Abstract. This paper studies the convergence of solutions of a nonlocal interaction equation to
the solution of the quadratic porous medium equation 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
the gradient flow structure of the equations 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 could be slightly modified in order to extend our result to PDEs with
no gradient flow structure. 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. However, the latter excludes particle solutions of the nonlocal interaction
equation.

1. Introduction

In this manuscript we deal with the connection between the quadratic porous medium equation
(PME) and the nonlocal interaction equation (NLIE), for a suitable choice of the interaction po-
tential. 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. More precisely,
let $W_1 := V_1 \ast V_1$, for a function $V_1$ satisfying some assumptions that will be clarified later, cf. (V).
For any $\varepsilon > 0$, consider the scaling $W_\varepsilon(x) = \varepsilon^{-d}W_1(x/\varepsilon)$, whence $W_\varepsilon = V_\varepsilon \ast V_\varepsilon$. We prove that, as
$\varepsilon \to 0^+$, a sequence of weak measure solutions to
\[
\partial_t \rho^\varepsilon = \nabla \cdot (\rho^\varepsilon \nabla W_\varepsilon \ast \rho^\varepsilon)
\] (NLIE)
converges to a weak solution of
\[
\partial_t \rho = \frac{1}{2} \Delta (\rho^2) = \nabla \cdot (\rho \nabla \rho).
\] (PME)

The main motivation for this work is to provide further insights into the derivation of the porous
medium equation from a system of interacting particles, rather than a physical derivation as in
continuum mechanics. We remind the reader to [36] 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 rigourous
derivation and way to construct solutions of PDEs, leading to well-posedness, as well as powerful
numerical methods. We mention here the seminal works [28, 26, 17] and the review [20]. In
case of transport equations (with no diffusion), e.g. (NLIE), deterministic approaches represent
a reasonable choice since weak measure solutions may exist, in particular particle solutions of the
form
\[
\rho^N_t = \frac{1}{N} \sum_{i=1}^N \delta_{X_i(t)}.
\]
where, for any \(i = 1, \ldots, N\), \(X_i(t)\) 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 [8, 5], and to [15, 14] in case of systems of PDEs. The problem is substantially different when diffusion is present, as particles do not remain particles. 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 [31, 21] for one dimensional linear and nonlinear diffusion, respectively, and [6] in any dimension.

A successful attempt to overtake the aforementioned difficulty is given by stochastic particles undergoing a Brownian motion. We start mentioning the inspiring work for our paper, [19], 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, [27, 25], and Philipowski, cf. [19], the latter concerning only the case \(m = 2\). As byproduct of their analysis, the authors of [19] prove propagation of chaos, thus providing a connection between microscopic and macroscopic description. In [29] 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 [35], 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.

In the recent article [6], the authors provide a deterministic particle method for linear and nonlinear diffusion equation, interpreted as 2-Wasserstein gradient flow, as counterpoint to stochastic methods. Their approach is inspired by the blob method for aggregation equations in [10]. In particular, in [6], Carrillo et al. 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, [32, 34]. 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-Galliac, [23], on a numerical scheme for (PME) on a bounded domain with periodic boundary conditions. In case \(m > 2\) or more general initial data, it is an open problem to check and apply the stability in [6, Theorem 5.8].

Our result is related to [6] as it provides a rigorous procedure to derive the quadratic porous medium equation from the nonlocal interaction equation. We observe that the regularisation of the energy in [6], for \(m = 2\), 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 may be used even if a gradient flow structure is not exhibited (see below and section 6). More precisely, we construct solutions of (NLIE) by means of the JKO scheme, [22], 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 \(v^\varepsilon := V_\varepsilon \ast \rho^\varepsilon\). 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 [6, Remark 6.3], though numerical simulations in [6, Section 6] give confidence that deterministic approximation could be achieved. Interacting stochastic particle systems represent by now a solid method. We notice that our approach does not exploit neither \(\lambda\)-convexity of the energies involved, as in [6], nor any equivalent gradient flow formulation of the equations such as \textit{evolution variational inequality} or \textit{curve of maximal slope}. Indeed, 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, [4]. The latter issue is also relevant for the extension of our result to cross-diffusion systems, since geodesic convexity is valid in few cases, \textit{cf.} [38]. We 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. As previously mentioned, the extension to the non-viscous and \textit{non-quadratic} porous medium equation is still an open problem, \textit{cf.} [6, 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 6.

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, \textit{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 \(\varepsilon \to 0^+\) limit in Theorem 5.1 in section 5. We conclude the paper with some remarks on possible extensions of our result in section 6.

2. Notation and preliminaries

The interaction potential we consider in this work is the (rescaled) convolution \(W_1 := V_1 \ast V_1\), being \(V_1 : \mathbb{R}^d \to \mathbb{R}\) such that the following conditions hold:

(V) \(V_1 \in C_b(\mathbb{R}^d; [0, +\infty)) \cap C^1(\mathbb{R}^d \setminus \{0\})\), \(|V_1|_{L^1} = 1\), \(V_1(x) = V_1(-x)\), \(\int_{\mathbb{R}^d} |x|V_1(x) \, dx < +\infty\), \(\nabla V_1 \in L^1(\mathbb{R}^d)\), 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 considered in previous results. The assumption (V) implies that the interaction potential satisfies

(W) \(W_1 \in C(\mathbb{R}^d; [0, \infty))\), \(W_1(x) = W_1(-x)\) for all \(x \in \mathbb{R}^d\), \(W_1 \in C^1(\mathbb{R}^d \setminus \{0\})\) such that \(\nabla W_1 = \nabla V_1 \ast V_1\) 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 \(W_\varepsilon = V_\varepsilon \ast V_\varepsilon\).

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 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 \(p\)-th order moment of \(\rho\), for \(1 \leq p < \infty\). We shall use \(\mathcal{P}_p(\mathbb{R}^d)\) 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
\[
d_W^2(\mu_1, \mu_2) := \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \left\{ \int_{\mathbb{R}^{2d}} |x - y|^2 \, d\gamma(x, y) \right\},
\]
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 $T : \mathbb{R}^d \to \mathbb{R}^n$, $n \in \mathbb{N}$, the push-forward of $\rho$ through $T$ is defined by
\[
\int_{\mathbb{R}^n} f(y) \, d\# T \rho(y) = \int_{\mathbb{R}^d} f(T(x)) \, d\rho(x) \quad \text{for all } f \text{ Borel functions on } \mathbb{R}^n.
\]
Setting $\Gamma_0(\mu_1, \mu_2)$ as the class of optimal plans, i.e. minimizers 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, 37, 33] 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
\[
m_2(\rho_1) \leq 2m_2(\rho_0) + 2d_W^2(\rho_0, \rho_1), \quad \forall \rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d).
\]

In proposition 4.2 we use the 1-Wasserstein distance, denoted by $d_1$ and defined by
\[
d_1(\mu_1, \mu_2) := \min_{\gamma \in \Gamma(\mu_1, \mu_2)} \left\{ \int_{\mathbb{R}^{2d}} |x - y| \, d\gamma(x, y) \right\}.
\]

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
\[
\begin{align*}
\partial_t \rho &= \nabla \cdot (\rho \nabla \rho) \\
\rho(0, \cdot) &= \rho_0
\end{align*}
\]
on the time interval $[0, T]$ with initial datum $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ 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 $\varphi \in C^1_c(\mathbb{R}^d)$ 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 $\rho^\varepsilon : [0, T] \to \mathcal{P}_2(\mathbb{R}^d)$, mapping $t \in [0, T] \mapsto \rho^\varepsilon_t \in \mathcal{P}_2(\mathbb{R}^d)$, is a weak measure solution to (NLIE) if, for every $\varphi \in C^1_c(\mathbb{R}^d)$ and any $t \in [0, T]$, it holds
\[
\int_{\mathbb{R}^d} \varphi(x) \, d\rho^\varepsilon_t(x) - \int_{\mathbb{R}^d} \varphi(x) \, d\rho_0(x) = -\frac{1}{2} \int_0^t \int_{\mathbb{R}^{2d}} (\nabla \varphi(x) - \nabla \varphi(y)) \cdot \nabla W_\varepsilon(x - y) \, d\rho^\varepsilon_t(y) \, d\rho^\varepsilon_t(x) \, dr.
\]
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

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

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

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

where we used the growth condition on $|\nabla V_\varepsilon|$ and preservation of second order moments (cf. Lemma 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 overcomes the loss of regularity at 0 for $\nabla W_\varepsilon$, as already noticed in [8].

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 [8]. Furthermore, it is worth to mention [2], 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 [8], though not convex, with the aim of applying the JKO scheme, [22], in order to obtain a priori estimates on the solutions of (NLIE) and their smoothed version $v^\varepsilon = V_\varepsilon \ast \rho^\varepsilon$. In turn, we are able to show convergence towards (PME). The interaction potential we choose is the (rescaled) convolution $W_\varepsilon = V_\varepsilon \ast V_\varepsilon$, 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 $\nabla W_\varepsilon = \nabla V_\varepsilon \ast V_\varepsilon$ 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 $\rho_0 \in P_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, and the interaction energy functional $W_\varepsilon : P_2(\mathbb{R}^d) \to (-\infty, +\infty]$ is given by

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

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

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

\[
W_\varepsilon[\rho_0] = \frac{1}{2} \int_{\mathbb{R}^d} (W_\varepsilon * \rho_0)(x) \rho_0(x) \, dx \\
= \frac{1}{2} \int_{\mathbb{R}^d} |(V_\varepsilon * \rho_0)(x)|^2 \, dx \\
= \frac{1}{2} \|V_\varepsilon * \rho_0\|_{L^2(\mathbb{R}^d)}^2 \leq \frac{1}{2} \|V_\varepsilon\|_{L^1}^2 \|\rho_0\|_{L^2}^2 = \frac{1}{2} \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2 < \infty.
\]

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

\[
\rho_{\tau, \varepsilon}^{n+1} \in \text{argmin}_{\rho \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{d^2_W(\rho_{\tau, \varepsilon}^n, \rho)}{2\tau} + W_\varepsilon[\rho] \right\}.
\]

The above sequence is well-defined if, for fixed \( \bar{\rho} \in \mathcal{P}_2(\mathbb{R}^d) \), the penalised energy functional \( \rho \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \frac{d^2_W(\rho, \bar{\rho})}{2\tau} + W_\varepsilon[\rho] \) admits minimisers. This can be easily proven by applying the direct method of calculus of variations. For further details we refer to [8, 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 of 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

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

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

In the next proposition we prove narrow compactness (in \( \tau \)) for \( \rho_{\tau, \varepsilon}^k \) 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 \( \bar{\rho}^\varepsilon : [0, T] \to \mathcal{P}_2(\mathbb{R}^d) \) such that the piecewise constant interpolation \( \rho_{\tau, \varepsilon}^k \) admits a subsequence \( \rho_{\tau, \varepsilon}^k \) narrowly converging to \( \bar{\rho}^\varepsilon \) uniformly in \( t \in [0, T] \) as \( k \to +\infty \). Moreover, for any \( t \in [0, T] \), the following uniform bounds in \( \tau \) and \( \varepsilon \) hold

\[
W_\varepsilon[\bar{\rho}^\varepsilon(t)] \leq \frac{1}{2} \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2, \\
m_2(\bar{\rho}^\varepsilon) \leq 2m_2(\rho_0) + 2T \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2.
\]

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

\[
\frac{d^2_W(\rho_{\tau, \varepsilon}^n, \rho_{\tau, \varepsilon}^{n+1})}{2\tau} + W_\varepsilon[\rho_{\tau, \varepsilon}^{n+1}] \leq W_\varepsilon[\rho_{\tau, \varepsilon}^n],
\]

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

\[
\sup_n W_\varepsilon[\rho_{\tau, \varepsilon}^n] \leq W_\varepsilon[\rho_0] \leq \frac{1}{2} \|V_1\|_{L^1}^2 \|\rho_0\|_{L^2}^2.
\]
By summing up over \( k \) inequality (6), we obtain

\[
\sum_{k=m}^{n} \frac{d_{W}^{2}(ρ_{τ,ε}^{k}, ρ_{τ,ε}^{k+1})}{2τ} \leq W_{ε}[ρ_{m,τ,ε}^{m}] - W_{ε}[ρ_{τ,ε}^{n+1}] .
\] (8)

The non-negativity of \( W_{ε} \) and the energy inequality (7) allow us to improve the above inequality as follows

\[
\sum_{k=m}^{n} \frac{d_{W}^{2}(ρ_{τ,ε}^{k}, ρ_{τ,ε}^{k+1})}{2τ} \leq W_{ε}[ρ_{0}] .
\] (9)

This implies

\[
d_{W}(ρ_{0}, ρ_{τ}^{ε}(t)) \leq 2TW_{ε}[ρ_{0}] \leq T\|V_{1}\|_{L^{1}}\|ρ_{0}\|_{L^{2}}^{2},
\]

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

\[
m_{2}(ρ_{τ}^{ε}(t)) \leq 2m_{2}(ρ_{0}) + 2d_{W}^{2}(ρ_{0}, ρ_{τ}^{ε}(t)) \leq 2m_{2}(ρ_{0}) + 2T\|V_{1}\|_{L^{1}}\|ρ_{0}\|_{L^{2}}^{2}.
\] (10)

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

\[
d_{W}(ρ_{τ}^{ε}(s), ρ_{τ}^{ε}(t)) \leq \sum_{k=m}^{n-1} d_{W}(ρ_{τ,ε}^{k}, ρ_{τ,ε}^{k+1}) \leq \left( \sum_{k=m}^{n-1} \frac{d_{W}^{2}(ρ_{τ,ε}^{k}, ρ_{τ,ε}^{k+1})}{2τ} \right)^{1/2} \vert n - m \vert^{1/2}
\] (11)

\[
\leq c \left( \sqrt{|t - s|} + \sqrt{τ} \right),
\]

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

\[
\liminf_{k \to +∞} \int_{\mathbb{R}^{d}} |x|^{2} dρ_{τ,ε}^{k}(x) \geq \int_{\mathbb{R}^{d}} |x|^{2} d\tilde{ρ}^{ε}(x)
\]

\[
\liminf_{k \to +∞} W_{ε}[ρ_{τ,ε}^{k}] \geq W_{ε}[\tilde{ρ}^{ε}],
\]

whence the thesis follows by applying the above inequalities to (7) and (10).

Next, we show that \( \tilde{ρ}^{ε} \) 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 \( W_{ε} \) does not satisfy the convexity assumption required in \([1, 8]\), 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 \( \tilde{ρ}^{ε} \) is a weak measure solution to (NLIE) according to Definition 2.2.

**Proof.** Let us consider two consecutive elements of the sequence \( \{ρ_{τ,ε}^{n}\}_{n \in \mathbb{N}} \) defined in (4), i.e. \( ρ_{τ,ε}^{n} \) and \( ρ_{τ,ε}^{n+1} \). We perturb \( ρ_{τ,ε}^{n+1} \) by using a map \( P^{σ} = id + σζ \), for some \( ζ \in C_{c}^{∞}(\mathbb{R}^{d}; \mathbb{R}^{d}) \) and \( σ \geq 0 \), that is we consider the perturbation

\[
ρ^{σ} := P^{σ}_{\#}ρ_{τ,ε}^{n+1}.
\] (12)

Being \( ρ_{τ,ε}^{n+1} \) a minimiser of (4), we have

\[
\frac{1}{2τ} \left[ \frac{d_{W}^{2}(ρ_{τ,ε}^{n}, ρ^{σ}) - d_{W}^{2}(ρ_{τ,ε}^{n+1}, ρ_{τ,ε}^{n+1})}{σ} \right] + \frac{W_{ε}[ρ^{σ}] - W_{ε}[ρ_{τ,ε}^{n+1}]}{σ} \geq 0.
\] (13)
First, we consider the interaction terms in (13)

\[
\frac{1}{2 \tau} \int_{\mathbb{R}^d} W_\varepsilon \ast \rho^\sigma(x) d\rho^\sigma(x) - \frac{1}{2 \tau} \int_{\mathbb{R}^d} W_\varepsilon \ast \rho^{\tau,\varepsilon}_{n+1}(x) d\rho^{\tau,\varepsilon}_{n+1}(x) \\
= \frac{1}{2} \int_{\mathbb{R}^d} \frac{\int_{\mathbb{R}^d} W_\varepsilon(x-y) d\rho^{\tau,\varepsilon}_{n+1}(y)}{\sigma} d\rho^{\tau,\varepsilon}_{n+1}(x) \\
= \frac{1}{2} \int_{\mathbb{R}^d} \frac{\int_{\mathbb{R}^d} \left[ W_\varepsilon(x-y) + \sigma(\zeta(x) - \zeta(y)) - W_\varepsilon(x-y) \right]}{\sigma} d\rho^{\tau,\varepsilon}_{n+1}(y) d\rho^{\tau,\varepsilon}_{n+1}(x)
\]  

(14)

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

\[
\frac{W_\varepsilon(x-y+\sigma(\zeta(x) - \zeta(y))) - W_\varepsilon(x-y)}{\sigma} \rightarrow \nabla W_\varepsilon(x-y) \cdot (\zeta(x) - \zeta(y)).
\]

(15)

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

\[
\int_{B_\eta} d\rho^{\tau,\varepsilon}_{n+1}(y) d\rho^{\tau,\varepsilon}_{n+1}(x) < \eta
\]

and the convergence (15) is uniform on \( \mathbb{R}^d \setminus B_\eta \). 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

\[
\frac{1}{2} \int_{\mathbb{R}^d} \frac{\int_{\mathbb{R}^d} \left[ W_\varepsilon(x-y) + \sigma(\zeta(x) - \zeta(y)) - W_\varepsilon(x-y) \right]}{\sigma} d\rho^{\tau,\varepsilon}_{n+1}(y) d\rho^{\tau,\varepsilon}_{n+1}(x) \\
\rightarrow \int_{\mathbb{R}^d} \nabla W_\varepsilon(x-y) \cdot (\zeta(x) - \zeta(y)) d\rho^{\tau,\varepsilon}_{n+1}(y) d\rho^{\tau,\varepsilon}_{n+1}(x).
\]

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

\[
\frac{1}{2\tau} \left[ d_{W}^{\rho^n_{\tau,\varepsilon}, \rho^{n+1}_{\tau,\varepsilon}}(\sigma) - d_{W}^{\rho^n_{\tau,\varepsilon}, \rho^{n+1}_{\tau,\varepsilon}}(\sigma) \right] = \frac{1}{2\tau} \int_{\mathbb{R}^d} (|x - P^\sigma(y)|^2 - |x - y|^2) d\gamma^n_{\tau,\varepsilon}(x,y) \\
= \frac{1}{2\tau} \int_{\mathbb{R}^d} (|x - y - \sigma \zeta(y)|^2 - |x - y|^2) d\gamma^n_{\tau,\varepsilon}(x,y) \\
= -\frac{1}{\tau} \int_{\mathbb{R}^d} (x - y) \cdot (\zeta(y)) d\gamma^n_{\tau,\varepsilon}(x,y) + o(\sigma),
\]

where in the last equality we applied a first order Taylor expansion. By sending \( \sigma \) to 0 it holds

\[
\frac{1}{\tau} \int_{\mathbb{R}^d} (x - y) \cdot (\zeta(y)) d\gamma^n_{\tau,\varepsilon}(x,y) \leq \frac{1}{2} \int_{\mathbb{R}^d} \nabla W_\varepsilon(x-y) \cdot (\zeta(x) - \zeta(y)) d\rho^{\tau,\varepsilon}_{n+1}(y) d\rho^{\tau,\varepsilon}_{n+1}(x).
\]

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

\[
\frac{1}{\tau} \int_{\mathbb{R}^d} (x - y) \cdot \nabla \varphi(y) d\gamma^n_{\tau,\varepsilon}(x,y) = \frac{1}{2} \int_{\mathbb{R}^d} \nabla W_\varepsilon(x-y) \cdot (\nabla \varphi(x) - \nabla \varphi(y)) d\rho^{\tau,\varepsilon}_{n+1}(y) d\rho^{\tau,\varepsilon}_{n+1}(x).
\]

(16)

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

\[
\frac{1}{\tau} \int_{\mathbb{R}^d} (x - y) \cdot \nabla \varphi(y) d\gamma^n_{\tau,\varepsilon}(x,y) = \frac{1}{\tau} \int_{\mathbb{R}^d} \varphi(x) d(\rho^n_{\tau,\varepsilon} - \rho^{n+1}_{\tau,\varepsilon})(x) + O(\tau).
\]

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

\[ h = \left[ \frac{s}{\tau} \right] + 1 \quad \text{and} \quad k = \left[ \frac{t}{\tau} \right]. \]
Taking into account the last equality, by summing in (16) over \( j \) from \( h \) to \( k \), we obtain

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

which is equivalent to

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

Up to pass to a subsequence, the thesis follows by considering the limit as \( \tau \to 0^+ \) 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, \( V_\varepsilon \ast \tilde{\rho}^\varepsilon \), in the \( \varepsilon \to 0^+ \) limit. In this section we deal with the compactness for both the sequences. For 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 \( \tilde{\rho} : [0, T] \to \mathcal{P}_2(\mathbb{R}^d) \) 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 \( k \to +\infty \).

**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 \( \rho^\varepsilon_{\tau} \), cf. Eq. (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 \( \gamma^\varepsilon_{\tau} \in \Gamma_0(\rho^\varepsilon_{\tau}(s), \rho^\varepsilon_{\tau}(t)) \) such that

\[
d^2_{W}(\rho^\varepsilon_{\tau}(s), \rho^\varepsilon_{\tau}(t)) = \iint_{\mathbb{R}^{2d}} |x - y|^2 \, d\gamma^\varepsilon_{\tau}(x, y).
\]

By stability of optimal transport plans, cf. [37, Corollary 5.21], we get \( \gamma^\varepsilon_{\tau} \to \gamma^\varepsilon \) as \( \tau \to 0^+ \), and

\[
\liminf_{\tau \to 0} d^2_{W}(\rho^\varepsilon_{\tau}(s), \rho^\varepsilon_{\tau}(t)) = \liminf_{\tau \to 0} \iint_{\mathbb{R}^{2d}} |x - y|^2 \, d\gamma^\varepsilon_{\tau}(x, y) \\
\geq \iint_{\mathbb{R}^{2d}} |x - y|^2 \, d\gamma^\varepsilon(x, y) \\
\geq d^2_{W}(\rho^\varepsilon(s), \rho^\varepsilon(t)).
\]

In particular, from Eq. (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 thesis 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 [24], cf. also [16, 13, 7] 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, [22], 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 \( S^\varepsilon_0 \), given by the heat semigroup (cf. [1, 12, 16]). 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 \( S^\varepsilon_\tau : [0, +\infty] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathcal{P}_2(\mathbb{R}^d) \) is a \( \lambda \)-flow for a functional \( \mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\} \) with respect to the distance \( d_W \); if, for an arbitrary \( \rho \in \mathcal{P}_2(\mathbb{R}^d) \), the curve \( t \mapsto S^\varepsilon_\tau \rho \) is absolutely continuous on \( [0, +\infty[ \) and it satisfies the evolution variational inequality (EVI)

\[
\frac{1}{2} \frac{d^2}{dt} d_W^2(S^\varepsilon_\tau \rho, \bar{\rho}) + \frac{\lambda}{2} d_W^2(S^\varepsilon_\tau \rho, \bar{\rho}) \leq \mathcal{F}(\bar{\rho}) - \mathcal{F}(S^\varepsilon_\tau \rho)
\]

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, that is

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

**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 [22, Proposition 4.1] it is shown that

\[
\mathcal{H}(\rho) \geq -C(m_2(\rho) + 1)^2,
\]

for every \( \rho \in \mathcal{P}^0_2(\mathbb{R}^d), \beta \in (\frac{d}{d+2}, 1) \) 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 \( S^\varepsilon_0 \nu \) 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 \( S^\varepsilon_\tau \) by

\[
D^\varepsilon S^\varepsilon_\tau \rho := \limsup_{s \downarrow 0} \left\{ \frac{W^\varepsilon[\rho] - W^\varepsilon[S^\varepsilon_\tau \rho]}{s} \right\}.
\]

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

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

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

**Proof.** From proposition 3.1 we infer the uniform bound in \( \tau \) and \( \varepsilon \)

\[
\|V_\tau * \rho^\varepsilon\|_{L^2([0, T]; L^2(\mathbb{R}^d))}^2 = \int_0^T \int_{\mathbb{R}^d} |(V_\tau * \rho^\varepsilon(t))(x)|^2 \, dx \, dt = 2 \int_0^T W^\varepsilon[\rho^\varepsilon(t)] \, dt \leq 2T W^\varepsilon[\rho_0] \leq T \|V_1\|_{L^1(\mathbb{R}^d)}^2 \|\rho_0\|_{L^2(\mathbb{R}^d)}^2,
\]
Thus, there exists a subsequence $\tau_k$ such that $V_\varepsilon * \rho^{n}_{\tau_k} \to w^\varepsilon$ in $L^2([0,T];L^2(\mathbb{R}^d))$ as $\tau_k \to 0$. The limit $w^\varepsilon \equiv v^\varepsilon$ due to uniqueness of limit and proposition 4.1. Up to pass to a subsequence, we have
\[
\|v^\varepsilon\|_{L^2([0,T];L^2(\mathbb{R}^d))} \leq T\|V_1\|_{L^1(\mathbb{R}^d)}^2\|\rho_0\|_{L^2(\mathbb{R}^d)}^2,
\] (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 $S_\varepsilon^{1}(\rho^{n+1}_{r,\varepsilon}, \rho^{n}_{r,\varepsilon})$ as competitor of $\rho^{n+1}_{r,\varepsilon}$ in the minimisation problem (4), as direct consequence of the definition of the sequence $\{\rho^{n}_{r,\varepsilon}\}_{n \in \mathbb{N}}$ we have
\[
\frac{1}{2\tau} d_W(S_\varepsilon^{n+1}_{r,\varepsilon}, \rho^{n}_{r,\varepsilon}) + W_\varepsilon[\rho^{n+1}_{r,\varepsilon}] \leq \frac{1}{2\tau} d_W(S_\varepsilon^{n}_{r,\varepsilon}, \rho^{n}_{r,\varepsilon}) + W_\varepsilon[S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon})],
\]
whence, dividing by $s > 0$ and passing to the lim sup as $s \downarrow 0$,
\[
\tau D_\varepsilon W_\varepsilon(\rho^{n+1}_{r,\varepsilon}) \leq \frac{1}{2} \frac{d}{dt} \left( d_W(S_\varepsilon^{n+1}_{r,\varepsilon}, \rho^{n}_{r,\varepsilon}) \right) \bigg|_{t=0} \leq \mathcal{H}[\rho^{n}_{r,\varepsilon}] - \mathcal{H}[\rho^{n+1}_{r,\varepsilon}].
\] (21)
In the last inequality we used that $S_\varepsilon$ is a 0-flow. Now, let us focus on the left hand side of (21). First of all, note that
\[
D_\varepsilon W_\varepsilon(\rho^{n+1}_{r,\varepsilon}) = \limsup_{s \downarrow 0} \frac{W_\varepsilon[\rho^{n+1}_{r,\varepsilon}] - W_\varepsilon[S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon})]}{s}
\]
\[
= \limsup_{s \downarrow 0} \int_0^1 \left( - \frac{d}{dz} |_{z=st} W_\varepsilon[S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon})] \right) dt.
\] (22)
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:
\[
\frac{d}{dt} W_\varepsilon[S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon})] = - \int_{\mathbb{R}^d} \nabla (W_\varepsilon * S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon}))(x) \nabla S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon})(x) dx
\]
\[
= - \int_{\mathbb{R}^d} |\nabla (V_\varepsilon * S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon}))(x)|^2 dx.
\] (23)
By substituting (23) into (22), from (21) we obtain
\[
\tau \liminf_{s \downarrow 0} \int_0^1 \int_{\mathbb{R}^d} |\nabla (V_\varepsilon * S_\varepsilon^{n+1}(\rho^{n}_{r,\varepsilon}))(x)|^2 dx dt \leq \mathcal{H}[\rho^{n}_{r,\varepsilon}] - \mathcal{H}[\rho^{n+1}_{r,\varepsilon}],
\]
whence, by $L^2$ lower semi-continuity of the $H^1$ seminorm,
\[
\tau \int_{\mathbb{R}^d} |\nabla (V_\varepsilon * \rho^{n+1}_{r,\varepsilon})|^2 dx dt \leq \mathcal{H}[\rho^{n}_{r,\varepsilon}] - \mathcal{H}[\rho^{n+1}_{r,\varepsilon}].
\]
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
\[
\int_0^T \int_{\mathbb{R}^d} |\nabla (V_\varepsilon * \rho^{n+1}_{r,\varepsilon})(t)|^2 dx dt \leq \mathcal{H}[\rho_0] - \mathcal{H}[\rho^{n+1}_{r,\varepsilon}] \leq \|\rho_0\|_{L^2(\mathbb{R}^d)}^2 + C(\rho_0, V_1, T).
\]
In particular, using weak lower semi-continuity of the norm,
\[
\|\nabla v^\varepsilon\|_{L^2([0,T];L^2(\mathbb{R}^d))}^2 = \int_0^T \int_{\mathbb{R}^d} |\nabla v^\varepsilon(t)|^2 dx dt \leq \|\rho_0\|_{L^2(\mathbb{R}^d)}^2 + C(\rho_0, V_1, T).
\] (24)
The bounds in eqs. (20) and (24) give the first result of the statement, and an application of the Banach–Alaoglu Theorem concludes the proof.
Remark 4.2. Let us notice that, for any \( t \in [0, T] \), the first order moment of \( v^\varepsilon \) 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:
\[
\int_{\mathbb{R}^d} |x|v^\varepsilon(x) \, dx = \int_{\mathbb{R}^d} |x|V_\varepsilon(x) \, d\rho_\varepsilon(y) \, dx \\
\leq \int_{\mathbb{R}^d} V_\varepsilon(x-y)|x-y| \, d\rho_\varepsilon(y) \, dx + \int_{\mathbb{R}^d} V_\varepsilon(y-x)|y| \, d\rho_\varepsilon(y) \, dx \\
= \varepsilon \int_{\mathbb{R}^d} V_1(z)|z| \, dz + \int_{\mathbb{R}^d} V_1(z) \, dz \int_{\mathbb{R}^d} |y| \, d\rho_\varepsilon(y) \\
\leq \varepsilon \int_{\mathbb{R}^d} V_1(z)|z| \, dz + \sqrt{m_2(\rho^\varepsilon)} \int_{\mathbb{R}^d} V_1(z) \, dz < +\infty.
\]

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 [30, 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. [30, Theorem 2] Let \( X \) be a separable Banach space. Consider

- a lower semicontinuous functional \( \mathcal{F} : X \to [0, +\infty] \) with relatively compact sublevels in \( X \);
- a pseudo-distance \( g : X \times X \to [0, +\infty] \), i.e., \( g(\rho, \eta) = 0 \) 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 \( u : (0, T) \to X \), with a fixed \( T > 0 \). Assume further that
\[
\sup_{u \in U} \int_0^T \mathcal{F}(u(t)) \, dt < \infty \quad \text{and} \quad \lim_{h \downarrow 0} \sup_{u \in U} \int_0^{T-h} g(u(t + h), u(t)) \, dt = 0. \tag{25}
\]

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 \( \tilde{u} : (0, T) \to X \), i.e.,
\[
\lim_{n \to \infty} \{ t \in (0, T) : \| u_n(t) - \tilde{u}(t) \|_X \geq \sigma \} = 0, \quad \forall \sigma > 0.
\]

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 \( U := \{u^\varepsilon\}_{0 \leq \varepsilon \leq 1} \) 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 \( \mathcal{F} : L^2(\mathbb{R}^d) \to [0, +\infty] \) defined by
\[
\mathcal{F}[v] = \begin{cases} 
\| v \|_{H^1(\mathbb{R}^d)} + \int_{\mathbb{R}^d} |x|v(x) \, dx, & \text{if } v \in \mathcal{P}_1(\mathbb{R}^d) \cap H^1(\mathbb{R}^d); \\
+\infty & \text{otherwise}.
\end{cases}
\]

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 \( v_n \to v \) in \( L^2(\mathbb{R}^d) \) and \( \mathcal{F}[v_n] \leq +\infty \), 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
\[
\mathcal{F}^1[v_n] \geq \int_{\mathbb{R}^d} |\nabla v(x)|^2 \, dx + 2 \int_{\mathbb{R}^d} (\nabla v_{n_k}(x) - \nabla v) \cdot \nabla v(x) \, dx,
\]
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 \( v_n \to v \) 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 relatively 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

\[
\int_{\mathbb{R}^d} \varepsilon^2 \, dx = \int_{\mathbb{R}^d} \left| \int_0^1 \frac{d}{dt} (v + t\varepsilon) \, dt \right|^2 \, dx \leq |h|^2 \int_{\mathbb{R}^d} \left| \nabla v + t\varepsilon \right|^2 \, dx = |h|^2 \left\| \nabla v \right\|_{L^2(\mathbb{R}^d)}^2,
\]

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

\[
\|v\|_{L^2(\mathbb{R}^d \setminus B_R)}^2 = \int_{|x| > R} |v(x)|^2 \, dx \leq \frac{1}{R^d} \int_{\mathbb{R}^d} |x|^{2\delta} |v(x)|^2 \, dx \leq \frac{1}{R^d} \left( \int_{\mathbb{R}^d} |x|v(x) \, dx \right)^\delta \left( \int_{\mathbb{R}^d} |v(x)|^{2\frac{\delta}{1-\delta}} \, dx \right)^{1-\delta},
\]

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

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

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 \( v^\varepsilon : [0, T] \to L^2(\mathbb{R}^d) \) the sequence defined above by \( v^\varepsilon = V_\varepsilon \ast \rho^\varepsilon \), which satisfies lemma 4.1. For any \( 0 \leq \varepsilon \leq 1 \), it holds

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

where we also used Remark 4.2 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 equicontinuity 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 [33, 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 $v^{\varepsilon_k} \rightarrow v$ in $L^2([0,T];H^1(\mathbb{R}^d))$, in view of lemma 4.1. According to proposition 4.2, there exists a subsequence $v^{\varepsilon_k}$ such that $v^{\varepsilon_k}$ converges in measure (with respect to time with values in $X = L^2(\mathbb{R}^d)$) to a curve $\bar{v} \equiv v$, 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 $\sup_k \|v^{\varepsilon_k}(t)\|_{L^2(\mathbb{R}^d)}^2 \leq \|V\|_{L^1(\mathbb{R}^d)}^2 \|\rho_0\|_{L^2}^2$, 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.

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 $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ with $\mathcal{H}[\rho_0] < \infty$.

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

$$
\int_{\mathbb{R}^d} \varphi(x) d\rho^\varepsilon_t(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) = -\frac{1}{2} \int_0^T \int_{\mathbb{R}^d} (\nabla \varphi(x) - \nabla \varphi(y)) \cdot \nabla W_\varepsilon(x-y) d\rho^\varepsilon_t(y) d\rho^\varepsilon_t(x) dt,
$$

which can be rewritten as

$$
\int_{\mathbb{R}^d} \varphi(x) d\tilde{\rho}^\varepsilon_t(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) = -\int_0^T \int_{\mathbb{R}^d} V_\varepsilon * (\rho^\varepsilon_t \nabla \varphi)(x) \cdot \nabla \tilde{v}^\varepsilon_t(x) dx dt
$$

$$
= -\int_0^T \int_{\mathbb{R}^d} \tilde{v}^\varepsilon_t(x) \nabla \varphi(x) \cdot \nabla \tilde{v}^\varepsilon_t(x) dx dt
$$

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

$$z^\varepsilon_t(x) := V_\varepsilon * (\rho^\varepsilon_t \nabla \varphi)(x) - (V_\varepsilon * \rho^\varepsilon_t)(x) \nabla \varphi(x) = V_\varepsilon * (\rho^\varepsilon_t \nabla \varphi)(x) - \tilde{v}^\varepsilon_t(x) \nabla \varphi(x).
$$

Remark 5.1. Note that the integral after the second equality in (27) makes sense since $\nabla V_\varepsilon * \tilde{v}^\varepsilon_t \in C(\mathbb{R}^d)$. 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 $\check{\rho}$ 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 $z^\varepsilon \rightarrow 0$ 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 $\varphi \in C^2_c(\mathbb{R}^d)$. 

14
Proof. For any $t \in [0, T]$ and $\varphi \in C^2_0(\mathbb{R}^d)$ we obtain
\[
\int_{\mathbb{R}^d} |z^\varepsilon_t(x)| \, dx \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V_\varepsilon(x - y) |\nabla \varphi(y) - \nabla \varphi(x)| d\rho^\varepsilon_t(y) \, dx
\]
\[
\leq \|D^2 \varphi\|_\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V_\varepsilon(x - y) |y - x| d\rho^\varepsilon_t(y) \, dx
\]
\[
= \varepsilon \|D^2 \varphi\|_\infty \int_{\mathbb{R}^d} |z| V_1(z) \, dz,
\]
by means of the change of variable $z = \frac{x - y}{\varepsilon}$. The thesis follows by taking the supremum over $t \in [0, T]$. \hfill \square

**Lemma 5.2.** There exists a constant $C$ only depending on $\varphi$ and $V_1$ such that for all $\varepsilon > 0$
\[
\|z^\varepsilon\|_{L^3([0, T] \times \mathbb{R}^d)} \leq C.
\]

**Proof.** For almost every $x \in \mathbb{R}^d$ and $t \in [0, T]$, for $i = 1, \ldots, d$, the non-negativity of $V_1$ and $\rho^\varepsilon_t$ gives
\[
\left| \int_{\mathbb{R}^d} V_\varepsilon(x - y) \partial_{x_i} \varphi(y) d\rho^\varepsilon_t(y) \right| \leq \int_{\mathbb{R}^d} V_\varepsilon(x - y) |\partial_{x_i} \varphi(y)| d\rho^\varepsilon_t(y) \leq \|\partial_{x_i} \varphi\|_\infty v^\varepsilon_t(x).
\]
In particular, this implies
\[
|z_t(x)| \leq 2\|\nabla \varphi\|_\infty |v^\varepsilon_t(x)|,
\]
thus $\|z^\varepsilon_t\|_{L^p(\mathbb{R}^d)} \leq 2\|\nabla \varphi\|_\infty \|v^\varepsilon_t\|_{L^p}$ 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
\[
\sup_{t \in [0, T]} \|v^\varepsilon_t\|_{L^2(\mathbb{R}^d)} \leq c \quad \text{and} \quad \int_0^T \|\nabla v^\varepsilon_t\|_{L^2}^2 \, dt \leq c.
\]
The embedding $H^1(\mathbb{R}^d) \hookrightarrow L^4(\mathbb{R}^d)$ further implies
\[
\int_0^T \|v^\varepsilon_t\|_{L^4(\mathbb{R}^d)}^2 \, dt \leq c,
\]
for some constant still denoted by $c$. As consequence of the interpolation inequality for $L^p$ functions, we estimate
\[
\int_0^T \|v^\varepsilon_t\|_{L^3(\mathbb{R}^d)}^3 \, dt \leq \int_0^T \|v^\varepsilon_t\|_{L^2(\mathbb{R}^d)} \|v^\varepsilon_t\|_{L^4(\mathbb{R}^d)}^2 \, dt,
\]
which is uniformly bounded due to the previous estimates. \hfill \square

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

**Proof.** The proof is a simple consequence of the interpolation inequality for $L^p$ functions and the previous lemmas. More precisely, it holds
\[
\int_0^T \|z^\varepsilon_t\|_{L^2(\mathbb{R}^d)}^2 \, dt \leq \int_0^T \|z^\varepsilon_t\|_{L^1(\mathbb{R}^d)} \|z^\varepsilon_t\|_{L^3(\mathbb{R}^d)}^2 \, dt
\]
\[
\leq \left( \int_0^T \|z^\varepsilon_t\|_{L^1(\mathbb{R}^d)} \, dt \right)^{\frac{1}{2}} \left( \int_0^T \|z^\varepsilon_t\|_{L^3(\mathbb{R}^d)}^3 \, dt \right)^{\frac{1}{2}} \leq (\varepsilon T)^{\frac{1}{2}} C(V_1, \varphi),
\]
which gives the result by letting $\varepsilon$ to 0. \hfill \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 \( \varphi \in C^1_c(\mathbb{R}^d) \) we have

\[
\lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^d} \varphi(x)v^\varepsilon_t(x) \, dx = \int_{\mathbb{R}^d} \varphi(x) d\tilde{\rho}(t).
\]

**Proof.** For any \( t \in [0, T] \) and any \( \varphi \in C^1_c(\mathbb{R}^d) \), by using the definition of \( v^\varepsilon_t \) we obtain:

\[
\left| \int_{\mathbb{R}^d} \varphi(x)v^\varepsilon_t(x) \, dx - \int_{\mathbb{R}^d} \varphi(x) d\rho^\varepsilon_t(x) \right| = \left| \int_{\mathbb{R}^d} \varphi(x)(V^\varepsilon_t + \rho^\varepsilon_t)(x) \, dx - \int_{\mathbb{R}^d} \varphi(x) d\rho^\varepsilon_t(x) \right|
\]

\[
= \left| \int_{\mathbb{R}^d} (\varphi \ast V^\varepsilon_t)(x) \, d\rho^\varepsilon_t(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho^\varepsilon_t(x) \right|
\]

\[
= \left| \int_{\mathbb{R}^d} [(\varphi \ast V^\varepsilon_t)(x) - \varphi(x)] d\rho^\varepsilon_t(x) \right|
\]

\[
\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\varphi(x) - \varphi(y)| V^\varepsilon_t(y) \, dy \, d\rho^\varepsilon_t(x)
\]

\[
\leq \|\varphi\|_{\infty} \int_{\mathbb{R}^d} |y| V^\varepsilon_t(y) \, dy
\]

\[
= \varepsilon \|\varphi\|_{\infty} \int_{\mathbb{R}^d} |x| V_1(x) \, dx,
\]

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

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

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

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

\[
\int_{\mathbb{R}^d} \varphi(x) d\rho^\varepsilon_x(x) - \int_{\mathbb{R}^d} \varphi(x) d\rho_0(x) = - \int_0^t \int_{\mathbb{R}^d} v^\varepsilon_t(x) \nabla \varphi(x) \cdot \nabla v^\varepsilon_t(x) \, dx \, dt
\]

\[
- \int_0^t \int_{\mathbb{R}^d} z^\varepsilon_t(x) \cdot \nabla v^\varepsilon_t(x) \, dx \, dt,
\]

as explained in Eqs. (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

\[
v^{\varepsilon_k} \to \tilde{\rho} \quad \text{in} \quad L^2([0, T]; L^2(\mathbb{R}^d));
\]

\[
\nabla v^{\varepsilon_k}_t \to \nabla \tilde{\rho}(t) \quad \text{in} \quad L^2([0, T]; L^2(\mathbb{R}^d)).
\]

Before letting \( \varepsilon \to 0^+ \) and obtaining the result we need to further regularise the test function, \( \varphi \), since corollary 5.1 holds for test functions in \( C^2_c(\mathbb{R}^d) \). In this regard, we consider a standard mollifier \( \eta \in C^\infty_c(\mathbb{R}^d) \) and the corresponding sequence \( \varphi^\sigma := \varphi^\sigma \ast \varphi \in C^\infty_c(\mathbb{R}^d) \), being \( \varphi^\sigma(x) = \sigma^{-d} \eta(x/\sigma^d) \) for any \( x \in \mathbb{R}^d \) and \( \sigma > 0 \). As consequence of the observations above and corollary 5.1, by letting
\(\varepsilon \to 0^+\) 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 \(\sigma \to 0\) 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. [11, 36]. \(\square\)

6. Further perspectives

The main contribution of this work is to provide a rigorous analytical derivation of the quadratic porous medium equation. Our strategy relies on an appropriate time-discretisation of a nonlocal interaction equation in the 2-Wasserstein space. This is relevant for both the well-posedness of (PME) and the numerical study of the equation. 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 however to provide an analysis that works without geodesic \(\lambda\)-convexity techniques, having in mind cases where only a JKO-approach may be feasible.

**Systems and cross-diffusion.** An important possible application of our strategy is to nonlocal systems of the form
\[
\partial_t \rho_i = \sum_{j=1}^M \text{div} (\rho_i A_{ij} \nabla W_{\varepsilon} * \rho_j),
\]
in order to obtain in the local limit cross-diffusion systems
\[
\partial_t \rho_i = \sum_{j=1}^M \text{div} (\rho_i A_{ij} \nabla \rho_j),
\]
under suitable assumptions on the matrix of the coefficients. If \(A\) is symmetric we obtain a Wasserstein-gradient flow structure, and if \(A\) is positive definite we can gain similar estimates as in this paper by the flow interchange technique. It also seems reasonable to look at diffusion matrices \(A = B^{-1}C\) with \(B\) and \(C\) positive definite, which allows to define a transport metric for the system with weight matrix \(B\) and possibly to carry out a JKO-type approach. The flow interchange will work at least if \(B\) is diagonal. The extension of our result to cross-diffusion systems will be developed in a future paper, also including different interaction kernels among species.

We observe that using the theory of \(\lambda\)-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 \(j \neq i\), see [38, 13]. It is worth to mention that a stochastic approach (based on moderate limits) to derive cross-diffusion systems from particle systems is done in the recent work [9].

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. [15]. However, this does not exclude to apply a time-discretisation of the system to get uniform bounds and existence of a (sequence) of solutions.

**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 [4].
The equation above is also significant in the context of networks, where Wasserstein-type metrics have been derived recently (cf. [3, 18]).

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 [19] or [6, 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. Both this paper and [6], 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 [6, Section 6]. The main challenge is then to relax the initial assumption on the logarithmic entropy.

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. Moreover, the authors thankfully acknowledge support by the German Science Foundation (DFG) through CRC TR 154 “Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks”.

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
[2] A. L. Bertozzi, T. Laurent, and J. Rosado. \( L^p \) theory for the multidimensional aggregation equation. Comm. Pure Appl. Math., 64(1):45–83, 2011.
[3] M Burger, I. Humpert, and J.-F. Pietschmann. Dynamic optimal transport on networks. arXiv preprint arXiv:2101.03415, 2021.
[4] G. Carlier and M. Laborde. A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts. Nonlinear Anal., 150:1–18, 2017.
[5] J. A. Carrillo, Y.-P. Choi, and M. Hauray. The derivation of swarming models: mean-field limit and wasserstein distances. In Collective dynamics from bacteria to crowds, pages 1–46. Springer, 2014.
[6] J. A. Carrillo, K. Craig, and F. S. Patacchini. A blob method for diffusion. Calc. Var. Partial Differential Equations, 58(2):Paper No. 53, 53, 2019.
[7] J. A. Carrillo, M. Di Francesco, A. Esposito, S. Fagioli, and M. Schmidtchen. Measure solutions to a system of continuity equations driven by Newtonian nonlocal interactions. Discrete Contin. Dyn. Syst., 40(2):1191–1231, 2020.
[8] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepcev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. Duke Math. J., 156(2):229–271, 2011.
[9] L. Chen, E. S. Daus, A. Holzinger, and A. Jüngel. Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems. J. Nonlinear Sci., 31(6):Paper No. 94, 38, 2021.
[10] K. Craig and A. L. Bertozzi. A blob method for the aggregation equation. Math. Comp., 85(300):1681–1717, 2016.
[11] B. E. J. Dahlberg and C. E. Kenig. Nonnegative solutions of the porous medium equation. Comm. Partial Differential Equations, 9(5):409–437, 1984.
[12] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the wasserstein distance. SIAM J. Math. Anal., 40(3):1104–1122, 2008.
[13] M. Di Francesco, A. Esposito, and S. Fagioli. Nonlinear degenerate cross-diffusion systems with nonlocal interaction. Nonlinear Analysis, 169:94–117, 2018.
[14] M. Di Francesco, A. Esposito, and M. Schmidtchen. Many-particle limit for a system of interaction equations driven by Newtonian potentials. Calc. Var. Partial Differential Equations, 60(2):Paper No. 68, 44, 2021.
M. Di Francesco and S. Fagioli. Measure solutions for nonlocal interaction pdes with two species. *Nonlinearity*, 26:2777–2808, 2013.

M. Di Francesco and 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):199–230, 2014.

R. L. Dobrušin. Vlasov equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 96, 1979.

M. Erbar, D. Forkert, J. Maas, and D. Mugnolo. Gradient flow formulation of diffusion equations in the wasserstein space over a metric graph. *arXiv preprint arXiv:2105.05677*, 2021.

A. Figalli and R. Philipowski. Convergence to the viscous porous medium equation and propagation of chaos. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:185–203, 2008.

F. Golse. The mean-field limit for the dynamics of large particle systems. In *Journées “Équations aux Dérivées Partielles”*, pages Exp. No. IX, 47. Univ. Nantes, Nantes, 2003.

L. Gosse and G. Toscani. Lagrangian numerical approximations to one-dimensional convolution-diffusion equations. *SIAM J. Sci. Comput.*, 28(4):1203–1227, 2006.

R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

P.-L. Lions and S. Mas-Gallic. Une méthode particulaire déterministe pour des équations diffusives non linéaires. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(4):369–376, 2001.

D. Matthes, R.J. McCann, and G. Savaré. A family of fourth order equations of gradient flow type. *Comm. P.D.E.*, 34(11):1352–1397, 2009.

D. Morale, V. Capasso, and K. Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations. *Journal of mathematical biology*, 50(1):49–66, 2005.

Charles B. Morrey, Jr. On the derivation of the equations of hydrodynamics from statistical mechanics. *Comm. Pure Appl. Math.*, 8:279–326, 1955.

K. Oelschläger. A sequence of integro-differential equations approximating a viscous porous medium equation. *Z. Anal. Anwendungen*, 20(1):55–91, 2001.

L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev. (2)*, 65:117–149, 1944.

R. Philipowski. Interacting diffusions approximating the porous medium equation and propagation of chaos. *Stochastic Process. Appl.*, 117(4):526–538, 2007.

R. Rossi and G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395—431, 2003.

G. Russo. Deterministic diffusion of particles. *Comm. Pure Appl. Math.*, 43(6):697–733, 1990.

E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser Verlag, Basel, 2015.

S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011.

J. L. Vázquez. An introduction to the mathematical theory of the porous medium equation. In *Shape optimization and free boundaries (Montreal, PQ, 1990)*, volume 380 of *NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci.*, pages 347–389. Kluwer Acad. Publ., Dordrecht, 1992.

J. L. Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007. Mathematical theory.

C. Villani. *Optimal transport : old and new*. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 2009.

J. Zinsl and D. Matthes. Transport distances and geodesic convexity for systems of degenerate diffusion equations. *Calc. Var. Partial Differential Equations*, 54(4):3397–3438, 2015.

M. Burger, A. Esposito – Department Mathematik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany.

Email address: martin.burger@fau.de

Email address: antonio.esposito@fau.de