

Degenerate Cahn–Hilliard systems: From nonlocal to local

José A. Carrillo 

*Mathematical Institute, University of Oxford
Woodstock Road, Oxford OX2 6GG, UK
carrillo@maths.ox.ac.uk*

Charles Elbar 

*Laboratoire Jacques-Louis Lions (LJLL)
Faculté des Sciences et Ingénierie
Sorbonne Université, 4 place Jussieu
F-75005 Paris, France
charles.elbar@sorbonne-universite.fr*

Jakub Skrzeczkowski *

*Mathematical Institute, University of Oxford
Woodstock Road, Oxford OX2 6GG, UK
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw, Stefana Banacha 2
02-097 Warsaw, Poland
jakub.skrzeczkowski@maths.ox.ac.uk*

Received 13 July 2023

Revised 17 June 2024

Accepted 24 June 2024

Published 31 August 2024

We provide a rigorous mathematical framework to establish the limit of a nonlocal model of cell–cell adhesion system to a local model. When the parameter of the nonlocality goes to 0, the system tends to a Cahn–Hilliard system with degenerate mobility and cross-interaction forces. Our analysis relies on *a-priori* estimates and compactness properties.

Keywords: Degenerate Cahn–Hilliard equation; nonlocal Cahn–Hilliard equation; aggregation–diffusion; singular limit.

Mathematics Subject Classification 2020: 35B40, 35D30, 35K25, 35K55

*Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC BY) License which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

1. Introduction

We consider the nonlocal system of Cahn–Hilliard equations with degenerate mobility derived in [36],

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla (\kappa B[\rho] + \alpha B[\eta] - \gamma \rho - \beta \eta)) \quad \text{in } (0, +\infty) \times \mathbb{T}^d, \tag{1.1}$$

$$\frac{\partial \eta}{\partial t} = \nabla \cdot (\eta \nabla (\alpha B[\rho] + B[\eta] - \beta \rho - \eta)) \quad \text{in } (0, +\infty) \times \mathbb{T}^d, \tag{1.2}$$

equipped with an initial datum $(\rho_0, \eta_0) \in \mathcal{P}(\mathbb{T}^d)^2$ where $\mathcal{P}(\mathbb{T}^d)$ is the space of probability measures on the flat torus of dimension d . Here, $\kappa, \alpha, \gamma > 0$, $\beta \in \mathbb{R}$ while B is the nonlocal operator $B = B_\varepsilon$ defined with

$$B_\varepsilon[u_\varepsilon](x) = \frac{1}{\varepsilon^2} (u_\varepsilon(x) - \omega_\varepsilon * u_\varepsilon(x)) = \frac{1}{\varepsilon^2} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (u_\varepsilon(x) - u_\varepsilon(x - y)) \, dy \tag{1.3}$$

for ε small enough and ω_ε is the usual radial mollification kernel, $\omega_\varepsilon(x) = \frac{1}{\varepsilon^d} \omega(\frac{x}{\varepsilon})$, with ω compactly supported in the unit ball satisfying

$$\int_{\mathbb{R}^d} \omega(y) \, dy = 1, \quad \int_{\mathbb{R}^d} y \omega(y) \, dy = 0, \quad \int_{\mathbb{R}^d} y_i y_j \omega \, dy = \delta_{i,j} \frac{2}{d}. \tag{1.4}$$

Our target is to prove that as $\varepsilon \rightarrow 0$, the constructed solutions of

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla (\kappa B_\varepsilon[\rho] + \alpha B_\varepsilon[\eta] - \gamma \rho - \beta \eta)) \quad \text{in } (0, +\infty) \times \mathbb{T}^d, \tag{1.5}$$

$$\frac{\partial \eta}{\partial t} = \nabla \cdot (\eta \nabla (\alpha B_\varepsilon[\rho] + B_\varepsilon[\eta] - \beta \rho - \eta)) \quad \text{in } (0, +\infty) \times \mathbb{T}^d \tag{1.6}$$

tend to the weak solution of the local system of degenerate Cahn–Hilliard equations

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla (-\kappa \Delta \rho - \alpha \Delta \eta - \gamma \rho - \beta \eta)), \tag{1.7}$$

$$\frac{\partial \eta}{\partial t} = \nabla \cdot (\eta \nabla (-\alpha \Delta \rho - \Delta \eta - \beta \rho - \eta)). \tag{1.8}$$

Here, $\kappa > 0$ and $\gamma > 0$ represent the relative self-adhesion strength of ρ with respect to η ; while $\alpha > 0$ and $\beta \in \mathbb{R}$ give the relative strength of the cross-attraction forces.

We also denote by

$$\mu_{\rho,\varepsilon} = \kappa B_\varepsilon[\rho] + \alpha B_\varepsilon[\eta] - \gamma \rho - \beta \eta,$$

$$\mu_{\eta,\varepsilon} = \alpha B_\varepsilon[\rho] + B_\varepsilon[\eta] - \beta \rho - \eta$$

the chemical potentials (from the Cahn–Hilliard terminology) related to the nonlocal system. The nonlocal system is associated with the following formal energy/entropy structure:

$$E_\varepsilon[\rho, \eta] := \frac{1}{4} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (\kappa |\rho(x) - \rho(x - y)|^2 + |\eta(x) - \eta(x - y)|^2) \, dx \, dy$$

$$\begin{aligned}
 & + \frac{\alpha}{2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (\rho(x) - \rho(x-y))(\eta(x) - \eta(x-y)) \, dx \, dy \\
 & - \int_{\mathbb{T}^d} \frac{\gamma}{2} \rho^2 + \frac{1}{2} \eta^2 + \beta \rho \eta \, dx,
 \end{aligned} \tag{1.9}$$

$$\Phi[\rho, \eta] := \int_{\mathbb{T}^d} \rho(\log(\rho) - 1) + \eta(\log(\eta) - 1) \, dx. \tag{1.10}$$

Their dissipation is formally controlled by the identities

$$E_\varepsilon[\rho, \eta](t) + \int_0^t \int_{\mathbb{T}^d} \rho |\nabla \mu_{\rho, \varepsilon}|^2 + \int_0^t \int_{\mathbb{T}^d} \eta |\nabla \mu_{\eta, \varepsilon}|^2 \leq E_\varepsilon[\rho_0, \eta_0], \tag{1.11}$$

$$\Phi[\rho, \eta](t) + \mathcal{D}\Phi[\rho, \eta](t) \leq \Phi[\rho_0, \eta_0], \tag{1.12}$$

where $\mathcal{D}\Phi[\rho, \eta](t)$ is the dissipation of the entropy defined as

$$\begin{aligned}
 & \mathcal{D}\Phi[\rho, \eta](t) \\
 & = \frac{1}{2} \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (\kappa |\nabla \rho(x) - \nabla \rho(x-y)|^2 + |\nabla \eta(x) - \nabla \eta(x-y)|^2) \, dx \, dy \, ds \\
 & + \alpha \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (\nabla \rho(x) - \nabla \rho(x-y)) \cdot (\nabla \eta(x) - \nabla \eta(x-y)) \, dx \, dy \, ds \\
 & - \int_0^t \int_{\mathbb{T}^d} \gamma |\nabla \rho|^2 + |\nabla \eta|^2 + 2\beta \nabla \rho \cdot \nabla \eta \, dx \, ds.
 \end{aligned}$$

It turns out (see Proposition 2.1) that for $\kappa > \alpha^2$, $\kappa > 0$ they provide strong compactness in the space of ρ_ε , η_ε , $\nabla \rho_\varepsilon$, $\nabla \eta_\varepsilon$.

Our first result states that we can construct solutions to (1.5)–(1.6) satisfying additional uniform estimates which will be relevant in the sequel.

Theorem 1.1 (Existence of solutions for the nonlocal system). *Suppose that $\kappa > 0$, $\kappa > \alpha^2$. Let $\varepsilon_0 = \min(\varepsilon_0^A, \varepsilon_0^B)$ be given by Proposition 2.1 and Lemma 3.1. Let $u_0 = (\rho_0, \eta_0) \in \mathcal{P}(\mathbb{T}^d)^2$ be an initial datum with finite energy and entropy $E_\varepsilon(u^0), \Phi(u^0) \leq C$ defined in (1.9)–(1.10), where C is independent of ε . Let $\varepsilon \leq \varepsilon_0$. Then there exists a global weak solution $u_\varepsilon \in \mathcal{P}(\mathbb{T}^d)^2$ of (1.5)–(1.6) as defined in Definition 3.1. Moreover, it satisfies*

$$\int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\rho_\varepsilon(x) - \rho_\varepsilon(x-y)|^2 + |\eta_\varepsilon(x) - \eta_\varepsilon(x-y)|^2) \, dx \, dy \leq C, \tag{1.13}$$

$$\begin{aligned}
 & \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\nabla \rho_\varepsilon(x) - \nabla \rho_\varepsilon(x-y)|^2 \\
 & + |\nabla \eta_\varepsilon(x) - \nabla \eta_\varepsilon(x-y)|^2) \, dx \, dy \, dt \leq C
 \end{aligned} \tag{1.14}$$

for a constant C that depends on the parameters and the initial condition u^0 but not on ε .

Remark 1.1. The condition $E_\varepsilon(u^0) \leq C$ where C is independent of ε can be ensured by assuming that the initial conditions belong to $H^1(\mathbb{T}^d)$. We refer to [11, Theorem 1].

The proof of this result follows from the argument of [29]. We decided to include the proof to demonstrate the estimates (1.13)–(1.14) which are essential for our main result which reads as follows.

Theorem 1.2 (Convergence of nonlocal to local Cahn–Hilliard equation on the torus). Suppose that $\kappa > 0$, $\kappa > \alpha^2$. Let $u_0 \geq 0$ be an initial datum with finite energy and entropy $E_\varepsilon(u^0), \Phi(u^0) \leq C$ defined in (1.9) and (1.10), where C is independent of ε . Let $\{u_\varepsilon\}$ be a sequence of solutions of the degenerate nonlocal Cahn–Hilliard equations (1.5)–(1.6) as defined in Definition 3.1. Then, up to a subsequence (not relabeled),

$$u_\varepsilon \rightarrow u \text{ in } L^2(0, T; H^1(\mathbb{T}^d)) \text{ as } \varepsilon \rightarrow 0,$$

where u is a weak solution of the degenerate Cahn–Hilliard system (1.7)–(1.8) as in Definition 4.1.

Remark 1.2. The case of the strong cross-diffusion effect $0 < \kappa \leq \alpha^2$ seems out of reach as we are not able to obtain the necessary *a-priori* estimates for the compactness arguments, as in Proposition 2.1.

Remark 1.3. Theorems 1.1 and 1.2 are proved for the particular case of the degenerate mobility $m(u) = u$. Concerning the existence of solutions to the Cahn–Hilliard-type equations with general mobilities, we refer to [45] where the case of concave mobility has been considered by exploiting the modified Wasserstein distances introduced in [30]. The convergence $\varepsilon \rightarrow 0$ for equations with general mobility is an ongoing research direction.

We conclude with a short discussion of the applied techniques and considered problems.

Existence of weak solutions. The strategy to prove the existence of weak solutions is based on the gradient flow structure of the equation in the Wasserstein space. We use the JKO scheme first introduced in [42] by Jordan *et al.* in the context of the Fokker–Planck equation. The main idea is to use an implicit time discretization of the associated variational problems. The sequence created minimizes the movements. When the time step goes to 0, the sequence converges to the associated gradient flow. This strategy has been used for instance in [43, 44] in the context of Cahn–Hilliard equation and in [29] in the context of cross-diffusion systems with nonlocal interaction. The proof of Theorem 1.1 is in fact closely related to the one in [29] and we follow to some extent their proof. The difference is the control of the positiveness of the energy and the transport of uniform bounds independently of ε needed in the last section. Also, our settings are periodic in space. The proof uses two main components:

- The “gradient flow” structure of the scheme which provides classical discrete energy estimates and Hölder continuity in time and allows narrow convergences of the scheme.
- With the weak convergences being not enough, we use the flow interchange lemma [44, 45]. The idea is that we usually obtain better estimates with the entropy in the Cahn–Hilliard equation. This entropy generates a heat flow. Then the flow interchange lemma allows to exchange the dissipation of one functional along the gradient flow of another one, and thus improves the regularity of the scheme.

Nevertheless, since we assume that $\kappa > \alpha^2$ and $\kappa > 0$, the system is strongly parabolic with respect to (ρ, η) . Therefore, to prove existence, one could approximate (1.5)–(1.6) by

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (T_\delta(\rho) \nabla (\kappa B_\varepsilon[\rho] + \alpha B_\varepsilon[\eta] - \gamma \rho - \beta \eta)) \quad \text{in } (0, +\infty) \times \mathbb{T}^d, \quad (1.15)$$

$$\frac{\partial \eta}{\partial t} = \nabla \cdot (T_\delta(\eta) \nabla (\alpha B_\varepsilon[\rho] + B_\varepsilon[\eta] - \beta \rho - \eta)) \quad \text{in } (0, +\infty) \times \mathbb{T}^d, \quad (1.16)$$

where T_δ is a function such that $\delta \leq T_\delta \leq \frac{1}{\delta}$ and $T_\delta(\varrho) \rightarrow \varrho$ as $\delta \rightarrow 0$. Then, one sends $\delta \rightarrow 0$ and obtains the solutions to (1.5)–(1.6). This method is a standard way of proving the existence of solutions to Cahn–Hilliard equation with degenerate mobility [28, 32, 34, 48]. Of course, one has to prove the existence of solutions to (1.15)–(1.16) by virtue of fixed point method and Schauder’s estimates for parabolic equations which is very technical. Therefore, we preferred to apply gradient flow techniques which are natural for our problem.

Passage to the limit $\varepsilon \rightarrow 0$ and nonlocal compactness results. We use the strategy developed by Elbar and Skrzeczkowski in [33] for the single Cahn–Hilliard equation. The main tool is the compactness result due to Bourgain *et al.* [11] and Ponce [49] which reads as follows.

Proposition 1.1. *Let $\{f_\varepsilon\}$ be a sequence bounded in $L^2(\mathbb{T}^d)$. Suppose that*

$$\int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{|f_\varepsilon(x) - f_\varepsilon(y)|^2}{|x - y|^2} \omega_\varepsilon(|x - y|) \, dx \, dy \leq C \quad (1.17)$$

for some constant C . Then, $\{f_\varepsilon\}$ is strongly compact in $L^2(\mathbb{T}^d)$ and the limit $f \in W^{1,2}(\mathbb{T}^d)$.

Proposition 1.1 is crucial because the difference quotient appearing in (1.17) can be controlled by the energy and the dissipation of the entropy of Cahn–Hilliard equation which yields compactness. The novelty in this paper comes from the treatment of cross-interaction terms. It turns out that we can control possibly the negative terms appearing in the energy and the dissipation of entropy by a simple interpolation inequality, see Lemma 2.1 and Proposition 2.1.

Aggregation–diffusion systems. System (1.5)–(1.6) is an example of aggregation–diffusion system [18] which is attracting a lot of mathematical attention nowadays [6, 21, 22, 35, 52]. To motivate, let us first start with an aggregation equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0, \quad u = -\nabla W * \rho, \tag{1.18}$$

where $W : \mathbb{T}^d \rightarrow \mathbb{R}$ is a symmetric interaction potential. Over the last years, Eq. (1.18) was applied in the contexts of biological aggregation [10, 47, 53], materials science [39], granular media [7, 24] and it attracted also a lot of mathematical interest, particularly for nonsmooth potentials. We only refer to [9] for the L^p theory, to [19] for the theory in spaces of measures and to [8] for blow-up conditions. Equation (1.18) can be derived from the particle system via mean-field limit [40, 41]: one can consider N particles having the positions $\{X_i\}_{i=1,\dots,N}$ satisfying the system of ODEs

$$X'_i = -\frac{1}{N} \sum_{j \neq i} \nabla W(X_i - X_j);$$

then, under appropriate assumptions on W , the empirical measure $\rho^N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)}$ converges (in the weak* topology of measures) to a solution of (1.18), see [15, Theorem 3.1] for the proof and [16] for some extensions.

To obtain the aggregation–diffusion equation from (1.18), one can consider interaction potentials of the form $W_\nu = W + 2\nu\delta_0$ modeling two effects: repulsion of strength ν and nonlocal attraction. Informally, we obtain

$$\frac{\partial \rho}{\partial t} = \nu \Delta \rho^2 + \nabla \cdot (\rho \nabla W * \rho), \tag{1.19}$$

but this can be made rigorous by approximating the Dirac mass with a sequence of mollifiers, see [12, 17] with $m = 2$. Of course, (1.19) can be extended to include the more general diffusion term $\Delta \rho^m$ instead of $\Delta \rho^2$ as well as an advection term, see [20].

One can also consider a system of particles $\{X_i\}_{i=1,\dots,N}$, $\{Y_i\}_{i=1,\dots,N}$ representing two populations. Then, repeating the derivation explained above, one can arrive at a system of equations of the form (1.1)–(1.2) which can be used to model cell–cell adhesion [2, 23, 25] to reproduce the Steinberg cell-sorting phenomena. Structure-preserving numerical schemes have been derived for equations and systems of aggregation–diffusion type [3, 4, 14] as well as Cahn–Hilliard equations and systems [5, 36] recovering the cell-sorting mechanism.

Gradient flows in the periodic setting. Due to the presence of a nonlocal operator B_ε in (1.5)–(1.6), we develop our theory on the d -dimensional torus $\mathbb{T}^d$ which makes nonlocal terms easy to be defined. As already explained, solutions to (1.5)–(1.6) will be constructed via the JKO scheme. In what follows, we briefly review the theory of optimal transport on $\mathbb{T}^d$ in comparison to the usual case of $\mathbb{R}^d$ or bounded domain $\Omega \subset \mathbb{R}^d$. We refer the reader to [51, Sec. 1.3.2].

First, we need to define a metric and the natural choice is

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

where $|\cdot|$ is the Euclidean distance. Then, the Wasserstein distance W_2 is defined as

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{T}^d \times \mathbb{T}^d} d(x, y)^2 \, d\pi(x, y), \quad \mu, \nu \in \mathcal{P}(\mathbb{T}^d),$$

where $\Pi(\mu, \nu)$ is the set of couplings between μ and ν . This already implies that, say in dimension $d = 1$, the optimal transport maps are not necessarily monotone on the torus. Nevertheless, the optimal map always exists if at least one of the measures is absolutely continuous with respect to the Lebesgue measure. Moreover, the optimal map is given by the gradient of some function which is differentiable almost everywhere (a.e.). We refer to [27] for the first proof of this fact and to [51, Theorem 1.25] for a modern presentation. We also refer to the general result of McCann [46] who proved the existence of the optimal map on the general manifold, including the case of torus.

To conclude, let us mention that optimal transport was used to study several PDEs on the torus via the JKO scheme, in particular the fractional porous medium equation [37], continuity equation with nonlocal velocity in one dimension [38], systems of continuity equations with nonlinear diffusion and nonlocal drifts [13] and certain fourth-order equation in one dimension [26].

2. Energy and Entropy for the Nonlocal System

As already mentioned, system (1.5)–(1.6) has the energy/entropy structure which can be used to obtain the compactness estimates. In order to do so, one has to assure the nonnegativity of the quantities of interest. In our case, we focus on the energy E_ε and the dissipation of entropy $\mathcal{D}\Phi$.

Proposition 2.1. *Suppose that $\kappa > 0$, $\kappa > \alpha^2$. Then there exists $\varepsilon_0^A > 0$ depending on κ , α , β , γ with the following property: for all $\varepsilon \in (0, \varepsilon_0^A)$ and $\rho, \eta \in \mathcal{P}(\mathbb{T}^d)$, up to a constant, the energy defined by (1.9) and the dissipation of the entropy defined in (1.12) are nonnegative and provide the estimates on the quantities*

$$\begin{aligned} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\rho(x) - \rho(x - y)|^2 + |\eta(x) - \eta(x - y)|^2) \, dx \, dy &\leq C + E_\varepsilon[\rho, \eta], \\ \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\nabla \rho(x) - \nabla \rho(x - y)|^2 + |\nabla \eta(x) - \nabla \eta(x - y)|^2) \, dx \, dy \, ds \\ &\leq C + \mathcal{D}\Phi[\rho, \eta](t), \end{aligned}$$

where C depends on κ , α , β and γ .

The main tool to establish the nonnegativity of the energy and the dissipation of the entropy (up to a constant) is the following nonlocal Poincaré inequality with parameter which allows to handle negative terms.

Lemma 2.1 (Poincaré inequality with a parameter). *For all $\delta > 0$, there exist $C(\delta)$ and $\varepsilon_0 = \varepsilon_0(\delta)$ such that for all $\varepsilon < \varepsilon_0$ and all $f \in H^1(\mathbb{T}^d)$, we have*

$$\|f\|_2^2 \leq \delta \int_{\mathbb{T}^d \times \mathbb{T}^d} \omega_\varepsilon(y) \frac{|f(x) - f(x-y)|^2}{\varepsilon^2} dx dy + C(\delta) \|f\|_1^2, \tag{2.1}$$

$$\|\nabla f\|_2^2 \leq \delta \int_{\mathbb{T}^d \times \mathbb{T}^d} \omega_\varepsilon(y) \frac{|\nabla f(x) - \nabla f(x-y)|^2}{\varepsilon^2} dx dy + C(\delta) \|f\|_1^2. \tag{2.2}$$

Proof. We prove (2.2) as (2.1) is in fact easier. Aiming at a contradiction, suppose that there exists $\delta > 0$ with the following property: there exist a sequence $\{\varepsilon_n\}$ with $0 < \varepsilon_n < \frac{1}{n}$ and a sequence $\{f_n\}$ such that

$$\|\nabla f_n\|_{L^2(\mathbb{T}^d)}^2 > \delta \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{|\nabla f_n(x) - \nabla f_n(y)|^2}{\varepsilon_n^2} \omega_{\varepsilon_n}(|x-y|) dx dy + n \|f_n\|_{L^1(\mathbb{T}^d)}^2.$$

As $\|\nabla f_n\|_{L^2(\mathbb{T}^d)} > 0$, we may define $g_n := \frac{f_n}{\|\nabla f_n\|_{L^2(\mathbb{T}^d)}}$. Note that $\|\nabla g_n\|_{L^2(\mathbb{T}^d)} = 1$ and

$$1 > \gamma \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{|\nabla g_n(x) - \nabla g_n(y)|^2}{\varepsilon_n^2} \omega_{\varepsilon_n}(|x-y|) dx dy + n \|g_n\|_{L^1(\mathbb{T}^d)}^2.$$

By Poincaré inequality with average, $\{g_n\}$ is bounded in $L^2(\mathbb{T}^d)$ and so in $H^1(\mathbb{T}^d)$. Moreover, the first term gives compactness of the gradients [because $\{g_n\}$ is bounded in $H^1(\mathbb{T}^d)$] so that, together with the Rellich–Kondrachov theorem, there exists a function g such that $g_n \rightarrow g$ in $H^1(\mathbb{T}^d)$ (after passing on to a subsequence). But then $g = 0$ because $n \|g_n\|_{L^1(\mathbb{T}^d)} < 1$. This is however a contradiction with $\|\nabla g\|_{L^2(\mathbb{T}^d)} = \lim_{n \rightarrow \infty} \|\nabla g_n\|_{L^2(\mathbb{T}^d)} = 1$. $\square$

Proof of Proposition 2.1. We first focus on the energy. We can estimate

$$-\int_{\mathbb{T}^d} \frac{\gamma}{2} \rho^2 + \frac{1}{2} \eta^2 + \beta \rho \eta \geq -\frac{\gamma + |\beta|}{2} \int_{\mathbb{T}^d} \rho^2 - \frac{1 + |\beta|}{2} \int_{\mathbb{T}^d} \eta^2.$$

Then, we use (2.1) in Lemma 2.1 with $\delta := \delta / \max(\frac{\gamma + |\beta|}{2}, \frac{1 + |\beta|}{2})$ and $\delta > 0$ to be chosen later [this also determines $\varepsilon_0^A = \varepsilon_0^A(\delta)$], so that we obtain for $\varepsilon \in (0, \varepsilon_0^A)$

$$\begin{aligned} -\int_{\mathbb{T}^d} \frac{\gamma}{2} \rho^2 + \frac{1}{2} \eta^2 + \beta \rho \eta &\geq -\frac{\delta}{4} \int_{\mathbb{T}^d \times \mathbb{T}^d} \omega_\varepsilon(y) \frac{|\rho(x) - \rho(x-y)|^2}{\varepsilon^2} dx dy \\ &\quad -\frac{\delta}{4} \int_{\mathbb{T}^d \times \mathbb{T}^d} \omega_\varepsilon(y) \frac{|\eta(x) - \eta(x-y)|^2}{\varepsilon^2} dx dy - C(\delta). \end{aligned}$$

Therefore, using (1.9), we can bound the energy as follows:

$$\begin{aligned} E_\varepsilon[\rho, \eta] &\geq \frac{1}{4\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) ((\kappa - \delta) |\rho(x) - \rho(x - y)|^2 \\ &\quad + (1 - \delta) |\eta(x) - \eta(x - y)|^2) \, dx \, dy \\ &\quad + \frac{\alpha}{2\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (\rho(x) - \rho(x - y)) (\eta(x) - \eta(x - y)) \, dx \, dy - C(\delta). \end{aligned}$$

Now, by continuity, we choose δ sufficiently small so that we have $\kappa - \delta > 0$ and $(\kappa - \delta)(1 - \delta) - \alpha^2 > 0$, such that the matrix $\begin{pmatrix} \kappa - \delta & \alpha \\ \alpha & 1 - \delta \end{pmatrix}$ is positively defined. It follows that the associated quadratic form is bounded from below, i.e. there exists a constant C (in fact, this constant is the smallest eigenvalue of the matrix) such that

$$E_\varepsilon[\rho, \eta] \geq \frac{C}{\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (|\rho(x) - \rho(x - y)|^2 + |\eta(x) - \eta(x - y)|^2) \, dx \, dy - C(\delta).$$

The proof for the entropy is the same: this time we use (2.2) in place of (2.1). $\square$

3. Existence of Weak Solutions to the Nonlocal Problem

To prove the existence of weak solutions for our system, we want to apply the JKO scheme, see [42]. As ε is fixed in this section, we write (ρ, η) for the solution, instead of $(\rho_\varepsilon, \eta_\varepsilon)$.

Definition 3.1 (Weak solutions). We say that $u = (\rho(\cdot), \eta(\cdot)) : [0, +\infty) \rightarrow \mathcal{P}(\mathbb{T}^d)^2$ is a weak solution of (1.5)–(1.6) with the initial condition (ρ_0, η_0) satisfying (3.4) if $\rho, \eta \in C([0, T]; \mathcal{P}(\mathbb{T}^d))$, $\rho, \eta \in L^2(0, T; H^1(\mathbb{T}^d)) \cap L^\infty(0, T; L^2(\mathbb{T}^d))$ for all $T > 0$, and if for all $\varphi, \phi \in C_c^\infty([0, +\infty) \times \mathbb{T}^d)$ we have

$$\begin{aligned} & - \int_0^\infty \int_{\mathbb{T}^d} \rho \partial_t \varphi \, dx \, dt - \int_{\mathbb{T}^d} \rho_0 \varphi(0) \, dx \\ & = -\kappa \int_0^\infty \int_{\mathbb{T}^d} \rho \nabla B_\varepsilon[\rho] \cdot \nabla \varphi \, dx \, dt - \alpha \int_0^\infty \int_{\mathbb{T}^d} \rho \nabla B_\varepsilon[\eta] \cdot \nabla \varphi \, dx \, dt \\ & \quad - \int_0^\infty \int_{\mathbb{T}^d} \rho (\gamma \nabla \rho - \beta \nabla \eta) \cdot \nabla \varphi \, dx \, dt, \\ & - \int_0^\infty \int_{\mathbb{T}^d} \eta \partial_t \phi \, dx \, dt - \int_{\mathbb{T}^d} \eta_0 \phi(0) \, dx \\ & = - \int_0^\infty \int_{\mathbb{T}^d} \eta \nabla B_\varepsilon[\eta] \cdot \nabla \phi \, dx \, dt - \alpha \int_0^\infty \int_{\mathbb{T}^d} \eta \nabla B_\varepsilon[\rho] \cdot \nabla \phi \, dx \, dt \\ & \quad - \int_0^\infty \int_{\mathbb{T}^d} \eta (\nabla \eta - \beta \nabla \rho) \cdot \nabla \phi \, dx \, dt. \end{aligned}$$

We first rewrite the system of equations (1.5)–(1.6) in the following form:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\rho \nabla (\tilde{\kappa} \rho - (\tilde{\kappa} + \gamma) \omega_\varepsilon * \rho + \tilde{\alpha} \eta - (\tilde{\alpha} + \beta) \omega_\varepsilon * \eta)), \quad (3.1)$$

$$\frac{\partial \eta}{\partial t} = \nabla \cdot (\eta \nabla (\tilde{\alpha} \rho - (\tilde{\alpha} + \beta) \omega_\varepsilon * \rho + \tilde{c} \eta - (\tilde{c} + 1) \omega_\varepsilon * \eta)), \quad (3.2)$$

where $\tilde{\kappa} = \frac{\kappa}{\varepsilon^2} - \gamma$, $\tilde{\alpha} = \frac{\alpha}{\varepsilon^2} - \beta$, $\tilde{c} = \frac{1}{\varepsilon^2} - 1$ and the equations are posed in $(0, +\infty) \times \mathbb{T}^d$. Using this notation, the energy can be written as

$$\begin{aligned} E_\varepsilon[\rho, \eta] = & \int_{\mathbb{T}^d} \frac{\tilde{\kappa}}{2} \rho^2 + \frac{\tilde{c}}{2} \eta^2 + \tilde{\alpha} \rho \eta \, dx - \frac{\tilde{\kappa} + \gamma}{2} \int_{\mathbb{T}^d} \omega_\varepsilon * \rho \, d\rho \\ & - \frac{\tilde{c} + 1}{2} \int_{\mathbb{T}^d} \omega_\varepsilon * \eta \, d\eta - (\tilde{\alpha} + \beta) \int_{\mathbb{T}^d} \omega_\varepsilon * \eta \, d\rho \end{aligned} \quad (3.3)$$

for $\rho, \eta \in \mathcal{P}(\mathbb{T}^d)$. We note that when we use ρ and η in expressions of the form $\rho^2, \eta^2, \rho\eta$, we mean in fact their Radon–Nikodym derivatives with respect to the Lebesgue measures. If they exist and belong to $L^2(\mathbb{T}^d)$, the energy above makes sense. Otherwise, we consider that these quantities are $+\infty$.

We denote the Wasserstein distance for vectors $u = (u_1, u_2)$ and $v = (v_1, v_2)$ as

$$\mathcal{W}_2^2(u, v) = W_2^2(u_1, v_1) + W_2^2(u_2, v_2)$$

for all $u, v \in \mathcal{P}(\mathbb{T}^d) \times \mathcal{P}(\mathbb{T}^d)$ and where W_2 is the usual Wasserstein distance of order 2. For the initial condition, we assume that $(\rho_0, \eta_0) \in \mathcal{P}(\mathbb{T}^d)^2$ are absolutely continuous with respect to the Lebesgue measure and satisfy

$$E_\varepsilon[\rho_0, \eta_0] \leq C, \quad \Phi[\rho_0, \eta_0] \leq C, \quad (3.4)$$

where C is independent of ε .

In what follows, it will be necessary to know when the function

$$(\rho, \eta) \mapsto \frac{\tilde{\kappa}}{2} \rho^2 + \frac{\tilde{c}}{2} \eta^2 + \tilde{\alpha} \rho \eta$$

is convex. Fortunately, when $\kappa > 0$, $\kappa > \alpha^2$, this is always the case for sufficiently small ε .

Lemma 3.1. *Suppose that $\kappa > 0$, $\kappa > \alpha^2$. Then there exists $\varepsilon_0^B > 0$ depending on $\kappa, \alpha, \gamma, \beta$ such that for all $\varepsilon \in (0, \varepsilon_0^B)$ the function $(\rho, \eta) \mapsto \frac{\tilde{\kappa}}{2} \rho^2 + \frac{\tilde{c}}{2} \eta^2 + \tilde{\alpha} \rho \eta$ is convex.*

Proof. We note that the Hessian matrix reads $\begin{pmatrix} \tilde{\kappa} & \tilde{\alpha} \\ \tilde{\alpha} & \tilde{c} \end{pmatrix}$, so by the Sylvester’s criterion, the desired convexity is equivalent with $\tilde{\kappa} > 0$, $\tilde{\kappa} \tilde{c} - \tilde{\alpha}^2 > 0$. Concerning the first condition, as $\kappa > 0$ and $\tilde{\kappa} = \frac{\kappa}{\varepsilon^2} - \gamma$, we can easily find ε such that $\tilde{\kappa} > 0$. Concerning the second condition, we observe that the term standing next to the highest-order term $\frac{1}{\varepsilon^4}$ equals $\kappa - \alpha^2$ and it is positive by assumption. Therefore, the conclusion follows. $\square$

3.1. Construction of weak solutions

Let $T \in (0, \infty)$ be a final time of existence and consider the following scheme: given a time step size $\tau > 0$ and an initial condition $u_0 = (\rho_0, \eta_0) \in \mathcal{P}(\mathbb{T}^d)^2$ with $E[u_0] < +\infty$, we define by induction

$$u_\tau^0 := u_0, \quad u_\tau^{n+1} := \operatorname{argmin}_{u \in \mathcal{P}(\mathbb{T}^d)^2} \left\{ \frac{1}{2\tau} \mathcal{W}_2^2(u_\tau^n, u) + E_\varepsilon[u] \right\} \quad (3.5)$$

and we set $E_\varepsilon[u] = +\infty$ if $u \notin L^2(\mathbb{T}^d)^2$.

Lemma 3.2 (Existence of minimizers). *Let $\tau > 0$ and $u_0 = (\rho_0, \eta_0) \in \mathcal{P}_2(\mathbb{T}^d)^2$ with $E_\varepsilon[u_0] < +\infty$. Then the scheme defined by (3.5) is well defined. Moreover, we have the following energy estimate:*

$$E_\varepsilon[u_\tau^N] + \frac{1}{2\tau} \sum_{n=1}^N \mathcal{W}_2^2(u_\tau^n, u_\tau^{n-1}) \leq E_\varepsilon[u_0]. \quad (3.6)$$

Proof

Step 1: The infimum is bounded. Provided u_τ^n is defined, we notice that $u_\tau^n \in \mathcal{P}(\mathbb{T}^d)^2$ satisfy $\frac{1}{2\tau} \mathcal{W}_2^2(u_\tau^n, u_\tau^n) + E_\varepsilon[u_\tau^n] = E_\varepsilon[u_\tau^n] < +\infty$. Therefore, the infimum is bounded from above. Since $E_\varepsilon[u]$ is bounded from below by Proposition 2.1 and $\mathcal{W}_2^2(u_\tau^n, u)$ is nonnegative for all u , we also know that the infimum is bounded from below.

Step 2: Candidate for a minimizer. We suppose that u_τ^n is defined and we want to define u_τ^{n+1} . Let $\{u^k\}_k$ be a minimizing sequence in $\mathcal{P}(\mathbb{T}^d)^2$ for the problem (3.5). Without loss of generality, it can be assumed that for k large enough we have

$$\frac{1}{2\tau} \mathcal{W}_2^2(u_\tau^n, u^k) + E_\varepsilon[u^k] \leq 2 E_\varepsilon[u_\tau^n], \quad (3.7)$$

since the infimum is bounded by $E_\varepsilon[u_\tau^n]$. In particular,

$$E_\varepsilon[u^k] \leq C(u_\tau^n),$$

with a constant C independent of k . By definition of the energy in (3.3), this proves that we can extract from $(u^k)_k$ a subsequence (still denoted by k) which converges weakly to some u in $L^2(\mathbb{T}^d)^2$. Now we consider the function $f(\rho, \eta) = \frac{\tilde{\kappa}}{2} \rho^2 + \frac{\tilde{\varepsilon}}{2} \eta^2 + \tilde{\alpha} \rho \eta$ which is convex by Lemma 3.1. It follows by Tonelli theorem that the functional defined by the first integral of (3.3) is lower semi-continuous with respect to the L^2 weak convergence. Since the other terms of (3.3) are defined with convolutions, these terms are continuous with respect to the L^2 weak convergence. In the end, u minimizes (3.5). Finally, $u \in \mathcal{P}(\mathbb{T}^d)^2$ as a direct consequence of the weak convergence.

Step 3: The energy estimate (3.6). Estimate (3.6) is a consequence of an induction of the inequality

$$E_\varepsilon[u_\tau^{n+1}] + \frac{1}{2\tau} \mathcal{W}_2^2(u_\tau^{n+1}, u_\tau^n) \leq E_\varepsilon[u_\tau^n],$$

by the definition of u_τ^n . □

We have constructed a discrete-in-time sequence. We want to prove that a time interpolation of this sequence converges to a solution of (3.1)–(3.2). Let $T > 0$ be fixed and $n = \lceil \frac{T}{\tau} \rceil$. We define the interpolation $u_\tau(t) = (\rho_\tau(t), \eta_\tau(t))$ by

$$\rho_\tau(t) = \rho_\tau^n, \quad \eta_\tau(t) = \eta_\tau^n, \quad t \in ((n-1)\tau, n\tau),$$

where $(\rho_\tau^n, \eta_\tau^n) = u_\tau^n$ defined in (3.5). We prove that this sequence is compact in the following lemma.

Lemma 3.3 (Compactness of the time interpolation sequence). *The sequence of curves $\{u_\tau\}_\tau$ is uniformly bounded in $L^\infty(0, T; L^2(\mathbb{T}^d))$. Moreover, for all $T > 0$, there exists an absolutely continuous curve $u : [0, T] \rightarrow \mathcal{P}(\mathbb{T}^d)^2$ such that up to a subsequence, $u_\tau(t, \cdot)$ converges to $u(t, \cdot)$ weakly in $L^2(\mathbb{T}^d)$ as $\tau \rightarrow 0$ for all $t \in [0, T]$ and u_τ converges to u in $C([0, T], \mathcal{P}(\mathbb{T}^d)^2)$ as $\tau \rightarrow 0$ for all $T > 0$. The curve u is globally $1/2$ -Hölder-continuous in time*

$$\mathcal{W}_2(u_t, u_s) \leq \sqrt{2E_\varepsilon[u_0]} \sqrt{|s - t|} \tag{3.8}$$

and we have the estimate

$$\|u\|_{L^\infty(0, T; L^2(\mathbb{T}^d))} \leq C(1 + E_\varepsilon[u_0]).$$

Proof. From (3.6) and by the definition of u_τ which takes discrete values in time, we have that

$$E_\varepsilon[u_\tau(t)] + \frac{1}{2} \int_0^{t-\tau} \left(\frac{\mathcal{W}_2(u_\tau(s + \tau), u_\tau(s))}{\tau} \right)^2 ds \leq E_\varepsilon[u_0]. \tag{3.9}$$

With Proposition 2.1, we obtain

$$\int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} |u_\tau(x) - u_\tau(x - y)|^2 dx dy \leq C(1 + E_\varepsilon[u_0]), \tag{3.10}$$

so by Lemma 2.1 we deduce the uniform estimate in $L^\infty(0, T; L^2(\Omega))$. To prove weak compactness for all times $t \in [0, T]$ and Hölder continuity in time for the limiting curve, it is sufficient to prove

$$\limsup_{\tau \rightarrow 0} \mathcal{W}_2(u_\tau(s), u_\tau(t)) \leq C \sqrt{|t - s|}$$

for some constant C and apply [1, Proposition 3.3.1] [with d being the $\mathcal{W}_2$ distance, σ being a weak topology on $L^2(\Omega)$ and K being the ball in $L^2(\mathbb{T}^d)$ such that $u_\tau(s) \in K$ for all $s \in [0, T]$, $\mathcal{C} = \emptyset$ and $\omega(t, s) = \sqrt{t - s}$]. To this end, we write for $0 \leq s < t$ such that $s \in ((m-1)\tau, m\tau]$ and $t \in ((n-1)\tau, n\tau]$,

$$\begin{aligned} \mathcal{W}_2(u_\tau(s), u_\tau(t)) &\leq \sum_{i=m}^{n-1} \mathcal{W}_2(u_\tau^i, u_\tau^{i+1}) \leq \left(\sum_{i=m}^{n-1} \mathcal{W}_2^2(u_\tau^i, u_\tau^{i+1}) \right)^{1/2} |n - m|^{1/2} \\ &\leq \sqrt{2E_\varepsilon[u_0]} \sqrt{|t - s| + \tau}, \end{aligned}$$

where in the last line we used (3.6) and $|n - m| < \frac{|t - s|}{\tau} + 1$. This concludes the proof. □

3.2. H^1 estimates for the JKO scheme via the flow interchange lemma

The weak convergence of the JKO scheme is not enough to pass on to the limit in the definition of weak solutions. We need to obtain better estimates on the solutions. In the Cahn–Hilliard equation, better estimates are derived from considering the functional $\mathcal{U}[\rho, \eta] = \int \rho \log \rho + \eta \log \eta$. This functional generates the heat flow with respect to the Wasserstein distance $\mathcal{W}_2$. To improve the regularity, we, therefore, use the flow interchange lemma which states that *the dissipation of one functional along the gradient flow of another functional equals the dissipation of the second functional along the gradient flow of the first one*.

The main result of this section reads as follows.

Proposition 3.1. *Each solution of the JKO scheme (3.5) satisfies*

$$u_\tau^n \in H^1(\mathbb{T}^d) \quad \text{for all } n \in \mathbb{N}, \tau > 0,$$

and the time interpolation $u_\tau = (\rho_\tau, \eta_\tau)$ satisfies

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\nabla \rho_\tau(t, x) - \nabla \rho_\tau(t, x - y)|^2 \\ & \quad + |\nabla \eta_\tau(t, x) - \nabla \eta_\tau(t, x - y)|^2) \, dx \, dy \, dt \leq CT \end{aligned}$$

for all $T > 0$. Moreover, for every sequence $\tau_k \downarrow 0$, we can extract a subsequence (still denoted by τ_k) such that for all $T > 0$,

$$\begin{aligned} u_{\tau_k} &\rightarrow u \quad \text{strongly in } L^2((0, T) \times \mathbb{T}^d), \\ u_{\tau_k} &\rightharpoonup u \quad \text{weakly in } L^2(0, T; H^1(\mathbb{T}^d)). \end{aligned}$$

The main tool to prove Proposition 3.1 will be the following lemma.

Lemma 3.4. *Let $v_0 = (\rho_0, \eta_0) \in L^2(\mathbb{T}^d)^2$ with $E_\varepsilon[v_0] < +\infty$. Let $v = (\rho, \eta) : [0, +\infty) \rightarrow L^2(\mathbb{T}^d)^2$ be a solution of the heat flow*

$$\begin{cases} \partial_t \rho_t = \Delta \rho_t, & \partial_t \eta_t = \Delta \eta_t \quad \text{in } (0, +\infty) \times \mathbb{T}^d, \\ (\rho(0), \eta(0)) = (\rho_0, \eta_0). \end{cases} \quad (3.11)$$

Suppose that

$$\liminf_{s \downarrow 0} \frac{1}{s} (E_\varepsilon[v_s] - E_\varepsilon[v_0]) > -\infty. \quad (3.12)$$

Then, $v_0 \in H^1(\mathbb{T}^d)$ and for some constant C depending only on the parameters $\kappa, \alpha, \gamma, \beta$, we have

$$\begin{aligned} & \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\nabla \rho_0(x) - \nabla \rho_0(x - y)|^2 + |\nabla \eta_0(x) - \nabla \eta_0(x - y)|^2) \, dx \, dy \\ & \leq -C \liminf_{s \downarrow 0} \frac{1}{s} (E_\varepsilon[v_s] - E_\varepsilon[v_0]). \end{aligned} \quad (3.13)$$

The plan is to initiate the heat flow at the solutions to the JKO scheme. If (3.12) is verified, then (3.13) will provide the desired H^1 estimate by Lemma 2.1. The crucial information here is that dissipation of Cahn–Hilliard energy along the heat flow is related to the dissipation of Cahn–Hilliard entropy. The technical assumption (3.12) will be verified with the flow interchange lemma which requires a definition of a λ -flow in $\mathcal{P}(\mathbb{T}^d)^2$.

Definition 3.2. Let $\mathcal{F} : \mathcal{P}(\mathbb{T}^d)^2 \rightarrow (-\infty, +\infty]$ be a proper lower semi-continuous functional and $\lambda \in \mathbb{R}$. A continuous semi-group $\mathbf{S}^t : \text{Dom}(\mathcal{F}) \rightarrow \text{Dom}(\mathcal{F})$, $t \geq 0$, is a λ -flow for $\mathcal{F}$ if it satisfies the Evolution Variational Inequality (EVI)

$$\frac{1}{2} \limsup_{h \downarrow 0} \left[\frac{\mathcal{W}_2^2(\mathbf{S}^h u, v) - \mathcal{W}_2^2(u, v)}{h} \right] + \frac{\lambda}{2} \mathcal{W}_2^2(u, v) + \mathcal{F}(u) \leq \mathcal{F}(v) \quad (3.14)$$

for all measures $u, v \in \text{Dom}(\mathcal{F})$ with $\mathcal{W}_2(u, v) < +\infty$.

Lemma 3.5 (Flow interchange lemma [44, 45]). Assume that $\mathbf{S}_{\mathcal{F}}$ is a λ -flow for the proper, lower semi-continuous functional $\mathcal{F}$ in $\mathcal{P}(\mathbb{T}^d)^2$ and let u_τ^n be an n th step approximation constructed by the minimizing movement scheme (3.5). If $u_\tau^n \in \text{Dom}(\mathcal{F})$, then

$$\mathcal{F}[u_\tau^n] - \mathcal{F}[u_\tau^{n-1}] \leq \tau \liminf_{h \downarrow 0} \left(\frac{E_\varepsilon[\mathbf{S}_{\mathcal{F}}^h(u_\tau^n)] - E_\varepsilon[u_\tau^n]}{h} \right) - \frac{\lambda}{2} \mathcal{W}_2^2(u_\tau^n, u_\tau^{n-1}). \quad (3.15)$$

Remark 3.1. Based on [45, Lemma 3.2], Lemma 3.5 was formulated for the case of a bounded domain. However, it also works in the case of torus as it is just a combination of scheme (3.5) and the definition of λ -flow (3.14).

We are concerned with the functional

$$\mathcal{U}[\rho, \eta] = \int_{\mathbb{T}^d} \rho(x) \log \rho(x) + \eta(x) \log \eta(x) \, dx,$$

which generates a gradient flow in the product space $(\mathcal{P}(\mathbb{T}^d)^2, \mathcal{W}_2)$ defined by (3.11). We admit the fairly classical result that the entropy $\mathcal{U}$ possesses a 0-flow given by the heat semi-group.

Proof of Lemma 3.4. We recall that E_ε can be written as in (1.9),

$$\begin{aligned} E_\varepsilon[\rho, \eta] &= \frac{1}{4\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (\kappa |\rho(x) - \rho(x-y)|^2 + |\eta(x) - \eta(x-y)|^2) \, dx \, dy \\ &\quad + \frac{\alpha}{2\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (\rho(x) - \rho(x-y)) (\eta(x) - \eta(x-y)) \, dx \, dy \\ &\quad - \int_{\mathbb{T}^d} \frac{\gamma}{2} \rho^2 + \frac{1}{2} \eta^2 + \beta \rho \eta \, dx. \end{aligned}$$

From parabolic theory, we know that the solution of (3.11) is smooth for $t > 0$ and thus we have after integration by parts

$$\begin{aligned} & \frac{d}{dt} E_\varepsilon[\rho(t, \cdot), v(t, \cdot)] \\ &= -\frac{1}{2\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (\kappa |\nabla \rho(x) - \nabla \rho(x-y)|^2 + |\nabla \eta(x) - \nabla \eta(x-y)|^2) dx dy \\ & \quad - \frac{\alpha}{\varepsilon^2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (\nabla \rho(x) - \nabla \rho(x-y)) \cdot (\nabla \eta(x) - \nabla \eta(x-y)) dx dy \\ & \quad + \int_{\mathbb{T}^d} \gamma |\nabla \rho|^2 + |\nabla \eta|^2 + 2\beta \nabla \rho \cdot \nabla \eta dx. \end{aligned}$$

Note that we recognize the dissipation of the entropy (1.10). Since the map $t \rightarrow v(t, \cdot)$ is continuous in $L^2(\mathbb{T}^d)^2$, we get that the map $t \rightarrow E_\varepsilon[v(t, \cdot)]$ is continuous at $t = 0$. By (3.12) and Proposition 2.1, there exists C such that for all $t \leq t_0$ for t_0 sufficiently small,

$$\begin{aligned} -C &\leq \frac{E_\varepsilon[v(t, \cdot)] - E_\varepsilon[v_0]}{t} = \frac{1}{t} \int_0^1 \frac{d}{ds} E_\varepsilon[v(st, \cdot)] ds \\ &\leq -\frac{C}{2\varepsilon^2} \int_0^1 \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (|\nabla \rho(st, x) - \nabla \rho(st, x-y)|^2) dx dy ds \\ &\quad - \frac{C}{2\varepsilon^2} \int_0^1 \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) (|\nabla \eta(st, x) - \nabla \eta(st, x-y)|^2) dx dy ds. \end{aligned} \tag{3.16}$$

Therefore, using Lemma 2.1, the family $\{\nabla v(t, \cdot)\}_{t \leq t_0}$ is bounded in $L^2((0, 1) \times \mathbb{T}^d)$. Now, as $v_0 \in L^2(\mathbb{T}^d)$ and v is the solution of heat equation with the initial condition v_0 , $v \in C([0, T]; L^2(\mathbb{T}^d))$. Therefore, there exists a bounded modulus of continuity $\pi : [0, T] \rightarrow \mathbb{R}^+$ such that $\lim_{t \rightarrow 0} \pi(t) = 0$ and

$$\|v(t, \cdot) - v_0\|_{L^2(\mathbb{T}^d)} \leq \pi(t).$$

In particular,

$$\|v(st, \cdot) - v_0\|_{L^2(\mathbb{T}^d)} \leq \pi(st),$$

so the dominated convergence theorem implies that $v(t, \cdot) \rightarrow v_0$ in $L^2((0, 1) \times \mathbb{T}^d)$ when $t \rightarrow 0$. Finally, choosing a subsequence such that $\nabla v(t, \cdot) \rightharpoonup \xi$ converges weakly in $L^2((0, 1) \times \mathbb{T}^d)$, using the identity

$$\int_0^1 \int_{\mathbb{T}^d} v(ts, x) \operatorname{div} \varphi(s, x) ds dx = - \int_0^1 \int_{\mathbb{T}^d} \nabla v(ts, x) \varphi(s, x) ds dx$$

for a smooth test function φ and passing on to the limit $t \rightarrow 0$, we obtain that $\nabla v_0 = \xi$ and so $v_0 \in H^1(\mathbb{T}^d)$.

To obtain a uniform estimate on v_0 in $H^1(\mathbb{T}^d)$, we want to pass on to the limit $t \rightarrow 0$ in (3.16). For this, we observe that

$$\begin{aligned} & \int_0^1 \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) |\nabla \rho(st, x) - \nabla \rho(st, x - y)|^2 dx dy ds \\ &= 2 \int_0^1 \int_{\mathbb{T}^d} |\nabla \rho(st, x)|^2 dx ds - 2 \int_0^1 \int_{\mathbb{T}^d} \nabla \rho(st, x) \cdot \nabla (\omega_\varepsilon * \rho(st, x)) dx ds. \end{aligned}$$

The first term is lower semi-continuous with respect to the L^2 weak topology, while the second converges to

$$\int_0^1 \int_{\mathbb{T}^d} \nabla \rho(st, x) \cdot \nabla (\omega_\varepsilon * \rho(st, x)) dx ds \rightarrow \int_{\mathbb{T}^d} \nabla \rho_0(x) \cdot \nabla (\omega_\varepsilon * \rho_0(x)) dx$$

as a product of the weakly and strongly convergent sequences. □

With this lemma, we are finally able to prove Proposition 3.1.

Proof of Proposition 3.1. First, we want to apply the flow interchange lemma to the functional $\mathcal{U}$. This is possible since $\mathcal{U}$ is bounded from below on $\mathcal{P}(\mathbb{T}^d)$. Applying Lemma 3.5 with $\mathcal{F} = \mathcal{U}$ and Lemma 3.4 with $v_0 = u_\tau^n$,

$$\begin{aligned} & C\tau \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\nabla \rho_\tau^n(x) - \nabla \rho_\tau^n(x - y)|^2 + |\nabla \eta_\tau^n(x) - \nabla \eta_\tau^n(x - y)|^2) dx dy \\ & \leq \mathcal{U}[u_\tau^{n-1}] - \mathcal{U}[u_\tau^n]. \end{aligned}$$

Summing from $n = 1$ to $n = N$, we obtain

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} (|\nabla \rho_\tau(t, x) - \nabla \rho_\tau(t, x - y)|^2 \\ & \quad + |\nabla \eta_\tau(t, x) - \nabla \eta_\tau(t, x - y)|^2) dx dy dt \leq CT. \end{aligned}$$

From Lemma 2.1, we obtain uniform estimate in $L^2(0, T; H^1(\mathbb{T}^d))$ and so the weak compactness in this space.

It remains to prove the strong compactness in $L^2(0, T; L^2(\mathbb{T}^d))$. First, we want to apply Theorem B.1 in Appendix B with the Banach space $X = L^2(\mathbb{T}^d) \times L^2(\mathbb{T}^d)$, set $U = \{u_\tau\}_{\tau>0}$, pseudometric $g(u_1, u_2) = \mathcal{W}_2^2(u_1, u_2)$ (extended to $+\infty$ in case u_1 or u_2 is not a probability measure) and functional $\mathcal{F}$ defined as

$$\mathcal{F}(u) = \begin{cases} \|u\|_{H^1(\mathbb{T}^d)}^2 & \text{if } u \in H^1(\mathbb{T}^d) \times H^1(\mathbb{T}^d) \cap \mathcal{P}(\mathbb{T}^d) \times \mathcal{P}(\mathbb{T}^d), \\ +\infty & \text{if } u \notin H^1(\mathbb{T}^d) \times H^1(\mathbb{T}^d) \cap \mathcal{P}(\mathbb{T}^d) \times \mathcal{P}(\mathbb{T}^d). \end{cases}$$

We can do this as $\mathcal{F}$ is lower semi-continuous and its level sets are compact in $L^2(\mathbb{T}^d)$ by the Rellich–Kondrachov theorem. Furthermore, g is lower semi-continuous [55, Chap. 6]. Finally, (B.1) in Appendix B follows from the uniform estimates in $L^2(0, T; H^1(\mathbb{T}^d))$ and the estimate (3.9).

Therefore, Theorem B.1 in Appendix B gives us a subsequence (not relabeled) such that

$$\|u_n(t, \cdot) - u(t, \cdot)\|_{L^2(\mathbb{T}^d)}^2 \rightarrow 0 \quad \text{for a.e. } t \in [0, T].$$

As the sequence $\{u_n\}$ is bounded in $L^\infty(0, T; L^2(\mathbb{T}^d))$, the conclusion follows by dominated convergence theorem. $\square$

3.3. Weak formulation

To derive the weak formulation, we follow [54, Sec. 8.4.2]. Since (3.5) was derived from the Lagrangian point of view, the idea is to investigate its first variation and prove that it is a time-discrete scheme of (3.1). For this, we introduce a suitable perturbation of $u_\tau^{n+1} = (\rho_\tau^{n+1}, \eta_\tau^{n+1})$ as follows: let ξ be a smooth periodic vector field and $T_\delta = \text{Id} + \delta\xi$. It is classical to prove that for δ small enough, T_δ is a C^1 diffeomorphism and $\det(\nabla T_\delta) > 0$. Then, we define

$$\tilde{u}_\delta = (T_\delta \# \rho_\tau^{n+1}, \eta_\tau^{n+1}).$$

Note carefully that only the first component was perturbed. This will result in the first equation in Definition 3.1. To obtain the second one, it is sufficient to introduce a similar perturbation on the second component of u_τ^{n+1} . This results in analogous computations as outlined below and we do not repeat them.

Using the standard properties of push-forward measure, we obtain from (3.3)

$$\begin{aligned} E_\varepsilon[\tilde{u}_\delta] &= \int_{\mathbb{T}^d} \frac{\tilde{\kappa}}{2} \frac{(\rho_\tau^{n+1}(x))^2}{\det(\text{Id} + \delta\nabla\xi)} + \frac{\tilde{c}}{2} (\eta_\tau^{n+1}(x))^2 + \tilde{\alpha} \rho_\tau^{n+1}(x) \eta(x + \delta\xi) \, dx \\ &\quad - \frac{\tilde{\kappa} + \gamma}{2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(x - y + \delta(\xi(x) - \xi(y))) \rho_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx \\ &\quad - (\tilde{\alpha} + \beta) \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(x + \delta\xi(x) - y) \eta_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx \\ &\quad - \frac{\tilde{c} + 1}{2} \int_{\mathbb{T}^d} \omega_\varepsilon * \eta_\tau^{n+1}(x) \eta_\tau^{n+1}(x) \, dx. \end{aligned} \quad (3.17)$$

Using the minimizing property of u_τ^{n+1} in (3.5) gives

$$0 \leq \frac{1}{2\tau} [\mathcal{W}_2^2(u_\tau^n, \tilde{u}_\delta) - \mathcal{W}_2^2(u_\tau^n, u_\tau^{n+1})] + E_\varepsilon[\tilde{u}_\delta] - E_\varepsilon[u_\tau^{n+1}]. \quad (3.18)$$

The plan is to expand this inequality in terms of δ , send $\delta \rightarrow 0$ and then $\tau \rightarrow 0$ which will provide the weak formulation in Definition 3.1. We consider three types of terms separately.

Step 1: The nonlocal terms in $E_\varepsilon[\tilde{u}_\delta] - E_\varepsilon[u_\tau^{n+1}]$. When taking the difference $E_\varepsilon[\tilde{u}_\delta] - E_\varepsilon[u_\tau^{n+1}]$, there are two types of nonlocal terms. The first one reads

$$-\frac{\tilde{\kappa} + \gamma}{2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} [\omega_\varepsilon(x - y + \delta(\xi(x) - \xi(y))) - \omega_\varepsilon(x - y)] \rho_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx.$$

Similarly to [29, Proof of Theorem 3.3], we perform Taylor's expansion and using uniform L^2 estimates, we obtain that this term is equal to

$$-\delta \frac{\tilde{\kappa} + \gamma}{2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot (\xi(x) - \xi(y)) \rho_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx + o(\delta).$$

The second term comes from the cross-interaction potentials and we obtain similarly

$$-\delta(\tilde{\alpha} + \beta) \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot \xi(x) \eta_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx + o(\delta).$$

Step 2: The local terms. As in Step 1, there are only two differences. The first one reads

$$\frac{\tilde{\kappa}}{2} \int_{\mathbb{T}^d} \frac{(\rho_\tau^{n+1})^2}{\det(\text{Id} + \delta \nabla \xi)} - (\rho_\tau^{n+1})^2 \, dx.$$

Using $\det(\text{Id} + \delta \nabla \xi) = 1 + \delta \operatorname{div} \xi(x) + o(\delta)$, we obtain that this term is equal to

$$-\delta \frac{\tilde{\kappa}}{2} \int_{\mathbb{T}^d} (\rho_\tau^{n+1})^2 \operatorname{div} \xi(x) \, dx + o(\delta).$$

The second term reads

$$\int_{\mathbb{T}^d} \tilde{\alpha} \rho_\tau^{n+1}(x) (\eta_\tau^{n+1}(x + \delta \xi) - \eta_\tau^{n+1}(x)) \, dx.$$

As $\nabla \eta_\tau^{n+1} \in L^2(\mathbb{T}^d)$, the second term is equal to

$$\delta \tilde{\alpha} \int_{\mathbb{T}^d} \rho_\tau^{n+1}(x) \nabla \eta_\tau^{n+1}(x) \cdot \xi \, dx + o(\delta).$$

Step 3: The Wasserstein terms. Since ρ_τ^n and ρ_τ^{n+1} are absolutely continuous measures, we know that there exists an optimal map $\nabla \varphi$ such that $\nabla \varphi \# \rho_\tau^n = \rho_\tau^{n+1}$ [51, Theorem 1.25] and

$$\mathcal{W}_2^2(\rho_\tau^n, \rho_\tau^{n+1}) = \int_{\mathbb{T}^d} \rho_\tau^n(x) |x - \nabla \varphi(x) + k(x)|^2 \, dx, \quad (3.19)$$

where $k(x) \in \mathbb{Z}^d$. Moreover, we have $T_\delta \# \rho_\tau^{n+1} = [(\text{Id} + \delta \xi) \circ \nabla \varphi] \# \rho_\tau^n$. Therefore, by the definition of the Wasserstein distance, and for δ small enough, we have

$$\mathcal{W}_2^2(\rho_\tau^n, T_\delta \# \rho_\tau^{n+1}) \leq \int_{\mathbb{T}^d} \rho_\tau^n(x) |x - \nabla \varphi(x) - \delta \xi \circ \nabla \varphi(x) + k(x)|^2 \, dx. \quad (3.20)$$

Using (3.19)–(3.20) and performing Taylor's expansion, we obtain

$$\begin{aligned} & \frac{1}{2\tau} [\mathcal{W}_2^2(u_\tau^n, \tilde{u}_\delta) - \mathcal{W}_2^2(u_\tau^n, u_\tau^{n+1})] \\ & \leq \frac{\delta}{\tau} \int_{\mathbb{T}^d} \rho_\tau^n(x) (\nabla \varphi(x) - x + k(x)) \cdot (\xi \circ \nabla \varphi(x)) \, dx + o(\delta). \end{aligned}$$

Weak formulation. We plug the inequalities from Steps 1–3 into (3.18). As ξ can be replaced with $-\xi$, we obtain the equality

$$\begin{aligned} & \frac{1}{\tau} \int_{\mathbb{T}^d} \rho_\tau^n(x) (\nabla \varphi(x) - x + k(x)) \cdot (\xi \circ \nabla \varphi(x)) \, dx \\ &= \frac{\tilde{\kappa}}{2} \int_{\mathbb{T}^d} (\rho_\tau^{n+1}(x))^2 \operatorname{div} \xi(x) \, dx - \tilde{\alpha} \int_{\mathbb{T}^d} \rho_\tau^{n+1}(x) \nabla \eta_\tau^{n+1}(x) \cdot \xi \, dx \\ &+ (\tilde{\alpha} + \beta) \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot \xi(x) \eta_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx \\ &+ \frac{\tilde{\kappa} + \gamma}{2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot (\xi(x) - \xi(y)) \rho_\tau^{n+1}(y) \rho_\tau^{n+1}(x) \, dy \, dx. \quad (3.21) \end{aligned}$$

Now we consider the test function $\xi = \nabla \zeta$ for some $\zeta \in C^\infty(\mathbb{T}^d)$. By periodicity, we have $\zeta(\nabla \varphi(x)) - \zeta(x) = \zeta(\nabla \varphi(x)) - \zeta(x - k(x))$. Therefore, we have

$$\zeta(\nabla \varphi(x)) - \zeta(x) = (\nabla \varphi(x) - x + k(x)) \cdot (\nabla \zeta \circ \nabla \varphi(x)) + O(|x - k(x) - \nabla \varphi(x)|^2)$$

and $\rho_\tau^{n+1} = \nabla \varphi \# \rho_\tau^n$, so we obtain that the LHS of (3.21) is equal to

$$\frac{1}{\tau} \left(\int_{\mathbb{T}^d} \rho_\tau^{n+1} \zeta - \int_{\mathbb{T}^d} \rho_\tau^n \zeta \right) + O \left(\frac{\mathcal{W}_2^2(\rho_\tau^n, \rho_\tau^{n+1})}{\tau} \right).$$

Now, let t_1, t_2 be arbitrary. As the curve ρ_τ is piecewise constant, we can sum up from $n = n_1 = \lfloor t_1/\tau \rfloor$ to $n_2 = \lfloor t_2/\tau \rfloor + 1$ and obtain

$$\begin{aligned} & \int_{\mathbb{T}^d} \rho_\tau(t_2) \zeta \, dx - \int_{\mathbb{T}^d} \rho_\tau(t_1) \zeta \, dx + O \left(\sum_{n=n_1}^{n_2} \mathcal{W}_2^2(\rho_\tau^n, \rho_\tau^{n+1}) \right) \\ &= \frac{\tilde{\kappa}}{2} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \rho_\tau(t)^2 \Delta \zeta(x) \, dx \, dt - \tilde{\alpha} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \rho_\tau(t, x) \nabla \eta_\tau(t, x) \cdot \nabla \zeta \, dx \, dt \\ &+ (\tilde{\alpha} + \beta) \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot \nabla \zeta(x) \eta_\tau(t, y) \rho_\tau(t, x) \, dy \, dx \, dt \\ &+ \frac{\tilde{\kappa} + \gamma}{2} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \\ &\cdot (\nabla \zeta(x) - \nabla \zeta(y)) \rho_\tau(t, y) \rho_\tau(t, x) \, dy \, dx \, dt + O(\tau), \end{aligned}$$

where the term $O(\tau)$ appears because each term has at least one term (ρ_τ or η_τ) which is bounded in $L^\infty(0, T; L^2(\mathbb{T}^d))$. Using the energy estimate (3.6) on the Wasserstein distance, we can incorporate the term $O(\sum_{n=n_1}^{n_2} \mathcal{W}_2^2(\rho_\tau^n, \rho_\tau^{n+1}))$ into

$O(\tau)$. Sending $\tau \rightarrow 0$, using Proposition 3.1 and pointwise (in time) weak convergence (in space) from Lemma 3.3, yields

$$\begin{aligned} & \int_{\mathbb{T}^d} \rho(t_2, x) \zeta(x) \, dx - \int_{\mathbb{T}^d} \rho(t_1, x) \zeta(x) \, dx \\ &= \frac{\tilde{\kappa}}{2} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \rho(t, x)^2 \Delta \zeta(x) \, dx \, dt - \tilde{\alpha} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \rho(t, x) \nabla \eta(t, x) \cdot \nabla \zeta \, dx \, dt \\ &+ (\tilde{\alpha} + \beta) \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot \nabla \zeta(x) \eta(t, y) \rho(t, x) \, dy \, dx \, dt \\ &+ \frac{\tilde{\kappa} + \gamma}{2} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot (\nabla \zeta(x) - \nabla \zeta(y)) \rho(t, y) \rho(t, x) \, dy \, dx \, dt. \end{aligned} \tag{3.22}$$

Performing integration by parts, we obtain

$$\frac{\tilde{\kappa}}{2} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \rho(t, x)^2 \Delta \zeta(x) \, dx \, dt = \tilde{\kappa} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \rho(t, x) \nabla \rho(t, x) \cdot \nabla \zeta(x) \, dx \, dt,$$

while changing the variables we obtain

$$\begin{aligned} & \frac{\tilde{\kappa} + \gamma}{2} \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \nabla \omega_\varepsilon(x - y) \cdot (\nabla \zeta(x) - \nabla \zeta(y)) \rho(t, y) \rho(t, x) \, dy \, dx \, dt \\ &= -(\tilde{\kappa} + \gamma) \int_{t_1}^{t_2} \int_{\mathbb{T}^d} \omega_\varepsilon * \rho(t, x) \cdot \nabla \zeta(x) \, dx \, dt. \end{aligned}$$

Having these two observations in mind, we obtain the weak formulation with the test function $\zeta(x)$ depending only on x . The general weak formulation with the test functions depending on t and x as in Definition 3.1 follows from multiplying with $\partial_t \psi(t)$, integrating in time and using the classical density of functions of the form $\psi(t) \zeta(x)$ over the set of test functions $\varphi(t, x)$ (see [31, Theorem D.5]).

Regularity estimates (1.13)–(1.14). Clearly, it is sufficient to prove these estimates for the first component of $u = (\rho, \eta)$. For (1.13), we note that from Lemma 3.3 we have that $u_\tau(t, \cdot) \rightharpoonup u(t, \cdot)$ for all fixed $t \in [0, T]$. Moreover, we can write the quantity of interest as

$$\begin{aligned} & \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega_\varepsilon(y) |\rho_\tau(t, x) - \rho_\tau(t, x - y)|^2 \, dx \, dy \\ &= 2 \int_{\mathbb{T}^d} |\rho_\tau(t, x)|^2 \, dx - 2 \int_{\mathbb{T}^d} \rho_\tau(t, x) \omega_\varepsilon * \rho_\tau(t, x) \, dx, \end{aligned}$$

so that the first term is weakly lower semi-continuous while the second is the product of weakly and strongly converging sequences. Applying $\liminf_{\varepsilon \rightarrow 0}$ to (3.10) we deduce (1.13). Concerning (1.14), the proof is carried out in a similar way: this time we use strong compactness of ρ_τ and weak compactness of $\nabla \rho_\tau$ in $L^2((0, T) \times \mathbb{T}^d)$ from Lemma 3.1 which allow to handle the integral with respect to time.

4. Limit $\varepsilon \rightarrow 0$

We want to send $\varepsilon \rightarrow 0$ and obtain the convergence of weak solutions of the nonlocal Cahn–Hilliard system (1.5)–(1.6) to weak solutions of the local version of the Cahn–Hilliard system. We define the weak solutions of the latter.

Definition 4.1. We say that $u = (\rho(\cdot), \eta(\cdot)) : [0, \infty) \rightarrow \mathcal{P}(\mathbb{T}^d)^2$ is a weak solution of (1.7)–(1.8) with the initial condition (ρ_0, η_0) if $\rho, \eta \in L^2(0, T; H^2(\mathbb{T}^d))$ for all $T > 0$, and if for all $\varphi, \phi \in C_c^\infty([0, +\infty) \times \mathbb{T}^d)$ we have

$$\begin{aligned} & - \int_0^\infty \int_{\mathbb{T}^d} \rho \partial_t \varphi \, dx \, dt - \int_{\mathbb{T}^d} \rho_0 \varphi(0, x) \, dx = -\kappa \int_0^\infty \int_{\mathbb{T}^d} \Delta \rho \nabla \rho \cdot \nabla \varphi \, dx \, dt \\ & \quad - \kappa \int_0^\infty \int_{\mathbb{T}^d} \rho \Delta \rho \Delta \varphi \, dx \, dt - \alpha \int_0^\infty \int_{\mathbb{T}^d} D^2 \eta : (\nabla \rho \otimes \nabla \varphi + \rho D^2 \varphi) \, dx \, dt \\ & \quad - \int_0^\infty \int_{\mathbb{T}^d} \rho (\gamma \nabla \rho - \beta \nabla \eta) \cdot \nabla \varphi \, dx \, dt, \\ & - \int_0^\infty \int_{\mathbb{T}^d} \eta \partial_t \phi \, dx \, dt - \int_{\mathbb{T}^d} \eta_0 \phi(0, x) \, dx = - \int_0^\infty \int_{\mathbb{T}^d} \Delta \eta \nabla \eta \cdot \nabla \phi \, dx \, dt \\ & \quad - \int_0^\infty \int_{\mathbb{T}^d} \eta \Delta \eta \Delta \phi \, dx \, dt - \alpha \int_0^\infty \int_{\mathbb{T}^d} D^2 \rho : (\nabla \eta \otimes \nabla \phi + \eta D^2 \phi) \, dx \, dt \\ & \quad - \int_0^\infty \int_{\mathbb{T}^d} \eta (\nabla \eta - \beta \nabla \rho) \cdot \nabla \phi \, dx \, dt. \end{aligned}$$

As we will see (Lemma 4.2), we have bounds at most on the gradient of $\nabla \rho_\varepsilon, \nabla \eta_\varepsilon$, and the limit equation has four derivatives. That means, we need to mimic at the epsilon-level integration by parts for the nonlocal operators. For that purpose, we define the operator

$$S_\varepsilon[\varphi](x, y) := \frac{\sqrt{\omega_\varepsilon(y)}}{\sqrt{2\varepsilon}} (\varphi(x - y) - \varphi(x)), \quad (4.1)$$

which has the following properties, see [33, Lemma 3.4].

Lemma 4.1. *The operator S_ε satisfies the following:*

- (S1) S_ε is a linear operator that commutes with derivatives with respect to x ;
- (S2) for all functions $f, g : \mathbb{T}^d \rightarrow \mathbb{R}$, we have

$$\begin{aligned} & S_\varepsilon[fg](x, y) - S_\varepsilon[f](x, y)g(x) - S_\varepsilon[g](x, y)f(x) \\ & = \frac{\sqrt{\omega_\varepsilon(y)}}{\sqrt{2\varepsilon}} [(f(x - y) - f(x))(g(x - y) - g(x))]; \end{aligned}$$

- (S3) for all $u, \varphi \in L^2(\mathbb{T}^d)$,

$$\langle B_\varepsilon[u](\cdot), \varphi(\cdot) \rangle_{L^2(\mathbb{T}^d)} = \langle S_\varepsilon[u](\cdot, \cdot), S_\varepsilon[\varphi](\cdot, \cdot) \rangle_{L^2(\mathbb{T}^d \times \mathbb{T}^d)};$$

(S4) if $\{u_\varepsilon\}$ is strongly compact in $L^2(0, T; H^1(\mathbb{T}^d))$ and $\varphi \in L^\infty((0, T) \times \mathbb{T}^d)$, we have

$$\int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} (S_\varepsilon[u_\varepsilon])^2 \varphi(t, x) \rightarrow \int_0^T \int_{\mathbb{T}^d} |\nabla u(t, x)|^2 \varphi(t, x).$$

4.1. Uniform estimates in ε and compactness

We collect here the estimates for the solutions of (1.5)–(1.6) which are uniform in $\varepsilon > 0$.

Lemma 4.2. *Let $(\rho_\varepsilon, \eta_\varepsilon)$ be the solution of (1.5)–(1.6) constructed in Theorem 1.1. Then the following sequences are bounded:*

- (1) $\{\rho_\varepsilon\}, \{\eta_\varepsilon\}$ in $L^\infty(0, T; L^2(\Omega))$ and $L^2(0, T; H^1(\Omega))$;
- (2) $\{\partial_t \rho_\varepsilon\}, \{\partial_t \eta_\varepsilon\}$ in $L^2(0, T; H^{-2-\frac{d}{2}}(\Omega))$.

Moreover, the sequences $\{\rho_\varepsilon\}, \{\eta_\varepsilon\}$ are strongly compact in $L^2(0, T; H^1(\Omega))$ and the limits belong to $L^2(0, T; H^2(\Omega))$.

Remark 4.1. We note that we cannot use the estimates coming from the dissipation of energy (1.11) which can provide better estimates on time derivatives because this information is lost in the JKO scheme. We proceed with a different approach.

Proof of Lemma 4.2. The first estimate follows from (1.13)–(1.14) and Lemma 2.1. To see the second, we consider the equation for ρ_ε and test it with a smooth and compactly supported function $\varphi(t, x)$. We need to control the following terms:

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^d} \rho_\varepsilon \nabla B_\varepsilon[\rho_\varepsilon] \cdot \nabla \varphi \, dx \, dt, \quad \int_0^T \int_{\mathbb{T}^d} \rho_\varepsilon \nabla B_\varepsilon[\eta_\varepsilon] \cdot \nabla \varphi \, dx \, dt, \\ & \int_0^T \int_{\mathbb{T}^d} \rho_\varepsilon (\gamma \nabla \rho_\varepsilon - \beta \nabla \eta_\varepsilon) \cdot \nabla \varphi \, dx \, dt. \end{aligned}$$

The first two terms are controlled in the same way, so we focus on the second. Using S3 in Lemma 4.1 and $\nabla B_\varepsilon[\eta_\varepsilon] = B_\varepsilon[\nabla \eta_\varepsilon]$, we have

$$\int_0^T \int_{\mathbb{T}^d} \rho_\varepsilon \nabla B_\varepsilon[\eta_\varepsilon] \cdot \nabla \varphi \, dx \, dt = \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} S_\varepsilon[\nabla \eta_\varepsilon] \cdot S_\varepsilon[\rho_\varepsilon \nabla \varphi] \, dy \, dx \, dt,$$

which is bounded if $\varphi \in L^2(0, T; H^2(\mathbb{T}^d))$. Concerning the third term, it is controlled when $\nabla \varphi \in L^2(0, T; L^\infty(\mathbb{T}^d))$ which is implied by $\varphi \in L^2(0, T; H^{2+\frac{d}{2}}(\mathbb{T}^d))$ by Sobolev embedding. The conclusion follows.

The strong compactness follows from Lemma B.1 in Appendix B and the estimates (1.13)–(1.14). □

4.2. Convergence $\varepsilon \rightarrow 0$

Proof of Theorem 1.2. As Eqs. (1.5) and (1.6) have a similar structure, we focus only on (1.5). More precisely, we pass on to the limit in the term $\int_0^\infty \int_{\mathbb{T}^d} \operatorname{div}(\rho_\varepsilon \nabla \mu_{\rho, \varepsilon}) \varphi \, dx \, dt$ where $\varphi \in C^3([0, \infty) \times \mathbb{T}^d)$. Integrating by parts, we obtain

$$\begin{aligned} \int_0^\infty \int_{\mathbb{T}^d} \operatorname{div}(\rho_\varepsilon \nabla \mu_{\rho, \varepsilon}) \varphi \, dx \, dt &= - \int_0^\infty \int_{\mathbb{T}^d} \rho_\varepsilon \nabla \mu_{\rho, \varepsilon} \cdot \nabla \varphi \, dx \, dt \\ &= \kappa \int_0^\infty \int_{\mathbb{T}^d} B_\varepsilon[\rho_\varepsilon] \nabla \rho_\varepsilon \cdot \nabla \varphi \, dx \, dt + \kappa \int_0^\infty \int_{\mathbb{T}^d} B_\varepsilon[\rho_\varepsilon] \rho_\varepsilon \Delta \varphi \, dx \, dt \\ &\quad - \alpha \int_0^\infty \int_{\mathbb{T}^d} \rho_\varepsilon B_\varepsilon[\nabla \eta_\varepsilon] \cdot \nabla \varphi \, dx \, dt - \int_0^\infty \int_{\mathbb{T}^d} \rho_\varepsilon (\gamma \nabla \rho_\varepsilon - \beta \nabla \eta_\varepsilon) \cdot \nabla \varphi \, dx \, dt \\ &=: I_1 + I_2 + I_3 + I_4. \end{aligned} \quad (4.2)$$

Concerning the term I_4 , its convergence is straightforward because all the sequences $\{\rho_\varepsilon\}$, $\{\eta_\varepsilon\}$, $\{\nabla \rho_\varepsilon\}$ are compact in $L^2((0, T) \times \mathbb{T}^d)$. For the passage to the limit in I_1, I_2 , we refer to [33, Steps 1 and 2, Proof of Theorem 1.8] (these are exactly the terms that appear for the analysis of a single equation). We now prove the convergence of the term I_3 . Due to Lemma 4.1(S2), we have (omitting the constant α)

$$\begin{aligned} -I_3 &= \int_0^\infty \int_{\mathbb{T}^d} S_\varepsilon[\rho_\varepsilon] S_\varepsilon[\nabla \eta_\varepsilon] \cdot \nabla \varphi \, dx \, dt + \int_0^\infty \int_{\mathbb{T}^d} \rho_\varepsilon S_\varepsilon[\nabla \eta_\varepsilon] \cdot S_\varepsilon[\nabla \varphi] \, dx \, dt + R_\varepsilon \\ &=: J_1 + J_2 + R_\varepsilon, \end{aligned}$$

where R_ε is defined as

$$\begin{aligned} R_\varepsilon &= \int_0^\infty \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} S_\varepsilon[\nabla \eta_\varepsilon] \cdot \frac{\sqrt{w_\varepsilon(y)}}{\sqrt{2\varepsilon}} [(\nabla \varphi(x-y) - \nabla \varphi(x)) \\ &\quad \cdot (\rho_\varepsilon(x-y) - \rho_\varepsilon(x))] \, dx \, dy \, dt. \end{aligned}$$

After a change of variables $y \rightarrow \frac{y}{\varepsilon}$ and using the definition of w_ε , we obtain

$$\begin{aligned} J_1 &= \frac{1}{2} \int_0^\infty \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega(y) \frac{\rho_\varepsilon(x) - \rho_\varepsilon(x - \varepsilon y)}{\varepsilon} \frac{\nabla \eta_\varepsilon(x) - \nabla \eta_\varepsilon(x - \varepsilon y)}{\varepsilon} \\ &\quad \cdot \nabla \varphi(x) \, dx \, dy \, dt. \end{aligned} \quad (4.3)$$

From Lemmas A.1 and A.2 in Appendix A, we obtain the strong convergence of $\frac{\rho_\varepsilon(x) - \rho_\varepsilon(x - \varepsilon y)}{\varepsilon} \rightarrow \nabla \rho(x) \cdot y$ and the weak convergence of $\sqrt{\omega(y)} \frac{\nabla \eta_\varepsilon(x) - \nabla \eta_\varepsilon(x - \varepsilon y)}{\varepsilon} \rightharpoonup \sqrt{\omega(y)} D^2 \eta(x) \cdot y$ [both in $L^2((0, T) \times \mathbb{T}^d \times \mathbb{T}^d)$] so that we can easily conclude

$$J_1 \rightarrow \frac{1}{2} \int_{\mathbb{T}^d} \omega(y) \int_0^\infty \int_{\mathbb{T}^d} (\nabla \rho(x) \cdot y) (D^2 \eta(x) y) \cdot \nabla \varphi(x) \, dy \, dx \, dt,$$

where $D^2\eta$ denotes the Hessian matrix $(\partial_{ij}\eta)_{i,j}$. For J_2 , we similarly write

$$J_2 = \frac{1}{2} \int_0^\infty \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \omega(y) \rho_\varepsilon(x) \frac{\nabla \eta_\varepsilon(x) - \nabla \eta_\varepsilon(x - \varepsilon y)}{\varepsilon} \cdot \frac{\nabla \varphi(x) - \nabla \varphi(x - \varepsilon y)}{\varepsilon} dx dy dt \quad (4.4)$$

and the same argument as for J_1 shows that

$$J_2 \rightarrow \frac{1}{2} \int_{\mathbb{T}^d} \omega(y) \int_0^\infty \int_{\mathbb{T}^d} \rho(x) (D^2\eta(x)y) \cdot (D^2\varphi(x)y) dy dx dt.$$

By the properties of ω , we obtain

$$J_1 + J_2 = \int_0^\infty \int_{\mathbb{T}^d} D^2\eta : (\nabla \rho \otimes \nabla \varphi + \rho D^2\varphi) dx dt.$$

It remains to show $R_\varepsilon \rightarrow 0$. By Cauchy–Schwarz inequality (in time and space) as well as the bounds on $S_\varepsilon[\nabla \eta_\varepsilon]$ from the entropy, it remains to prove that

$$\int_0^\infty \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} |\rho_\varepsilon(x - y) - \rho_\varepsilon(x)|^2 |\nabla \varphi(x - y) - \nabla \varphi(x)|^2 dy dx dt \rightarrow 0.$$

Using Taylor’s expansion, we can estimate this integral with

$$\varepsilon \|D^2\varphi\|_{L^\infty} \int_0^\infty \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{\omega_\varepsilon(y)}{\varepsilon^2} |\rho_\varepsilon(x - y) - \rho_\varepsilon(x)|^2 dy dx dt,$$

which converges to zero by the bound from the entropy so that $R_\varepsilon \rightarrow 0$. □

Acknowledgments

José A. Carrillo was supported by the Advanced Grant Nonlocal-CPD (Nonlocal PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronization) of the European Research Council Executive Agency (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement No. 883363). José A. Carrillo was also partially supported by the EPSRC Grant Nos. EP/T022132/1 and EP/V051121/1. José A. Carrillo was also partially supported by the “Maria de Maeztu” Excellence Unit IMAG, Reference CEX2020-001105-M, funded by MCIN/AEI/10.13039/501100011033/. Jakub Skrzeczkowski was supported by the National Science Center, Poland through the Project No. 2019/35/N/ST1/03459.

Appendix A. Results on Difference Quotients

Lemma A.1. *Let $\{u_\varepsilon\}$ be a sequence strongly compact in $L^2(0, T; H^1(\mathbb{T}^d))$. Then, for fixed $y \in \mathbb{T}^d$,*

$$\frac{u_\varepsilon(t, x - \varepsilon y) - u_\varepsilon(t, x)}{\varepsilon} \rightarrow -\nabla u(t, x) \cdot y \quad \text{strongly in } L^2((0, T) \times \mathbb{T}^d \times \mathbb{T}^d).$$

Proof. Clearly, the sequence converges in $L^2((0, T) \times \mathbb{T}^d)$, i.e.

$$\int_0^T \int_{\mathbb{T}^d} \left| \frac{u_\varepsilon(t, x - \varepsilon y) - u_\varepsilon(t, x)}{\varepsilon} - \nabla u(t, x) \cdot y \right|^2 dx dt \rightarrow 0.$$

To see the convergence in $L^2((0, T) \times \mathbb{T}^d \times \mathbb{T}^d)$, it is sufficient to apply the dominated convergence theorem as we have the estimate

$$\int_0^T \int_{\mathbb{T}^d} \left| \frac{u_\varepsilon(t, x - \varepsilon y) - u_\varepsilon(t, x)}{\varepsilon} \right|^2 \leq \int_0^T \int_{\mathbb{T}^d} |\nabla u^\varepsilon \cdot y|^2 dx dt \leq |y| \sup_\varepsilon \|\nabla u_\varepsilon\|_{L^2_{t,x}}^2. \quad \square$$

Lemma A.2. Let $\varphi \in L^\infty(\mathbb{T}^d)$ and $\{\eta_\varepsilon\}$ be a sequence such that

- $\varphi(y) \frac{\nabla \eta_\varepsilon(x) - \nabla \eta_\varepsilon(x - \varepsilon y)}{\varepsilon}$ is bounded in $L^2((0, T) \times \mathbb{T}^d \times \mathbb{T}^d)$;
- $\eta_\varepsilon(t, x) \rightharpoonup \eta(t, x)$ in $L^2(0, T; H^1(\mathbb{T}^d))$ and $\eta \in L^2(0, T; H^2(\mathbb{T}^d))$.

Then,

$$\varphi(y) \frac{\nabla \eta_\varepsilon(x) - \nabla \eta_\varepsilon(x - \varepsilon y)}{\varepsilon} \rightharpoonup \varphi(y) D^2 \eta(t, x) \cdot y \quad \text{weakly in } L^2((0, T) \times \mathbb{T}^d \times \mathbb{T}^d).$$

Proof. Clearly, after passing to a subsequence, the limit exists and we only need to identify it. For this, we consider a smooth and compactly supported test function $\psi(t, x, y)$ and compute

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \varphi(y) \frac{\nabla \eta_\varepsilon(t, x) - \nabla \eta_\varepsilon(t, x - \varepsilon y)}{\varepsilon} \psi(t, x, y) dy dx dt \\ &= \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \varphi(y) \frac{\psi(t, x, y) - \psi(t, x + \varepsilon y, y)}{\varepsilon} \nabla \eta_\varepsilon(t, x) dy dx dt, \end{aligned}$$

which converges to

$$\begin{aligned} & - \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \varphi(y) \nabla \psi(t, x, y) \cdot y \nabla \eta(t, x) dy dx dt \\ &= \int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \varphi(y) \psi(t, x, y) D^2 \eta(t, x) \cdot y dy dx dt \end{aligned}$$

because $\eta \in L^2(0, T; H^2(\mathbb{T}^d))$. □

Appendix B. Compactness Results

B.1. A version of the Lions–Aubin lemma for JKO scheme

We recall here from [29, Theorem 2.1] a version of Lions–Aubin lemma useful for establishing the compactness of a sequence of solutions to the JKO scheme. For the proof, we refer to [50, Theorem 2].

Theorem B.1. Let $(X, \|\cdot\|_X)$ be a Banach space. We consider

- a lower semi-continuous functional $\mathcal{F} : X \rightarrow [0, +\infty]$ with relatively compact sublevels in X ;

- a pseudodistance $g : X \times X \rightarrow [0, +\infty]$, i.e. g is lower semi-continuous and $g(\rho, \eta) = 0$ for some $\rho, \eta \in X$ such that $\mathcal{F}(\rho), \mathcal{F}(\eta) < \infty$ implies $\rho = \eta$.

Let U be a set of measurable functions $u : (0, T) \times X$ with $T > 0$ fixed. Assume further that

$$\sup_{u \in U} \int_0^T \mathcal{F}(u(t)) \, dt < \infty, \quad \lim_{h \rightarrow 0} \sup_{u \in U} g(u(t+h), u(t)) \, dt = 0. \quad (\text{B.1})$$

Then U contains a sequence $\{u_n\}$ converging in measure to some $u \in X$, i.e.

$$\forall \varepsilon > 0 \, |\{t \in [0, T] : \|u_n - u\|_X > \varepsilon\}| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

In particular, there exists a subsequence (not relabeled) such that

$$u_n(t) \rightarrow u(t) \quad \text{in } X \quad \text{for a.e. } t \in [0, T].$$

B.2. A nonlocal version of the Lions–Aubin lemma

The following result was proved in [33, Theorem B.1] based on [11, 49]. In fact, the proof of Theorem B.2 is the proof of Proposition 1.1 integrated in time.

Theorem B.2. *Let $d \geq 2$. Let $\{f_\varepsilon\}$ be a sequence bounded in $L^p((0, T) \times \mathbb{T}^d)$. Suppose that there exists a sequence $\{\rho_\varepsilon\}$ as above such that*

$$\int_0^T \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{|f_\varepsilon(t, x) - f_\varepsilon(t, y)|^p}{\varepsilon^p} \omega_\varepsilon(|x - y|) \, dx \, dy \, dt \leq C \quad (\text{B.2})$$

for some constant C . Then $\{f_\varepsilon\}$ is compact in space in $L^p((0, T) \times \mathbb{T}^d)$, i.e.

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \int_0^T \int_{\mathbb{T}^d} |f_\varepsilon * \varphi_\delta(t, x) - f_\varepsilon(t, x)|^p \, dx \, dt = 0 \quad (\text{B.3})$$

for all families of mollifiers $\{\varphi_\delta\}_{0 < \delta < 1}$.

We prove here the following.

Lemma B.1. *Suppose that $\{f_\varepsilon\}$ is a sequence bounded in $L^2((0, T) \times \mathbb{T}^d)$ such that*

- $\{\partial_t f_\varepsilon\}$ is uniformly bounded in $L^2(0, T; H^{-k}(\mathbb{T}^d))$ for some $k \in \mathbb{N}$;
- $\{f_\varepsilon\}$ is compact in space in $L^2((0, T) \times \mathbb{T}^d)$, i.e.

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \int_0^T \int_{\mathbb{T}^d} |f_\varepsilon * \varphi_\delta(t, x) - f_\varepsilon(t, x)|^2 \, dx \, dt = 0 \quad (\text{B.4})$$

for some family of mollifiers $\{\varphi_\delta\}_{0 < \delta < 1}$.

Then $\{f_\varepsilon\}$ is compact in time in $L^2((0, T) \times \mathbb{T}^d)$, i.e.

$$\lim_{h \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, x) - f_\varepsilon(t, x)|^2 \, dx \, dt \rightarrow 0 \quad \text{as } h \rightarrow 0, \quad (\text{B.5})$$

and so it is compact in $L^2((0, T) \times \mathbb{T}^d)$.

Proof. Using the mollifiers with $\delta = \delta(h)$ depending on h to be specified later in the way that $\delta(h) \rightarrow 0$ as $h \rightarrow 0$, we first split

$$\begin{aligned} & \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, x) - f_\varepsilon(t, x)|^2 dx dt \\ & \leq 4 \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t, x) - f_\varepsilon(t, \cdot) * \varphi_\delta(x)|^2 dx dt \\ & \quad + 4 \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, x) - f_\varepsilon(t+h, \cdot) * \varphi_\delta(x)|^2 dx dt \\ & \quad + 4 \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, \cdot) * \varphi_\delta(x) - f_\varepsilon(t, \cdot) * \varphi_\delta(x)|^2 dx dt. \end{aligned}$$

When we apply the limit $\lim_{h \rightarrow 0} \limsup_{\varepsilon \rightarrow 0}$, the first and second terms vanish due to (B.4). It remains to study the third term. For this, suppose first that $f_\varepsilon(t, \cdot)$ is smooth in the time variable. Then,

$$\begin{aligned} & \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, \cdot) * \varphi_\delta(x) - f_\varepsilon(t, \cdot) * \varphi_\delta(x)|^2 dx dt \\ & = \int_0^{T-h} \int_{\mathbb{T}^d} \left| \int_t^{t+h} \partial_t f_\varepsilon(s, \cdot) * \varphi_\delta(x) ds \right|^2 dx dt. \end{aligned}$$

Now, we can estimate the convolution as follows:

$$\begin{aligned} \partial_t f_\varepsilon(s, \cdot) * \varphi_\delta(x) &= \int_{\mathbb{T}^d} \partial_t f_\varepsilon(s, y) \varphi_\delta(x-y) dy \\ &\leq \|\partial_t f_\varepsilon(s, \cdot)\|_{H^{-k}(\mathbb{T}^d)} \|\varphi_\delta(x - \cdot)\|_{H^k(\mathbb{T}^d)}. \end{aligned}$$

Using this and applying invariance in the $H^k(\mathbb{T}^d)$ norm, we obtain

$$\begin{aligned} & \int_0^{T-h} \int_{\mathbb{T}^d} \left| \int_t^{t+h} \partial_t f_\varepsilon(s, \cdot) * \varphi_\delta(x) ds \right|^2 dx dt \\ & = h^2 \|\varphi_\delta\|_{H^k(\mathbb{T}^d)}^2 \int_0^{T-h} \left| \frac{1}{h} \int_t^{t+h} \|\partial_t f_\varepsilon(s, \cdot)\|_{H^{-k}(\mathbb{T}^d)} ds \right|^2 dt. \end{aligned}$$

Applying Jensen's inequality, we obtain

$$\begin{aligned} & h^2 \|\varphi_\delta\|_{H^k(\mathbb{T}^d)}^2 \int_0^{T-h} \left| \frac{1}{h} \int_t^{t+h} \|\partial_t f_\varepsilon(s, \cdot)\|_{H^{-k}(\mathbb{T}^d)} ds \right|^2 dt \\ & \leq Th \|\varphi_\delta\|_{H^k(\mathbb{T}^d)} \|\partial_t f_\varepsilon(s, \cdot)\|_{L^2(0, T; H^{-k}(\mathbb{T}^d))}^2. \end{aligned}$$

Using that $\|\varphi_\delta\|_{H^k(\mathbb{T}^d)} \leq \frac{C}{\delta^{k+d/2}}$, we finally obtain

$$\begin{aligned} & \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, \cdot) * \varphi_\delta(x) - f_\varepsilon(t, \cdot) * \varphi_\delta(x)|^2 dx dt \\ & \leq CT \frac{h}{\delta^{k+d/2}} \|\partial_t f_\varepsilon(s, \cdot)\|_{L^2(0,T;H^{-k}(\mathbb{T}^d))}^2. \end{aligned}$$

Now, if $f_\varepsilon(t, x)$ is not smooth in time, we extend it with $f_\varepsilon(0, x)$ for $t < 0$, $f_\varepsilon(T, x)$ for $t > T$ and apply the usual regularization to obtain the same estimate. Hence, if we choose $h(\delta) = \delta^{2k+d}$, we conclude

$$\lim_{h \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \int_0^{T-h} \int_{\mathbb{T}^d} |f_\varepsilon(t+h, x) - f_\varepsilon(t, x)|^2 dx dt \leq \theta(h).$$

Combined with the compactness in space (B.5) and the Fréchet–Kolmogorov theorem, we obtain the compactness of $\{f_\varepsilon\}$ in $L^2((0, T) \times \mathbb{T}^d)$. $\square$

ORCID

José A. Carrillo  <https://orcid.org/0000-0001-8819-4660>

Charles Elbar  <https://orcid.org/0000-0002-2261-4172>

Jakub Skrzeczkowski  <https://orcid.org/0000-0003-3328-4428>

References

- [1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, 2nd edn., Lectures in Mathematics ETH Zürich (Birkhäuser Verlag, Basel, 2008).
- [2] N. J. Armstrong, K. J. Painter and J. A. Sherratt, A continuum approach to modelling cell–cell adhesion, *J. Theoret. Biol.* **243**(1) (2006) 98–113.
- [3] R. Bailo, J. A. Carrillo and J. Hu, Fully discrete positivity-preserving and energy-dissipating schemes for aggregation-diffusion equations with a gradient-flow structure, *Commun. Math. Sci.* **18**(5) (2020) 1259–1303.
- [4] R. Bailo, J. A. Carrillo and J. Hu, Bound-preserving finite-volume schemes for systems of continuity equations with saturation, *SIAM J. Appl. Math.* **83**(3) (2023) 1315–1339.
- [5] R. Bailo, J. A. Carrillo, S. Kalliadasis and S. P. Perez, Unconditional bound-preserving and energy-dissipating finite-volume schemes for the Cahn–Hilliard equation, *Commun. Comput. Phys.* **34**(3) (2023) 713–748.
- [6] R. Bailo, J. A. Carrillo, H. Murakawa and M. Schmidtchen, Convergence of a fully discrete and energy-dissipating finite-volume scheme for aggregation-diffusion equations, *Math. Models Methods Appl. Sci.* **30**(13) (2020) 2487–2522.
- [7] D. Benedetto, E. Caglioti and M. Pulvirenti, A kinetic equation for granular media, *RAIRO Modél. Math. Anal. Numér.* **31**(5) (1997) 615–641.
- [8] A. L. Bertozzi, J. A. Carrillo and T. Laurent, Blow-up in multidimensional aggregation equations with mildly singular interaction kernels, *Nonlinearity* **22**(3) (2009) 683–710.
- [9] A. L. Bertozzi, T. Laurent and J. Rosado, L^p theory for the multidimensional aggregation equation, *Comm. Pure Appl. Math.* **64**(1) (2011) 45–83.

- [10] S. Boi, V. Capasso and D. Morale, Modeling the aggregative behavior of ants of the species *Polyergus rufescens*, *Nonlinear Anal. Real World Appl.* **1** (2000) 163–176 [published also in *Proc. Conf. Spatial Heterogeneity in Ecological Models* (1998)].
- [11] J. Bourgain, H. Brezis and P. Mironescu, Another look at Sobolev spaces, in *Optimal Control and Partial Differential Equations* (IOS Press, Amsterdam, 2001), pp. 439–455.
- [12] M. Burger and A. Esposito, Porous medium equation and cross-diffusion systems as limit of nonlocal interaction, *Nonlinear Anal.* **235** (2023) 113347.
- [13] G. Carlier and M. Laborde, On systems of continuity equations with nonlinear diffusion and nonlocal drifts, preprint, hal-01147666, version 1 (2015), <https://inria.hal.science/hal-01147666/>.
- [14] J. A. Carrillo, A. Chertock and Y. Huang, A finite-volume method for nonlinear nonlocal equations with a gradient flow structure, *Commun. Comput. Phys.* **17**(1) (2015) 233–258.
- [15] J. A. Carrillo, Y.-P. Choi and M. Hauray, The derivation of swarming models: Mean-field limit and Wasserstein distances, in *Collective Dynamics from Bacteria to Crowds*, CISM International Centre for Mechanical Sciences, Vol. 553 (Springer, Vienna, 2014), pp. 1–46.
- [16] J. A. Carrillo, Y.-P. Choi, M. Hauray and S. Salem, Mean-field limit for collective behavior models with sharp sensitivity regions, *J. Eur. Math. Soc. (JEMS)* **21**(1) (2019) 121–161.
- [17] J. A. Carrillo, K. Craig and F. S. Patacchini, A blob method for diffusion, *Calc. Var. Partial Differential Equations* **58**(2) (2019) 53.
- [18] J. A. Carrillo, K. Craig and Y. Yao, Aggregation-diffusion equations: Dynamics, asymptotics, and singular limits, in *Active Particles, Volume 2: Advances in Theory, Models, and Applications*, Modeling and Simulation in Science, Engineering and Technology (Birkhäuser, Cham, 2019), pp. 65–108.
- [19] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent and D. Slepčev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations, *Duke Math. J.* **156**(2) (2011) 229–271.
- [20] J. A. Carrillo, A. Esposito and J. S.-H. Wu, Nonlocal approximation of nonlinear diffusion equations, preprint, arXiv:2302.08248 [math.AP] (2023).
- [21] J. A. Carrillo and R. S. Gvalani, Phase transitions for nonlinear nonlocal aggregation-diffusion equations, *Comm. Math. Phys.* **382**(1) (2021) 485–545.
- [22] J. A. Carrillo, S. Hittmeir, B. Volzone and Y. Yao, Nonlinear aggregation-diffusion equations: Radial symmetry and long time asymptotics, *Invent. Math.* **218**(3) (2019) 889–977.
- [23] J. A. Carrillo, Y. Huang and M. Schmidtchen, Zoology of a nonlocal cross-diffusion model for two species, *SIAM J. Appl. Math.* **78**(2) (2018) 1078–1104.
- [24] J. A. Carrillo, R. J. McCann and C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media, *Arch. Ration. Mech. Anal.* **179**(2) (2006) 217–263.
- [25] J. A. Carrillo, H. Murakawa, M. Sato, H. Togashi and O. Trush, A population dynamics model of cell-cell adhesion incorporating population pressure and density saturation, *J. Theoret. Biol.* **474** (2019) 14–24.
- [26] J. A. Carrillo and D. Slepčev, Example of a displacement convex functional of first order, *Calc. Var. Partial Differential Equations* **36**(4) (2009) 547–564.
- [27] D. Cordero-Erausquin, Sur le transport de mesures périodiques, *C. R. Acad. Sci. Paris Sér. I Math.* **329**(3) (1999) 199–202.

- [28] S. Dai and Q. Du, Weak solutions for the Cahn-Hilliard equation with degenerate mobility, *Arch. Ration. Mech. Anal.* **219**(3) (2016) 1161–1184.
- [29] M. Di Francesco, A. Esposito and S. Fagioli, Nonlinear degenerate cross-diffusion systems with nonlocal interaction, *Nonlinear Anal.* **169** (2018) 94–117.
- [30] J. Dolbeault, B. Nazaret and G. Savaré, A new class of transport distances between measures, *Calc. Var. Partial Differential Equations* **34** (2009) 193–231.
- [31] C. Düll, P. Gwiazda, A. Marciniak-Czochra and J. Skrzeczkowski, *Spaces of Measures and Their Applications to Structured Population Models*, Cambridge Monographs on Applied and Computational Mathematics, Vol. 36 (Cambridge University Press, Cambridge, 2022).
- [32] C. Elbar, B. Perthame and A. Poulain, Degenerate Cahn-Hilliard and incompressible limit of a Keller-Segel model, *Commun. Math. Sci.* **20**(7) (2022) 1901–1926.
- [33] C. Elbar and J. Skrzeczkowski, Degenerate Cahn-Hilliard equation: From nonlocal to local, *J. Differential Equations* **364** (2023) 576–611.
- [34] C. M. Elliott and H. Garcke, On the Cahn-Hilliard equation with degenerate mobility, *SIAM J. Math. Anal.* **27**(2) (1996) 404–423.
- [35] S. Fagioli and E. Radici, Solutions to aggregation-diffusion equations with nonlinear mobility constructed via a deterministic particle approximation, *Math. Models Methods Appl. Sci.* **28**(9) (2018) 1801–1829.
- [36] C. Falcó, R. E. Baker and J. A. Carrillo, A local continuum model of cell-cell adhesion, *SIAM J. Appl. Math.* **84** (2024) S17–S42, doi:10.1137/22M1506079.
- [37] L. C. F. Ferreira, M. C. Santos and J. C. Valencia-Guevara, Minimizing movement for a fractional porous medium equation in a periodic setting, *Bull. Sci. Math.* **153** (2019) 86–117.
- [38] L. C. F. Ferreira and J. C. Valencia-Guevara, Periodic solutions for a 1D-model with nonlocal velocity via mass transport, *J. Differential Equations* **260**(10) (2016) 7093–7114.
- [39] D. Holm and V. Putkaradze, Formation of clumps and patches in self-aggregation of finite-size particles, *Physica D* **220** (2005) 183–196.
- [40] P.-E. Jabin, A review of the mean field limits for Vlasov equations, *Kinet. Relat. Models* **7**(4) (2014) 661–711.
- [41] P.-E. Jabin and Z. Wang, Mean field limit and propagation of chaos for Vlasov systems with bounded forces, *J. Funct. Anal.* **271**(12) (2016) 3588–3627.
- [42] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.* **29**(1) (1998) 1–17.
- [43] M. Kroemer and T. Laux, The Hele-Shaw flow as the sharp interface limit of the Cahn-Hilliard equation with disparate mobilities, *Comm. Partial Differential Equations* **47**(12) (2022) 2444–2486.
- [44] 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) (2012) 814–850.
- [45] D. Matthes, R. J. McCann and G. Savaré, A family of nonlinear fourth order equations of gradient flow type, *Comm. Partial Differential Equations* **34**(10–12) (2009) 1352–1397.
- [46] R. J. McCann, Polar factorization of maps on Riemannian manifolds, *Geom. Funct. Anal.* **11**(3) (2001) 589–608.
- [47] A. Mogilner and L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Biol.* **38**(6) (1999) 534–570.

- [48] B. Perthame and A. Poulain, Relaxation of the Cahn–Hilliard equation with singular single-well potential and degenerate mobility, *European J. Appl. Math.* **32**(1) (2021) 89–112.
- [49] A. C. Ponce, An estimate in the spirit of Poincaré’s inequality, *J. Eur. Math. Soc. (JEMS)* **6**(1) (2004) 1–15.
- [50] R. Rossi and G. Savaré, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **2**(2) (2003) 395–431.
- [51] F. Santambrogio, *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*, Progress in Nonlinear Differential Equations and Their Applications (Birkhäuser, Cham, 2015).
- [52] R. Shu, Equilibration of aggregation-diffusion equations with weak interaction forces, *SIAM J. Math. Anal.* **53**(6) (2021) 6562–6616.
- [53] C. M. Topaz, A. L. Bertozzi and M. A. Lewis, A nonlocal continuum model for biological aggregation, *Bull. Math. Biol.* **68**(7) (2006) 1601–1623.
- [54] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58 (American Mathematical Society, Providence, 2003).
- [55] C. Villani, *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] (Springer-Verlag, Berlin, 2009).