n, the solution c_1^n to the minimization problem (65) is in general not unique since a prescribed amount of mass can be distributed on the level set $\{F^n = \ell\}$ in different ways without changing the value of the functional. But this lack of uniqueness does not affect the relations (70).

Recall that we still have the freedom to change the Kantorovich potential φ_1^n in the definition (31) of F^n by addition of a global constant. We choose that constant to enforce $\ell = 0$ in (70) above. With that normalization, (70) implies (32). □

Lemma 3.5. *The chemical potentials $\boldsymbol{\mu}$ satisfy the following τ-uniform estimates:*

$$\|\mu_{i,\tau}\|_{L^2((0,T);L^{qd}(\Omega))} \le C, \qquad \|\mu_{1,\tau} - \mu_{2,\tau}\|_{L^2(Q_T)} \le C, \qquad \left\|\sqrt{c_{i,\tau}}\boldsymbol{\nabla}\mu_{i,\tau}\right\|_{L^2(Q_T)} \le C. \tag{71}$$

Proof. Recall the definition of $\boldsymbol{\mu}$ in (33). It follows from (53) that

$$\iint_{Q_T} (\mu_{1,\tau} - \mu_{2,\tau})^2 \,\mathrm{d}\boldsymbol{x}\mathrm{d}t \le C. \tag{72}$$

The definition of $\overline{\mu}^n$ implies that

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

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

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

Since $\boldsymbol{c}^n \in \boldsymbol{\mathcal{A}}$, one has $\int_\Omega c_i^n \mathrm{d}\boldsymbol{x} = \int_\Omega c_i^0 \mathrm{d}\boldsymbol{x}$. Thus it follows from (27), (41), (45) and (72) that

$$\|\boldsymbol{\nabla}\overline{\mu}_\tau\|^2_{L^2((0,T);L^1(\Omega)^d)} \le \sum_{n=0}^{\lceil \frac{T}{\tau} \rceil} \tau \|\boldsymbol{\nabla}\overline{\mu}^n\|^2_{L^1(\Omega)^d} \le C. \tag{73}$$

Bearing (34) in mind, we can use the Poincaré-Sobolev inequality and get that

$$\|\overline{\mu}_\tau\|_{L^2((0,T);L^{d/(d-1)}(\Omega))} \le C. \tag{74}$$

Since

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

we deduce from (72) the desired $L^2((0,T);L^{q_d}(\Omega))$ estimates on the phase potentials μ_i^n with q_d defined in Definition 2.3. Finally, the combination of the relations (36), (27), (53) and (41) yields

$$\iint_{Q_T} c_{i,\tau} \left|\boldsymbol{\nabla}\mu_{i,\tau}\right|^2 \mathrm{d}\boldsymbol{x}\mathrm{d}t \leq C.$$

□

The following lemma is a first step towards the recovery of the weak formulation (38).

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

$$\left|\int_\Omega \left(c_i^n - c_i^{n-1}\right)\xi \mathrm{d}\boldsymbol{x} + \tau m_i \int_\Omega \left(c_i^n \boldsymbol{\nabla}\left(\mu_i^n + \Psi_i\right) + \theta_i \boldsymbol{\nabla} c_i^n\right)\cdot\boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\right| \leq \frac{1}{2} W_i^2(c_i^n, c_i^{n-1}) \|D^2\xi\|_\infty. \tag{75}$$

Proof. The optimal transport map

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

sending c_i^n to c_i^{n-1} maps Ω into itself because Ω is convex. Therefore, since $c_i^{n-1} = \boldsymbol{t}_i^n \# c_i^n$ and thanks to (36), one gets that

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

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

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

so that

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

which is exactly the desired result. □

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

Proposition 3.7. *There exist $\boldsymbol{c} \in C([0,T];L^2(\Omega)) \cap L^2((0,T);H^2(\Omega)) \cap L^\infty((0,T);H^1(\Omega))$ with $\boldsymbol{c}(t) \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$ for a.e. $t \in [0,T]$, and $\boldsymbol{\mu} \in L^2((0,T);L^{q_d}(\Omega))$ such that, up to the extraction of a*

subsequence, the following convergence properties hold:

$$c_{i,\tau} \underset{\tau\to 0}{\longrightarrow} c_i \quad \textit{in the } L^\infty((0,T);H^1(\Omega))\textit{-weak-}\star \textit{ sense}, \tag{76a}$$

$$\|c_{i,\tau}(\cdot,t) - c_i(\cdot,t)\|_{L^2(\Omega)} \underset{\tau\to 0}{\longrightarrow} 0 \quad \textit{for all } t\in[0,T], \tag{76b}$$

$$c_{i,\tau} \underset{\tau\to 0}{\longrightarrow} c_i \quad \textit{in } L^2((0,T);W^{1,d}(\Omega)), \tag{76c}$$

$$c_{i,\tau} \underset{\tau\to 0}{\longrightarrow} c_i \quad \textit{weakly in } L^2((0,T);H^2(\Omega)), \tag{76d}$$

$$\mu_{i,\tau} \underset{\tau\to 0}{\longrightarrow} \mu_i \quad \textit{for the weak topology of } L^2((0,T);L^{q_d}(\Omega)), \tag{76e}$$

$$c_{i,\tau}\boldsymbol{\nabla}\mu_{i,\tau} \underset{\tau\to 0}{\longrightarrow} c_i\boldsymbol{\nabla}\mu_i \quad \textit{weakly in } L^2(Q_T)^d. \tag{76f}$$

Proof. All the convergence properties stated below occur up to the extraction of a subsequence when τ tends to 0. We deduce from Estimate (45) that the family $(\boldsymbol{c}_\tau)_{\tau>0}$ is bounded in $L^\infty((0,T);H^1(\Omega))$. Hence we can assume that $\boldsymbol{c}_\tau$ tends to some $\boldsymbol{c}$ in the $L^\infty((0,T);H^1(\Omega))$-weak-$\star$ sense. Moreover, since $0\le c_{i,\tau}\le 1$, we also have that $0\le c_i\le 1$ a.e. in Q_T. We also infer from Estimate (54) that $\boldsymbol{c}_\tau$ converges weakly in $L^2((0,T);H^2(\Omega))$ towards $\boldsymbol{c}$.

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

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

(see more precisely [37, Lemma 3.4]). Therefore, we infer from (43) that

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

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

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

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

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

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

$$\boldsymbol{c}_\tau \underset{\tau\to 0}{\longrightarrow} \boldsymbol{c} \quad \text{in } L^p(Q_T)^2, \quad \forall p\in[1,+\infty).$$

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 $s\ge 1+\max(0,\frac{d-2}{2})$ (see for instance [17, Theorem 6.7]) ensures that

$$\boldsymbol{c}_\tau \underset{\tau\to 0}{\longrightarrow} \boldsymbol{c} \quad \text{in } L^2((0,T);W^{1,d}(\Omega))^2. \tag{77}$$

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

$$\mu_{i,\tau} \underset{\tau\to 0}{\longrightarrow} \mu_i \quad \text{weakly in } L^2((0,T);L^{q_d}(\Omega)). \tag{78}$$

In Lemma 3.5, we also established a (uniform w.r.t. τ) $L^2(Q_T)^d$ estimate on $\left(\sqrt{c_{i,\tau}}\boldsymbol{\nabla}\mu_{i,\tau}\right)_{\tau>0}$. Since $0\le c_{i,\tau}\le 1$, it implies a uniform $L^2(Q_T)^d$ estimate on $(c_{i,\tau}\boldsymbol{\nabla}\mu_{i,\tau})_{\tau>0}$. Therefore, there exists $\boldsymbol{\vartheta}_i\in L^2(Q_T)^d$ such that $c_{i,\tau}\boldsymbol{\nabla}\mu_{i,\tau}$ converges weakly in $L^2(Q_T)^d$ to $\boldsymbol{\vartheta}_i$ as τ tends to 0. It

remains to show that $\boldsymbol{\vartheta}_i = c_i \boldsymbol{\nabla} \mu_i$. First, the distributions c_i and $\boldsymbol{\nabla}\mu_i$ can be multiplied since c_i belongs to $L^2((0,T);W^{1,d}(\Omega))$ and $\boldsymbol{\nabla}\mu_i$ belongs to $L^2((0,T);W^{-1,q_d}(\Omega))$ with $q_d \leq \frac{d}{d-1}$. Moreover, for all $\boldsymbol{\phi} \in C_c^\infty(Q_T)^d$, one has

$$\iint_{Q_T} c_{i,\tau} \boldsymbol{\nabla} \mu_{i,\tau} \cdot \boldsymbol{\phi} \mathrm{d}\boldsymbol{x}\mathrm{d}t = - \iint_{Q_T} \mu_{i,\tau} \left(\boldsymbol{\nabla} c_{i,\tau} \cdot \boldsymbol{\phi} + c_{i,\tau} \boldsymbol{\nabla} \cdot \boldsymbol{\phi} \right) \mathrm{d}\boldsymbol{x}\mathrm{d}t. \tag{79}$$

Thanks to (77) and (78), we can pass in the limit in the right-hand side of the above expression. This leads to

$$\iint_{Q_T} c_{i,\tau} \boldsymbol{\nabla} \mu_{i,\tau} \cdot \boldsymbol{\phi} \mathrm{d}\boldsymbol{x}\mathrm{d}t \underset{\tau \to 0}{\longrightarrow} - \iint_{Q_T} \mu_i \left(\boldsymbol{\nabla} c_i \cdot \boldsymbol{\phi} + c_i \boldsymbol{\nabla} \cdot \boldsymbol{\phi} \right) \mathrm{d}\boldsymbol{x}\mathrm{d}t = \langle c_i \boldsymbol{\nabla} \mu_i \, , \, \boldsymbol{\phi} \rangle_{\mathcal{D}',\mathcal{D}}. \tag{80}$$

As a consequence, $\boldsymbol{\vartheta}_i = c_i \boldsymbol{\nabla} \mu_i$ in the distributional sense, thus also in $L^2(Q_T)$.

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

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

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

We have all the necessary convergence properties to pass to the limit $\tau \to 0$ and to identify the limit $(\boldsymbol{c}, \boldsymbol{\mu})$ exhibited in Proposition 3.7 as a weak solution in the sense of Definition 2.3.

Proposition 3.8. *Let $(\boldsymbol{c}, \boldsymbol{\mu})$ be as in Proposition 3.7, then $(\boldsymbol{c}, \boldsymbol{\mu})$ is a weak solution to the problem* (17), (2)–(3) *in the sense of Definition 2.3.*

Proof. Since $c_{i,\tau}$ and $\mu_{i,\tau}$ tend weakly in $L^2((0,T);H^2(\Omega))$ and $L^2((0,T);L^{q_d}(\Omega))$ towards c_i and μ_i respectively, we can pass to the limit in (35). Moreover, one can also pass to the limit in the relation $0 = \boldsymbol{\nabla} c_{1,\tau} \cdot \boldsymbol{n}$ established in Lemma 3.2, leading to $\boldsymbol{\nabla} c_1 \cdot \boldsymbol{n} = 0$ on $\partial\Omega$.

It only remains to recover the weak formulation (38). Let $t_1, t_2 \in [0,T]$ with $t_2 \geq t_1$, then summing (75) over $n \in \left\{ \lceil \frac{t_1}{\tau} \rceil + 1, \dots, \lceil \frac{t_2}{\tau} \rceil \right\}$ yields

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

the last inequality being the consequence of the squared distance estimate (41). We can pass to the limit $\tau \to 0$ in the above relation thanks to Proposition 3.7. □

We have finally proved Theorem 2.4 that is a combination of Propositions 3.7 and 3.8.

Acknowledgements. The authors warmly thank the referees for their valuable remarks and suggestions. This research was supported by the DFG Collaborative Research Center TRR 109, "Discretization in Geometry and Dynamics" and by the French National Research Agency (ANR) through grants ANR-13-JS01-0007-01 (project GEOPOR) and ANR-11-LABX-0007-01 (Labex CEMPI).

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

[6] 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.

[7] 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.

[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] 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.

[10] C. Cancès, T. O. Gallouët, M. Laborde, and L. Monsaingeon. Simulation of multiphase porous media flows with minimizing movement and finite volume schemes. HAL: hal-01700952, to appear in European J. Appl. Math., 2018.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] P. G. de Gennes. Dynamics of fluctuations and spinodal decomposition in polymer blends. *J. Chem. Phys.*, 72:4756–4763, 1980.

[17] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhikers guide to the fractional Sobolev spaces. *Bull. Sci. math.*, 136:521573, 2012.

[18] 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.

[19] W. E and P. Palffy-Muhoray. Phase separation in incompressible systems. *Phys. Rev. E*, 55:R3844–R3846, Apr 1997.

[20] C. M. Elliott and H. Garcke. On the Cahn-Hilliard equation with degenerate mobility. *SIAM J. Math. Anal.*, 27(2):404–423, 1996.

[21] 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.

[22] R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *M2AN Math. Model. Numer. Anal.*, 37(6):937–972, 2003.

[23] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[24] M. E. Gurtin. Generalized Ginzburg-Landau and Cahn-Hilliard equations based on a microforce balance. *Physica D: Nonlinear Phenomena*, 92(3):178 – 192, 1996.

[25] 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.

[26] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
[27] 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.
[28] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.
[29] 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.
[30] 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.
[31] D. Matthes and H. Osberger. A convergent Lagrangian discretization for a nonlinear fourth-order equation. *Found. Comput. Math.*, 17(1):73–126, 2017.
[32] 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.
[33] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.
[34] L. Modica and S. Mortola. Un esempio di γ-convergenza. *Boll. Un. Mat. Ital. B*, 14(1):285299, 1980.
[35] F. Otto and W. E. Thermodynamically driven incompressible fluid mixtures. *J. Chem. Phys.*, 107(23):10177–10184, 1997.
[36] M. A. Peletier. Variational modelling: Energies, gradient flows, and large deviations. Lecture Notes, Würzburg. Available at `http://www.win.tue.nl/~mpeletie`, Feb. 2014.
[37] 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.
[38] 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`)

Zentrum für Mathematik, Technische Universität München, 85747 Garching, Germany (`matthes@ma.tum.de`)

CMAP, Centre de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France (`flore.nabet@polytechnique.edu`)