



# Porous medium equation and cross-diffusion systems as limit of nonlocal interaction



Martin Burger<sup>a,b</sup>, Antonio Esposito<sup>c,\*</sup>

<sup>a</sup> Computational Imaging Group and Helmholtz Imaging Deutsches Elektronen Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany

<sup>b</sup> Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany

<sup>c</sup> Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, United Kingdom

## ARTICLE INFO

### Article history:

Received 4 November 2022

Accepted 26 June 2023

Communicated by Francesco Maggi

### MSC:

35Q70

35A15

35D30

35K45

35Q82

### Keywords:

Porous medium equation

Nonlocal interaction

Cross-diffusion systems

Local limit

Variational time discretisation

## ABSTRACT

This paper studies the derivation of the quadratic porous medium equation and a class of cross-diffusion systems from nonlocal interactions. We prove convergence of solutions of a nonlocal interaction equation, resp. system, to solutions of the quadratic porous medium equation, resp. cross-diffusion system, in the limit of a localising interaction kernel. The analysis is carried out at the level of the (nonlocal) partial differential equations and we use gradient flow techniques to derive bounds on energy, second order moments, and logarithmic entropy. The dissipation of the latter yields sufficient regularity to obtain compactness results and pass to the limit in the localised convolutions. The strategy we propose relies on a discretisation scheme, which can be slightly modified in order to extend our result to PDEs without gradient flow structure. In particular, it does not require convexity of the associated energies. Our analysis allows to treat the case of limiting weak solutions of the non-viscous porous medium equation at relevant low regularity, assuming the initial value to have finite energy and entropy.

© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In this manuscript we deal with the connection between nonlocal interaction and the quadratic porous medium equation (PME), as well as a class of cross-diffusion systems (CDS), for a suitable choice of the interaction potentials. We show that a weak solution of the quadratic porous medium equation can be obtained as limit of a sequence of weak measure solutions of a nonlocal interaction equation; this can be extended to a class of cross-diffusion systems. More precisely, starting with the case of (PME) for ease of presentation, let  $W_1 := V_1 * V_1$ , for a function  $V_1$  satisfying some assumptions that will be clarified later, cf. (V). For any  $\varepsilon > 0$ , consider the scaling

$$W_\varepsilon(x) = \varepsilon^{-d} W_1(x/\varepsilon), \quad \text{i.e.} \quad W_\varepsilon = V_\varepsilon * V_\varepsilon.$$

\* Corresponding author.

E-mail addresses: [martin.burger@desy.de](mailto:martin.burger@desy.de) (M. Burger), [antonio.esposito@maths.ox.ac.uk](mailto:antonio.esposito@maths.ox.ac.uk) (A. Esposito).

We prove that, as  $\varepsilon \rightarrow 0^+$ , a sequence of weak measure solutions to

$$\partial_t \rho^\varepsilon = \nabla \cdot (\rho^\varepsilon \nabla W_\varepsilon * \rho^\varepsilon) \tag{NLIE}$$

converges to a weak solution of

$$\partial_t \rho = \frac{1}{2} \Delta(\rho^2) = \nabla \cdot (\rho \nabla \rho). \tag{PME}$$

An analogous result in the case of multi-species leads to a class of cross diffusion systems

$$\partial_t \rho_i = \sum_{j=1}^M \operatorname{div} (\rho_i A_{ij} \nabla \rho_j), \tag{CDS}$$

for  $i, j = 1, \dots, M$ ,  $M \in \mathbb{N}$ , under suitable assumptions on the matrix of the coefficients, as well as interaction kernels in the nonlocal version.

The main motivation for this work is to provide further insights into the derivation of diffusion-type equations from a system of interacting particles, rather than a direct derivation as in continuum mechanics. We refer the reader to [46] for a complete overview on the analysis of the porous medium equation. Obtaining a particle approximation for the partial differential equation under study is a fascinating and useful result for the analysis of PDEs, as it provides a rigorous derivation and way to construct solutions of PDEs, leading to well-posedness, as well as powerful numerical methods. We mention here the seminal works [22,33,37] and the review [26]. In case of transport equations (without diffusion), e.g. (NLIE), deterministic approaches represent a reasonable choice since weak measure solutions may exist, in particular *particle solutions* in the form of an empirical measure

$$\rho_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)},$$

where, for any  $i = 1, \dots, N$ ,  $X_i(t)$  solves a suitable ODE. For instance, for (NLIE) we would have the ODEs

$$\dot{X}_i(t) = -\frac{1}{N} \sum_j \nabla W_\varepsilon(X_i(t) - X_j(t)).$$

For further details we refer the reader to [6,9], and to [19,20] in case of systems of nonlocal PDEs. The problem is substantially different when diffusion is present, initial values in the form of an empirical measure disperse. More precisely, starting from a Dirac delta as initial datum, we will see an immediate smoothing effect which excludes measure solutions. For this reason, deterministic particle approximations are challenging, even though numerical methods have been proposed in this direction. We mention [27,40] for one dimensional linear and nonlinear diffusion, respectively, and [7,17] in any dimension.

A successful attempt to overtake the aforementioned difficulty is given by stochastic particles undergoing a Brownian motion. We start mentioning an inspiring work for our paper, [24], where Figalli and Philipowski deal with the viscous porous medium equation with exponent  $m > 1$ . They obtain (very weak) solutions as limit of a sequence of distributions of the solutions to nonlinear stochastic differential equations. This generalises previous results by Oelschläger, [32,36], and Philipowski, cf. [24], the latter concerning only the case  $m = 2$ . As byproduct of their analysis, the authors of [24] prove propagation of chaos, thus providing a connection between microscopic and macroscopic description. In [38] the quadratic porous medium equation is derived from a stochastic mean field interacting particle system with the addition of a vanishing Brownian motion. The concept of solution used is that of strong  $L^1$ , following [45], which is not the one used in this work, cf. Definition 2.1 below. We point out that our strategy is different from the aforementioned papers since it does not require the addition of higher regularity induced by (vanishing) viscosity, but is based on an optimal transport theory approach, using the 2-Wasserstein gradient flow structure of the two equations.

As counterpoint to stochastic methods, in [35], Oelschläger proves for the first time a particle approximation for classical and positive solutions of (PME) in  $\mathbb{R}^d$ , and for weak solutions in one dimension. Still in one space dimension, [15] presents a deterministic particle approximation for aggregation-diffusion equations, including the porous medium equation. In the recent article [7], the authors provide a deterministic particle method for linear and nonlinear diffusion equations, interpreted as 2-Wasserstein gradient flows. Their approach is inspired by the blob method for aggregation equations in [12]. In particular, in [7], Carrillo, Craig, and Patacchini proceed by regularising the associated internal energy and prove  $\Gamma$ -convergence towards the unregularised energy, both for linear and nonlinear diffusion, that is  $m \geq 1$ . With the addition of a confining drift or interaction potential, they also show stability of minimisers, ensured by the additional potentials. In case  $m \geq 2$ , they provide stability of gradient flows under sufficient regularity conditions, using the approach of Sandier and Serfaty, [41,43]. For the quadratic porous medium equation, i.e.  $m = 2$ , the regularity conditions needed are satisfied for an initial datum with bounded second order moments and entropy — as in our case. This generalises a previous result by Lions and Mas-Gallic, [30], on a numerical scheme for (PME) on a bounded domain with periodic boundary conditions. In the case  $m > 2$  or more general initial data, it is an open problem to check and apply the stability in [7, Theorem 5.8]. Recently in [13], Craig et al. use the blob method to obtain a deterministic particle approximation for weighted (quadratic) porous medium equations, relevant in sampling methods, control theory, and in models of two-layer neural networks.

In the case of cross-diffusion systems, a stochastic approach has been recently considered by Chen et al. in [10], extending the mean-field limit studied in [11], the latter differing from (CDS) by the addition of linear diffusion in each species. The key idea in [10] is to consider interacting diffusion coefficients in the systems of stochastic differential equations and to perform the mean-field limit using an intermediate nonlocal cross-diffusion system. In [34], Moussa shows the nonlocal-to-local limit for the triangular SKT model on a torous, with bounded coefficients. The latter work partially addresses a question raised by Fonbona and Méléard in [25], where they consider a nonlocal version of the SKT model, [44]. Let us stress that (CDS) is different from the SKT model. To the best of our knowledge, a deterministic derivation has not been proven yet.

Our result is strictly related to [7,13] as it provides a rigorous procedure to derive the quadratic porous medium equation from the nonlocal interaction equation, upon choosing a suitable localising kernel. Indeed, one can observe the two associated energies are strictly connected if  $W = V * V$ , cf. Remark 3.1. This assumption is technical and used in the aforementioned results. We observe that the regularisation of the energy in [7,13], for  $m = 2$  and  $\bar{\rho} \equiv 1$ , corresponds to our choice for the interaction potential in (NLIE), though we relax the regularity assumption on the kernel in the convolution, so that to include Morse type potentials, cf. (V). Moreover, we propose an alternative approach that can be used without  $\lambda$ -convexity and even if a gradient flow structure is not exhibited (see below and Section 7). More precisely, we construct solutions of (NLIE) by means of the JKO scheme, [28], in order to obtain uniform estimates on the sequence of solutions  $\{\rho^\varepsilon\}_\varepsilon$ , which are nevertheless only measures. This issue is solved by considering a smoothed version of  $\rho^\varepsilon$ , given by  $v^\varepsilon := V_\varepsilon * \rho^\varepsilon$ . Indeed, starting from an initial probability density in  $L^2(\mathbb{R}^d)$  with finite second order moment and logarithmic entropy, we are able to prove a uniform bound in  $H^1$  for  $v_\varepsilon$ , entailing the right compactness to pass to the limit in the weak formulations of the equations and recover a weak solution of (PME). Our analysis mirrors that convergence from deterministic particle system might not work due to infinite entropy. This problem is also observed in [7, Remark6.3], though numerical simulations in [7, Section6] give confidence that deterministic approximation could be achieved. Indeed,  $\lambda$ -convexity of the associated energy allows to exploit stability estimates with the 2-Wasserstein distance so that to obtain a particle approximation when the number of particles involved depends on  $\varepsilon$ , i.e.  $N = N(\varepsilon)$ , in a way the approximation of the initial datum converges zero in Wasserstein fast enough — this is proven in [13, Theorem 1.4] assuming more regularity on the kernel, at least  $V \in C^2$ . A result for the number of particles independent of the localisation scaling is still open. Interacting stochastic particle systems still represent a solid method.

Our approach neither exploits  $\lambda$ -convexity of the energies involved, as in [7,13], nor any equivalent gradient flow formulation of the equations such as *evolution variational inequality* or *curve of maximal slope*. This is indeed an advantage, since using the JKO scheme at the level of the nonlocal interaction equation allows to extend our strategy to the case of equations which are not gradient flows by means of a suitable splitting scheme, [5]. The latter issue is also relevant for the extension of our result to cross-diffusion systems, since geodesic convexity is valid in few cases, if any. Indeed,  $\lambda$ -convex gradient flows may be too restrictive since the corresponding assumption on the diffusion matrix effectively leads to diagonal diffusion, i.e.  $A_{ij} = 0$  for  $j \neq i$ , cf. [2,48]. Furthermore, in order to include a large class of cross-diffusion systems, we note that the corresponding nonlocal interaction system does not necessarily exhibit a Wasserstein gradient flow structure, cf. [20]. However, this does not exclude to apply a time-discretisation of the system to get uniform bounds and existence of a (sequence) of solutions, as proven in Section 6.

We also observe that our strategy can be applied to linear Fokker–Planck equations, for suitable assumptions on the external potentials, since we anyway need to assume finite logarithmic entropy initially. Similarly, one can add linear diffusion for each species in (CDS). As previously mentioned, the extension to the non-viscous and *non-quadratic* porous medium equation is still an open problem, cf. [7, Theorem 5.8]. It is then natural to see whether our approach can be used for  $m \neq 2$ , using a different nonlocal equation. For a better understanding of these problems we provide more details in Section 7.

### 1.1. Structure of the paper

First, in Section 2 we specify the notation and preliminary concepts used throughout the paper. In Section 3 we focus on the nonlocal interaction equation (NLIE). We provide existence of a sequence of weak measure solutions, cf. Definition 2.2, by means of the JKO scheme, which is useful to derive uniform estimates on the associated energy and second order moments. Section 4 is devoted to obtain the suitable compactness for the sequence of weak measure solutions to (NLIE). In order to pass to the limit in the weak formulation of (NLIE) to obtain the weak solution of (PME) we derive a uniform  $H^1$  bound (in space) on a suitable smoothed sequence associated, by taking advantage of the time-discretisation of (NLIE). In view of this analysis, we recover the weak solution of (PME) in the  $\varepsilon \rightarrow 0^+$  limit in Theorem 5.1 in Section 5. The results obtained in the previous sections are extended to the multi-species case in Section 6. We conclude the paper with some remarks on possible extensions of our result in Section 7.

## 2. Notation and preliminaries

The interaction potential we consider in the one-species case is the (rescaled) convolution  $W_1 := V_1 * V_1$ , being  $V_1 : \mathbb{R}^d \rightarrow \mathbb{R}$  such that the following conditions hold:

$$(V) \quad V_1 \in C_b(\mathbb{R}^d; [0, +\infty)) \cap C^1(\mathbb{R}^d \setminus \{0\}), \|V_1\|_{L^1} = 1, V_1(x) = V_1(-x), \int_{\mathbb{R}^d} |x|V_1(x) dx < +\infty, \nabla V_1 \in L^1(\mathbb{R}^d), \text{ and } |\nabla V_1(x)| \leq C(1 + |x|).$$

Among possible examples of kernels  $V_1$ , we mention Gaussians or pointy potentials such as Morse, the latter not covered by previous results. The List (V) implies that the interaction potential satisfies

$$(W) \quad W_1 \in C(\mathbb{R}^d; [0, \infty)), W_1(x) = W_1(-x) \text{ for all } x \in \mathbb{R}^d, W_1 \in C^1(\mathbb{R}^d \setminus \{0\}) \text{ such that } \nabla W_1 = \nabla V_1 * V_1 \text{ and } \nabla W_1 \in L^1(\mathbb{R}^d).$$

More precisely, we consider the rescaled functions  $W_\varepsilon(x) = \varepsilon^{-d}W_1(x/\varepsilon)$  and  $V_\varepsilon(x) = \varepsilon^{-d}V_1(x/\varepsilon)$ , hence  $W_\varepsilon = V_\varepsilon * V_\varepsilon$ .

**Remark 2.1.** We observe that regularity for  $W_1$  is inferred by  $V_1$ , following a standard proof where the lack of compact support can be overtaken by using boundedness from above of  $V_1$  and Egorov’s theorem.

Continuity of partial derivatives is then obtained from  $L^1$  integrability of  $\nabla V_1$  and Lebesgue’s dominated convergence theorem.

Throughout the manuscript we will denote by  $\mathcal{P}(\mathbb{R}^d)$  the set of probability measures on  $\mathbb{R}^d$ , for  $d \in \mathbb{N}$ , and by  $\mathcal{P}_p(\mathbb{R}^d) := \{\rho \in \mathcal{P}(\mathbb{R}^d) : m_p(\rho) < +\infty\}$ , being  $m_p(\rho) := \int_{\mathbb{R}^d} |x|^p d\rho(x)$  the  $p^{\text{th}}$ -order moment of  $\rho$ , for  $1 \leq p < \infty$ . We shall use  $\mathcal{P}_p^a(\mathbb{R}^d)$  for elements in  $\mathcal{P}_p(\mathbb{R}^d)$  which are absolutely continuous with respect to the Lebesgue measure. For  $p = 2$ , the 2-Wasserstein distance between  $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$  is

$$d_W^2(\mu_1, \mu_2) := \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \left\{ \int_{\mathbb{R}^{2d}} |x - y|^2 d\gamma(x, y) \right\}, \tag{1}$$

where  $\Gamma(\mu_1, \mu_2)$  is the class of all transport plans between  $\mu_1$  and  $\mu_2$ , that is the class of measures  $\gamma \in \mathcal{P}(\mathbb{R}^{2d})$  such that, denoting by  $\pi_i$  the projection operator on the  $i$ th component of the product space, the marginality condition

$$(\pi_i)_\# \gamma = \mu_i \quad \text{for } i = 1, 2$$

is satisfied. In the expression above marginals are the push-forward of  $\gamma$  through  $\pi_i$ . For a measure  $\rho \in \mathcal{P}(\mathbb{R}^d)$  and a Borel map  $T : \mathbb{R}^d \rightarrow \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , the push-forward of  $\rho$  through  $T$  is defined by

$$\int_{\mathbb{R}^n} f(y) dT_\# \rho(y) = \int_{\mathbb{R}^d} f(T(x)) d\rho(x) \quad \text{for all } f \text{ Borel functions on } \mathbb{R}^n.$$

Setting  $\Gamma_0(\mu_1, \mu_2)$  as the class of optimal plans, i.e. minimisers of (1), the 2-Wasserstein distance can be written as

$$d_W^2(\mu_1, \mu_2) = \int_{\mathbb{R}^{2d}} |x - y|^2 d\gamma(x, y), \quad \gamma \in \Gamma_0(\mu_1, \mu_2).$$

We refer the reader to [1,42,47] for further details on optimal transport theory and Wasserstein spaces.

**Remark 2.2.** From the Definition of the 2-Wasserstein distance and the inequality  $|y|^2 \leq 2|x|^2 + 2|x - y|^2$  it follows that

$$m_2(\rho_1) \leq 2m_2(\rho_0) + 2d_W^2(\rho_0, \rho_1), \quad \forall \rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d).$$

In Proposition 4.2 we use the 1-Wasserstein distance, denoted by  $d_1$  and defined by

$$d_1(\mu_1, \mu_2) := \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \left\{ \int_{\mathbb{R}^{2d}} |x - y| d\gamma(x, y) \right\}. \tag{2}$$

Below we specify the concept of solution to the quadratic porous medium we consider, as well as that of weak measure solutions of the nonlocal interaction equation.

**Definition 2.1** (*Weak Solution To (PME)*). A weak solution to the porous medium equation

$$\begin{cases} \partial_t \rho = \nabla \cdot (\rho \nabla \rho) \\ \rho(0, \cdot) = \rho_0 \end{cases} \tag{PME}$$

on the time interval  $[0, T]$  with initial datum  $\rho_0 \in \mathcal{P}_2^a(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  such that  $\int_{\mathbb{R}^d} \rho_0(x) \log \rho_0(x) dx < \infty$  is a curve  $\rho \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d))$  satisfying the following properties:

- (1) for almost every  $t \in [0, T]$  the measure  $\rho(t)$  has a density with respect to the Lebesgue measure, still denoted by  $\rho(t)$ , and  $\rho \in L^2([0, T]; H^1(\mathbb{R}^d))$ ;
- (2) for any  $\varphi \in C_c^1(\mathbb{R}^d)$  and all  $t \in [0, T]$  it holds

$$\int_{\mathbb{R}^d} \varphi(x) \rho(t, x) dx = \int_{\mathbb{R}^d} \varphi(x) \rho_0(x) dx - \int_0^t \int_{\mathbb{R}^d} \rho(s, x) \nabla \varphi(x) \cdot \nabla \rho(s, x) dx ds.$$

**Definition 2.2** (*Weak Measure Solution To (NLIE)*). A narrowly continuous curve  $\rho^\varepsilon : [0, T] \rightarrow \mathcal{P}_2(\mathbb{R}^d)$ , mapping  $t \in [0, T] \mapsto \rho_t^\varepsilon \in \mathcal{P}_2(\mathbb{R}^d)$ , is a weak measure solution to (NLIE) if, for every  $\varphi \in C_c^1(\mathbb{R}^d)$  and any  $t \in [0, T]$ , it holds

$$\int_{\mathbb{R}^d} \varphi(x) d\rho_t^\varepsilon(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) = -\frac{1}{2} \int_0^t \iint_{\mathbb{R}^{2d}} (\nabla\varphi(x) - \nabla\varphi(y)) \cdot \nabla W_\varepsilon(x - y) d\rho_r^\varepsilon(y) d\rho_r^\varepsilon(x) dr. \tag{3}$$

**Remark 2.3.** Our choice for the definition of weak measure solution to (NLIE) strongly depends on the aim of our paper, that is showing convergence of solutions of (NLIE) to weak solutions of (PME), according to Definition 2.1. Note that (NLIE) is a continuity equation of the form

$$\begin{cases} \partial_t \rho_t + \nabla \cdot (\rho_t w_t^\varepsilon) = 0 \\ w_t^\varepsilon = -\nabla W_\varepsilon * \rho_t, \end{cases}$$

with a Borel velocity field such that, for  $\varepsilon > 0$ ,

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} |w_t^\varepsilon(x)| d\rho_t(x) dt &= \int_0^T \int_{\mathbb{R}^d} |\nabla W_\varepsilon * \rho_t(x)| d\rho_t(x) dt \\ &\leq \frac{CT}{\varepsilon^d} \|V_1\|_{L^1(\mathbb{R}^d)} + \frac{2C}{\varepsilon^{d+1}} \|V_1\|_{L^1(\mathbb{R}^d)} \int_0^T \int_{\mathbb{R}^d} |x| d\rho_t(x) dx dt \\ &\quad + \frac{CT}{\varepsilon^d} \int_{\mathbb{R}^d} |x| V_1(x) dx < +\infty, \end{aligned}$$

where we used the growth condition on  $|\nabla V_\varepsilon|$  and preservation of second order moments (cf. Proposition 3.1). In turn, [1, Lemma 8.2.1] provides the existence of a continuous representative for distributional solutions of continuity equations with velocity fields in  $L^1([0, T]; L^1(\rho_t))$ . In particular, for test functions time-independent, we get formulation (3), where we also used that  $\nabla W_\varepsilon$  is odd. Note that this formulation overtakes the loss of regularity at 0 for  $\nabla W_\varepsilon$ , as already noticed in [9].

For the reader’s convenience we postpone notations and preliminaries on the multi-species case to Section 6.

### 3. Results on the nonlocal interaction equation

The nonlocal interaction equation has been intensively studied, especially in the context of 2-Wasserstein gradient flows. In [1], the authors deal with (NLIE) for convex potentials that do not produce a blow-up in finite time. In case of more singular convex potentials, a well-posedness theory for weak measure solutions is given by [9]. Furthermore, it is worth to mention [3] and the references therein, where  $L^p$  theory for the aggregation equation is provided.

In this paper we consider an interaction potential satisfying assumptions similar to [9], though not convex, with the aim of applying the JKO scheme, [28], in order to obtain a priori estimates on the solutions of (NLIE) and their smoothed version  $v^\varepsilon = V_\varepsilon * \rho^\varepsilon$ . In turn, we are able to show convergence towards (PME). The interaction potential we choose is the (rescaled) convolution  $W_\varepsilon = V_\varepsilon * V_\varepsilon$ , for  $W_\varepsilon(x) = \varepsilon^{-d} W_1(x/\varepsilon)$ , satisfying (W) that we recall here for convenience:

(W)  $W_\varepsilon \in C(\mathbb{R}^d; [0, \infty))$ ,  $W_\varepsilon(x) = W_\varepsilon(-x)$  for all  $x \in \mathbb{R}^d$ ,  $W_\varepsilon \in C^1(\mathbb{R}^d \setminus \{0\})$  such that  $\nabla W_\varepsilon = \nabla V_\varepsilon * V_\varepsilon$  and  $\nabla W_\varepsilon \in L^1(\mathbb{R}^d)$ .

Let us emphasise that in this section  $\varepsilon > 0$  is fixed and finite. We assume the initial datum  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ , and the interaction energy functional  $W_\varepsilon : \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$  is given by

$$W_\varepsilon[\rho] = \frac{1}{2} \int_{\mathbb{R}^d} (W_\varepsilon * \rho)(x) d\rho(x).$$

**Remark 3.1.** Using that  $V_\varepsilon$  is even, we observe that the nonlocal interaction energy is nothing but the  $L^2$  norm of the smoothed solution  $v^\varepsilon$ , since

$$\int_{\mathbb{R}^d} [V_\varepsilon * (V_\varepsilon * \rho)](x) d\rho(x) = \int_{\mathbb{R}^d} |(V_\varepsilon * \rho)(x)|^2 dx.$$

In view of the equivalence above, for  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  we have a uniform bound from above for the nonlocal interaction energy at the initial datum. More precisely,

$$\begin{aligned} \mathcal{W}_\varepsilon[\rho_0] &= \frac{1}{2} \int_{\mathbb{R}^d} (W_\varepsilon * \rho_0)(x) \rho_0(x) dx \\ &= \frac{1}{2} \int_{\mathbb{R}^d} |(V_\varepsilon * \rho_0)(x)|^2 dx \\ &= \frac{1}{2} \|V_\varepsilon * \rho_0\|_{L^2(\mathbb{R}^d)}^2 \leq \frac{1}{2} \|V_\varepsilon\|_{L^1}^2 \|\rho_0\|_{L^2}^2 = \frac{1}{2} \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2 < \infty. \end{aligned}$$

We now proceed with the JKO scheme. First, we define a sequence recursively as follows:

- fix a time step  $\tau > 0$  such that  $\rho_{\tau,\varepsilon}^0 := \rho_0$ ;
- for a given  $\rho_{\tau,\varepsilon}^n \in \mathcal{P}_2(\mathbb{R}^d)$ , choose

$$\rho_{\tau,\varepsilon}^{n+1} \in \operatorname{argmin}_{\rho \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{d_W^2(\rho_{\tau,\varepsilon}^n, \rho)}{2\tau} + \mathcal{W}_\varepsilon[\rho] \right\}. \tag{4}$$

The above sequence is well-defined if, for fixed  $\bar{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$ , the penalised energy functional  $\rho \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \frac{d_W^2(\bar{\rho}, \rho)}{2\tau} + \mathcal{W}_\varepsilon[\rho]$  admits minimisers. This can be easily proven by applying the direct method of calculus of variations. For further details we refer to [9, Lemma 2.3 and Proposition 2.5], although we notice that in our case lower semi-continuity is easier. More precisely, the penalised energy functional is bounded from below and lower semicontinuous w.r.t. the narrow convergence by noticing that  $W_\varepsilon$  is continuous and bounded from below, and the 2-Wasserstein distance is lower semicontinuous.

Let  $T > 0$  be fixed, and define a piecewise constant interpolation as follows: assume  $N := \lceil \frac{T}{\tau} \rceil$  and set

$$\rho_\tau^\varepsilon(t) = \rho_{\tau,\varepsilon}^n \quad t \in ((n-1)\tau, n\tau],$$

being  $\rho_{\tau,\varepsilon}^n$  defined in (4).

In the next Proposition we prove narrow compactness (in  $\tau$ ) for  $\rho_\tau^\varepsilon$  and two crucial estimates for its limiting curve, which we shall see it is a solution to (NLIE). More precisely we prove uniform bounds in  $\tau$  and  $\varepsilon$  for the interaction energy and second order moments.

**Proposition 3.1** (Narrow Compactness, Energy & Moments Bound). *There exists an absolutely continuous curve  $\tilde{\rho}^\varepsilon : [0, T] \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  such that the piecewise constant interpolation  $\rho_\tau^\varepsilon$  admits a subsequence  $\rho_{\tau_k}^\varepsilon$  narrowly converging to  $\tilde{\rho}^\varepsilon$  uniformly in  $t \in [0, T]$  as  $k \rightarrow +\infty$ . Moreover, for any  $t \in [0, T]$ , the following uniform bounds in  $\tau$  and  $\varepsilon$  hold*

$$\mathcal{W}_\varepsilon[\tilde{\rho}^\varepsilon(t)] \leq \frac{1}{2} \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2, \tag{5a}$$

$$m_2(\tilde{\rho}^\varepsilon) \leq 2m_2(\rho_0) + 2T \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2. \tag{5b}$$

**Proof.** From the Definition of the sequence  $\{\rho_{\tau,\varepsilon}^n\}_{n \in \mathbb{N}}$  it holds

$$\frac{d_W^2(\rho_{\tau,\varepsilon}^n, \rho_{\tau,\varepsilon}^{n+1})}{2\tau} + \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^{n+1}] \leq \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^n], \tag{6}$$



which implies  $\mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^{n+1}] \leq \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^n]$ , and, in particular, the following bound for the interaction energy (cf. Remark 3.1)

$$\sup_n \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^n] \leq \mathcal{W}_\varepsilon[\rho_0] \leq \frac{1}{2} \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2. \tag{7}$$

By summing up over  $k$  inequality (6), we obtain

$$\sum_{k=m}^n \frac{d_W^2(\rho_{\tau,\varepsilon}^k, \rho_{\tau,\varepsilon}^{k+1})}{2\tau} \leq \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^m] - \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^{n+1}]. \tag{8}$$

The non-negativity of  $W_\varepsilon$  and the energy inequality (7) allow us to improve the above inequality to

$$\sum_{k=m}^n \frac{d_W^2(\rho_{\tau,\varepsilon}^k, \rho_{\tau,\varepsilon}^{k+1})}{2\tau} \leq \mathcal{W}_\varepsilon[\rho_0]. \tag{9}$$

This implies

$$d_W^2(\rho_0, \rho_\tau^\varepsilon(t)) \leq 2T\mathcal{W}_\varepsilon[\rho_0] \leq T\|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2,$$

whence we obtain that second order moments are uniformly bounded on  $[0, T]$  in view of Remark 2.2, i.e.

$$m_2(\rho_\tau^\varepsilon(t)) \leq 2m_2(\rho_0) + 2d_W^2(\rho_0, \rho_\tau^\varepsilon(t)) \leq 2m_2(\rho_0) + 2T\|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2. \tag{10}$$

Now, let us consider  $0 \leq s < t$  such that  $s \in ((m - 1)\tau, m\tau]$  and  $t \in ((n - 1)\tau, n\tau]$  (which implies  $|n - m| < \frac{|t-s|}{\tau} + 1$ ); by Cauchy–Schwarz inequality and (9), we obtain

$$\begin{aligned} d_W(\rho_\tau^\varepsilon(s), \rho_\tau^\varepsilon(t)) &\leq \sum_{k=m}^{n-1} d_W(\rho_{\tau,\varepsilon}^k, \rho_{\tau,\varepsilon}^{k+1}) \leq \left( \sum_{k=m}^{n-1} d_W^2(\rho_{\tau,\varepsilon}^k, \rho_{\tau,\varepsilon}^{k+1}) \right)^{\frac{1}{2}} |n - m|^{\frac{1}{2}} \\ &\leq c \left( \sqrt{|t - s|} + \sqrt{\tau} \right), \end{aligned} \tag{11}$$

where  $c$  is a positive constant. Thus  $\rho_\tau^\varepsilon$  is  $\frac{1}{2}$ -Hölder equi-continuous, up to a negligible error of order  $\sqrt{\tau}$ . By using a refined version of Ascoli–Arzelà’s theorem, [1, Proposition 3.3.1], we obtain  $\rho_\tau^\varepsilon$  admits a subsequence narrowly converging to a limit  $\tilde{\rho}^\varepsilon$  as  $\tau \rightarrow 0^+$  uniformly on  $[0, T]$ . Since  $|\cdot|^2$  and  $W_\varepsilon$  are lower semicontinuous and bounded from below, we actually have for any  $t \in [0, T]$

$$\begin{aligned} \liminf_{k \rightarrow +\infty} \int_{\mathbb{R}^d} |x|^2 d\rho_{\tau_k}^\varepsilon(x) &\geq \int_{\mathbb{R}^d} |x|^2 d\tilde{\rho}^\varepsilon(x) \\ \liminf_{k \rightarrow +\infty} \mathcal{W}[\rho_{\tau_k}^\varepsilon] &\geq \mathcal{W}_\varepsilon[\tilde{\rho}^\varepsilon], \end{aligned}$$

whence the assertion follows.  $\square$

Next, we show that  $\tilde{\rho}^\varepsilon$  provided by Proposition 3.1 is indeed a solution to (NLIE). We stress that this result is not surprising and it is not the main purpose of this paper. Nevertheless, our interaction potential  $W_\varepsilon$  does not satisfy the convexity assumption required in [1,9], where there is a rigorous theory for weak measure solutions to (NLIE). Therefore, for the sake of completeness we show that the lack of convexity does not affect existence of solutions to (NLIE). In fact, we can pass to the limit in the Euler–Lagrange equation associated to (4).

**Theorem 3.1.** *The curve  $\tilde{\rho}^\varepsilon$  is a weak measure solution to (NLIE) according to Definition 2.2.*

**Proof.** Let us consider two consecutive elements of the sequence  $\{\rho_{\tau,\varepsilon}^n\}_{n \in \mathbb{N}}$  defined in (4), i.e.  $\rho_{\tau,\varepsilon}^n$  and  $\rho_{\tau,\varepsilon}^{n+1}$ . We perturb  $\rho_{\tau,\varepsilon}^{n+1}$  by using a map  $P^\sigma = id + \sigma\zeta$ , for some  $\zeta \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$  and  $\sigma \geq 0$ , that is we consider the perturbation

$$\rho^\sigma := P_{\#}^\sigma \rho_{\tau,\varepsilon}^{n+1}. \tag{12}$$



Being  $\rho_{\tau,\varepsilon}^{n+1}$  a minimiser of (4), we have

$$\frac{1}{2\tau} \left[ \frac{d_W^2(\rho_{\tau,\varepsilon}^n, \rho^\sigma) - d_W^2(\rho_{\tau,\varepsilon}^n, \rho_{\tau,\varepsilon}^{n+1})}{\sigma} \right] + \frac{\mathcal{W}_\varepsilon[\rho^\sigma] - \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^{n+1}]}{\sigma} \geq 0. \tag{13}$$

First, we consider the interaction terms in (13)

$$\begin{aligned} & \frac{1}{2\sigma} \int_{\mathbb{R}^d} W_\varepsilon * \rho^\sigma(x) d\rho^\sigma(x) - \frac{1}{2\sigma} \int_{\mathbb{R}^d} W_\varepsilon * \rho_{\tau,\varepsilon}^{n+1}(x) d\rho_{\tau,\varepsilon}^{n+1}(x) \\ &= \frac{1}{2} \iint_{\mathbb{R}^{2d}} \left[ \frac{W_\varepsilon(P^\sigma(x) - P^\sigma(y)) - W_\varepsilon(x - y)}{\sigma} \right] d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x) \\ &= \frac{1}{2} \iint_{\mathbb{R}^{2d}} \left[ \frac{W_\varepsilon(x - y + \sigma(\zeta(x) - \zeta(y))) - W_\varepsilon(x - y)}{\sigma} \right] d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x). \end{aligned} \tag{14}$$

Since the interaction potential satisfies  $W_\varepsilon \in C(\mathbb{R}^d) \cap C^1(\mathbb{R}^d \setminus \{0\})$ , for all  $(x, y) \in \mathbb{R}^{2d}$  it holds

$$\frac{W_\varepsilon(x - y + \sigma(\zeta(x) - \zeta(y))) - W_\varepsilon(x - y)}{\sigma} \xrightarrow{\sigma \rightarrow 0} \nabla W_\varepsilon(x - y) \cdot (\zeta(x) - \zeta(y)). \tag{15}$$

By means of Egorov’s theorem, for every  $\eta > 0$  there exists  $B_\eta \subset \mathbb{R}^{2d}$  measurable such that

$$\iint_{B_\eta} d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x) < \eta$$

and the convergence (15) is uniform on  $\mathbb{R}^{2d} \setminus B_\eta$ . The integral on  $B_\eta$  can be neglected in the limit-integral interchange since the sequence in (15) is uniformly bounded in  $\sigma$ . Thus, we obtain

$$\begin{aligned} & \frac{1}{2} \iint_{\mathbb{R}^{2d}} \left( \frac{W_\varepsilon(x - y + \sigma(\zeta(x) - \zeta(y))) - W_\varepsilon(x - y)}{\sigma} \right) d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x) \\ & \xrightarrow{\sigma \rightarrow 0} \frac{1}{2} \iint_{\mathbb{R}^{2d}} \nabla W_\varepsilon(x - y) \cdot (\zeta(x) - \zeta(y)) d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x). \end{aligned}$$

Regarding the terms involving the 2-Wasserstein distance, let us consider an optimal transport plan  $\gamma_{\tau,\varepsilon}^n \in \Gamma_0(\rho_{\tau,\varepsilon}^n, \rho_{\tau,\varepsilon}^{n+1})$  between  $\rho_{\tau,\varepsilon}^n$  and  $\rho_{\tau,\varepsilon}^{n+1}$ . By Definition of  $d_W$ , we have

$$\begin{aligned} \frac{1}{2\tau} \left[ \frac{d_W^2(\rho_{\tau,\varepsilon}^n, \rho^\sigma) - d_W^2(\rho_{\tau,\varepsilon}^n, \rho_{\tau,\varepsilon}^{n+1})}{\sigma} \right] &\leq \frac{1}{2\tau\sigma} \iint_{\mathbb{R}^{2d}} (|x - P^\sigma(y)|^2 - |x - y|^2) d\gamma_{\tau,\varepsilon}^n(x, y) \\ &= \frac{1}{2\tau\sigma} \iint_{\mathbb{R}^{2d}} (|x - y - \sigma\zeta(y)|^2 - |x - y|^2) d\gamma_{\tau,\varepsilon}^n(x, y) \\ &= -\frac{1}{\tau} \iint_{\mathbb{R}^{2d}} (x - y) \cdot \zeta(y) d\gamma_{\tau,\varepsilon}^n(x, y) + o(\sigma), \end{aligned}$$

where in the last equality we applied a first order Taylor expansion (note that  $\zeta \in C^\infty$ ). By sending  $\sigma$  to 0 it holds

$$\frac{1}{\tau} \iint_{\mathbb{R}^{2d}} (x - y) \cdot \zeta(y) d\gamma_{\tau,\varepsilon}^n(x, y) \leq \frac{1}{2} \iint_{\mathbb{R}^{2d}} \nabla W_\varepsilon(x - y) \cdot (\zeta(x) - \zeta(y)) d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x).$$

Repeating the same computation for  $\sigma \leq 0$ , we actually obtain an equality, that is, for  $\zeta = \nabla\varphi$

$$\frac{1}{\tau} \iint_{\mathbb{R}^{2d}} (x - y) \cdot \nabla\varphi(y) d\gamma_{\tau,\varepsilon}^n(x, y) = \frac{1}{2} \iint_{\mathbb{R}^{2d}} \nabla W_\varepsilon(x - y) \cdot (\nabla\varphi(x) - \nabla\varphi(y)) d\rho_{\tau,\varepsilon}^{n+1}(y) d\rho_{\tau,\varepsilon}^{n+1}(x). \tag{16}$$

Note that the Hölder estimate (11) and  $(x - y) \cdot \nabla\varphi(y) = \varphi(x) - \varphi(y) + o(|x - y|^2)$  imply

$$\frac{1}{\tau} \iint_{\mathbb{R}^{2d}} (x - y) \cdot \nabla\varphi(y) d\gamma_{\tau,\varepsilon}^n(x, y) = \frac{1}{\tau} \int_{\mathbb{R}^d} \varphi(x) d(\rho_{\tau,\varepsilon}^n - \rho_{\tau,\varepsilon}^{n+1})(x) + O(\tau).$$

Now, let  $0 \leq s < t$  be fixed, with

$$h = \left\lceil \frac{s}{\tau} \right\rceil + 1 \quad \text{and} \quad k = \left\lceil \frac{t}{\tau} \right\rceil.$$

Taking into account the last equality, by summing in (16) over  $j$  from  $h$  to  $k$ , we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d\rho_{\tau,\varepsilon}^{k+1} - \int_{\mathbb{R}^d} \varphi(x) d\rho_{\tau,\varepsilon}^h + O(\tau^2) = \\ - \sum_{j=h}^k \frac{\tau}{2} \iint_{\mathbb{R}^{2d}} \nabla W_\varepsilon(x-y) \cdot (\nabla\varphi(x) - \nabla\varphi(y)) d\rho_{\tau,\varepsilon}^{j+1}(y) d\rho_{\tau,\varepsilon}^{j+1}(x), \end{aligned}$$

which is equivalent to

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d\rho_\tau^\varepsilon(t)(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_\tau^\varepsilon(s)(x) + O(\tau^2) = \\ - \frac{1}{2} \int_s^t \iint_{\mathbb{R}^{2d}} \nabla W_\varepsilon(x-y) \cdot (\nabla\varphi(x) - \nabla\varphi(y)) d\rho_\tau^\varepsilon(r)(y) d\rho_\tau^\varepsilon(r)(x) dr. \end{aligned}$$

Up to pass to a subsequence, the result follows by considering the limit as  $\tau \rightarrow 0^+$  and choosing  $s = 0$ .  $\square$

#### 4. Compactness for $\rho^\varepsilon$ and $v^\varepsilon$

The sequence of solutions  $\{\tilde{\rho}^\varepsilon\}_{\varepsilon>0}$  to (NLIE) constructed in Section 3 is the candidate approximating weak solution of (PME), if we use higher regularity of its smoothed version,  $V_\varepsilon * \tilde{\rho}^\varepsilon$ , in the limit  $\varepsilon \rightarrow 0^+$ . In this section we deal with the compactness for both the sequences. For the ease of presentation, from this point on we drop the symbol tilde used in the previous section to denote the sequence of solutions to (NLIE) in Theorem 3.1.

First, we prove that  $\{\rho^\varepsilon\}_{\varepsilon>0}$  is relatively compact in  $C([0, T], \mathcal{P}_2(\mathbb{R}^d))$ , again by means of a refined version of the Ascoli–Arzelà theorem, [1, Proposition 3.3.1].

**Proposition 4.1.** *There exists an absolutely continuous curve  $\tilde{\rho} : [0, T] \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  such that the sequence  $\{\rho^\varepsilon\}_{\varepsilon>0}$  admits a subsequence  $\{\rho^{\varepsilon_k}\}$  such that  $\rho^{\varepsilon_k}(t)$  narrow converges to  $\tilde{\rho}(t)$  for any  $t \in [0, T]$  as  $k \rightarrow +\infty$ .*

**Proof.** Firstly, a subset  $K \subset \mathcal{P}_2(\mathbb{R}^d)$  is relatively compact if and only if it is tight, due to Prokhorov’s theorem. The sequence  $\rho^\varepsilon$  is tight since its second order moments are uniformly bounded according to Proposition 3.1. Secondly, the equi-continuity of  $\rho^\varepsilon$  follows from that of  $\rho_\tau^\varepsilon$ , cf. (11), by lower semi-continuity of the 2-Wasserstein distance. More precisely, for any  $\varepsilon > 0$  and  $s, t \in [0, T]$ , let us consider a sequence of optimal transport plans  $\gamma_\tau^\varepsilon \in \Gamma_0(\rho_\tau^\varepsilon(s), \rho_\tau^\varepsilon(t))$  such that

$$d_W^2(\rho_\tau^\varepsilon(s), \rho_\tau^\varepsilon(t)) = \iint_{\mathbb{R}^{2d}} |x - y|^2 d\gamma_\tau^\varepsilon(x, y).$$

By stability of optimal transport plans, cf. [47, Corollary 5.21], we get  $\gamma_\tau^\varepsilon \rightarrow \gamma^\varepsilon$  as  $\tau \rightarrow 0^+$ , and

$$\begin{aligned} \liminf_{\tau \rightarrow 0} d_W^2(\rho_\tau^\varepsilon(s), \rho_\tau^\varepsilon(t)) &= \liminf_{\tau \rightarrow 0} \iint_{\mathbb{R}^{2d}} |x - y|^2 d\gamma_\tau^\varepsilon(x, y) \\ &\geq \iint_{\mathbb{R}^{2d}} |x - y|^2 d\gamma^\varepsilon(x, y) \\ &\geq d_W^2(\rho^\varepsilon(s), \rho^\varepsilon(t)). \end{aligned}$$

In particular, from (11) in Proposition 3.1 we obtain

$$d_W(\rho^\varepsilon(s), \rho^\varepsilon(t)) \leq c|t - s|,$$

for a positive constant  $c$ . Finally, the assertion follows by applying the aforementioned version of the Ascoli–Arzelà theorem.  $\square$

Next we consider the corresponding smoothed (sub)sequence  $\{v^\varepsilon\}_\varepsilon$ , being  $v^\varepsilon(t) := V_\varepsilon * \rho^\varepsilon(t)$  for any  $t \in [0, T]$ . Note that we removed the subscript  $k$  for ease of presentation. For the latter sequence we obtain an  $H^1$  estimate by using the *flow interchange technique*, developed by Matthes, McCann and Savaré in [31], cf. also [8,18,21] for further details. The strategy is to compute the dissipation of the interaction energy functional  $W_\varepsilon$  along a solution of an *auxiliary gradient flow*, in order to use the *Evolution Variational Inequality (EVI)* to obtain the desired estimate, leading to compactness.

Since the seminal work by Jordan, Kinderlehrer, and Otto, [28], it is known that the heat equation can be regarded as a 2-Wasserstein steepest descent of the opposite of the Boltzmann entropy, i.e.  $\mathcal{H}[\rho] = \int_{\mathbb{R}^d} \rho(x) \log \rho(x) dx$ . The entropy functional is 0-convex along geodesics and it possesses a unique 0-flow, denoted by  $S_{\mathcal{H}t}$ , given by the heat semigroup (cf. [1,16,21]). For the reader’s convenience we recall the Definition of  $\lambda$ -flow for a general functional  $\mathcal{F}$ .

**Definition 4.1** ( $\lambda$ -flow). A semigroup  $S_{\mathcal{F}} : [0, +\infty] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  is a  $\lambda$ -flow for a functional  $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  with respect to the distance  $d_W$  if, for an arbitrary  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ , the curve  $t \mapsto S_{\mathcal{F}t}^t \rho$  is absolutely continuous on  $[0, +\infty[$  and it satisfies the evolution variational inequality (EVI)

$$\frac{1}{2} \frac{d^+}{dt} d_W^2(S_{\mathcal{F}t}^t \rho, \bar{\rho}) + \frac{\lambda}{2} d_W^2(S_{\mathcal{F}t}^t \rho, \bar{\rho}) \leq \mathcal{F}(\bar{\rho}) - \mathcal{F}(S_{\mathcal{F}t}^t \rho) \tag{17}$$

for all  $t > 0$ , with respect to every reference measure  $\bar{\rho} \in \mathcal{P}_2(\mathbb{R}^d)$  such that  $\mathcal{F}(\bar{\rho}) < \infty$ .

Below we use the flow interchange by considering the heat equation as auxiliary flow, and the entropy as auxiliary functional, i.e.

$$\mathcal{H}[\rho] = \begin{cases} \int_{\mathbb{R}^d} \rho(x) \log \rho(x) dx, & \rho \log \rho \in L^1(\mathbb{R}^d); \\ +\infty & \text{otherwise.} \end{cases} \tag{18}$$

**Remark 4.1.** We remind the reader that the entropy is controlled from below by the second order moment of  $\rho$ , denoted by  $m_2(\rho)$ . More precisely, in [28, Proposition 4.1] it is shown that

$$\mathcal{H}(\rho) \geq -C(m_2(\rho) + 1)^\beta,$$

for every  $\rho \in \mathcal{P}_2^a(\mathbb{R}^d)$ ,  $\beta \in (\frac{d}{d+2}, 1)$  and  $C < +\infty$ , depending only on the space dimension  $d$ . We use this bound in order to have a uniform control from below for the entropy.

In the following, for any  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$  such that  $\mathcal{H}(\nu) < +\infty$ , we denote by  $S_{\mathcal{H}t}^t \nu$  the solution at time  $t$  of the heat equation coupled with an initial value  $\nu$  at  $t = 0$ . Moreover, for every  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ , we define the dissipation of  $W_\varepsilon$  along  $S_{\mathcal{H}t}$  by

$$D_{\mathcal{H}t} W_\varepsilon(\rho) := \limsup_{s \downarrow 0} \left\{ \frac{W_\varepsilon[\rho] - W_\varepsilon[S_{\mathcal{H}t}^s \rho]}{s} \right\}.$$

We can now prove a uniform bound for  $\{v^\varepsilon\}_\varepsilon$  in  $L^2([0, T]; H^1(\mathbb{R}^d))$ .

**Lemma 4.1.** *Let  $\rho_0 \in \mathcal{P}_2^g(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  such that  $\mathcal{H}[\rho_0] < \infty$ . There exists a constant  $C = C(\rho_0, V_1, T)$  such that, for any  $\varepsilon > 0$ ,*

$$\|v^\varepsilon\|_{L^2([0, T]; H^1(\mathbb{R}^d))} \leq C(\rho_0, V_1, T). \tag{19}$$

*Therefore, there exists a subsequence  $\{v^{\varepsilon_k}\}_k$  and a curve  $v \in L^2([0, T]; H^1(\mathbb{R}^d))$  such that  $v^{\varepsilon_k} \rightharpoonup v$  in  $L^2([0, T]; H^1(\mathbb{R}^d))$ .*

**Proof.** From Proposition 3.1 we infer the uniform bound in  $\tau$  and  $\varepsilon$

$$\begin{aligned} \|V_\varepsilon * \rho_\tau^\varepsilon\|_{L^2([0,T];L^2(\mathbb{R}^d))}^2 &= \int_0^T \int_{\mathbb{R}^d} |[V_\varepsilon * \rho_\tau^\varepsilon(t)](x)|^2 dx dt = 2 \int_0^T \mathcal{W}_\varepsilon[\rho_\tau^\varepsilon(t)] dt \\ &\leq 2T\mathcal{W}_\varepsilon[\rho_0] \leq T\|V_1\|_{L^1(\mathbb{R}^d)}^2 \|\rho_0\|_{L^2(\mathbb{R}^d)}^2. \end{aligned}$$

Thus, there exists a subsequence  $\tau_k$  such that  $V_\varepsilon * \rho_{\tau_k}^\varepsilon \rightharpoonup w^\varepsilon$  in  $L^2([0, T]; L^2(\mathbb{R}^d))$  as  $\tau_k \rightarrow 0$ . The limit  $w^\varepsilon \equiv v^\varepsilon$  due to uniqueness of limit and Proposition 4.1. Up to pass to a subsequence, we have

$$\|v^\varepsilon\|_{L^2([0,T];L^2(\mathbb{R}^d))}^2 \leq T\|V_1\|_{L^1(\mathbb{R}^d)}^2 \|\rho_0\|_{L^2(\mathbb{R}^d)}^2, \tag{20}$$

since the norm is weakly lower semicontinuous. Now, we obtain a uniform bound for  $\nabla v^\varepsilon$ . For all  $s > 0$ , if we consider  $S_{\mathcal{H}}^s \rho_{\tau,\varepsilon}^{n+1}$  as competitor of  $\rho_{\tau,\varepsilon}^{n+1}$  in the minimisation problem (4), as direct consequence of the Definition of the sequence  $\{\rho_{\tau,\varepsilon}^n\}_{n \in \mathbb{N}}$  we have

$$\frac{1}{2\tau} d_W^2(\rho_{\tau,\varepsilon}^{n+1}, \rho_{\tau,\varepsilon}^n) + \mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^{n+1}] \leq \frac{1}{2\tau} d_W^2(S_{\mathcal{H}}^s \rho_{\tau,\varepsilon}^{n+1}, \rho_{\tau,\varepsilon}^n) + \mathcal{W}_\varepsilon[S_{\mathcal{H}}^s \rho_{\tau,\varepsilon}^{n+1}],$$

whence, dividing by  $s > 0$  and passing to the lim sup as  $s \downarrow 0$ ,

$$\tau D_{\mathcal{H}} \mathcal{W}_\varepsilon(\rho_{\tau,\varepsilon}^{n+1}) \leq \frac{1}{2} \frac{d^+}{dt} \left( d_W^2(S_{\mathcal{H}}^t \rho_{\tau,\varepsilon}^{n+1}, \rho_{\tau,\varepsilon}^n) \right) \Big|_{t=0} \stackrel{(E.V.I.)}{\leq} \mathcal{H}[\rho_{\tau,\varepsilon}^n] - \mathcal{H}[\rho_{\tau,\varepsilon}^{n+1}]. \tag{21}$$

In the last inequality we used that  $S_{\mathcal{H}}$  is a 0-flow. Now, let us focus on the left hand side of (21). First of all, note that

$$\begin{aligned} D_{\mathcal{H}} \mathcal{W}_\varepsilon(\rho_{\tau,\varepsilon}^{n+1}) &= \limsup_{s \downarrow 0} \left\{ \frac{\mathcal{W}_\varepsilon[\rho_{\tau,\varepsilon}^{n+1}] - \mathcal{W}_\varepsilon[S_{\mathcal{H}}^s \rho_{\tau,\varepsilon}^{n+1}]}{s} \right\} \\ &= \limsup_{s \downarrow 0} \int_0^1 \left( -\frac{d}{dz} \Big|_{z=st} \mathcal{W}_\varepsilon[S_{\mathcal{H}}^z \rho_{\tau,\varepsilon}^{n+1}] \right) dt. \end{aligned} \tag{22}$$

Thus, we now compute the time derivative inside the above integral, by using integration by parts and keeping in mind the regularity of the solution to the heat equation:

$$\begin{aligned} \frac{d}{dt} \mathcal{W}_\varepsilon[S_{\mathcal{H}}^t \rho_{\tau,\varepsilon}^{n+1}] &= - \int_{\mathbb{R}^d} \nabla(W_\varepsilon * S_{\mathcal{H}}^t \rho_{\tau,\varepsilon}^{n+1})(x) \nabla S_{\mathcal{H}}^t \rho_{\tau,\varepsilon}^{n+1}(x) dx \\ &= - \int_{\mathbb{R}^d} |\nabla(V_\varepsilon * S_{\mathcal{H}}^t \rho_{\tau,\varepsilon}^{n+1})(x)|^2 dx. \end{aligned} \tag{23}$$

By substituting (23) into (22), from (21) we obtain

$$\tau \liminf_{s \downarrow 0} \int_0^1 \int_{\mathbb{R}^d} |\nabla(V_\varepsilon * S_{\mathcal{H}}^{st} \rho_{\tau,\varepsilon}^{n+1})(x)|^2 dx dt \leq \mathcal{H}[\rho_{\tau,\varepsilon}^n] - \mathcal{H}[\rho_{\tau,\varepsilon}^{n+1}],$$

whence, by  $L^2$  lower semi-continuity of the  $H^1$  seminorm,

$$\tau \int_{\mathbb{R}^d} |\nabla(V_\varepsilon * \rho_{\tau,\varepsilon}^{n+1})(x)|^2 dx dt \leq \mathcal{H}[\rho_{\tau,\varepsilon}^n] - \mathcal{H}[\rho_{\tau,\varepsilon}^{n+1}].$$

By summing up over  $n$  from 0 to  $N - 1$ , taking into account that  $x \log x \leq x^2$  for any  $x \geq 0$ , Remark 4.1 and that second order moments are uniformly bounded (see Proposition 3.1), we get

$$\int_0^T \int_{\mathbb{R}^d} |\nabla(V_\varepsilon * \rho_\tau^\varepsilon(t))(x)|^2 dx dt \leq \mathcal{H}[\rho_0] - \mathcal{H}[\rho_{\tau,\varepsilon}^N] \leq \|\rho_0\|_{L^2(\mathbb{R}^d)}^2 + C(\rho_0, V_1, T).$$

In particular, using weak lower semi-continuity of the norm,

$$\|\nabla v^\varepsilon\|_{L^2([0,T];L^2(\mathbb{R}^d))}^2 = \int_0^T \int_{\mathbb{R}^d} |\nabla v^\varepsilon(t)(x)|^2 dx dt \leq \|\rho_0\|_{L^2(\mathbb{R}^d)}^2 + C(\rho_0, V_1, T). \tag{24}$$

The bounds in (20) and (24) give the first result of the statement, and an application of the Banach–Alaoglu Theorem concludes the proof.  $\square$

**Remark 4.2.** We observe that  $\rho_0 \in \mathcal{P}_2^a(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  does imply  $\mathcal{H}[\rho_0] < \infty$ , since  $x \log x \leq x^2$  for any  $x \geq 0$ . We prefer to keep this assumption as this is the main issue when dealing with the particle approximation using this strategy.

**Remark 4.3.** Let us notice that, for any  $t \in [0, T]$ , the first order moment of  $v_t^\varepsilon$  is finite. In fact, using that  $V_\varepsilon$  is even, the assumption  $\int_{\mathbb{R}^d} |x|V_1(x) dx < +\infty$ , and [Proposition 3.1](#):

$$\begin{aligned} \int_{\mathbb{R}^d} |x|v_t^\varepsilon(x) dx &= \iint_{\mathbb{R}^{2d}} |x|V_\varepsilon(x - y) d\rho_t^\varepsilon(y) dx \\ &\leq \iint_{\mathbb{R}^{2d}} V_\varepsilon(y - x)|x - y| d\rho_t^\varepsilon(y) dx + \iint_{\mathbb{R}^{2d}} V_\varepsilon(y - x)|y| d\rho_t^\varepsilon(y) dx \\ &= \varepsilon \int_{\mathbb{R}^d} V_1(z)|z| dz + \int_{\mathbb{R}^d} V_1(z) dz \int_{\mathbb{R}^d} |y| d\rho_t^\varepsilon(y) \\ &\leq \varepsilon \int_{\mathbb{R}^d} V_1(z)|z| dz + \sqrt{m_2(\rho^\varepsilon)} \int_{\mathbb{R}^d} V_1(z) dz < +\infty. \end{aligned}$$

The strong  $L^2$  compactness in time and space follows by applying a refined version of the Aubin–Lions Lemma due to Rossi and Savaré [[39](#), Theorem 2]. For the reader’s convenience we recall the latter result below, before presenting the compactness result for  $\{v^{\varepsilon_k}\}_k$ .

**Proposition 4.2** ([\[39, Theorem 2\]](#)). *Let  $X$  be a separable Banach space. Consider*

- a lower semicontinuous functional  $\mathcal{F} : X \rightarrow [0, +\infty]$  with relatively compact sublevels in  $X$ ;
- a pseudo-distance  $g : X \times X \rightarrow [0, +\infty]$ , i.e.,  $g$  is lower semicontinuous and such that  $g(\rho, \eta) = 0$  for any  $\rho, \eta \in X$  with  $\mathcal{F}(\rho) < \infty, \mathcal{F}(\eta) < \infty$  implies  $\rho = \eta$ .

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

$$\sup_{u \in U} \int_0^T \mathcal{F}(u(t)) dt < \infty \quad \text{and} \quad \limsup_{h \downarrow 0} \sup_{u \in U} \int_0^{T-h} g(u(t+h), u(t)) dt = 0. \tag{25}$$

Then  $U$  contains an infinite sequence  $(u_n)_{n \in \mathbb{N}}$  that converges in measure, with respect to  $t \in (0, T)$ , to a measurable  $\tilde{u} : (0, T) \rightarrow X$ , i.e.

$$\lim_{n \rightarrow \infty} |\{t \in (0, T) : \|u_n(t) - \tilde{u}(t)\|_X \geq \sigma\}| = 0, \quad \forall \sigma > 0.$$

The two conditions in [\(25\)](#) are called *tightness* and *weak integral equicontinuity*, respectively.

**Proposition 4.3.** *Let  $\varepsilon \leq 1$ . The sequence  $\{v^{\varepsilon_k}\}_k$  obtained in [Lemma 4.1](#) converges strongly to the curve  $v$  in  $L^2([0, T]; L^2(\mathbb{R}^d))$ , for any  $T > 0$ .*

**Proof.** The proof of the result is obtained by applying [Proposition 4.2](#) to a subset of  $U := \{v^\varepsilon\}_{0 \leq \varepsilon \leq 1}$  for  $X := L^2(\mathbb{R}^d)$  and  $g := d_1$  being the 1-Wasserstein distance — extended to  $+\infty$  outside of  $\mathcal{P}_1(\mathbb{R}^d) \times \mathcal{P}_1(\mathbb{R}^d)$ . As for the functional, we consider  $\mathcal{F} : L^2(\mathbb{R}^d) \rightarrow [0, +\infty]$  defined by

$$\mathcal{F}[v] = \begin{cases} \|v\|_{H^1(\mathbb{R}^d)}^2 + \int_{\mathbb{R}^d} |x|v(x) dx, & \text{if } v \in \mathcal{P}_1(\mathbb{R}^d) \cap H^1(\mathbb{R}^d); \\ +\infty & \text{otherwise.} \end{cases}$$

Note that elements in the domain of the functional  $\mathcal{F}$  belong to  $\mathcal{P}_1(\mathbb{R}^d)$ , thus  $0 = g(\rho, \eta) = d_1(\rho, \eta)$  implies  $\rho = \eta$ . Next we show that  $\mathcal{F}$  is an admissible functional and later on we check the conditions in [\(25\)](#). In order to improve the readability we split the remainder of the proof in four steps.

*Step 1:  $\mathcal{F}$  is lower semicontinuous* Let  $\{v_n\}_n \subset L^2(\mathbb{R}^d)$  such that  $v_n \rightarrow v$  in  $L^2(\mathbb{R}^d)$  and  $F[v_n] < +\infty$ , otherwise it is trivial. We prove that  $\mathcal{F}^1[v] := \|\nabla v\|_{L^2(\mathbb{R}^d)}^2$  and  $\mathcal{F}^2[v] := \int_{\mathbb{R}^d} |x|v(x) dx$  are lower semicontinuous, since  $\|v\|_{L^2(\mathbb{R}^d)}^2$  obviously is. Note that  $\|v_n\|_{H^1(\mathbb{R}^d)}^2 \leq \sup_n \|v_n\|_{H^1(\mathbb{R}^d)}^2 =: \bar{\mathcal{F}} < +\infty$ . Thus, there exists a subsequence such that  $\nabla v_{n_k} \rightharpoonup \nabla v$  in  $L^2(\mathbb{R}^d)$ , since the limit is unique. A straightforward computation shows that

$$\mathcal{F}^1[v_n] \geq \int_{\mathbb{R}^d} |\nabla v(x)|^2 dx + 2 \int_{\mathbb{R}^d} (\nabla v_n(x) - \nabla v) \cdot \nabla v(x) dx,$$

which gives  $\liminf_n \mathcal{F}^1[v_n] \geq \mathcal{F}^1[v]$ . Regarding  $\mathcal{F}^2$ , let us consider  $B_R$  a ball of radius  $R$ . Since  $v_n \rightharpoonup v$  in  $L^2(\mathbb{R}^d)$  and  $|\cdot| \in L^2(B_R)$ , we have

$$\lim_n \int_{B_R} |x|v_n(x) dx = \int_{B_R} |x|v(x) dx,$$

whence

$$\liminf_n \int_{\mathbb{R}^d} |x|v_n(x) dx \geq \liminf_n \int_{B_R} |x|v_n(x) dx = \int_{B_R} |x|v(x) dx$$

The monotone convergence theorem gives the desired result.

*Step 2: sublevels of  $\mathcal{F}$  are relatively compact in  $L^2(\mathbb{R}^d)$*  Let  $A_c := \{v \in L^2(\mathbb{R}^d) : \mathcal{F}[v] \leq c\}$  be a sublevel of  $\mathcal{F}$ , where  $c$  is a positive constant. The Riesz–Fréchet–Kolmogorov theorem provides relative compactness in  $L^2(\mathbb{R}^d)$  of  $A_c$ . In fact, elements of  $A_c$  are bounded in  $L^2(\mathbb{R}^d)$  and it holds the uniform continuity estimate

$$\begin{aligned} \int_{\mathbb{R}^d} |v(x+h) - v(x)|^2 dx &= \int_{\mathbb{R}^d} \left| \int_0^1 \frac{d}{d\tau} v(x+\tau h) d\tau \right|^2 dx = \int_{\mathbb{R}^d} \left| \int_0^1 h \cdot \nabla v(x+\tau h) d\tau \right|^2 dx \\ &\leq |h|^2 \int_{\mathbb{R}^d} \int_0^1 |\nabla v(x+\tau h)|^2 d\tau dx = |h|^2 \|\nabla v\|_{L^2(\mathbb{R}^d)}^2, \end{aligned} \tag{26}$$

which implies  $\|v(\cdot+h) - v(\cdot)\|_{L^2(\mathbb{R}^d)} \rightarrow 0$  as  $h \rightarrow 0^+$ . Moreover, we have uniform integrability at infinity by means of Hölder and Gagliardo–Nirenberg inequalities. In particular,

$$\begin{aligned} \|v\|_{L^2(\mathbb{R}^d \setminus B_R)}^2 &= \int_{|x| \geq R} |v(x)|^2 dx \leq \frac{1}{R^\delta} \int_{\mathbb{R}^d} |x|^\delta |v(x)|^2 dx \\ &\leq \frac{1}{R^\delta} \left( \int_{\mathbb{R}^d} |x|v(x) dx \right)^\delta \left( \int_{\mathbb{R}^d} |v(x)|^{\frac{2-\delta}{1-\delta}} dx \right)^{1-\delta}, \end{aligned}$$

where  $\delta$  can be chosen in  $(0, 1)$  such a way the exponent  $p := (2-\delta)/(1-\delta)$  satisfies  $p \in (2, +\infty)$  for  $d = 1, 2$ , and  $2 < p < \frac{2d}{d-2}$  for  $d > 2$ . The latter requirements are implied by the Gagliardo–Nirenberg inequality

$$\|v\|_{L^p(\mathbb{R}^d)} \leq C \|\nabla v\|_{L^2(\mathbb{R}^d)}^\theta \|v\|_{L^2(\mathbb{R}^d)}^{1-\theta}, \quad \theta = \frac{(p-2)d}{2p},$$

which guarantees that  $\|v\|_{L^p(\mathbb{R}^d)}$  is finite, thus the uniform integrability at infinity.

*Step 3: tightness and weak integral equicontinuity* Let us set  $U := \{v^\varepsilon\}_{0 \leq \varepsilon \leq 1}$ , being  $v^\varepsilon : [0, T] \rightarrow L^2(\mathbb{R}^d)$  the sequence defined above by  $v^\varepsilon = V_\varepsilon * \rho^\varepsilon$ , which satisfies [Lemma 4.1](#). For any  $0 \leq \varepsilon \leq 1$ , it holds

$$\begin{aligned} \int_0^T \mathcal{F}[v^\varepsilon(t)] dt &= \int_0^T \|v^\varepsilon(t)\|_{H^1(\mathbb{R}^d)}^2 dt + \int_0^T \int_{\mathbb{R}^d} |x|v_t^\varepsilon(x) dx dt \\ &\leq C(\rho_0, V_1, T) + T \int_{\mathbb{R}^d} V_1(z)|z| dz < +\infty, \end{aligned}$$

where we also used [Remark 4.3](#) and that  $\varepsilon \leq 1$  — note that the bound for  $\varepsilon$  is arbitrary as we could choose any constant. Taking the supremum in  $U$  we have tightness. The weak integral equi-continuity is a

consequence of the equi-continuity of  $\rho^\varepsilon$  proven in Proposition 4.1. More precisely, for any  $\varepsilon \geq 0$  and  $h > 0$  it holds

$$\int_0^{T-h} d_1(v^\varepsilon(t+h), v^\varepsilon(t)) dt \leq \int_0^{T-h} d_W(v^\varepsilon(t+h), v^\varepsilon(t)) dt \leq \int_0^{T-h} d_W(\rho^\varepsilon(t+h), \rho^\varepsilon(t)) dt \leq c|h|T,$$

where in the intermediate inequalities we used well known properties of Wasserstein distances, cf. for example [42, Section 5.1].

*Step 4: relatively compactness in  $L^2([0, T]; L^2(\mathbb{R}^d))$*  By abuse of notation we denote by  $U := \{v^{\varepsilon_k}\}_k$  a subsequence of  $\{v^\varepsilon\}_{0 \leq \varepsilon \leq 1}$  such that  $v^{\varepsilon_k} \rightharpoonup v$  in  $L^2([0, T]; H^1(\mathbb{R}^d))$ , in view of Lemma 4.1. According to Proposition 4.2, there exists a subsubsequence  $v^{\varepsilon'_k}$  such that  $v^{\varepsilon'_k}$  converges in measure (with respect to time with values in  $X = L^2(\mathbb{R}^d)$ ) to a curve  $\tilde{v} \equiv v$ , due to the weakly convergence of  $v^{\varepsilon_k}$ . By standard arguments we can conclude that  $\{v^{\varepsilon'_k}\}$  converges to  $v$  in measure, thus pointwise almost everywhere (up to pass to a subsequence). Since  $\sup_t \|v^\varepsilon(t)\|_{L^2(\mathbb{R}^d)}^2 \leq \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2$ , we infer strong convergence of  $v^{\varepsilon_k}$  to  $v$  in  $L^2([0, T]; L^2(\mathbb{R}^d))$  by applying Lebesgue’s dominated convergence theorem.  $\square$

### 5. Towards the quadratic porous medium equation

In view of the analysis carried out in the previous sections, we are now able to prove convergence of solutions of (NLIE) to the solution of (PME), as  $\varepsilon \rightarrow 0^+$ . The key issue is to pass to the limit in the weak formulation, which is not straightforward since  $\rho^\varepsilon$  is only a measure in general. As already explained earlier in the paper, we use the higher regularity of  $v^\varepsilon$  and that  $\rho^\varepsilon - v^\varepsilon$  converges to zero in the sense of distributions, starting from  $\rho_0 \in \mathcal{P}_2^a(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  with  $\mathcal{H}[\rho_0] < \infty$ .

According to Definition 2.2, for any  $\varepsilon > 0$  and any  $\varphi \in C_c^1(\mathbb{R}^d)$ ,  $\rho^\varepsilon$  satisfies

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d\rho_T^\varepsilon(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) &= -\frac{1}{2} \int_0^T \iint_{\mathbb{R}^{2d}} (\nabla\varphi(x) - \nabla\varphi(y)) \cdot \nabla W_\varepsilon(x-y) d\rho_t^\varepsilon(y) d\rho_t^\varepsilon(x) dt, \\ &= -\int_0^T \int_{\mathbb{R}^d} \nabla\varphi(x) \cdot \nabla V_\varepsilon * v_t^\varepsilon(x) d\rho_t^\varepsilon(x) dt, \end{aligned} \tag{27}$$

which can be rewritten as

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d\rho_T^\varepsilon(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) &= -\int_0^T \int_{\mathbb{R}^d} V_\varepsilon * (\rho_t^\varepsilon \nabla\varphi)(x) \cdot \nabla v_t^\varepsilon(x) dx dt \\ &= -\int_0^T \int_{\mathbb{R}^d} v_t^\varepsilon(x) \nabla\varphi(x) \cdot \nabla v_t^\varepsilon(x) dx dt \\ &\quad - \int_0^T \int_{\mathbb{R}^d} z_t^\varepsilon(x) \cdot \nabla v_t^\varepsilon(x) dx dt, \end{aligned} \tag{28}$$

where for any  $t \in [0, T]$  and  $x \in \mathbb{R}^d$  the excess term is given by

$$z_t^\varepsilon(x) := V_\varepsilon * (\rho_t^\varepsilon \nabla\varphi)(x) - (V_\varepsilon * \rho_t^\varepsilon)(x) \nabla\varphi(x) = V_\varepsilon * (\rho_t^\varepsilon \nabla\varphi)(x) - v_t^\varepsilon(x) \nabla\varphi(x).$$

**Remark 5.1.** Note that the integral after the second equality in (27) makes sense since  $\nabla V_\varepsilon * v_t^\varepsilon \in C(\mathbb{R}^d)$ . This can be easily verified by applying Lebesgue dominated convergence theorem using that  $\nabla V_\varepsilon \in L^1(\mathbb{R}^d)$ , for  $\varepsilon > 0$ .

Lemma 4.1 and Proposition 4.3 entail to pass to the limit in the first term on the right-hand side of (28), upon considering a subsequence, since  $v^\varepsilon$  converges strongly in  $L^2([0, T]; L^2(\mathbb{R}^d))$  and  $\nabla v^\varepsilon$  converges weakly in  $L^2([0, T]; L^2(\mathbb{R}^d))$ . We will show that  $v^\varepsilon$  converges to the same limit of  $\rho^\varepsilon$  in the sense of distributions, whence we infer that the limit  $\tilde{\rho}$  from Proposition 4.1 attains the same regularity, namely  $L^2([0, T]; H^1(\mathbb{R}^d))$ . Furthermore, we prove that the excess term  $z^\varepsilon$  converges to 0 in  $L^2([0, T] \times \mathbb{R}^d)$ .



5.1. Convergence of the excess term

**Lemma 5.1.** *The excess term  $z^\varepsilon$  satisfies*

$$\|z^\varepsilon\|_{L^\infty([0,T];L^1(\mathbb{R}^d))} \leq \varepsilon C(V_1, \varphi),$$

for any  $\varphi \in C_c^2(\mathbb{R}^d)$ .

**Proof.** For any  $t \in [0, T]$  and  $\varphi \in C_c^2(\mathbb{R}^d)$  we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} |z_t^\varepsilon(x)| dx &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V_\varepsilon(x-y) |\nabla\varphi(y) - \nabla\varphi(x)| d\rho_t^\varepsilon(y) dx \\ &\leq \|D^2\varphi\|_\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V_\varepsilon(x-y) |y-x| d\rho_t^\varepsilon(y) dx \\ &= \varepsilon \|D^2\varphi\|_\infty \int_{\mathbb{R}^d} |z| V_1(z) dz, \end{aligned}$$

by means of the change of variable  $z = \frac{x-y}{\varepsilon}$ . The assertion follows by taking the supremum over  $t \in [0, T]$ .  $\square$

**Lemma 5.2.** *There exists a constant  $C$  only depending on  $\varphi$  and  $V_1$  such that for all  $\varepsilon > 0$  and a  $\delta > 0$  small enough*

$$\|z^\varepsilon\|_{L^2([0,T];L^{2+\delta}(\mathbb{R}^d))} \leq C.$$

**Proof.** For almost every  $x \in \mathbb{R}^d$  and  $t \in [0, T]$ , for  $i = 1, \dots, d$ , the non-negativity of  $V_1$  and  $\rho_t^\varepsilon$  gives

$$\left| \int_{\mathbb{R}^d} V_\varepsilon(x-y) \partial_{x_i} \varphi(y) d\rho_t^\varepsilon(y) \right| \leq \int_{\mathbb{R}^d} V_\varepsilon(x-y) |\partial_{x_i} \varphi(y)| d\rho_t^\varepsilon(y) \leq \|\partial_{x_i} \varphi\|_\infty v_t^\varepsilon(x).$$

In particular, this implies

$$|z_t^\varepsilon(x)| \leq 2 \|\nabla\varphi\|_\infty |v_t^\varepsilon(x)|,$$

thus  $\|z_t^\varepsilon\|_{L^p(\mathbb{R}^d)} \leq 2 \|\nabla\varphi\|_\infty \|v_t^\varepsilon\|_{L^p}$  for almost every  $t \in [0, T]$ , and  $p \geq 1$ .

**Proposition 3.1** and **Lemma 4.1** entail existence of a constant  $c$ , independent of  $\varepsilon$ , such that

$$\sup_{t \in [0, T]} \|v_t^\varepsilon\|_{L^2(\mathbb{R}^d)} \leq c \quad \text{and} \quad \int_0^T \|\nabla v_t^\varepsilon\|_{L^2}^2 dt \leq c.$$

Sobolev embedding theorems provide the estimate since

$$\int_0^T \|v_t^\varepsilon\|_{L^{2+\delta}(\mathbb{R}^d)}^2 dt \leq c,$$

for some constant still denoted by  $c$ .  $\square$

**Corollary 5.1.** *The excess term converges to zero in  $L^2([0, T] \times \mathbb{R}^d)$  as  $\varepsilon \rightarrow 0^+$ .*

**Proof.** The proof is a simple consequence of the interpolation inequality for  $L^p$  functions and the previous Lemmas. More precisely, for  $\alpha = \delta/(2(1 + \delta))$  it holds

$$\begin{aligned} \int_0^T \|z_t^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 dt &\leq \int_0^T \|z_t^\varepsilon\|_{L^1(\mathbb{R}^d)}^{2\alpha} \|z_t^\varepsilon\|_{L^{2+\delta}(\mathbb{R}^d)}^{2(1-\alpha)} dt \\ &\leq \left( \int_0^T \|z_t^\varepsilon\|_{L^1(\mathbb{R}^d)}^2 dt \right)^\alpha \left( \int_0^T \|z_t^\varepsilon\|_{L^{2+\delta}(\mathbb{R}^d)}^2 dt \right)^{1-\alpha} \leq \varepsilon^{2\alpha} T^\alpha C(V_1, \varphi), \end{aligned} \tag{29}$$

which gives the result by letting  $\varepsilon$  tend to 0.  $\square$

5.2. Convergence to the quadratic porous medium equation

Let us consider the subsequence from Proposition 4.1, still denoted by  $\{\rho^\varepsilon\}_\varepsilon$ , which narrowly converges to the curve  $\tilde{\rho}$ . In the next Lemma we show that the corresponding smoothed subsequence, still denoted by  $\{v^\varepsilon\}_\varepsilon$ , converges to  $\tilde{\rho}$  in the sense of distributions.

**Lemma 5.3.** For any  $t \in [0, T]$  and any  $\varphi \in C_c^1(\mathbb{R}^d)$  we have

$$\lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} \varphi(x) v_t^\varepsilon(x) dx = \int_{\mathbb{R}^d} \varphi(x) d\tilde{\rho}(t).$$

**Proof.** For any  $t \in [0, T]$  and any  $\varphi \in C_c^1(\mathbb{R}^d)$ , by using the Definition of  $v_t^\varepsilon$  we obtain:

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \varphi(x) v_t^\varepsilon(x) dx - \int_{\mathbb{R}^d} \varphi(x) d\rho_t^\varepsilon(x) \right| &= \left| \int_{\mathbb{R}^d} \varphi(x) (V_\varepsilon * \rho_t^\varepsilon)(x) dx - \int_{\mathbb{R}^d} \varphi(x) d\rho_t^\varepsilon(x) \right| \\ &= \left| \int_{\mathbb{R}^d} (\varphi * V_\varepsilon)(x) d\rho_t^\varepsilon(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_t^\varepsilon(x) \right| \\ &= \left| \int_{\mathbb{R}^d} [(\varphi * V_\varepsilon)(x) - \varphi(x)] d\rho_t^\varepsilon(x) \right| \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\varphi(x - y) - \varphi(x)| V_\varepsilon(y) dy d\rho_t^\varepsilon(x) \\ &\leq \|\varphi\|_\infty \int_{\mathbb{R}^d} |y| V_\varepsilon(y) dy \\ &= \varepsilon \|\varphi\|_\infty \int_{\mathbb{R}^d} |x| V_1(x) dx, \end{aligned}$$

which converges to 0 as  $\varepsilon \rightarrow 0^+$  since  $\int_{\mathbb{R}^d} |x| V_1(x) dx < +\infty$ .  $\square$

We have now all the information to prove our first main result.

**Theorem 5.1.** Let  $\rho_0 \in \mathcal{P}_2^a(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  such that  $\mathcal{H}[\rho_0] < \infty$ . The sequence  $\{\rho^\varepsilon\}_{0 < \varepsilon \leq 1}$  of solutions to (NLIE) narrowly converges to the unique weak solution  $\tilde{\rho}$  of (PME). as  $\varepsilon \rightarrow 0$

**Proof.** Since  $\rho^\varepsilon$  is a weak solution to (NLIE), for any  $\varphi \in C_c^1(\mathbb{R}^d)$  and  $t \in [0, T]$  it satisfies

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x) d\rho_t^\varepsilon(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) &= - \int_0^t \int_{\mathbb{R}^d} v_t^\varepsilon(x) \nabla \varphi(x) \cdot \nabla v_t^\varepsilon(x) dx dt \\ &\quad - \int_0^t \int_{\mathbb{R}^d} z_t^\varepsilon(x) \cdot \nabla v_t^\varepsilon(x) dx dt, \end{aligned}$$

as explained in (27) and (28). In view of Proposition 4.1, Lemmas 4.1, 5.3, and Proposition 4.3, we know there exists a subsequence of  $\rho^\varepsilon(t)$  narrowly converging to  $\tilde{\rho} \in L^2([0, T]; H^1(\mathbb{R}^d))$ , and, in particular,  $\{v^\varepsilon\}_\varepsilon$  admits a subsequence such that

$$\begin{aligned} v^{\varepsilon_k} &\rightarrow \tilde{\rho} && \text{in } L^2([0, T]; L^2(\mathbb{R}^d)); \\ \nabla v^{\varepsilon_k} &\rightarrow \nabla \tilde{\rho} && \text{in } L^2([0, T]; L^2(\mathbb{R}^d)). \end{aligned}$$

Before letting  $\varepsilon \rightarrow 0^+$  and obtaining the result we need to further regularise the test function,  $\varphi$ , since Corollary 5.1 holds for test functions in  $C_c^2(\mathbb{R}^d)$ . In this regard, we consider a standard mollifier  $\eta \in C_c^\infty(\mathbb{R}^d)$  and the corresponding sequence  $\varphi^\sigma := \eta^\sigma * \varphi \in C_c^\infty(\mathbb{R}^d)$ , being  $\eta^\sigma(x) = \sigma^{-d} \eta(x/\sigma^d)$  for any  $x \in \mathbb{R}^d$  and

$\sigma > 0$ . As consequence of the observations above and [Corollary 5.1](#), by letting  $\varepsilon \rightarrow 0^+$  we obtain, for any  $\sigma > 0$  and  $t \in [0, T]$ ,

$$\int_{\mathbb{R}^d} \varphi^\sigma(x) \tilde{\rho}(t, x) dx = \int_{\mathbb{R}^d} \varphi^\sigma(x) \rho_0(x) dx - \int_0^t \int_{\mathbb{R}^d} \tilde{\rho}(s, x) \nabla \varphi^\sigma(x) \cdot \nabla \tilde{\rho}(s, x) dx ds.$$

Since  $\varphi^\sigma$  converges uniformly to  $\varphi$  on compact sets, we can let  $\sigma \rightarrow 0$  and obtain that  $\tilde{\rho}$  is a weak solution to [\(PME\)](#) in the sense of [Definition 2.1](#). Uniqueness of weak solutions of [\(PME\)](#) is a known result, cf. e.g. [\[14,46\]](#). By a standard contradiction argument we can prove the whole sequence  $\rho^\varepsilon$  narrowly converges to  $\tilde{\rho}$ , by further using uniqueness of weak solutions.  $\square$

### 6. Extension to cross-diffusion systems

A key motivation for our approach is that it does not rely on geodesic convexity of the associated energy, hence it can be extended to scenarios when convexity fails. This is indeed the case of cross-diffusion systems, as shown in [\[2\]](#). In what follows we extend the previous analysis to cross-diffusion systems, taking into account related works in literature, such as [\[18,48\]](#). More precisely, we derive the class of systems

$$\begin{cases} \partial_t \rho_1 = \operatorname{div}(\rho_1 A_{1,1} \nabla \rho_1 + \rho_1 A_{1,2} \nabla \rho_2), \\ \partial_t \rho_2 = \operatorname{div}(\rho_2 A_{2,2} \nabla \rho_2 + \rho_2 A_{2,1} \nabla \rho_1), \end{cases} \tag{CDS}$$

which can be rewritten in matrix form as

$$\begin{pmatrix} \partial_t \rho_1 \\ \partial_t \rho_2 \end{pmatrix} = \operatorname{div} \left[ \begin{pmatrix} \rho_1 A_{1,1} & \rho_1 A_{1,2} \\ \rho_2 A_{2,1} & \rho_2 A_{2,2} \end{pmatrix} \begin{pmatrix} \nabla \rho_1 \\ \nabla \rho_2 \end{pmatrix} \right]. \tag{CDS-M}$$

Hereafter,  $\rho_1$  and  $\rho_2$  are two probability density accounting for *two population species*; we consider two species for simplicity, though the results hold true for  $M$  species,  $M \in \mathbb{N}$ . The coefficients  $A_{i,j}$ , for  $i, j = 1, 2$ , are so that

$$\min\{A_{1,1}, A_{2,2}\} > \frac{A_{1,2} + A_{2,1}}{2} \geq 0. \tag{A}$$

These coefficients can be, e.g., constant second order derivatives of a function  $A$  depending on both species, i.e.  $A = A(\rho_1, \rho_2)$ . A prototype function  $A(\rho_1, \rho_2)$  we can consider is  $A(\rho_1, \rho_2) = (c_1 \rho_1 + c_2 \rho_2)^2 + c_3 \rho_1^2$ , for some constants  $c_i > 0$ ,  $i = 1, 2, 3$ . Note that cross-diffusion is present in the case  $A_{1,2} \neq 0$  and  $A_{2,1} \neq 0$ . In [\[18\]](#) the authors consider a general class of cross-diffusion systems, with the addition of nonlocal interaction terms, where  $A(\rho_1, \rho_2)$  is a nonlinear function modelling degenerate diffusion, for example of the form  $A(\rho_1, \rho_2) = \rho_1^{m_1} + \rho_2^{m_2} + p(\rho_1 + \rho_2)$ , being  $m_1, m_2 > 1$  and  $p$  regular enough. [\[18\]](#) provides existence of weak solutions by exploiting a semi-implicit version of the JKO scheme [\[29\]](#) in Wasserstein spaces.

Let us consider the functions  $V_i : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $U_{ij} : \mathbb{R}^d \rightarrow \mathbb{R}$  (or even a measure) for  $i, j = 1, 2$  and  $i \neq j$  such that  $V_i$  satisfy [\(V\)](#) and  $V_i \in W^{1,2}(\mathbb{R}^d)$ , while

[\(U\)](#)  $U_{ij} \in \mathcal{P}_1(\mathbb{R}^d)$  and it is even.

Let us define  $H_i := V_i * V_i$  and  $K_i := V_i * U_{ij} * V_j$  for  $i, j = 1, 2$  and  $i \neq j$ . For any  $\varepsilon > 0$ , consider the scaling  $H_i^\varepsilon(x) = \varepsilon^{-d} H_i(\frac{x}{\varepsilon})$  and  $K_i^\varepsilon(x) = \varepsilon^{-d} K_i(\frac{x}{\varepsilon})$ , whence  $H_i^\varepsilon = V_i^\varepsilon * V_i^\varepsilon$  and  $K_i^\varepsilon = V_i^\varepsilon * U_{ij}^\varepsilon * V_j^\varepsilon$ . Our goal is to show that, as  $\varepsilon \rightarrow 0^+$ , weak-measure solutions  $(\rho_1^\varepsilon, \rho_2^\varepsilon)$  to

$$\begin{cases} \partial_t \rho_1 = \operatorname{div}(\rho_1 A_{1,1} \nabla H_1^\varepsilon * \rho_1 + \rho_1 A_{1,2} \nabla K_1^\varepsilon * \rho_2) \\ \partial_t \rho_2 = \operatorname{div}(\rho_2 A_{2,2} \nabla H_2^\varepsilon * \rho_2 + \rho_2 A_{2,1} \nabla K_2^\varepsilon * \rho_1). \end{cases} \tag{NLIS}$$

converge to weak solutions to system [\(CDS\)](#). In order to achieve such a goal we will use the smoother version of  $(\rho_1^\varepsilon, \rho_2^\varepsilon)$  given by  $(v_1^\varepsilon, v_2^\varepsilon) = (V_1^\varepsilon * \rho_1^\varepsilon, V_2^\varepsilon * \rho_2^\varepsilon)$ . In [\(NLIS\)](#), the kernels  $H_i$  are the so-called self-interaction potentials since they model intra-specific interaction (among same species), whereas  $K_i$  are known as cross-interaction potentials as they take into account inter-specific interaction (among different species).

**Remark 6.1.** The kernels  $U_{ij}$  have been introduced to show a possible generalisation of this method, as this does not add other technical difficulties. Indeed, our assumptions allow to use  $U_{ij}$  as a Dirac delta — this is equivalent to excluding  $U_{ij}$ . However, we prefer to include a further regularisation as in other application it may be worth to introduce it in order to gain more regularity if  $U_{ij}$  is absolutely continuous with respect to the Lebesgue measure.

Below we state the Definitions of solutions used in this section.

**Definition 6.1** (*Weak Solution To (CDS)*). A weak solution to the cross-diffusion system (CDS) on the time interval  $[0, T]$  with initial datum  $\rho^0 \in (\mathcal{P}_2^g(\mathbb{R}) \cap L^2(\mathbb{R}^d))^2$  is a curve  $\rho \in C([0, T]; \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d))$  satisfying the following properties:

1. for almost every  $t \in [0, T]$  the measure  $\rho(t)$  has a density with respect to the Lebesgue measure, still denoted by  $\rho(t)$ , and  $\rho \in L^2([0, T]; H^1(\mathbb{R}^d)) \times L^2([0, T]; H^1(\mathbb{R}^d))$ ;
2. for any  $\varphi, \phi \in C_c^1(\mathbb{R}^d)$  and all  $t \in [0, T]$  it holds

$$\int_{\mathbb{R}^d} \varphi(x)\rho_1(t, x) dx = \int_{\mathbb{R}^d} \varphi(x)\rho_1^0(x) dx - A_{1,1} \int_0^t \int_{\mathbb{R}^d} \rho_1(s, x)\nabla\varphi(x) \cdot \nabla\rho_1(s, x) dx ds - A_{1,2} \int_0^t \int_{\mathbb{R}^d} \rho_1(s, x)\nabla\varphi(x) \cdot \nabla\rho_2(s, x) dx ds,$$

$$\int_{\mathbb{R}^d} \phi(x)\rho_2(t, x) dx = \int_{\mathbb{R}^d} \phi(x)\rho_2^0(x) dx - A_{2,2} \int_0^t \int_{\mathbb{R}^d} \rho_2(s, x)\nabla\phi(x) \cdot \nabla\rho_2(s, x) dx ds - A_{2,1} \int_0^t \int_{\mathbb{R}^d} \rho_2(s, x)\nabla\phi(x) \cdot \nabla\rho_1(s, x) dx ds.$$

**Definition 6.2** (*Weak Measure Solution To (NLIS)*). A narrowly continuous curve  $\rho^\varepsilon : [0, T] \rightarrow \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d)$ , mapping  $t \in [0, T] \mapsto \rho_t^\varepsilon \in \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d)$ , is a weak measure solution to (NLIS) if, for every  $\varphi, \phi \in C_c^1(\mathbb{R}^d)$  and any  $t \in [0, T]$ , it holds

$$\int_{\mathbb{R}^d} \varphi(x)d(\rho_{1,t}^\varepsilon - \rho_1^0)(x) = -\frac{A_{1,1}}{2} \int_0^t \iint_{\mathbb{R}^{2d}} (\nabla\varphi(x) - \nabla\varphi(y)) \cdot \nabla H_1^\varepsilon(x - y)d\rho_{1,r}^\varepsilon(y)d\rho_{1,r}^\varepsilon(x)dr - A_{1,2} \int_0^t \iint_{\mathbb{R}^{2d}} \nabla\varphi(x) \cdot \nabla K_1^\varepsilon(x - y)d\rho_{2,r}^\varepsilon(y)d\rho_{1,r}^\varepsilon(x)dr \tag{30}$$

$$\int_{\mathbb{R}^d} \phi(x)d(\rho_{2,t}^\varepsilon - \rho_2^0)(x) = -\frac{A_{2,2}}{2} \int_0^t \iint_{\mathbb{R}^{2d}} (\nabla\phi(x) - \nabla\phi(y)) \cdot \nabla H_2^\varepsilon(x - y)d\rho_{2,r}^\varepsilon(y)d\rho_{2,r}^\varepsilon(x)dr - A_{2,1} \int_0^t \iint_{\mathbb{R}^{2d}} \nabla\phi(x) \cdot \nabla K_2^\varepsilon(x - y)d\rho_{1,r}^\varepsilon(y)d\rho_{2,r}^\varepsilon(x)dr \tag{31}$$

**Remark 6.2.** We denote elements of a product space by using bold symbols, e.g.

$$\rho = (\rho_1, \rho_2) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d), \quad \text{or} \quad \mathbf{x} = (x_1, x_2) \in \mathbb{R}^d \times \mathbb{R}^d.$$

The Wasserstein distance of order two in the product space is defined as follows

$$\mathcal{W}_2^2(\boldsymbol{\mu}, \boldsymbol{\nu}) = d_W^2(\mu_1, \nu_1) + d_W^2(\mu_2, \nu_2)$$

for all  $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d)$ .

### 6.1. Nonlocal interaction system

Following the reference paper for nonlocal interaction systems [20], we apply a *semi-implicit* version of the JKO scheme, [29], in order to obtain a priori estimates on solutions to (NLIS) and their smoothed version  $v_i^\varepsilon = V_i^\varepsilon * \rho_i^\varepsilon$ , for  $i = 1, 2$ . The semi-implicit JKO scheme allows to prove existence of solutions to a class of systems of nonlocal interaction PDEs that do not exhibit a gradient flow structure, which is indeed the case when the cross-interaction potentials are not proportional, i.e.  $K_1 \neq \alpha K_2$  for a positive  $\alpha$ , cf. [20] for further details.

Let  $\varepsilon > 0$  be fixed and finite. We consider an initial datum  $\rho^0 \in (\mathcal{P}_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d))^2$  and we introduce the *relative energy* functional  $\mathcal{F}_\varepsilon : \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$  defined as follows: let  $\nu \in \mathcal{P}_2(\mathbb{R}^d)^2$  be a fixed (time independent) measure, for all  $\mu \in \mathcal{P}_2(\mathbb{R}^d)^2$  we set

$$\begin{aligned} \mathcal{F}_\varepsilon[\mu|\nu] := & \frac{A_{1,1}}{2} \int_{\mathbb{R}^d} H_1^\varepsilon * \mu_1 d\mu_1 + A_{1,2} \int_{\mathbb{R}^d} K_1^\varepsilon * \nu_2 d\mu_1 \\ & + \frac{A_{2,2}}{2} \int_{\mathbb{R}^d} H_2^\varepsilon * \mu_2 d\mu_2 + A_{2,1} \int_{\mathbb{R}^d} K_2^\varepsilon * \nu_1 d\mu_2. \end{aligned} \tag{32}$$

The above functional is referred to as *relative energy* since it accounts for the energy at the state  $\mu$  given the state  $\nu$ , which only affects the cross-interaction part of the functional. The latter observation suggests to rewrite the functional  $\mathcal{F}_\varepsilon$  as sum of two contributions, i.e. “self” and “cross” interactions. This will not only simplify notations, but also single out the two parts in the semi-implicit JKO scheme. Let us set

$$\mathcal{H}_\varepsilon[\mu] := \frac{A_{1,1}}{2} \int_{\mathbb{R}^d} H_1^\varepsilon * \mu_1 d\mu_1 + \frac{A_{2,2}}{2} \int_{\mathbb{R}^d} H_2^\varepsilon * \mu_2 d\mu_2,$$

and

$$\mathcal{K}_\varepsilon[\mu|\nu] := A_{1,2} \int_{\mathbb{R}^d} K_1^\varepsilon * \nu_2 d\mu_1 + A_{2,1} \int_{\mathbb{R}^d} K_2^\varepsilon * \nu_1 d\mu_2.$$

Then, we can rewrite  $\mathcal{F}_\varepsilon$  as the sum of the previous functionals, i.e.

$$\mathcal{F}_\varepsilon[\mu|\nu] = \mathcal{H}_\varepsilon[\mu] + \mathcal{K}_\varepsilon[\mu|\nu].$$

Note that the part  $\mathcal{H}_\varepsilon[\mu]$  is treated implicitly in the JKO scheme as usual, whereas  $\mathcal{K}[\mu|\nu]$  contains terms that are treated explicitly.

**Remark 6.3.** Note that for  $\rho^0 \in (\mathcal{P}_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d))^2$  the relative energy at the initial datum is finite. More precisely, for any  $i, j = 1, 2$  we have

$$\begin{aligned} \int_{\mathbb{R}^d} (H_i^\varepsilon * \rho_i^0)(x) \rho_i^0(x) dx &= \int_{\mathbb{R}^d} |(V_i^\varepsilon * \rho_i^0)(x)|^2 dx = \|V_i^\varepsilon * \rho_i^0\|_{L^2(\mathbb{R}^d)}^2 \\ &\leq \|V_i^\varepsilon\|_{L^1(\mathbb{R}^d)}^2 \|\rho_i^0\|_{L^2(\mathbb{R}^d)}^2 = \|V_i\|_{L^1(\mathbb{R}^d)}^2 \|\rho_i^0\|_{L^2(\mathbb{R}^d)}^2 < \infty, \end{aligned} \tag{33}$$

which implies

$$\begin{aligned} \mathcal{H}_\varepsilon[\rho^0] &\leq \frac{A_{1,1}}{2} \|V_1\|_{L^1(\mathbb{R}^d)}^2 \|\rho_1^0\|_{L^2(\mathbb{R}^d)}^2 + \frac{A_{2,2}}{2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \|\rho_2^0\|_{L^2(\mathbb{R}^d)}^2 \\ &\leq \max \left\{ \frac{A_{1,1}}{2} \|V_1\|_{L^1(\mathbb{R}^d)}^2, \frac{A_{2,2}}{2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \right\} \|\rho^0\|_{L^2}^2, \end{aligned} \tag{34}$$

and

$$\begin{aligned} \int_{\mathbb{R}^d} (K_i^\varepsilon * \nu_j)(x) \rho_i^0(x) dx &= \int_{\mathbb{R}^d} (V_i^\varepsilon * \rho_i^0)(x) (U_{ij}^\varepsilon * (V_j^\varepsilon * \nu_j))(x) dx \\ &\leq \|V_i^\varepsilon * \rho_i^0\|_{L^2(\mathbb{R}^d)} \|U_{ij}^\varepsilon * (V_j^\varepsilon * \nu_j)\|_{L^2(\mathbb{R}^d)} \\ &\leq \|V_i^\varepsilon\|_{L^1(\mathbb{R}^d)} \|\rho_i^0\|_{L^2(\mathbb{R}^d)} \|V_j^\varepsilon\|_{L^2(\mathbb{R}^d)} \nu_j(\mathbb{R}^d) \\ &= \frac{1}{\varepsilon^{d/2}} \|V_i\|_{L^1(\mathbb{R}^d)} \|\rho_i^0\|_{L^2(\mathbb{R}^d)} \|V_j\|_{L^2(\mathbb{R}^d)} < +\infty. \end{aligned} \tag{35}$$

The sequence is defined via the semi-implicit JKO scheme:

- fix a time step  $\tau > 0$  such that  $\rho_\tau^0 := \rho^0$ ;
- for a given  $\rho_\tau^{\varepsilon,n} \in (\mathcal{P}_2(\mathbb{R}^d))^2$ , choose

$$\rho_\tau^{\varepsilon,n+1} \in \operatorname{argmin}_{\rho \in (\mathcal{P}_2(\mathbb{R}^d))^2} \left\{ \frac{\mathcal{W}_2^2(\rho_\tau^{\varepsilon,n}, \rho)}{2\tau} + \mathcal{F}_\varepsilon[\rho | \rho_\tau^{\varepsilon,n}] \right\}. \tag{36}$$

Let  $T > 0$ ,  $N := \lceil \frac{T}{\tau} \rceil$ , and consider the piecewise constant interpolation

$$\rho_\tau^\varepsilon(t) = \rho_\tau^{\varepsilon,n} \quad t \in ((n-1)\tau, n\tau],$$

being  $\rho_\tau^{\varepsilon,n} = (\rho_{1,\tau}^{\varepsilon,n}, \rho_{2,\tau}^{\varepsilon,n})$  defined in (36).

**Proposition 6.1** (*Narrow Compactness – Energy & Moments Bound*). *There exists an absolutely continuous curve  $\tilde{\rho}^\varepsilon : [0, T] \rightarrow \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d)$  such that the piecewise constant interpolation  $\rho_\tau^\varepsilon$  admits a subsequence  $\rho_{\tau_k}^\varepsilon$  narrowly converging to  $\tilde{\rho}^\varepsilon$  uniformly in  $t \in [0, T]$  as  $k \rightarrow +\infty$ . Moreover, for any  $t \in [0, T]$ , the following uniform bounds in  $\tau$  and  $\varepsilon$  hold*

$$\mathcal{H}_\varepsilon[\tilde{\rho}^\varepsilon(t)] \leq c \|\rho^0\|_{L^2}^2, \tag{37a}$$

$$m_2(\tilde{\rho}^\varepsilon(t)) \leq 2m_2(\rho^0) + \tilde{c}, \tag{37b}$$

where  $c = \max \left\{ \frac{A_{1,1}}{2} \|V_1\|_{L^1(\mathbb{R}^d)}^2, \frac{A_{2,2}}{2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \right\}$  and  $\tilde{c} = 4cT \|\rho^0\|_{L^2}^2$ .

**Proof.** From the Definition of the sequence  $\{\rho_\tau^{\varepsilon,n}\}_{n \in \mathbb{N}}$  it holds

$$\begin{aligned} & \frac{1}{2\tau} \mathcal{W}_2^2(\rho_\tau^{\varepsilon,n}, \rho_\tau^{\varepsilon,n+1}) \leq \mathcal{F}_\varepsilon[\rho_\tau^{\varepsilon,n} | \rho_\tau^{\varepsilon,n}] - \mathcal{F}_\varepsilon[\rho_\tau^{\varepsilon,n+1} | \rho_\tau^{\varepsilon,n}] \\ & = \sum_{i=1}^2 \frac{A_{i,i}}{2} \left( \int_{\mathbb{R}^d} H_i^\varepsilon * \rho_{i,\tau}^{\varepsilon,n} d\rho_{i,\tau}^{\varepsilon,n} - \int_{\mathbb{R}^d} H_i^\varepsilon * \rho_{i,\tau}^{\varepsilon,n+1} d\rho_{i,\tau}^{\varepsilon,n+1} \right) \\ & \quad + \sum_{i \neq j} A_{i,j} \left( \int_{\mathbb{R}^d} K_i^\varepsilon * \rho_{j,\tau}^{\varepsilon,n} d\rho_{i,\tau}^{\varepsilon,n} - \int_{\mathbb{R}^d} K_i^\varepsilon * \rho_{j,\tau}^{\varepsilon,n+1} d\rho_{i,\tau}^{\varepsilon,n+1} \right) \\ & = \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n}] - \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n+1}] + \sum_{i \neq j} A_{i,j} \left( \int_{\mathbb{R}^d} K_i^\varepsilon * \rho_{j,\tau}^{\varepsilon,n} d\rho_{i,\tau}^{\varepsilon,n} - \int_{\mathbb{R}^d} K_i^\varepsilon * \rho_{j,\tau}^{\varepsilon,n+1} d\rho_{i,\tau}^{\varepsilon,n+1} \right). \end{aligned} \tag{38}$$

In order to have an estimate on the energy, we make use of the  $L^\infty$  bound  $\|\nabla V_i^\varepsilon * V_j^\varepsilon\|_{L^\infty(\mathbb{R}^d)} \leq \varepsilon^{-(d+1)} \|\nabla V_i\|_{L^2(\mathbb{R}^d)} \|V_j\|_{L^2(\mathbb{R}^d)}$  as follows for  $i, j = 1, 2, i \neq j$ . First note that

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} K_i^\varepsilon * \rho_{j,\tau}^{\varepsilon,n} d\rho_{i,\tau}^{\varepsilon,n} - \int_{\mathbb{R}^d} K_i^\varepsilon * \rho_{j,\tau}^{\varepsilon,n+1} d\rho_{i,\tau}^{\varepsilon,n+1} \right| \\ & = \left| \iint_{\mathbb{R}^d \times \mathbb{R}^d} K_i^\varepsilon(x-y) d\rho_{j,\tau}^{\varepsilon,n}(y) d\rho_{i,\tau}^{\varepsilon,n}(x) - \iint_{\mathbb{R}^d \times \mathbb{R}^d} K_i^\varepsilon(t-y) d\rho_{j,\tau}^{\varepsilon,n}(y) d\rho_{i,\tau}^{\varepsilon,n+1}(t) \right| \\ & = \left| \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} (K_i^\varepsilon(x-y) - K_i^\varepsilon(t-y)) d\gamma_{i,\tau}^{\varepsilon,n}(x,t) d\rho_{j,\tau}^{\varepsilon,n}(y) \right| \end{aligned}$$

where  $\gamma_{i,\tau}^{\varepsilon,n} \in \Gamma_o(\rho_{i,\tau}^{\varepsilon,n}, \rho_{i,\tau}^{\varepsilon,n+1})$  is an optimal transport plan connecting  $\rho_{i,\tau}^{\varepsilon,n}$  to  $\rho_{i,\tau}^{\varepsilon,n+1}$ . Now, due to the  $L^\infty$  bound on  $\nabla V_i^\varepsilon * V_j^\varepsilon$  we get

$$\begin{aligned} & \left| \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} (K_i^\varepsilon(x-y) - K_i^\varepsilon(t-y)) d\gamma_{i,\tau}^{\varepsilon,n}(x,t) d\rho_{j,\tau}^{\varepsilon,n}(y) \right| \\ &= \left| \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} \left\{ \int_{\mathbb{R}^d} [(V_i^\varepsilon * V_j^\varepsilon)(x-y-z) - (V_i^\varepsilon * V_j^\varepsilon)(t-y-z)] U_{ij}^\varepsilon(z) \right\} d\gamma_{i,\tau}^{\varepsilon,n}(x,t) d\rho_{j,\tau}^{\varepsilon,n}(y) \right| \\ &\leq \|\nabla V_i^\varepsilon * V_j^\varepsilon\|_{L^\infty(\mathbb{R}^d)} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x-t| d\gamma_{i,\tau}^{\varepsilon,n}(x,t) \\ &\leq \frac{1}{\varepsilon^{d+1}} \|\nabla V_i\|_{L^2(\mathbb{R}^d)} \|V_j\|_{L^2(\mathbb{R}^d)} dW(\rho_{i,\tau}^{\varepsilon,n}, \rho_{i,\tau}^{\varepsilon,n+1}) \leq \frac{1}{4\tau} d_W^2(\rho_{i,\tau}^{\varepsilon,n}, \rho_{i,\tau}^{\varepsilon,n+1}) + C \frac{\tau}{\varepsilon^{2(d+1)}}, \end{aligned}$$

where  $C$  is a positive constant independent of  $\tau$ . By using the latter estimate in (38) we obtain

$$\frac{1}{4\tau} \mathcal{W}_2^2(\rho_\tau^{\varepsilon,n}, \rho_\tau^{\varepsilon,n+1}) \leq \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n}] - \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n+1}] + C \frac{\tau}{\varepsilon^{2(d+1)}}, \tag{39}$$

which implies  $\mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n+1}] \leq \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n}] + C \frac{\tau}{\varepsilon^{2(d+1)}}$ , and, in particular, the following bound for the self-interaction part,  $\mathcal{H}_\varepsilon$ , of the relative energy  $\mathcal{F}_\varepsilon$ :

$$\mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n}] \leq \mathcal{H}_\varepsilon[\rho^0] + C \frac{n\tau}{\varepsilon^{2(d+1)}} \leq \mathcal{H}_\varepsilon[\rho^0] + C \frac{T}{\varepsilon^{2(d+1)}}, \quad \forall n \in \mathbb{N}. \tag{40}$$

By summing up over  $k$  inequality (39), we obtain

$$\sum_{k=m}^n \frac{\mathcal{W}_2^2(\rho_\tau^{\varepsilon,k}, \rho_\tau^{\varepsilon,k+1})}{4\tau} \leq \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,m}] - \mathcal{H}_\varepsilon[\rho_\tau^{\varepsilon,n+1}] + C \frac{\tau}{\varepsilon^{2(d+1)}}(n-m+1). \tag{41}$$

The non-negativity of  $\mathcal{H}_\varepsilon$  and the energy inequality (40) allow us to improve the above inequality as follows

$$\sum_{k=m}^n \frac{\mathcal{W}_2^2(\rho_\tau^{\varepsilon,k}, \rho_\tau^{\varepsilon,k+1})}{4\tau} \leq \mathcal{H}_\varepsilon[\rho^0] + C \frac{T}{\varepsilon^{2(d+1)}} + C \frac{\tau}{\varepsilon^{2(d+1)}}(n-m+1). \tag{42}$$

In particular, by using the bound (34) in Remark 6.3, the above inequality implies

$$\begin{aligned} \mathcal{W}_2^2(\rho^0, \rho_\tau^\varepsilon(t)) &\leq 4T\mathcal{H}_\varepsilon[\rho^0] + C \frac{T^2}{\varepsilon^{2(d+1)}} \\ &\leq 2T \max \left\{ A_{1,1} \|V_1\|_{L^1(\mathbb{R}^d)}^2, A_{2,2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \right\} \|\rho^0\|_{L^2}^2 + C \frac{T^2}{\varepsilon^{2(d+1)}}, \end{aligned}$$

whence we obtain the second order moments are uniformly bounded in  $\tau$  on  $[0, T]$  in view of Remark 2.2, i.e.

$$\begin{aligned} m_2(\rho_\tau^\varepsilon(t)) &\leq 2m_2(\rho^0) + 2\mathcal{W}_2^2(\rho^0, \rho_\tau^\varepsilon(t)) \\ &\leq 2m_2(\rho^0) + 2T \max \left\{ A_{1,1} \|V_1\|_{L^1(\mathbb{R}^d)}^2, A_{2,2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \right\} \|\rho^0\|_{L^2}^2 + C \frac{T^2}{\varepsilon^{2(d+1)}}. \end{aligned} \tag{43}$$

Now, let us consider  $0 \leq s < t$  such that  $s \in ((m-1)\tau, m\tau)$  and  $t \in ((n-1)\tau, n\tau)$  (which implies  $|n-m| < \frac{|t-s|}{\tau} + 1$ ); by the Cauchy-Schwarz inequality, (42) and again the bound (34) in Remark 6.3,



we obtain

$$\begin{aligned} \mathcal{W}_2(\rho_\tau^\varepsilon(s), \rho_\tau^\varepsilon(t)) &\leq \sum_{k=m}^{n-1} \mathcal{W}_2(\rho_\tau^{\varepsilon,k}, \rho_\tau^{\varepsilon,k+1}) \leq \left( \sum_{k=m}^{n-1} \mathcal{W}_2^2(\rho_\tau^{\varepsilon,k}, \rho_\tau^{\varepsilon,k+1}) \right)^{\frac{1}{2}} |n - m|^{\frac{1}{2}} \\ &\leq c \left( \sqrt{1 + \frac{T}{\varepsilon^{2(d+1)}}} \right) \left( \sqrt{|t - s|} + \sqrt{\tau} \right), \end{aligned} \tag{44}$$

where  $c = c(A_{1,1}, A_{2,2}, \rho^0, T)$  is a positive constant. Thus  $\rho_\tau^\varepsilon$  is  $\frac{1}{2}$ -Hölder equi-continuous, up to a negligible error of order  $\sqrt{\tau}$ . By using a refined version of Ascoli–Arzelà’s theorem (see [1], Section 3), we obtain  $\rho_\tau^\varepsilon$  admits a subsequence narrowly converging to a limit  $\tilde{\rho}^\varepsilon$  as  $\tau \rightarrow 0^+$  uniformly on  $[0, T]$ . Since  $|\cdot|^2$  and  $H_i^\varepsilon$  are lower semi-continuous and bounded from below, a refined analysis gives for any  $t \in [0, T]$

$$\mathcal{H}_\varepsilon[\tilde{\rho}^\varepsilon] \leq \liminf_{k \rightarrow +\infty} \mathcal{H}_\varepsilon[\rho_{\tau_k}^\varepsilon] \leq \mathcal{H}_\varepsilon[\rho^0] \leq \max \left\{ \frac{A_{1,1}}{2} \|V_1\|_{L^1(\mathbb{R}^d)}^2, \frac{A_{2,2}}{2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \right\} \|\rho^0\|_{L^2}^2,$$

and

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^2 d\tilde{\rho}_i^\varepsilon(t)(x) &\leq \liminf_{k \rightarrow +\infty} \int_{\mathbb{R}^d} |x|^2 d\rho_{\tau_k,i}^\varepsilon(t)(x) \\ &\leq 2m_2(\rho^0) + 2T \max \left\{ A_{1,1} \|V_1\|_{L^1(\mathbb{R}^d)}^2, A_{2,2} \|V_2\|_{L^1(\mathbb{R}^d)}^2 \right\} \|\rho^0\|_{L^2}^2 \end{aligned}$$

whence the assertion follows by applying the above inequalities to (40) and (43).  $\square$

As direct consequence of the previous Lemma, we actually have narrow compactness in  $\varepsilon$  by following the same argument used in Proposition 4.1; hence we omit the proof.

**Corollary 6.1.** *There exists an absolutely continuous curve  $\tilde{\rho} : [0, T] \rightarrow \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d)$  such that  $\tilde{\rho}^\varepsilon$  admits a subsequence  $\tilde{\rho}^{\varepsilon_k}$  narrowly converging to  $\tilde{\rho}$  uniformly in  $t \in [0, T]$  as  $k \rightarrow +\infty$ .*

As in Section 3, the limiting curve  $\tilde{\rho}^\varepsilon$  obtained in Proposition 6.1 is a weak measure solution of the nonlocal interaction system (NLIS) in the sense of Definition 6.2. The proof of this result is similar to 3.1, and more details can be found in [18, Theorem 3.3]. For this reason we omit the proof. Let us stress that the cross-interaction potentials,  $K_i^\varepsilon$  are  $C^1(\mathbb{R}^d)$  since continuity at zero is obtained using Lebesgue dominated convergence theorem, exploiting that  $U_{ij} \in \mathcal{P}_1(\mathbb{R}^d)$  and  $\nabla V_i^\varepsilon \in L^1(\mathbb{R}^d)$ . This is needed as we cannot cope with a possible discontinuity of  $\nabla K_i^\varepsilon$  at zero, as for the self-interaction kernels, cf. [18,20].

**Theorem 6.1.** *The curve  $\tilde{\rho}^\varepsilon$  is a weak measure solution to the system (NLIS) according to Definition 6.2.*

### 6.2. Compactness in $\varepsilon$

As in the one-species case, stronger convergence can be obtained for the sequence  $\{v^\varepsilon\}_\varepsilon$ , being  $v_i^\varepsilon(t) := V_i^\varepsilon * \rho_i^\varepsilon(t)$  for any  $t \in [0, T]$ . We use the flow interchange technique by considering a decoupled system of heat equations as auxiliary flow, i.e.

$$\begin{cases} \partial_t \eta_1 = \Delta \eta_1 \\ \partial_t \eta_2 = \Delta \eta_2, \end{cases} \tag{45}$$

and the entropy as auxiliary functional, that is

$$E[\eta_1, \eta_2] = \begin{cases} \int_{\mathbb{R}^d} [\eta_1(x) \log \eta_1(x) + \eta_2(x) \log \eta_2(x)] dx, & \eta_1 \log \eta_1, \eta_2 \log \eta_2 \in L^1(\mathbb{R}^d); \\ +\infty & \text{otherwise.} \end{cases} \tag{46}$$

For any  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$  such that  $E(\nu) < +\infty$ , we denote by  $\mathbf{S}_E^t \nu := (S_{E,1}^t \nu_1, S_{E,2}^t \nu_2)$  the solution at time  $t$  to system (45) coupled with an initial value  $\nu$  at  $t = 0$ . Moreover, for every  $\rho \in (\mathcal{P}_2(\mathbb{R}^d))^2$  and a given  $\mu \in (\mathcal{P}_2(\mathbb{R}^d))^2$ , the dissipation of  $\mathcal{F}_\varepsilon$  along  $\mathbf{S}_E$  by

$$D_E \mathcal{F}_\varepsilon(\rho|\mu) := \limsup_{s \downarrow 0} \left\{ \frac{\mathcal{F}_\varepsilon[\rho|\mu] - \mathcal{F}_\varepsilon[\mathbf{S}_E^s \rho|\mu]}{s} \right\}.$$

Below we prove an  $L_t^2 H_x^1$  bound crucial for the application of the Rossi–Savaré version of the Aubin–Lions Lemma.

**Lemma 6.1.** *Let  $\rho^0 \in (\mathcal{P}_2^a(\mathbb{R}^d) \cap L^2(\mathbb{R}^d))^2$  such that  $E[\rho^0] < \infty$ . There exists a constant  $C = C(\rho^0, A_{1,1}, A_{2,2}, V_1, V_2, T)$  such that, for any  $\varepsilon > 0$ ,*

$$\sum_{i=1}^2 \|v_i^\varepsilon\|_{L^2([0,T]; H^1(\mathbb{R}^d))}^2 \leq C. \tag{47}$$

**Proof.** First of all we obtain a uniform  $L^2$  bound in time and space from Proposition 6.1 by noticing that

$$\begin{aligned} \sum_{i=1}^2 A_{i,i} \|V_i^\varepsilon * \rho_{i,\tau}^\varepsilon\|_{L^2([0,T]; L^2(\mathbb{R}^d))}^2 &= \sum_{i=1}^2 A_{i,i} \int_0^T \int_{\mathbb{R}^d} |[V_i^\varepsilon * \rho_{i,\tau}^\varepsilon(t)](x)|^2 dx \\ &= \sum_{i=1}^2 A_{i,i} \int_0^T \int_{\mathbb{R}^d} [H_i^\varepsilon * \rho_{i,\tau}^\varepsilon(t)](x) d\rho_{i,\tau}^\varepsilon(t)(x) \\ &= 2 \int_0^T \mathcal{H}_\varepsilon[\rho_\tau^\varepsilon(t)] dt \leq C(\rho^0, A_{1,1}, A_{2,2}, T). \end{aligned} \tag{48}$$

Since  $A_{i,i} \neq 0$  for any  $i$ , we can divide and get the  $L^2$  bound, arguing as in Lemma 4.1. Let us now focus on the bound for the gradient. For all  $s > 0$ , using the minimising property of  $\rho_\tau^{\varepsilon,n+1}$  in (36) we have

$$\frac{1}{2\tau} \mathcal{W}_2^2(\rho_\tau^{\varepsilon,n+1}, \rho_\tau^{\varepsilon,n}) + \mathcal{F}_\varepsilon[\rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}] \leq \frac{1}{2\tau} \mathcal{W}_2^2(\mathbf{S}_E^s \rho_\tau^{\varepsilon,n+1}, \rho_\tau^{\varepsilon,n}) + \mathcal{F}_\varepsilon[\mathbf{S}_E^s \rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}],$$

whence, dividing by  $s > 0$  and passing to the lim sup as  $s \downarrow 0$ ,

$$\tau D_E \mathcal{F}_\varepsilon[\rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}] \leq \frac{1}{2} \frac{d^+}{dt} \left( \mathcal{W}_2^2(\mathbf{S}_E^t \rho_\tau^{\varepsilon,n+1}, \rho_\tau^{\varepsilon,n}) \right) \Big|_{t=0} \stackrel{(\mathbf{E.V.I.})}{\leq} E[\rho_\tau^{\varepsilon,n}] - E[\rho_\tau^{\varepsilon,n+1}]. \tag{49}$$

In the last inequality we used that  $\mathbf{S}_E$  is a 0-flow. Focusing on the left hand side of (49), we note

$$\begin{aligned} D_E \mathcal{F}_\varepsilon[\rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}] &= \limsup_{s \downarrow 0} \left\{ \frac{\mathcal{F}_\varepsilon[\rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}] - \mathcal{F}_\varepsilon[\mathbf{S}_E^s \rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}]}{s} \right\} \\ &= \limsup_{s \downarrow 0} \int_0^1 \left( -\frac{d}{dz} \Big|_{z=st} \mathcal{F}_\varepsilon[\mathbf{S}_E^z \rho_\tau^{\varepsilon,n+1}|\rho_\tau^{\varepsilon,n}] \right) dt. \end{aligned} \tag{50}$$

Thus, we now compute the time derivative inside the above integral, by using integration by parts and keeping in mind the  $C^\infty$  regularity of the solution to the heat equation:

$$\begin{aligned}
 \frac{d}{dt} \mathcal{F}_\varepsilon[\mathbf{S}_E^t \boldsymbol{\rho}_\tau^{\varepsilon, n+1} | \boldsymbol{\rho}_\tau^{\varepsilon, n}] &= -A_{1,1} \int_{\mathbb{R}^d} \nabla(H_1^\varepsilon * S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1})(x) \nabla S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1}(x) dx \\
 &\quad - A_{2,2} \int_{\mathbb{R}^d} \nabla(H_2^\varepsilon * S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1})(x) \nabla S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1}(x) dx \\
 &\quad - A_{1,2} \int_{\mathbb{R}^d} \nabla(K_1^\varepsilon * \rho_{2,\tau}^{\varepsilon, n})(x) \nabla S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1}(x) dx \\
 &\quad - A_{2,1} \int_{\mathbb{R}^d} \nabla(K_2^\varepsilon * \rho_{1,\tau}^{\varepsilon, n})(x) \nabla S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1}(x) dx \\
 &= -A_{1,1} \int_{\mathbb{R}^d} |\nabla V_1^\varepsilon * S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\quad - A_{2,2} \int_{\mathbb{R}^d} |\nabla V_2^\varepsilon * S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\quad - A_{1,2} \int_{\mathbb{R}^d} \nabla(V_1^\varepsilon * S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1})(x) [\nabla(U_{12}^\varepsilon * V_2^\varepsilon * \rho_{2,\tau}^{\varepsilon, n})](x) dx \\
 &\quad - A_{2,1} \int_{\mathbb{R}^d} \nabla(V_2^\varepsilon * S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1})(x) [\nabla(U_{21}^\varepsilon * V_1^\varepsilon * \rho_{1,\tau}^{\varepsilon, n})](x) dx \\
 &\leq -A_{1,1} \int_{\mathbb{R}^d} |\nabla V_1^\varepsilon * S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\quad - A_{2,2} \int_{\mathbb{R}^d} |\nabla V_2^\varepsilon * S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\quad + A_{1,2} \|\nabla(V_1^\varepsilon * S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1})\|_{L^2(\mathbb{R}^d)} \|\nabla(U_{12}^\varepsilon * V_2^\varepsilon * \rho_{2,\tau}^{\varepsilon, n})\|_{L^2(\mathbb{R}^d)} \\
 &\quad + A_{2,1} \|\nabla(V_2^\varepsilon * S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1})\|_{L^2(\mathbb{R}^d)} \|\nabla(U_{21}^\varepsilon * V_1^\varepsilon * \rho_{1,\tau}^{\varepsilon, n})\|_{L^2(\mathbb{R}^d)} \\
 &\leq -\left(A_{1,1} - \frac{A_{1,2}}{2}\right) \int_{\mathbb{R}^d} |\nabla V_1^\varepsilon * S_{E,1}^t \rho_{1,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\quad - \left(A_{2,2} - \frac{A_{2,1}}{2}\right) \int_{\mathbb{R}^d} |\nabla V_2^\varepsilon * S_{E,2}^t \rho_{2,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\quad + \frac{A_{1,2}}{2} \int_{\mathbb{R}^d} |\nabla(U_{12}^\varepsilon * V_2^\varepsilon * \rho_{2,\tau}^{\varepsilon, n})(x)|^2 dx \\
 &\quad + \frac{A_{2,1}}{2} \int_{\mathbb{R}^d} |\nabla(U_{21}^\varepsilon * V_1^\varepsilon * \rho_{1,\tau}^{\varepsilon, n})(x)|^2 dx.
 \end{aligned}$$

The above inequality, together with (49) and (50), implies

$$\begin{aligned}
 \tau \liminf_{s \downarrow 0} \int_0^1 \int_{\mathbb{R}^d} &\left(A_{1,1} - \frac{A_{1,2}}{2}\right) |\nabla V_1^\varepsilon * S_{E,1}^{st} \rho_{1,\tau}^{\varepsilon, n+1}(x)|^2 + \left(A_{2,2} - \frac{A_{2,1}}{2}\right) |\nabla V_2^\varepsilon * S_{E,2}^{st} \rho_{2,\tau}^{\varepsilon, n+1}(x)|^2 dx dt \\
 &\leq \tau \frac{A_{1,2}}{2} \int_{\mathbb{R}^d} |\nabla(U_{12}^\varepsilon * V_2^\varepsilon * \rho_{2,\tau}^{\varepsilon, n})(x)|^2 dx \\
 &\quad + \tau \frac{A_{2,1}}{2} \int_{\mathbb{R}^d} |\nabla(U_{21}^\varepsilon * V_1^\varepsilon * \rho_{1,\tau}^{\varepsilon, n})(x)|^2 dx + \mathbb{E}[\boldsymbol{\rho}_\tau^{\varepsilon, n}] - \mathbb{E}[\boldsymbol{\rho}_\tau^{\varepsilon, n+1}],
 \end{aligned}$$

thus, by  $L^2$  lower semi-continuity of the  $H^1$  seminorm,

$$\begin{aligned}
 \tau \int_{\mathbb{R}^d} &\left(A_{1,1} - \frac{A_{1,2}}{2}\right) |\nabla V_1^\varepsilon * \rho_{1,\tau}^{\varepsilon, n+1}(x)|^2 + \left(A_{2,2} - \frac{A_{2,1}}{2}\right) |\nabla V_2^\varepsilon * \rho_{2,\tau}^{\varepsilon, n+1}(x)|^2 dx \\
 &\leq \tau \frac{A_{1,2}}{2} \int_{\mathbb{R}^d} |\nabla(U_{12}^\varepsilon * V_2^\varepsilon) * \rho_{2,\tau}^{\varepsilon, n}(x)|^2 dx \\
 &\quad + \tau \frac{A_{2,1}}{2} \int_{\mathbb{R}^d} |\nabla(U_{21}^\varepsilon * V_1^\varepsilon) * \rho_{1,\tau}^{\varepsilon, n}(x)|^2 dx + \mathbb{E}[\boldsymbol{\rho}_\tau^{\varepsilon, n}] - \mathbb{E}[\boldsymbol{\rho}_\tau^{\varepsilon, n+1}],
 \end{aligned}$$

By summing up over  $n$  from 0 to  $N - 1$ , taking into account that  $x \log x \leq x^2$  for any  $x \geq 0$ , Remark 4.1 and that second order moments are uniformly bounded (see Proposition 6.1), we get

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} \left( A_{1,1} - \frac{A_{1,2}}{2} \right) |(\nabla V_1^\varepsilon * \rho_{1,\tau}^\varepsilon(t))(x)|^2 + \left( A_{2,2} - \frac{A_{2,1}}{2} \right) |(\nabla V_2^\varepsilon * \rho_{2,\tau}^\varepsilon(t))(x)|^2 dx \\ & \leq E[\rho^0] - E[\rho_\tau^{\varepsilon,N}] + \frac{A_{1,2}}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla(V_2^\varepsilon * \rho_{2,\tau}^\varepsilon(t)(x))|^2 dx dt \\ & \quad + \frac{A_{2,1}}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla(V_1^\varepsilon * \rho_{1,\tau}^\varepsilon(t))(x)|^2 dx dt \\ & \leq \|\rho^0\|_{L^2}^2 + C(\rho^0, V_1, V_2, A_{1,1}, A_{2,2}, T) + \frac{A_{1,2}}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla(V_2^\varepsilon * \rho_{2,\tau}^\varepsilon(t)(x))|^2 dx dt \\ & \quad + \frac{A_{2,1}}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla(V_1^\varepsilon * \rho_{1,\tau}^\varepsilon(t))(x)|^2 dx dt. \end{aligned}$$

Weak lower semi-continuity of the norm and (A) give the  $H^1$  bound

$$\begin{aligned} & \left( A_{1,1} - \frac{A_{1,2}}{2} - \frac{A_{2,1}}{2} \right) \int_0^T \int_{\mathbb{R}^d} |\nabla v_1^\varepsilon(t)(x)|^2 dx dt + \left( A_{2,2} - \frac{A_{1,2}}{2} - \frac{A_{2,1}}{2} \right) \int_0^T \int_{\mathbb{R}^d} |\nabla v_2^\varepsilon(t)(x)|^2 dx dt \quad \square \\ & \leq \|\rho^0\|_{L^2}^2 + C(\rho^0, V_1, V_2, A_{1,1}, A_{2,2}, T). \end{aligned}$$

An application of Proposition 4.2 provides strong convergence in  $L^2$ , needed to prove convergence to the cross-diffusion system (CDS), see Section 6.3.

**Proposition 6.2.** *Let  $\varepsilon \leq 1$ . There exists a subsequence  $\{v^{\varepsilon_k}\}_k$  strongly converging to  $v$  in  $L^2([0, T]; L^2(\mathbb{R}^d)) \times L^2([0, T]; L^2(\mathbb{R}^d))$ , for any  $T > 0$ .*

**Proof.** The proof is similar to the one of Proposition 4.3, see also [18], applied to  $X := L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ ,  $g = d_1$  being the 1-Wasserstein distance in the product space, and the functional  $\mathcal{B} : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \rightarrow [0, +\infty]$  defined as

$$\mathcal{B}[v] = \begin{cases} \sum_{i=1}^2 \|v_i\|_{H^1(\mathbb{R}^d)}^2 + \int_{\mathbb{R}^d} |x| v_i(x) dx, & \text{if } v_i \in \mathcal{P}_1(\mathbb{R}^d) \cap H^1(\mathbb{R}^d); \\ +\infty & \text{otherwise.} \end{cases} \quad \square$$

### 6.3. Towards cross-diffusion systems

Overall the strategy is similar to Section 5, though we have to clarify how to cope with the cross-interaction terms, leading to cross-diffusion. The right hand side in the Definition of weak measure solution of (NLIS), for the first component, can be written as

$$- A_{1,1} \int_0^T \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot \nabla V_1^\varepsilon * v_{1,t}^\varepsilon(x) d\rho_{1,t}^\varepsilon(x) dt - A_{1,2} \int_0^T \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot (V_1^\varepsilon * U_{12}^\varepsilon) * \nabla v_{2,t}^\varepsilon(x) d\rho_{1,t}^\varepsilon(x) dt. \quad (51)$$

While for the first integral above we can follow the argument in Section 5, applied of course to both components  $\rho_1^\varepsilon$  and  $\rho_2^\varepsilon$ , the cross-interaction part needs a slightly different approach, since  $K_1$  is a convolution of three functions. In particular,

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot (V_1^\varepsilon * U_{12}^\varepsilon) * \nabla v_{2,t}^\varepsilon d\rho_{1,t}^\varepsilon(x) dt &= \int_0^T \int_{\mathbb{R}^d} (U_{12}^\varepsilon * v_{1,t}^\varepsilon)(x) \nabla \varphi(x) \cdot \nabla v_{2,t}^\varepsilon(x) dx dt \\ & \quad + \int_0^T \int_{\mathbb{R}^d} z_{12,t}^\varepsilon(x) \cdot \nabla v_{2,t}^\varepsilon(x) dx dt, \end{aligned} \quad (52)$$

being, for any  $t \in [0, T]$  and  $x \in \mathbb{R}^d$ ,

$$z_{12,t}^\varepsilon(x) := (U_{12}^\varepsilon * V_1^\varepsilon) * (\rho_{1,t}^\varepsilon \nabla \varphi)(x) - (U_{12}^\varepsilon * v_{1,t}^\varepsilon)(x) \nabla \varphi(x). \tag{53}$$

The excess term converges to 0 strongly in  $L^2([0, T] \times \mathbb{R}^d)$  by following Section 5.1, applying the arguments to  $P_{12}^\varepsilon := V_1^\varepsilon * U_{12}^\varepsilon$  instead of  $V_\varepsilon$  — bearing in mind (V) and (U). For the readers convenience we remind that the  $L^2([0, T]; H^1(\mathbb{R}^d))$  bound for  $v_i^\varepsilon$  holds true for  $U_{ij}^\varepsilon * v_i^\varepsilon$  since, for  $i \neq j = 1, 2$ ,

$$\begin{aligned} \|U_{ij}^\varepsilon * v_{i,t}^\varepsilon\|_{L^2(\mathbb{R}^d)} &\leq \|v_{i,t}^\varepsilon\|_{L^2(\mathbb{R}^d)} = \|v_{i,t}^\varepsilon\|_{L^2(\mathbb{R}^d)}; \\ \|U_{ij}^\varepsilon * \nabla v_{i,t}^\varepsilon\|_{L^2(\mathbb{R}^d)} &\leq \|\nabla v_{i,t}^\varepsilon\|_{L^2(\mathbb{R}^d)} = \|\nabla v_{i,t}^\varepsilon\|_{L^2(\mathbb{R}^d)}. \end{aligned} \tag{54}$$

As in Lemma 5.3, one can prove that the sequence  $U_{ij}^\varepsilon * v_i^\varepsilon$  has the same distributional limit of the sequence  $v_i^\varepsilon$ , i.e.  $\rho_i$ , for  $i \neq j = 1, 2$ .

**Lemma 6.2.** For any  $t \in [0, T]$  and any  $\varphi \in C_c^1(\mathbb{R}^d)$  it holds, for  $i \neq j = 1, 2$ ,

$$\lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} \varphi(x) (U_{ij}^\varepsilon * v_{i,t}^\varepsilon)(x) dx = \int_{\mathbb{R}^d} \varphi(x) d\tilde{\rho}_i(t).$$

**Proof.** For  $t \in [0, T]$  and  $\varphi \in C_c^1(\mathbb{R}^d)$ , by using that  $U_{ij}$  is even and the Definition of  $v_{i,t}^\varepsilon$  we obtain:

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \varphi(x) (U_{ij}^\varepsilon * v_{i,t}^\varepsilon)(x) dx - \int_{\mathbb{R}^d} \varphi(x) v_{i,t}^\varepsilon(x) dx \right| &= \left| \int_{\mathbb{R}^d} (\varphi * U_{ij}^\varepsilon)(x) v_{i,t}^\varepsilon(x) dx - \int_{\mathbb{R}^d} \varphi(x) v_{i,t}^\varepsilon(x) dx \right| \\ &= \left| \int_{\mathbb{R}^d} [(\varphi * U_{ij}^\varepsilon)(x) - \varphi(x)] v_{i,t}^\varepsilon(x) dx \right| \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\varphi(x-y) - \varphi(x)| v_{i,t}^\varepsilon(x) dU_{ij}^\varepsilon(y) dx \\ &\leq \|\varphi\|_\infty \int_{\mathbb{R}^d} |y| dU_{ij}^\varepsilon(y) \\ &= \varepsilon \|\varphi\|_\infty \int_{\mathbb{R}^d} |x| dU_{ij}(x), \end{aligned}$$

which converges to 0 as  $\varepsilon \rightarrow 0^+$  since  $\int_{\mathbb{R}^d} |x| dU_{ij}(x) < +\infty$ .  $\square$

Below we state the main result for this section, stressing that in this case uniqueness is missing.

**Theorem 6.2.** Let  $\varepsilon \leq 1$  and  $\rho^0 \in (\mathcal{P}_2^a(\mathbb{R}^d) \cap L^2(\mathbb{R}^d))^2$  such that  $E[\rho^0] < \infty$ . The sequence  $\{\rho^\varepsilon\}_\varepsilon$  of solutions to (NLIE) admits a subsequence narrowly converging to a weak solution  $\tilde{\rho}$  of (CDS).

**Proof.** The main difference with respect to the proof of Theorem 5.1 is with the cross-interaction term as pointed out in (51) and (52). More precisely, in (52) we need to make sure that a subsequence of  $U_{12}^\varepsilon * v_{1,t}^\varepsilon$  strongly converges to  $\tilde{\rho}_1$  in  $L^2([0, T] \times \mathbb{R}^d)$ . In view of (54), we can use Proposition 4.2 applied to a subsequence of  $(U_{12}^\varepsilon * v_{1,t}^\varepsilon, U_{21}^\varepsilon * v_{2,t}^\varepsilon)$ , exactly as in Proposition 6.2. For each component, the strong  $L^2$  limit coincides with  $\rho_i$  due to Lemma 6.2.  $\square$

### 7. Further perspectives

The main contribution of this work is to provide a rigorous analytical derivation of the quadratic porous medium equation and a class of cross-diffusion systems. Our strategy relies on an appropriate time-discretisation of a nonlocal interaction equation (system) in the 2-Wasserstein space. This is relevant for

both the well-posedness of (CDS) and (PME), and their numerical study. We relaxed previous assumptions on the interaction kernel, allowing for pointy potentials, e.g. Morse. As mentioned above, a key motivation for our approach is to provide an analysis that works without geodesic  $\lambda$ -convexity techniques, having in mind cases where only a JKO-approach may be feasible. Therefore, we could prove a nonlocal-to-local limit for cross-diffusion systems.

### *Nonconservative forces*

In this paper, both equations considered have a 2-Wasserstein gradient flow structure, but our approach may be used even if the PDEs under study are not gradient flows — this is a considerable advantage of our result. A prototypical example is given by the PDE

$$\partial_t \rho = \frac{1}{2} \Delta \rho^2 + \nabla \cdot (\rho v),$$

being  $v \neq \nabla \varphi$ , for some function  $\varphi$ . The addition of the *non-gradient flow part* can be overtaken by considering a suitable splitting (JKO) scheme, as in [5].

The equation above is also significant in the context of networks, where Wasserstein-type metrics have been derived recently (cf. [4,23]).

### *Other exponents*

A natural question may arise is whether our approach can be extended to linear diffusion and  $m \neq 2$ . The first observation to be made is that the approximating equation should be different, for instance the non-viscous version of the one proposed in [24] or [7, Eq.(8)]. While the time-discretisation could be relatively “easy” to develop, it may be not trivial to obtain Sobolev bounds in order to obtain compactness. The analysis could be easier if one restricts to a torus, and using a different version of Aubin–Lions Lemma.

### *Deterministic particle methods*

Last but not least, it is still open to obtain an analytical proof of a deterministic particle approximation for the porous medium equation, as well as linear diffusion. Both this paper and [7], for  $m = 2$ , require initial data to have finite logarithmic entropy, thus excluding particle solutions of the nonlocal interaction equation. Anyway, numerical simulations show that this is not to be excluded, see [7, Section6]. The main challenge is then to relax the initial assumption on the logarithmic entropy. In this direction,  $\lambda$ -convexity of the energy plays a key role in [13] to rigorously prove a qualitative result when the number of particles,  $N$ , depends exponentially on  $\varepsilon$  and the approximation in Wasserstein of the initial datum. A similar result can be proven also in our case, even for systems, only using  $\lambda$ -convexity of the nonlocal energy. However, this would narrow the class of cross-diffusion systems obtained as one would need cross-interaction potentials to be proportional. Obtaining a proof for  $N$  independent on  $\varepsilon$  and quantitative estimates is still open. This is also relevant in the case of cross-diffusion systems as  $\lambda$ -convexity fails.

### **Acknowledgements**

The authors are grateful to Nadia Ansini (Sapienza University of Rome), Maria Bruna (University of Cambridge), José Antonio Carrillo (University of Oxford), and Francesco S. Patacchini (IFP Energies Nouvelles) for fruitful discussions. Furthermore, they would like to thank the anonymous reviewers for their valuable comments and suggestions. This work was carried out while AE was a postdoctoral researcher at

FAU Erlangen-Nürnberg. The authors thankfully acknowledge support by the German Science Foundation (DFG) through CRC TR 154 “Mathematical Modelling, Simulation and Optimisation Using the Example of Gas Networks”. AE also acknowledges support by the Advanced Grant Nonlocal-CPD (Nonlocal PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronisation) of the European Research Council Executive Agency (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 883363).

## References

- [1] L. Ambrosio, N. Gigli, G. Savaré, Gradient flows in metric spaces and in the space of probability measures, second ed., in: Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.
- [2] L. Beck, D. Matthes, M. Zizza, Exponential convergence to equilibrium for coupled systems of nonlinear degenerate drift diffusion equations, *SIAM J. Math. Anal.* 55 (3) (2023) 1766–1809.
- [3] A.L. Bertozzi, T. Laurent, J. Rosado,  $L^p$  theory for the multidimensional aggregation equation, *Comm. Pure Appl. Math.* 64 (1) (2011) 45–83.
- [4] M. Burger, I. Humpert, J.-F. Pietschmann, Dynamic optimal transport on networks, 2021, arXiv preprint arXiv:2101.03415.
- [5] G. Carlier, M. Laborde, A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts, *Nonlinear Anal.* 150 (2017) 1–18.
- [6] J.A. Carrillo, Y.-P. Choi, M. Hauray, The derivation of swarming models: Mean-field limit and Wasserstein distances, in: *Collective Dynamics from Bacteria to Crowds*, Springer, 2014, pp. 1–46.
- [7] J.A. Carrillo, K. Craig, F.S. Patacchini, A blob method for diffusion, *Calc. Var. Partial Differential Equations* 58 (2) (2019) Paper No. 53, 53.
- [8] J.A. Carrillo, M. Di Francesco, A. Esposito, S. Fagioli, M. Schmidtchen, Measure solutions to a system of continuity equations driven by Newtonian nonlocal interactions, *Discrete Contin. Dyn. Syst.* 40 (2) (2020) 1191–1231.
- [9] J.A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, D. Slepcev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations, *Duke Math. J.* 156 (2) (2011) 229–271.
- [10] L. Chen, E.S. Daus, A. Holzinger, A. Jüngel, Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems, *J. Nonlinear Sci.* 31 (6) (2021) Paper No. 94, 38.
- [11] L. Chen, E.S. Daus, A. Jüngel, Rigorous mean-field limit and cross-diffusion, *Z. Angew. Math. Phys.* 70 (4) (2019) Paper No. 122, 21.
- [12] K. Craig, A.L. Bertozzi, A blob method for the aggregation equation, *Math. Comp.* 85 (300) (2016) 1681–1717.
- [13] K. Craig, K. Elamvazhuthi, M. Haberland, O. Turanova, A blob method for inhomogeneous diffusion with applications to multi-agent control and sampling, 2022, arXiv preprint arXiv:2202.12927.
- [14] B.E.J. Dahlberg, C.E. Kenig, Nonnegative solutions of the porous medium equation, *Comm. Partial Differential Equations* 9 (5) (1984) 409–437.
- [15] S. Daneri, E. Radici, E. Runa, Deterministic particle approximation of aggregation-diffusion equations on unbounded domains, *J. Differential Equations* 312 (2022) 474–517.
- [16] S. Daneri, G. Savaré, Eulerian calculus for the displacement convexity in the Wasserstein distance, *SIAM J. Math. Anal.* 40 (3) (2008) 1104–1122.
- [17] P. Degond, F.-J. Mustieles, A deterministic approximation of diffusion equations using particles, *SIAM J. Sci. Stat. Comput.* 11 (2) (1990) 293–310.
- [18] M. Di Francesco, A. Esposito, S. Fagioli, Nonlinear degenerate cross-diffusion systems with nonlocal interaction, *Nonlinear Anal.* 169 (2018) 94–117.
- [19] M. Di Francesco, A. Esposito, M. Schmidtchen, Many-particle limit for a system of interaction equations driven by Newtonian potentials, *Calc. Var. Partial Differential Equations* 60 (2) (2021) Paper No. 68, 44.
- [20] M. Di Francesco, S. Fagioli, Measure solutions for nonlocal interaction PDEs with two species, *Nonlinearity* 26 (2013) 2777–2808.
- [21] M. Di Francesco, D. Matthes, Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations, *Calc. Var. Partial Differential Equations* 50 (1–2) (2014) 199–230.
- [22] R.L. Dobrušin, Vlasov equations, *Funktsional. Anal. I Prilozhen.* 13 (2) (1979) 48–58, 96.
- [23] M. Erbar, D. Forkert, J. Maas, D. Mugnolo, Gradient flow formulation of diffusion equations in the Wasserstein space over a metric graph, *Netw. Heterogeneous Media* 17 (5) (2022) 687–717.
- [24] A. Figalli, R. Philipowski, Convergence to the viscous porous medium equation and propagation of chaos, *ALEA Lat. Am. J. Probab. Math. Stat.* 4 (2008) 185–203.
- [25] J. Fontbona, S. Méléard, Non local Lotka-Volterra system with cross-diffusion in an heterogeneous medium, *J. Math. Biol.* 70 (4) (2015) 829–854.
- [26] F. Golse, The mean-field limit for the dynamics of large particle systems, in: *Journées “équations Aux Dérivées Partielles”*, Univ. Nantes, Nantes, 2003, Exp. No. IX, 47.
- [27] L. Gosse, G. Toscani, Lagrangian numerical approximations to one-dimensional convolution-diffusion equations, *SIAM J. Sci. Comput.* 28 (4) (2006) 1203–1227.
- [28] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker–Planck equation, *SIAM J. Math. Anal.* 29 (1) (1998) 1–17.



- [29] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.* 29 (1) (1998) 1–17.
- [30] P.-L. Lions, S. Mas-Gallic, Une méthode particulière déterministe pour des équations diffusives non linéaires, *C.R. Acad. Sci. Paris Sér. I Math.* 332 (4) (2001) 369–376.
- [31] D. Matthes, R. McCann, G. Savaré, A family of fourth order equations of gradient flow type, *Comm. Partial Differential Equations* 34 (11) (2009) 1352–1397.
- [32] D. Morale, V. Capasso, K. Oelschläger, An interacting particle system modelling aggregation behavior: From individuals to populations, *J. Math. Biol.* 50 (1) (2005) 49–66.
- [33] C.B. Morrey Jr., On the derivation of the equations of hydrodynamics from statistical mechanics, *Comm. Pure Appl. Math.* 8 (1955) 279–326.
- [34] A. Moussa, From nonlocal to classical Shigesada-Kawasaki-Teramoto systems: Triangular case with bounded coefficients, *SIAM J. Math. Anal.* 52 (1) (2020) 42–64.
- [35] K. Oelschläger, Large systems of interacting particles and the porous medium equation, *J. Differential Equations* 88 (2) (1990) 294–346.
- [36] K. Oelschläger, A sequence of integro-differential equations approximating a viscous porous medium equation, *Z. Anal. Anwend.* 20 (1) (2001) 55–91.
- [37] L. Onsager, Crystal statistics. I. A two-dimensional model with an order-disorder transition, *Phys. Rev.* (2) 65 (1944) 117–149.
- [38] R. Philipowski, Interacting diffusions approximating the porous medium equation and propagation of chaos, *Stochastic Process. Appl.* 117 (4) (2007) 526–538.
- [39] R. Rossi, 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.
- [40] G. Russo, Deterministic diffusion of particles, *Comm. Pure Appl. Math.* 43 (6) (1990) 697–733.
- [41] E. Sandier, S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg-Landau, *Comm. Pure Appl. Math.* 57 (12) (2004) 1627–1672.
- [42] F. Santambrogio, Optimal Transport for Applied Mathematicians, in: *Progress in Nonlinear Differential Equations and Their Applications*, vol. 87, Birkhäuser Verlag, Basel, 2015.
- [43] S. Serfaty, Gamma-convergence of gradient flows on Hilbert and metric spaces and applications, *Discrete Contin. Dyn. Syst.* 31 (4) (2011) 1427–1451.
- [44] N. Shigesada, K. Kawasaki, E. Teramoto, Spatial segregation of interacting species, *J. Theoret. Biol.* 79 (1) (1979) 83–99.
- [45] J.L. Vázquez, An introduction to the mathematical theory of the porous medium equation, in: *Shape Optimization and Free Boundaries*, Montreal, PQ, 1990, in: *NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci.*, vol. 380, Kluwer Acad. Publ., Dordrecht, 1992, pp. 347–389.
- [46] J.L. Vázquez, *The Porous Medium Equation*, in: *Oxford Mathematical Monographs*, The Clarendon Press, Oxford University Press, Oxford, 2007, p. xxii+624, Mathematical theory.
- [47] C. Villani, *Optimal Transport : Old and New*, in: *Grundlehren der mathematischen Wissenschaften*, Springer, Berlin, 2009.
- [48] J. Zinsl, D. Matthes, Transport distances and geodesic convexity for systems of degenerate diffusion equations, *Calc. Var. Partial Differential Equations* 54 (4) (2015) 3397–3438.