

# CONGESTED AGGREGATION VIA NEWTONIAN INTERACTION

KATY CRAIG, INWON KIM, AND YAO YAO

ABSTRACT. We consider a *congested aggregation* model that describes the evolution of a density through the competing effects of nonlocal Newtonian attraction and a hard height constraint. This provides a counterpoint to existing literature on repulsive-attractive nonlocal interaction models, where the repulsive effects instead arise from an interaction kernel or the addition of diffusion. We formulate our model as the *Wasserstein gradient flow* of an interaction energy, with a penalization to enforce the constraint on the height of the density. From this perspective, the problem can be seen as a singular limit of the Keller-Segel equation with degenerate diffusion. Two key properties distinguish our problem from previous work on height constrained equations: nonconvexity of the interaction kernel (which places the model outside the scope of classical gradient flow theory) and nonlocal dependence of the velocity field on the density (which causes the problem to lack a comparison principle). To overcome these obstacles, we combine recent results on gradient flows of nonconvex energies with viscosity solution theory. We characterize the dynamics of patch solutions in terms of a Hele-Shaw type free boundary problem and, using this characterization, show that in two dimensions patch solutions converge to a characteristic function of a disk in the long-time limit, with explicit rate of convergence. We believe that a key contribution of the present work is our blended approach, combining energy methods with viscosity solution theory.

## CONTENTS

1.	Introduction	2
2.	Convergence of gradient flows: drift diffusion to height constrained interaction	8
	2.1. Preliminary results	8
	2.2. Definitions of energies and discrete time sequences	12
	2.3. Distance between discrete time sequences of $E_\infty$ , $\tilde{E}_\infty$ , and $E_m$	17
3.	Convergence of viscosity solutions: drift diffusion pressure to free boundary problem	23
	3.1. Basic properties of viscosity solutions of $(P)_m$ and $(P)_\infty$	24
	3.2. Convergence of $(P)_m$ to $(P)_\infty$	28
4.	Long time behavior	37
	4.1. Evolution of the second moment	38
	4.2. Some rearrangement inequalities	40
	4.3. Convergence of energy functional as $t \rightarrow \infty$	43
	4.4. Convergence of $\rho_\infty(t)$ as $t \rightarrow \infty$	46
5.	Appendix	48

---

*Key words and phrases.* Wasserstein metric, gradient flow, discrete gradient flow, JKO scheme, Keller-Segel equation, aggregation equation, crowd motion, Hele-Shaw free boundary problem.

Katy Craig is supported by a UC President's Postdoctoral Fellowship and by the NSF grant DMS-1401867.

Inwon Kim is supported by the NSF grant DMS-1300445.

Yao Yao is supported by the NSF grant DMS-1565480.

5.1. Definition of viscosity solutions of $(P)_\infty$	48
5.2. Further properties of gradient flows of $E_\infty$ , $\tilde{E}_\infty$ , and $E_m$	49
References	55

## 1. INTRODUCTION

In recent years, there has been significant interest in physical and biological models with nonlocal interactions. These models describe the pairwise interactions of a large number of individual agents, for which, in the continuum limit, the nonnegative density  $\rho(x, t)$  satisfies the *aggregation equation with degenerate diffusion*

$$(1.1) \quad \rho_t = \nabla \cdot (\rho \nabla \mathcal{N} * \rho) + \Delta \rho^m,$$

for an interaction kernel  $\mathcal{N} : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $m \geq 1$ . This equation is mass-preserving and, provided that  $\mathcal{N}(x)$  possesses sufficient convexity and regularity, it is a *Wasserstein gradient flow* of the energy

$$E_m(\rho) = W(\rho) + S_m(\rho),$$

where the *interaction energy*  $W(\rho)$  and *Rényi entropy*  $S_m(\rho)$  are given by

$$W(\rho) = \frac{1}{2} \int (\mathcal{N} * \rho)(x) \rho(x) dx \quad \text{and} \quad S_m(\rho) := \frac{1}{m-1} \int \rho(x)^m dx.$$

See section 2.1 for further background on this gradient flow structure, including Remark 2.9 for the case when  $\int \rho \neq 1$ .

Depending on the choice of interaction kernel and diffusion parameter, equations similar to (1.1) arises in a range applications in physics and biology, including models of granular media [6, 18], biological swarming [12, 52], robotic swarming [19, 45], molecular self-assembly [21, 47, 55], and the evolution of vortex densities in superconductors [3, 37, 40, 46]. Of particular interest are kernels and diffusion parameters for which the model exhibits competing repulsive and attractive effects, causing solutions to blow up in finite time or form rich patterns in the asymptotic limit (c.f. [4, 5, 7–9, 23–25, 50]). For example, with  $m \geq 1$  and the interaction is given by the Newtonian interaction kernel

$$(1.2) \quad \mathcal{N}(x) = \begin{cases} \frac{1}{2\pi} \log |x| & \text{for } d = 2, \\ \frac{-1}{d(d-2)\alpha_d} |x|^{2-d} & \text{for } d \neq 2, \end{cases} \quad \text{with } \alpha_d \text{ the volume of the unit ball in } \mathbb{R}^d,$$

equation (1.1) corresponds to the Keller-Segel model for biological chemotaxis [10, 11, 29]

$$(1.3) \quad \rho_t = \nabla \cdot (\rho \nabla (\mathcal{N} * \rho)) + \Delta \rho^m.$$

In this case, the interaction kernel is purely attractive and competes with the repulsion induced by the degenerate diffusion. If  $m > 2 - 2/d$ , diffusion dominates at large density, and bounded solutions exist globally in time [49]. Otherwise, depending on the choice of initial data, solutions with bounded initial data may blow up in finite time.

In the present work, we consider a diffusion-aggregation model similar to the Keller-Segel equation, but with the role of diffusion instead played by a hard height constraint on the

density. Heuristically, the evolution of  $\rho(x, t)$  is given by the *congested aggregation equation*

$$(1.4) \quad \begin{cases} \rho_t = \nabla \cdot (\rho \nabla \mathbf{N}\rho) & \text{if } \rho(x, t) \leq 1, \\ \rho(x, t) \leq 1 & \text{always,} \end{cases}$$

where  $\mathbf{N}\rho := \mathcal{N} * \rho$  denotes the Newtonian potential of  $\rho$ . Informally, solutions of 1.4 seek to evolve according to the “desired velocity field”  $\nabla \mathbf{N}\rho$ , subject to a hard height constraint. More precisely, we define  $\rho(x, t)$  as the Wasserstein gradient flow of the *constrained interaction energy*

$$(1.5) \quad E_\infty(\rho) := \begin{cases} \frac{1}{2} \int \mathbf{N}\rho(x)\rho(x)dx & \text{if } \|\rho\|_\infty \leq 1, \\ +\infty & \text{otherwise.} \end{cases}$$

Our choice of hard height constraint is inspired by the work of Maury, Roudneff-Chupin, Santambrogio, and Venel [41, 42], who introduced such a constraint in their model of pedestrian crowd motion. They considered a *congested drift equation*

$$(1.6) \quad \begin{cases} \rho_t = \nabla \cdot (\rho \nabla V) & \text{if } \rho(x, t) \leq 1, \\ \rho(x, t) \leq 1 & \text{always,} \end{cases}$$

for a local drift  $V : \mathbb{R}^d \rightarrow \mathbb{R}$ , where  $\nabla V$  is the “desired velocity field” of the density. As in the present work, they rigorously defined the evolution of the density as the Wasserstein gradient flow of the *constrained potential energy*

$$V_\infty(\rho) := \begin{cases} \frac{1}{2} \int V(x)\rho(x)dx & \text{if } \|\rho\|_\infty \leq 1, \\ +\infty & \text{otherwise.} \end{cases}$$

They then showed that this gradient flow satisfies a formulation of the continuity equation, where the velocity field is given by the  $L^2$  projection of  $\nabla V$  onto the set of *admissible velocities* that do not increase the density in the saturated zone,  $\{\rho = 1\}$  [41]. Furthermore, when  $V(x)$  is semiconvex (e.g. when  $\nabla^2 V(x)$  is bounded below—see section 2.1) the energy  $V_\infty$  is likewise semiconvex and Wasserstein gradient flow theory ensures that this evolution is unique.

Building upon this work, Alexander, Kim, and Yao [1] showed that solutions of the congested drift equation could be approximated by solutions to a corresponding nonlinear diffusion equation

$$(1.7) \quad \rho_t = \nabla \cdot (\rho \nabla V) + \Delta \rho^m$$

as  $m \rightarrow +\infty$ , which are gradient flows of the energy

$$V_m(\rho) := \int V(x)\rho(x)dx + \frac{1}{m-1} \int \rho(x)^m dx,$$

(Note that, for a fixed  $\rho$ ,  $V_\infty(\rho)$  is the limit of  $V_m(\rho)$  as  $m \rightarrow \infty$ .) They then applied this result to characterize the dynamics of the congested drift equation: given a velocity field satisfying  $\Delta V > 0$  and initial data that is a characteristic function on a patch,  $\rho(x, 0) = \chi_{\Omega_0}(x)$  for

$$\chi_{\Omega_0}(x) := \begin{cases} 1 & \text{if } x \in \Omega_0, \\ 0 & \text{otherwise,} \end{cases}$$

the solution remains a characteristic function, and the evolution of the patch is given by a Hele-Shaw type free boundary problem.

In spite of the similarities between our congested aggregation equation (1.4) and the congested drift equation (1.6), two key differences prevent its analysis by the same methods. First, unlike  $V_\infty$ , the energy  $E_\infty$  does not satisfy the semiconvexity assumptions of classical gradient flow theory that ensure uniqueness. This lack of convexity also makes the equation inaccessible by classical approximation methods—specifically, quantitative approximation by the *discrete gradient flow* or *JKO scheme* for semiconvex energies—which was a key tool in Alexander, Kim, and Yao’s result on the convergence of the nonlinear diffusion equation (1.7) as  $m \rightarrow +\infty$  to the congested drift equation. The second major difference between the congested aggregation and congested drift equations is that the velocity field of the former depends nonlocally on the density. This prevents a direct adaptation of Maury, Roudneff-Chupin, and Santambrogio’s characterization of solutions in terms of a continuity equation, since their argument relies upon an Euler-Lagrange equation for the discrete gradient flow sequence, the proof of which strongly leverages the local nature of the drift. Finally, the nonlocal nature of the velocity field causes there to be no comparison principle, an important element in Alexander, Kim, and Yao’s analysis of the patch dynamics.

To overcome these difficulties, we combine new results on the Wasserstein gradient flow of non-semiconvex energies with a refined approximation of the congested aggregation equation by nonlinear diffusion equations to characterize the dynamics of patch solutions and study their asymptotic behavior. To address the lack of convexity, we appeal to recent work by the first author, inspired by the present problem, that proves well-posedness of Wasserstein gradient flows for energies that are merely  $\omega$ -convex and provides quantitative estimates on the convergence of the *discrete gradient flow*. (See section 2.1). We apply these results to conclude that if the initial data  $\rho_0$  satisfies  $\|\rho_0\|_\infty \leq 1$ , then there exists a unique Wasserstein gradient flow  $\rho_\infty$  of the constrained interaction energy  $E_\infty$ . However, due to the low regularity of  $E_\infty$ , gradient flow theory doesn’t provide a characterization of its evolution in terms of a partial differential equation.

Our goal in this paper is to study the dynamics and asymptotic behavior of  $\rho_\infty$ . We focus on the case when the initial data  $\rho_0$  is a *patch*, i.e.  $\rho_0 = \chi_{\Omega_0}$ , where  $\Omega_0 \subseteq \mathbb{R}^d$  is a bounded domain with Lipschitz boundary, and we seek to answer the following questions:

1. If  $\rho_0$  is a patch, does  $\rho_\infty(\cdot, t)$  remain a patch  $\chi_{\Omega(t)}$  for all  $t \geq 0$ ?
2. If so, what partial differential equation determines the evolution of the set  $\Omega(t)$ ?
3. What is the asymptotic behavior of  $\Omega(t)$  as  $t \rightarrow \infty$ ?

To answer these questions, we blend the gradient flow approach with viscosity solution theory. Due to the attractive nature of the Newtonian kernel (1.2), we show that the solution of the congested aggregation equation  $\rho_\infty(x, t)$  indeed remains a patch:  $\rho_\infty(x, t) = \chi_{\Omega(t)}(x)$  for a time dependent domain  $\Omega(t)$ . We then show that  $\Omega(t)$  evolves with outward normal velocity  $V = V(x, t)$  satisfying

$$V = -\nu \cdot (\nabla p + \nabla \mathbf{N} \rho_\infty) \quad \text{at } x \in \partial\Omega(t),$$

where  $\nu = \nu(x, t)$  is the outward unit normal at  $x \in \partial\Omega(t)$  and, for each  $t > 0$ ,  $p = p(x, t)$  solves

$$-\Delta p(\cdot, t) = 1 \text{ in } \Omega(t), \quad p(\cdot, t) = 0 \text{ outside of } \Omega(t).$$

Since,  $\Omega(t) = \{p(\cdot, t) > 0\}$ , this gives a Hele-Shaw type free boundary problem for the pressure variable  $p$ ,

$$(P) \quad \begin{cases} -\Delta p(\cdot, t) = 1 & \text{in } \{p > 0\}; \\ V = -\nu \cdot (\nabla p + \nabla \Phi) & \text{on } \partial\{p > 0\}; \\ \Phi = \mathbf{N}\chi_{\{p > 0\}}. \end{cases}$$

Provided that  $p$  is sufficiently regular (for example,  $p \in L^1([0, \infty); H^1(\mathbb{R}^d))$ ), this would imply that the solution of the congested aggregation is a weak solution of the continuity equation

$$(1.8) \quad \rho_t = \nabla \cdot (\rho(\nabla \mathbf{N}p + \nabla p)),$$

where  $\nabla p$  is the pressure generated by the height constraint that modifies the “desired velocity field”  $\nabla \mathbf{N}p$ . In terms of  $p$ ,  $\nu = -\nabla p/|\nabla p|$  and  $V = p_t/|\nabla p|$ , so in the smooth setting the second condition in (P) can be written as

$$p_t = |\nabla p|^2 + \nabla p \cdot \nabla \Phi \text{ on } \partial\{p > 0\}.$$

Even if  $\Omega_0$  has smooth boundary, the evolving set  $\Omega(t) = \{p(\cdot, t) > 0\}$  may undergo topological changes such as merging. Consequently, to describe the evolution of  $\Omega(t)$ , we require a notion of weak solution for (P). While viscosity solutions are a natural choice, given their utility in free boundary problems, because of the nonlocal dependence of the outward normal velocity  $V$  on  $p$  itself, (P) lacks a comparison principle. Instead, we consider an auxillary problem for a *fixed*, nonnegative function  $\rho(x, t) \in L^\infty(\mathbb{R}^d \times (0, \infty))$ ,

$$(P)_\infty \quad \begin{cases} -\Delta p(\cdot, t) = 1 & \text{in } \{p > 0\}; \\ V = -\nu \cdot (\nabla p + \nabla \Phi) & \text{on } \partial\{p > 0\}; \\ \Phi = \mathbf{N}\rho. \end{cases}$$

We show that the comparison principle holds for  $(P)_\infty$ , hence viscosity solution theory applies. We then define  $p$  to be a solution of (P) if it is a weak viscosity solution of  $(P)_\infty$  with  $\rho = \chi_{\{p > 0\}}$  almost everywhere.

We now state our first main result, which follows from Theorems 3.8 and 3.14.

**Theorem 1.1** (Characterization of dynamics of aggregation patches).

(a) Let  $\Omega_0 \subseteq \mathbb{R}^d$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty(\cdot, t) \in L^\infty(\mathbb{R}^n)$  be the gradient flow of  $E_\infty$  with initial data  $\chi_{\Omega_0}$ . Consider the free boundary problem  $(P)_\infty$  with  $\rho$  replaced by  $\rho_\infty$ , and the initial data  $p_0$  given by

$$(1.9) \quad -\Delta p_0(\cdot, 0) = 1 \text{ in } \Omega_0, \quad p_0(\cdot, 0) = 0 \text{ outside of } \Omega_0.$$

Then there is a unique minimal viscosity solution  $p(x, t)$  of  $(P)_\infty$  with initial data  $p_0$ .

(b) Let  $\Omega(t) = \{p(\cdot, t) > 0\}$ . Then  $\rho_\infty(\cdot, t)$  remains a patch for all times, and

$$\rho_\infty(\cdot, t) = \chi_{\Omega(t)} \text{ a.e. for all } t \geq 0.$$

(c) Therefore,  $p$  is a weak solution of (P) in the sense of Definition 3.2.

Next, we consider the asymptotic behavior of patch solutions as  $t \rightarrow +\infty$ . For any given mass and any dimension, the Riesz rearrangement inequality [35, Theorem 3.7] immediately gives that the global minimizer of the constrained interaction energy (1.5) must be a characteristic function of a ball. However, this does not guarantee that the gradient flow  $\rho_\infty(t)$  of the constrained interaction energy always converges to a translation of the global minimizer as  $t \rightarrow +\infty$ . In particular, the main obstacle is to show the mass of  $\rho_\infty(t)$  cannot escape to

infinity in the long time limit, which requires us to obtain some compactness estimates on  $\rho_\infty(t)$  uniformly in time.

For the Keller-Segel equation (1.3) with subcritical power  $m > 2 - 2/d$ , the situation is very similar. Again, there exists a unique (up to a translation) global minimizer of  $E_m$  for any given mass [36, 38], but it is unknown whether solutions converge to it as  $t \rightarrow +\infty$ . In dimension two, convergence has been recently shown by Carrillo, Hittmeir, Volzone, and the third author [16], where compactness is obtained via a uniform in time bound of the second moment, though no explicit convergence rate towards the global minimizer is given. For  $d \geq 3$ , the only available convergence result towards the global minimizer is work by the second two authors on radial solutions [31].

In our work, for dimension two, we not only prove convergence of solutions towards the global minimizer of the constrained interaction energy (1.5), but also provide explicit estimates on the rate of convergence. We accomplish this by again applying a blended approach, combining the gradient flow structure of the problem with viscosity solution theory and the characterization of patch dynamics from Theorem 1.1. We begin by using a rearrangement inequality of Talenti [51] to show that the second moment of  $\rho_\infty(t)$  is non-increasing in time and is strictly decreasing at time  $t$  unless  $\Omega(t)$  is a disk. Then, applying a quantitative version of the isoperimetric inequality due to Fusco, Maggi, and Pratelli [26] and our characterization of patch dynamics, Theorem 1.1, we provide explicit estimates on the rate that the second moment is decreasing, in terms of the symmetric difference between  $\Omega(t)$  and a disk. Finally, using the gradient flow structure of the problem, we show that as  $t \rightarrow +\infty$ ,  $\rho_\infty(t)$  strongly converges to a characteristic function of a disk in  $L^q$  for any  $1 \leq q < \infty$ , and its energy  $E_\infty(\rho_\infty)$  converges to its global minimizer with an explicit rate. This gives our second main result, which combines Theorems 4.9 and 4.12.

**Theorem 1.2** (Long time behavior in two dimensions). *Assume  $d = 2$ . Let  $\Omega_0 \subseteq \mathbb{R}^2$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty$  be the gradient flow of  $E_\infty$  with initial data  $\chi_{\Omega_0}$ . Then as  $t \rightarrow +\infty$ ,  $\rho_\infty(\cdot, t)$  converges to  $\chi_{B_0}$  in  $L^q$  for any  $1 \leq q < \infty$ , where  $B_0$  is the unique disk with the same area and center of mass as those of  $\Omega_0$ . Furthermore, we have the following rate of convergence in terms of the free energy,*

$$0 \leq E_\infty(\rho_\infty(\cdot, t)) - E_\infty(\chi_{B_0}) \leq C(|\Omega_0|, M_2[\Omega_0])t^{-1/6}.$$

*Remark 1.3.* Let us point out that our control for the second moment relies on the particular structure for the 2D Newtonian kernel, and we are unable to obtain similar compactness estimates for higher dimensions. For  $d \geq 3$ , whether  $\rho_\infty(t)$  converges to a ball as  $t \rightarrow \infty$  remains an interesting open question.

We now describe the key ingredients in our characterization of the dynamics of the congested aggregation equation. At the heart of our analysis is an approximation of this equation as the singular limit of a sequence of nonlinear diffusion equations. This provides the bridge between the gradient flow and viscosity solution approach. In particular, while the gradient flow of  $E_\infty$  is merely a curve in the space of measures, approximating it by a sequence of solutions to nonlinear diffusion equations allows us to bring to bear the tools of viscosity solution theory in the limit.

Following the analogy with Alexander, Kim, and Yao's previous work, one might hope to approximate the congested aggregation equation by the Keller-Segel equation (1.3), which also

has a gradient flow structure corresponding to the energy

$$(1.10) \quad E_{KS}(\rho) = \iint \rho(x) \mathbf{N} \rho(x) dx + \frac{1}{m-1} \int \rho(x)^m dx$$

Note that for a fixed  $\rho$ ,  $E_\infty(\rho)$  is the limit of  $E_{KS}(\rho)$  as  $m \rightarrow \infty$ . However,  $E_{KS}$  satisfies neither the classical assumptions for semiconvexity nor the weaker assumptions for  $\omega$ -convexity. Consequently, we lack the quantitative estimates on the rate of convergence of the discrete gradient flow that are an essential element of our approach. Instead, we replace the nonlocal potential  $\mathbf{N}\rho$  in  $E_{KS}$  with a local potential  $\Phi_{1/m}(x, t)$  that depends on time, the initial data  $\rho(x, 0)$ , and the diffusion parameter  $m \geq 1$ . (See Definition 2.10 for a precise definition of this potential.) This leads to the energy

$$E_{m,t}(\rho) := \int \rho(x) \Phi_{1/m}(x, t) dx + \frac{1}{m-1} \int \rho(x)^m dx,$$

which we can show is  $\omega$ -convex. We then prove that the (time dependent) gradient flow of this energy, which corresponds to a solution of

$$(PME-D)_m \quad \rho_t = \nabla \cdot (\rho \nabla \Phi_{1/m}) + \Delta \rho^m,$$

converges as  $m \rightarrow +\infty$  to a solution of the congested aggregation equation. (See section 2 for our construction of this time dependent gradient flow.) Then, rewriting  $(PME-D)_m$  in the form

$$(1.11) \quad \rho_t = \nabla \cdot (\rho (\nabla \Phi_{1/m} + \nabla p_m)), \quad \text{for } p_m = \frac{m-1}{m} \rho^{m-1},$$

we use viscosity solution theory to show that, as  $m \rightarrow +\infty$ ,  $p_m$  converges to a solution of the free boundary problem (P). By uniqueness of the limit, we conclude the characterization of dynamics of patch solutions of the congested aggregation equation, as stated in Theorem 1.1.

Our paper is organized as follows. In section 2, we prove that the solutions of the nonlinear diffusion equations  $(PME-D)_m$  converge as  $m \rightarrow +\infty$  to the gradient flow of  $E_\infty$  with an explicit rate depending on  $m$ . We also provide background on Wasserstein gradient flow, including recent results by the first author on the gradient flows of  $\omega$ -convex energies. In section 3, we show that the pressure  $p_m$  corresponding to the nonlinear diffusion equations, given in equation (1.11), converges as  $m \rightarrow +\infty$  to a solution of  $(P)_\infty$ . Combining these results, we show that the gradient flow of  $E_\infty$  is a characteristic function of the evolving set  $\Omega(t)$  and that  $\Omega(t)$  can be obtained from the viscosity solution of  $(P)_\infty$ . In section 4, we consider the asymptotic behavior of  $\rho_\infty$  in two dimensions, proving that it converges to a disk with explicit rate. Let us remark that the characterization of  $\rho_\infty$  by the pressure variable  $p$  plays a crucial role in the proof of this asymptotic result. Finally, we conclude with an appendix section 5, which contains proofs of several lemmas from section 2, as well as definitions of viscosity solutions for the limiting free boundary problem  $(P)_\infty$ .

There are several directions for future work. First, our analysis only addresses solutions that are initially a patch. Results for more general initial data could leverage recent work by Kim and Pozar [30] and Mellet, Perthame, and Quiros [44]. Second, in the light of work by Maury, Roudneff-Chupin, and Santambrogio [41], it would be interesting if one could characterize the modified velocity  $\nabla p + \nabla \mathbf{N}\rho$  in (1.8) as the projection of the original velocity  $\nabla \mathbf{N}\rho$  onto the space of admissible velocities under the height constraint. At the moment, this appears to be a difficult question, due to the highly nonlinear nature of the projection and its dependence on the solution. A third direction for future work would be to pursue to what extent our

analysis extends to nonlocal velocity field generated by kernels aside from the Newtonian  $\mathcal{N}$ , which arise in a range of biological and physical applications. While our result on the singular limit of the nonlinear diffusion equations extends to a range of kernels (see Remark 2.1), our analysis of the free boundary problem strongly leverages the structure of the Newtonian kernel. A final direction for future work would be to make rigorous the link between the congested aggregation equation and the Keller-Segel equation (1.3) as  $m \rightarrow +\infty$ , completing the analogy with previous work by Alexander, Kim, and Yao that found that the hard height constraint may be obtained as the limit of slow diffusion.

## 2. CONVERGENCE OF GRADIENT FLOWS: DRIFT DIFFUSION TO HEIGHT CONSTRAINED INTERACTION

In this section, we show that the gradient flow of the height constrained interaction energy  $E_\infty$ , defined in equation (1.5), may be approximated by solutions of the nonlinear diffusion equations (PME-D) $_m$  as  $m \rightarrow +\infty$ . This provides a link between the abstract Wasserstein gradient flow of  $E_\infty$ , which in general is merely a curve in the space of probability measures, and solutions to partial differential equations.

*Remark 2.1* (Choice of interaction kernel). For the sake of continuity with sections 3 and 4, we assume that the interaction kernel  $\mathcal{N}$  is Newtonian (1.2). However, our results in this section may be extended to any kernels that satisfy [20, Assumption 4.1] and the estimates of Proposition 2.3. In particular, this includes many repulsive-attractive potentials of interest in the literature.

**2.1. Preliminary results.** We begin by collecting some results on the Wasserstein gradient flow of  $\omega$ -convex energies that will be useful in what follows. For further background on the Wasserstein metric and gradient flows of semiconvex energies, we refer the reader to the books by Ambrosio, Gigli, and Savaré [2] and Villani [54]. For more details on gradient flows of  $\omega$ -convex energies, see recent work by the first author [20].

Let  $\mathcal{P}_2(\mathbb{R}^d)$  denote the set of probability measures on  $\mathbb{R}^d$  with finite second moment, i.e.  $\int |x|^2 d\mu < +\infty$ . If a measure  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  is absolutely continuous with respect to Lebesgue measure ( $\mu \ll \mathcal{L}^d$ ), we will identify  $\mu$  with its density, i.e.  $d\mu(x) = \mu(x)dx$ . In particular, we write  $\|\mu\|_{L^\infty} < +\infty$  if  $d\mu(x) = \mu(x)dx$  and  $\mu(x) \in L^\infty(\mathbb{R}^d)$ .

Given  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ , a measurable function  $\mathbf{t} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  transports  $\mu$  onto  $\nu$  in case  $\int f(\mathbf{t}(x))d\mu = \int f(y)d\nu$  for all  $f \in L^1(d\nu)$ . We then call  $\nu$  the *push-forward of  $\mu$  under  $\mathbf{t}$*  and write  $\nu = \mathbf{t}\#\mu$ . If  $\mu$  is absolutely continuous with respect to Lebesgue measure (as will be the case for all the measures we consider), then the *Wasserstein distance* from  $\mu$  to  $\nu$  is given by

$$(2.1) \quad W_2(\mu, \nu) = \inf \left\{ \left( \int |\mathbf{t} - \mathbf{id}|^2 d\mu \right)^{1/2} : \mathbf{t}\#\mu = \nu \right\},$$

where  $\mathbf{id}(x) = x$ . Furthermore, the infimum is attained by an *optimal transport map*  $\mathbf{t} = \mathbf{t}_\mu^\nu$ , which is unique  $\mu$ -almost everywhere.

The metric space  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$  is complete, and convergence can be characterized as

$$W_2(\mu_n, \mu) \rightarrow 0 \iff \int f d\mu_n \rightarrow \int f d\mu \text{ for all } f \in C(\mathbb{R}^d) \text{ such that} \\ \exists C > 0, x_0 \in \mathbb{R}^d \text{ so that } |f(x)| \leq C(1 + |x - x_0|^2).$$

We will refer to such  $f$  as *continuous functions with at most quadratic growth*. Furthermore, for any  $f \in C^1(\mathbb{R}^d)$  with uniformly bounded gradient, we can quantify the difference between the integral of  $f$  against  $\mu$  and the integral of  $f$  against  $\nu$  using the following elementary lemma.

**Lemma 2.2.** *For  $f \in C^1(\mathbb{R}^d)$  and  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ ,*

$$\left| \int f d\mu - \int f d\nu \right| \leq \|\nabla f\|_\infty W_2(\mu, \nu).$$

*Proof.* For simplicity, suppose that  $\mu \ll \mathcal{L}^d$ , so there exists an optimal transport map  $\mathbf{t}_\mu^\nu$ . (The proof is identical in the general case, using optimal transport plans.) By Jensen's inequality,

$$\left| \int f d\mu - \int f d\nu \right| \leq \int |f - f \circ \mathbf{t}_\mu^\nu| d\mu \leq \|\nabla f\|_\infty \left( \int |\mathbf{t}_\mu^\nu - \mathbf{id}|^2 d\mu \right)^{1/2} = \|\nabla f\|_\infty W_2(\mu, \nu).$$

□

Along with its metric structure,  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$  is a *geodesic space*, since any two measures  $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$  are connected by a *geodesic*  $\mu_\alpha \in \mathcal{P}_2(\mathbb{R}^d)$ ,  $\alpha \in [0, 1]$ , satisfying

$$W_2(\mu_\alpha, \mu_\beta) = |\beta - \alpha| W_2(\mu_0, \mu_1) \text{ for all } \alpha, \beta \in [0, 1].$$

If  $\mu_0 \ll \mathcal{L}^d$ , then the geodesic from  $\mu_0$  to any  $\mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$  is unique and of the form

$$\mu_\alpha = ((1 - \alpha)\mathbf{id} + \alpha\mathbf{t}_{\mu_0}^{\mu_1})\# \mu_0.$$

Unlike a square Hilbertian norm, the square Wasserstein distance is not convex along geodesics ( $\alpha \mapsto W_2^2(\nu, \mu_\alpha)$  is not convex) [2, Example 9.1.5]. Consequently, Ambrosio, Gigli, and Savaré introduced an expanded class of curves known as *generalized geodesics*, so that, between any two measures, there is always at least one curve along which the square distance is convex [2, Lemma 9.2.1, Definition 9.2.2]. Given  $\mu_0, \mu_1, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\nu \ll \mathcal{L}^d$ , the *generalized geodesic from  $\mu_0$  to  $\mu_1$  with base  $\nu$*  is

$$\mu_\alpha = ((1 - \alpha)\mathbf{t}_\nu^{\mu_0} + \alpha\mathbf{t}_\nu^{\mu_1})\#\nu,$$

and along such a curve we have

$$W_2^2(\nu, \mu_\alpha) = (1 - \alpha)W_2^2(\nu, \mu_0) + \alpha W_2^2(\nu, \mu_1) - \alpha(1 - \alpha)\|\mathbf{t}_\nu^{\mu_0} - \mathbf{t}_\nu^{\mu_1}\|_{L^2(d\nu)}^2.$$

An additional class of curves along which the square Wasserstein metric is convex are *linear interpolations* of measures,

$$\mu_\alpha := (1 - \alpha)\mu_0 + \alpha\mu_1.$$

For any  $\mu_0, \mu_1, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ , we have

$$(2.2) \quad W_2^2(\nu, \mu_\alpha) \leq (1 - \alpha)W_2^2(\nu, \mu_0) + \alpha W_2^2(\nu, \mu_1).$$

(See, for example, [48, Proposition 7.19].)

Due to the fact that  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$  is a geodesic space, it induces a natural notion of convexity on energy functionals  $E : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$ , i.e. given a geodesic  $\mu_\alpha$ , the function  $\alpha \mapsto E(\mu_\alpha)$  is convex. We recall both this standard notion of convexity, as well as two generalizations: semiconvexity and  $\omega$ -convexity.

- (i)  $E$  is *convex* along  $\mu_\alpha$  if  $E(\mu_\alpha) \leq (1 - \alpha)E(\mu_0) + \alpha E(\mu_1)$ .
- (ii)  $E$  is *semiconvex* along  $\mu_\alpha$  if there exists  $\lambda \in \mathbb{R}$  so that  $E(\mu_\alpha) \leq (1 - \alpha)E(\mu_0) + \alpha E(\mu_1) - \alpha(1 - \alpha)\frac{\lambda}{2}W_2^2(\mu_0, \mu_1)$ .

- (iii)  $E$  is  $\omega$ -convex along  $\mu_\alpha$  if there exists  $\lambda_\omega \in \mathbb{R}$  and a continuous, nondecreasing function  $\omega : [0, +\infty) \rightarrow [0, +\infty)$ , which vanishes only at  $x = 0$ , so that
- $$E(\mu_\alpha) \leq (1-\alpha)E(\mu_0) + \alpha E(\mu_1) - \frac{\lambda_\omega}{2} [(1-\alpha)\omega(\alpha^2 W_2^2(\mu_0, \mu_1)) + \alpha\omega((1-\alpha)W_2^2(\mu_0, \mu_1))].$$

If, for any  $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ , there exists a geodesic  $\mu_\alpha$  from  $\mu_0$  to  $\mu_1$  along which  $E$  satisfies one of the above inequalities, we say  $E$  is *convex/semiconvex/ $\omega$ -convex along geodesics*. On the other hand, if for any  $\mu_0, \mu_1, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ , there exists a generalized geodesic  $\mu_\alpha$  from  $\mu_0$  to  $\mu_1$  with base  $\nu$  along which  $E$  satisfies one of the above inequalities (replacing  $W_2(\mu_0, \mu_1)$  on the right hand side with  $\|\mathbf{t}_\nu^{\mu_0} - \mathbf{t}_\nu^{\mu_1}\|_{L^2(\nu)}^2$ ), we say  $E$  is *convex/semiconvex/ $\omega$ -convex along generalized geodesics*. We will also say that  $E$  is *proper* if the *domain* of the energy  $D(E) = \{\mu : E(\mu) < +\infty\}$  is nonempty.

A key element of our analysis is that the height constrained interaction energy  $E_\infty$  defined in equation (1.5) is  $\omega$ -convex along generalized geodesics. This follows from the following estimates on the Newtonian potential of a bounded, integrable function.

**Proposition 2.3** (c.f. [39, Theorem 2.7]). *Suppose  $\rho, \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\rho\|_\infty, \|\mu\|_\infty \leq 1$ . Then there exists  $C_d \geq 1$ , depending only on the dimension, so that*

$$\begin{aligned} \|\nabla \mathbf{N}\rho\|_\infty &\leq C_d, \quad \|\Delta \mathbf{N}\rho\|_\infty \leq 1, \quad \int \mathbf{N}\rho d\nu \geq -C_d, \\ |\nabla \mathbf{N}\rho(x) - \nabla \mathbf{N}\rho(y)| &\leq C_d \sigma(|x - y|), \quad \text{and} \quad \|\nabla \mathbf{N}\rho - \nabla \mathbf{N}\mu\|_{L^2(\mathbb{R}^d)} \leq W_2(\rho, \mu). \end{aligned}$$

where

$$(2.3) \quad \sigma(x) := \begin{cases} 2x|\log x| & \text{if } 0 \leq x \leq e^{(-1-\sqrt{2})/2}, \\ \sqrt{x^2 + 2(1+\sqrt{2})e^{-1-\sqrt{2}}} & \text{if } x > e^{(-1-\sqrt{2})/2}. \end{cases}$$

We defer the proof of this proposition to the appendix in Section 5.2.

By the above estimates and [20, Theorem 4.3, Proposition 4.4],  $E_\infty$  is  $\omega$ -convex along generalized geodesics with  $\lambda_\omega = -C_d$  and  $\omega(x)$  a log-Lipschitz modulus of convexity

$$(2.4) \quad \omega(x) = \begin{cases} x|\log x| & \text{if } 0 \leq x \leq e^{-1-\sqrt{2}}, \\ \sqrt{x^2 + 2(1+\sqrt{2})e^{-1-\sqrt{2}}} & \text{if } x > e^{-1-\sqrt{2}}. \end{cases}$$

The  $\omega$ -convexity of  $E_\infty$  then leads to the following result on the well-posedness of the gradient flow:

**Theorem 2.4** ([20, Theorem 4.3, Proposition 4.4]). *For any  $\rho_0 \in D(E_\infty)$  (that is,  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\rho_0\|_\infty \leq 1$ ), the gradient flow  $\rho_\infty(t)$  of  $E_\infty$  with initial data  $\rho_0$  is well-posed. Specifically  $\rho_\infty : (0, +\infty) \rightarrow \mathcal{P}_2(\mathbb{R}^d)$  is the unique curve that is locally absolutely continuous in time, with  $\rho_\infty(t) \xrightarrow{t \rightarrow 0} \rho_0$  and*

$$(2.5) \quad \frac{1}{2} \frac{d}{dt} W_2^2(\rho_\infty(t), \nu) + \frac{\lambda_\omega}{2} \omega(W_2^2(\rho_\infty(t), \nu)) \leq E(\nu) - E(\rho_\infty(t)), \quad \forall \nu \in D(E_\infty), \quad \text{a.e. } t > 0.$$

In order to provide a PDE characterization of  $\rho_\infty(x, t)$  in Section 3, we use the following higher regularity of  $\rho_\infty(x, t)$  and  $\nabla \mathbf{N}\rho_\infty(x, t)$ , which we prove in appendix Section 5.2.

**Proposition 2.5** (time regularity of the gradient flow of  $E_\infty$ ). *Suppose  $\rho_\infty(x, t)$ , with initial data  $\rho_\infty(x, 0) \in D(E_\infty)$ , is a gradient flow of  $E_\infty$ . Then  $W_2(\rho_\infty(t), \rho_\infty(s)) \leq 2C_d|t - s|$ , where  $C_d > 0$  is as in Proposition 2.3.*

**Proposition 2.6.** *Suppose  $\rho_\infty(x, t)$ , with initial data  $\rho_\infty(x, 0) \in D(E_\infty)$ , is a gradient flow of  $E_\infty$ . Then  $\nabla \mathbf{N} \rho_\infty(x, t)$  is log-Lipschitz in space and  $1/2d$ -Hölder continuous in time. In particular, with  $C_d > 0$  and  $\sigma(x)$  as in Proposition 2.3,*

$$|\nabla \mathbf{N} \rho_\infty(x, t) - \nabla \mathbf{N} \rho_\infty(y, t)| \leq C_d \sigma(|x - y|) \quad \text{for all } x, y \in \mathbb{R}^d, t \geq 0,$$

$$|\nabla \mathbf{N} \rho_\infty(x, t) - \nabla \mathbf{N} \rho_\infty(x, s)| \leq 10C_d |t - s|^{1/2d} \quad \text{for all } 0 < |t - s| < e^{(-1-\sqrt{2})/2}, x \in \mathbb{R}^d.$$

An important tool in the analysis of Wasserstein gradient flows is a discrete time approximation of gradient flows known as the *discrete gradient flow* or *JKO scheme* [27]. This scheme is analogous to the implicit Euler method for approximation of ordinary differential equations in Euclidean space. For any  $\mu \in D(E_\infty)$  and time step  $\tau > 0$ , the *discrete gradient flow* of  $E_\infty$  is given by

$$\rho_\tau^n \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho_\tau^{n-1}, \nu) + E_\infty(\nu) \right\} \quad \text{and } \rho_\tau^0 := \rho.$$

By [20, Theorem 4.3, Proposition 4.4], the discrete gradient flow of  $E_\infty$  exists for all  $\rho \in D(E_\infty)$  and  $\tau > 0$ , and if  $\tau = t/n$  for any  $t \geq 0$ , the discrete gradient flow converges to the continuous gradient flow,

$$\lim_{n \rightarrow +\infty} W_2(\rho_{t/n}^n, \rho_\infty(t)) = 0.$$

As demonstrated in previous work by the first author [20], well-posedness of the gradient flows of  $\omega$ -convex energies is closely related to the well-posedness of the ODE

$$(2.6) \quad \begin{cases} \frac{d}{dt} F_t(x) = -C_d \omega(F_t(x)), \\ F_0(x) = x. \end{cases}$$

For  $\omega(x)$  as in equation (2.4),  $0 \leq x \leq e^{-1-\sqrt{2}}$ , and  $t \geq 0$ , the solution is given by  $F_t(x) = x e^{C_d t}$ . Furthermore, for all  $x, t \geq 0$ ,  $F_t(x)$  is nondecreasing in space and nonincreasing in time.

In a similar way, analysis of the discrete gradient flow of  $E_\infty$  is closely related to a discrete time approximation of (2.6). In particular, we define

$$f_\tau(x) := \begin{cases} x - C_d \tau \omega(x) & \text{if } x \geq 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

so that  $f_\tau^{(m)}(x)$  is the  $m$ th step of the explicit Euler method with time step  $\tau$ . In the following proposition, we recall some properties of the function  $f_\tau(x)$  that will be useful in our estimates of the discrete time sequences.

**Proposition 2.7** (properties of  $f_\tau(x)$ ).

- (i) *If  $0 \leq x \leq y \leq r$ , there exists  $c_r > 0$  so that  $f_\tau(x) \leq f_\tau(y) + C_d^2 c_r^2 \tau^2$ .*
- (ii) *For all  $x, y \geq 0$ ,  $f_\tau(x + y) \leq f_\tau(x) + y$ .*
- (iii) *For all  $x, t \geq 0$ ,  $|F_t(x) - f_{t/n}^{(n)}(x)| \leq C_d \omega(x) t/n$ .*

*Proof.* (i) and (ii) are consequences of [20, Lemma 2.25]. (iii) is a consequence of [20, Proposition 2.24] and the fact that  $F_t(x)$  is nonincreasing in time.  $\square$

Finally, we recall a contraction inequality for the discrete gradient flow of an  $\omega$ -convex energy, which we use to conclude stability of the discrete gradient flow sequences.

**Proposition 2.8** (contraction inequality). *Let  $E : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R} \cup \{+\infty\}$  be proper, lower semi-continuous, bounded below, and  $\omega$ -convex along generalized geodesics, for  $\omega(x)$  as in equation (2.4) and  $\lambda_\omega \leq 0$ . Fix  $\rho, \mu \in D(E)$  and, for  $\tau > 0$ , choose  $\rho_\tau$  and  $\mu_\tau$  satisfying*

$$\rho_\tau \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho, \nu) + E(\nu) \right\} \quad \text{and} \quad \mu_\tau \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\mu, \nu) + E(\nu) \right\}.$$

*Then there exist positive constants  $C$  and  $\tau_*$  depending on  $W_2(\rho, \mu)$ ,  $\lambda_\omega$ ,  $E(\mu)$ , and  $E(\nu)$  so that for all  $0 < \tau < \tau_*$ ,*

$$f_\tau^{(2)}(W_2^2(\rho_\tau, \mu_\tau)) \leq W_2^2(\rho, \mu) + |\lambda_\omega| \tau \omega(CW_2(\mu, \mu_\tau)) + 2\tau(E(\rho) - E(\rho_\tau)) + C\tau^2.$$

*Proof.* This is a particular case of [20, Theorem 3.2]. □

*Remark 2.9* (Wasserstein gradient flow of measures with mass not equal to 1). We conclude by observing that the gradient flow theory can be easily extended to nonnegative measures whose integral is not equal to 1. For a fixed  $A > 0$ , let  $\mathcal{P}_{2,A}(\mathbb{R}^d)$  denote the set of non-negative measures that integrate to  $A$  and have finite second moment. For  $\mu, \nu \in \mathcal{P}_{2,A}(\mathbb{R}^d)$  (with the same  $A$ ), we can then define  $W_2(\mu, \nu)$  in the same way as in (2.1), and given initial data  $\rho_0 \in \mathcal{P}_{2,A}(\mathbb{R}^d)$  with  $\|\rho_0\|_\infty \leq 1$ , the same arguments lead to the well-posedness of a gradient flow  $\rho_\infty : (0, +\infty) \rightarrow \mathcal{P}_{2,A}(\mathbb{R}^d)$  of  $E_\infty$ . However, for the sake of simplicity, we will assume that  $\rho_0$  is a probability measure for the remainder of this section.

**2.2. Definitions of energies and discrete time sequences.** We now turn to the definitions of the energies and discrete time sequences that we will use to show that solutions of the nonlinear diffusion equations  $(\text{PME-D})_m$  converge as  $m \rightarrow +\infty$  to the the gradient flow of the height constrained interaction energy  $E_\infty$ . We begin by defining the local potential  $\Phi_{1/m}(x, t)$ , which induces the drift in  $(\text{PME-D})_m$ . As described in the introduction, previous work by Alexander, Kim, and Yao suggests that the gradient flow of  $E_\infty$  should be obtained as the limit of the gradient flows of the Keller-Segel energy  $E_{KS}$ , defined in equation (1.10). However, we lack sufficient convexity of  $E_{KS}$  to prove this rigorously. Instead, we replace the nonlocal potential  $\mathbf{N}\rho$  in  $E_{KS}$  with a local potential  $\Phi_{1/m}(x, t)$  that depends on time, the initial data  $\rho_0(x)$  of the gradient flow of  $E_\infty$ , and the diffusion parameter  $m \geq 1$ .

**Definition 2.10** (local potential  $\Phi_{1/m}(x, t)$ ). *Given initial data  $\rho_0$ , let  $\rho_\infty(x, t)$  be the gradient flow of the height constrained interaction energy  $E_\infty$ . Fix a mollifier  $\psi \in C_c^\infty(\mathbb{R}^d)$  satisfying  $\psi \geq 0$  and  $\int \psi = 1$ , and let  $\psi_{1/m}(x) = m^d \psi(mx)$ . Then, for any  $m > 1$ , define*

$$(2.7) \quad \Phi(x, t) = \mathbf{N}\rho_\infty(x, t) \quad \text{and} \quad \Phi_{1/m}(x, t) = \psi_{1/m} * \mathbf{N}\rho_\infty(x, t).$$

This definition is guided by the following intuition: given initial data  $\rho_0$ , one heuristically expects that the gradient flow of  $E_{KS}$  should converge to  $\rho_\infty$ . Consequently, if we replace  $\mathbf{N}\rho$  in the definition of  $E_{KS}$  by  $\mathbf{N}\rho_\infty$ , we expect that the gradient flow of this new energy will still converge to  $\rho_\infty$  as  $m \rightarrow +\infty$ . We include the extra mollification on the potential to leverage the existing theory on the porous medium equation with drift, which requires the potential to be twice continuously differentiable in space. By Proposition 2.6,  $\nabla\Phi(x, t)$  is log-Lipschitz in space, hence  $\nabla\Phi_{1/m} = \psi_{1/m} * \nabla\Phi$  converges to  $\nabla\Phi$  uniformly on  $\mathbb{R}^d \times [0, +\infty)$ . Furthermore, by Proposition 2.3,

$$(2.8) \quad \|\nabla\Phi_{1/m}(\cdot, t)\|_\infty \leq C_d, \quad \|\Delta\Phi_{1/m}(\cdot, t)\|_\infty \leq 1.$$

With this precise definition of the drift arising in  $(\text{PME-D})_m$  in hand, we now turn to the definitions of the the three energy functionals that we use in our analysis of the limit of  $(\text{PME-D})_m$  as  $m \rightarrow +\infty$ .

**Definition 2.11** (energy functionals). *Fix  $\psi \in C_c^\infty(\mathbb{R}^d)$  as in Definition 2.10 and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\mu\|_\infty \leq 1$ . For any  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ , define*

$$\begin{aligned} E_\infty(\rho) &:= \begin{cases} \frac{1}{2} \int \mathbf{N}\rho(x) d\rho(x) & \text{if } \|\rho\|_\infty \leq 1, \\ +\infty & \text{otherwise;} \end{cases} \\ \tilde{E}_\infty(\rho; \mu) &:= \begin{cases} \int \mathbf{N}\mu(x) d\rho(x) & \text{if } \|\rho\|_\infty \leq 1, \\ +\infty & \text{otherwise;} \end{cases} \\ E_m(\rho; \mu) &:= \begin{cases} \frac{1}{m-1} \int \rho(x)^m dx + \int \psi_{1/m} * \mathbf{N}\mu(x) d\rho(x) & \text{if } \rho \ll \mathcal{L}^d, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

As shown in previous work by the first author, the gradient flows of the above energies are well-posed [20, Theorem 4.3, Proposition 4.4]. In particular, while these energies fall outside the scope of the theory of gradients flows of semiconvex energies, all three energies are instead  $\omega$ -convex along generalized geodesics for  $\lambda_\omega = -C_d$ , as in Proposition 2.3, and  $\omega(x)$  a log-Lipschitz modulus of convexity, as in equation (2.4). The third energy is also  $\lambda$ -convex along generalized geodesics for  $\lambda = \lambda(m) \xrightarrow{m \rightarrow +\infty} -\infty$  [2, Proposition 9.3.2, Proposition 9.3.9].

Corresponding to these energies, we consider the following discrete time sequences.

**Definition 2.12** (discrete time sequences). *For a fixed time step  $\tau > 0$  and  $\rho \in D(E_\infty)$ , define*

(i) discrete gradient flow of  $E_\infty$ :

$$\rho_\tau^n \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho_\tau^{n-1}, \nu) + E_\infty(\nu) \right\} \text{ and } \rho_\tau^0 := \rho.$$

(ii) time varying discrete gradient flow of  $\tilde{E}_\infty$ : for  $\rho_\tau^n$  as in (i),

$$\tilde{\rho}_\tau^n \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\tilde{\rho}_\tau^{n-1}, \nu) + \tilde{E}_\infty(\nu; \rho_\tau^n) \right\} \text{ and } \tilde{\rho}_\tau^0 := \rho.$$

(iii) time varying discrete gradient flow of  $E_m$ : for  $\rho_\tau^n$  as in (i) and  $m > 1$ ,

$$\rho_{\tau,m}^n \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho_{\tau,m}^{n-1}, \nu) + E_m(\nu; \rho_\tau^n) \right\} \text{ and } \rho_{\tau,m}^0 := \rho.$$

The existence of the above sequences is guaranteed by [20, Theorem 4.3, Proposition 4.4]. However, they are not necessarily unique, and we use the notation  $\rho_\tau^n$ ,  $\tilde{\rho}_\tau^n$ , and  $\rho_{\tau,m}^n$  to denote any such sequence. Still, using Proposition 2.8, which provides a contraction inequality for  $\omega$ -convex functions, we can at least bound the Wasserstein distance between any two such sequences—for example, see Proposition 5.5 in the appendix for such an estimate for  $\tilde{E}_\infty$ .

If one takes  $\tau = t/n$  for  $t \geq 0$ , then as  $n \rightarrow +\infty$  the discrete gradient flow of  $E_\infty$  converges to the continuous gradient flow of  $E_\infty$  with initial data  $\rho_\infty(0) = \rho$  [20, Theorem 4.3, Proposition 4.4]. Likewise,  $\rho_{t/n,m}^n$  converges to a solution of the nonlinear diffusion equations  $(\text{PME-D})_m$ , which we denote by  $\rho_m(x, t)$ , with the same initial data (see Proposition 5.6). We refer to  $\rho_{\tau,m}^n$

as the ‘‘time varying’’ discrete gradient flow of  $E_m$  since we change the second argument of  $E_m(\cdot; \cdot)$  at each step of the sequence to accommodate the time dependent drift in  $(\text{PME-D})_m$ .

The main goal of this section is to show that  $\lim_{m \rightarrow +\infty} W_2(\rho_\infty(t), \rho_m(t)) = 0$ , which we accomplish by showing that the distance between the sequences  $\rho_\tau^n$  and  $\rho_{\tau,m}^n$  becomes arbitrarily small as  $m \rightarrow +\infty$ . We use the sequence  $\tilde{\rho}_\tau^n$ , defined in (ii) above, to serve as a bridge between the two. In what follows, we will often use the crude estimate  $\omega(x) \leq \sqrt{x}$ , for  $x \geq 0$  sufficiently small. Consequently, the rate of convergence we obtain for  $\rho_m(t) \rightarrow \rho_\infty(t)$  is certainly not sharp, but the inequalities are much simpler.

We close this introductory section with a few elementary estimates on the above discrete time sequences. In these estimates, as well as in what follows, it will be useful to consider one step of the above sequences:

**Definition 2.13** (one step minimizers). *For a fixed time step  $\tau > 0$ , we define*

(i) one step of discrete gradient flow of  $E_\infty$ : given  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ ,

$$\rho_\tau \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho, \nu) + E_\infty(\nu) \right\}$$

(ii) one step of discrete gradient flow of  $\tilde{E}_\infty(\cdot; \mu)$ : given  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\mu\|_\infty \leq 1$ ,

$$\tilde{\rho}_\tau \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho, \nu) + \tilde{E}_\infty(\nu; \mu) \right\}.$$

(iii) one step of discrete gradient flow of  $E_m(\cdot; \mu)$ : given  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ ,  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\mu\|_\infty \leq 1$ , and  $m > 1$ ,

$$\rho_{\tau,m} \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho, \nu) + E_m(\nu; \mu) \right\}.$$

As before, [20, Theorem 4.3, Proposition 4.4] ensures these minimization problems admit at least one solution. Again, these minimizers are not necessarily unique, and we use the notation  $\rho_\tau, \tilde{\rho}_\tau$ , and  $\rho_{\tau,m}$  to denote any such minimizer.

First, we estimate how the Wasserstein distance, energies, and  $L^m$  norms behave under one step of the discrete gradient flow.

**Lemma 2.14.** *Fix  $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\mu\|_\infty \leq 1$ . Then for  $C_d > 0$  as in Proposition 2.3 and any  $\tau > 0$  and  $m \geq 2$ ,*

(i) *If  $\|\rho\|_\infty \leq 1$ , then  $W_2(\rho_\tau, \rho) \leq 2C_d\tau$  and  $E_\infty(\rho) \leq E_\infty(\rho_\tau) + 2C_d^2\tau$ ;*

(ii) *If  $\|\rho\|_\infty \leq 1$ , then  $W_2(\tilde{\rho}_\tau, \rho) \leq 2C_d\tau$  and  $\tilde{E}_\infty(\rho; \mu) \leq \tilde{E}_\infty(\tilde{\rho}_\tau; \mu) + 2C_d^2\tau$ ;*

(iii) *For all  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ ,*

$$W_2(\rho_{\tau,m}, \rho) \leq \sqrt{\frac{2\tau}{m-1} (\|\rho\|_m^m - \|\rho_{\tau,m}\|_m^m)} + 2C_d\tau, \quad \frac{1}{m-1} \|\rho_{\tau,m}\|_m^m \leq \frac{1}{m-1} \|\rho\|_m^m + \frac{\tau}{2} C_d^2,$$

$$\text{and } E_m(\rho; \mu) \leq E_m(\rho_{\tau,m}; \mu) + (\|\rho\|_m^m - \|\rho_{\tau,m}\|_m^m) + C_d \sqrt{\frac{2\tau}{m-1} \|\rho\|_m^m} + 2C_d^2\tau.$$

*Proof.* We begin with (ii). Taking  $\nu = \rho$  in the definition of  $\tilde{\rho}_\tau$  and rearranging,

$$W_2^2(\tilde{\rho}_\tau, \rho) \leq 2\tau \left( \tilde{E}_\infty(\rho; \mu) - \tilde{E}_\infty(\tilde{\rho}_\tau; \mu) \right).$$

Thus, applying Lemma 2.2, with  $f = \mathbf{N}\mu$ , and Proposition 2.3

$$0 \leq \tilde{E}_\infty(\rho; \mu) - \tilde{E}_\infty(\tilde{\rho}_\tau; \mu) = \int \mathbf{N}\mu d\rho - \int \mathbf{N}\mu d\tilde{\rho}_\tau \leq C_d W_2(\tilde{\rho}_\tau, \rho).$$

Combining the above two inequalities gives the results.

Next, we show (i). Again, taking  $\nu = \rho$  in the definition of  $\rho_\tau$ ,

$$W_2^2(\rho_\tau, \rho) \leq 2\tau(E_\infty(\rho) - E_\infty(\rho_\tau)).$$

Thus, applying Lemma 2.2, with  $f = \mathbf{N}\rho$  and  $f = \mathbf{N}\rho_\tau$ , along with Proposition 2.3,

$$0 \leq E_\infty(\rho) - E_\infty(\rho_\tau) = \frac{1}{2} \left( \int \mathbf{N}\rho d\rho - \int \mathbf{N}\rho d\rho_\tau + \int \mathbf{N}\rho_\tau d\rho - \int \mathbf{N}\rho_\tau d\rho_\tau \right) \leq C_d W_2(\rho_\tau, \rho).$$

Combining the above two inequalities again give the results.

It remains to show (iii). For simplicity of notation, let  $\Phi_{1/m} = \psi_{1/m} * \mathbf{N}\mu$ . Taking  $\nu = \rho$  in the definition of  $\rho_{\tau,m}$ ,

$$(2.9) \quad \frac{1}{2\tau} W_2^2(\rho, \rho_{\tau,m}) + E_m(\rho_{\tau,m}; \mu) \leq E_m(\rho; \mu).$$

By definition of  $E_m$ , Lemma 2.2 with  $f = \Phi_{1/m}$ , and Proposition 2.3, this implies

$$\begin{aligned} \|\rho_{\tau,m}\|_m^m / (m-1) &\leq \|\rho\|_m^m / (m-1) + \left( \int \Phi_{1/m} \rho - \int \Phi_{1/m} \rho_{\tau,m} \right) - W_2^2(\rho, \rho_{\tau,m}) / (2\tau) \\ &\leq \|\rho\|_m^m / (m-1) + C_d W_2(\rho, \rho_{\tau,m}) - W_2^2(\rho, \rho_{\tau,m}) / (2\tau) \\ &= \|\rho\|_m^m / (m-1) - (W_2(\rho, \rho_{\tau,m}) - \tau C_d)^2 / (2\tau) + \tau C_d^2 / 2 \end{aligned}$$

Dropping the negative term shows the second inequality. Rearranging gives

$$(W_2(\rho, \rho_{\tau,m}) - \tau C_d)^2 / (2\tau) \leq \|\rho\|_m^m / (m-1) - \|\rho_{\tau,m}\|_m^m / (m-1) + \tau C_d^2 / 2,$$

which, by the subadditivity of  $\sqrt{\cdot}$ , gives the first inequality.

To show the third inequality, we combine (2.9) with Lemma 2.2 and use the previous estimate on the Wasserstein distance,

$$\begin{aligned} 0 \leq E_m(\rho; \mu) - E_m(\rho_{\tau,m}; \mu) &\leq \frac{1}{m-1} (\|\rho\|_m^m - \|\rho_{\tau,m}\|_m^m) + \int \mathbf{N}\mu d(\rho - \rho_{\tau,m}) \\ &\leq (\|\rho\|_m^m - \|\rho_{\tau,m}\|_m^m) + C_d \sqrt{\frac{2\tau}{m-1}} \|\rho\|_m^m + 2C_d^2 \tau. \end{aligned}$$

□

Iterating the above lemma provides bounds on the Wasserstein distance between the discrete time sequences of  $E_\infty$ ,  $\tilde{E}_\infty$ , and  $E_m$  and their initial data.

**Corollary 2.15.** *Under the assumptions in Lemma 2.14, given initial data  $\rho \in D(E_\infty)$ ,*

$$W_2(\rho_\tau^n, \rho) \leq 2C_d n\tau, \quad W_2(\tilde{\rho}_\tau^n, \rho) \leq 2C_d n\tau, \quad \text{and} \quad W_2(\rho_{m,\tau}^n, \rho) \leq \sqrt{4n\tau \|\rho\|_m^m + 8C_d^2 n^2 \tau^2}.$$

*Proof.* The first two inequalities are a direct consequence of Lemma 2.14 and triangle inequality, so it remains to show the third inequality. By Lemma 2.14 and  $(a+b)^2 \leq 2a^2 + 2b^2$ ,

$$W_2^2(\rho_{\tau,m}^i, \rho_{\tau,m}^{i-1}) \leq \frac{4\tau}{m-1} (\|\rho_{\tau,m}^{i-1}\|_m^m - \|\rho_{\tau,m}^i\|_m^m) + 8C_d^2 \tau^2.$$

The result then follows by the triangle inequality, Cauchy's inequality, and  $1/(m-1) \leq 1$ ,

$$W_2^2(\rho_{\tau,m}^n, \rho) \leq \left( \sum_{i=1}^n W_2(\rho_{\tau,m}^i, \rho_{\tau,m}^{i-1}) \right)^2 \leq n \sum_{i=1}^n W_2^2(\rho_{\tau,m}^i, \rho_{\tau,m}^{i-1}) \leq 4n\tau \|\rho\|_m^m + 8C_d^2 n^2 \tau^2.$$

□

In the next three lemmas, we estimate the size of  $\rho_{\tau,m}$ . These estimates are similar in some respects to the corresponding results in previous work by Alexander with the second and third authors [1]. However, the proofs must be adapted since the semiconvexity of the drift potential  $\psi_{1/m} * \mathbf{N}\mu$  in the energy  $E_m(\cdot; \mu)$  deteriorates as  $m \rightarrow +\infty$ , and we must instead use that  $E_m(\cdot; \mu)$  is  $\omega$ -convex uniformly in  $m$ .

Though we do not, in general, have  $\|\rho_{\tau,m}^1\|_\infty \leq 1$ , in the next lemma, we show that the mass of  $\rho_{\tau,m}$  above 1 becomes arbitrarily small as  $m \rightarrow +\infty$ .

**Lemma 2.16.** *Fix  $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$  with both  $\|\rho\|_\infty, \|\mu\|_\infty \leq 1$  and consider  $\rho_{\tau,m}$  as in Definition 2.13. Then for  $C_d > 0$  as in Proposition 2.3 and  $0 < \tau < 1$ ,  $m \geq 2$ ,*

$$\int (\rho_{\tau,m}(x) - 1)_+ dx \leq \sqrt{(2 + C_d^2)/m}.$$

*Proof.* By the Cauchy-Schwarz inequality and the fact that  $|\{\rho_{\tau,m} \geq 1\}| \leq \int \rho_{\tau,m} = 1$ ,

$$(2.10) \quad \int (\rho_{\tau,m} - 1)_+ \leq |\{\rho_{\tau,m} \geq 1\}|^{1/2} \left( \int (\rho_{\tau,m} - 1)_+^2 \right)^{1/2} \leq \left( \int (\rho_{\tau,m} - 1)_+^2 \right)^{1/2}.$$

Furthermore, for  $m \geq 2$ , the convexity of  $f(s) = s^m$  ensures  $s^m > 1 + m(s-1) + \frac{m(m-1)}{2}(s-1)^2$  for all  $s > 1$ , which yields  $(s-1)_+^2 \leq \frac{2}{m(m-1)} s^m$  for all  $s > 0$ . Consequently, (2.10) becomes

$$\int_{\mathbb{R}^d} (\rho_{\tau,m} - 1)_+ \leq \left( \frac{2}{m(m-1)} \int_{\mathbb{R}^d} \rho_{\tau,m}^m \right)^{1/2}.$$

Since  $\|\rho\|_\infty \leq 1$ ,  $m \geq 2$ , and  $\tau < 1$ , Lemma 2.14 (iii) ensures  $\frac{1}{m-1} \|\rho_{\tau,m}\|_m^m \leq 1 + C_d^2/2$ . Substituting this into the above inequality gives the result. □

Finally, we use the previous lemma to show that  $\rho_{\tau,m}$  is always close to a measure  $\nu$  that satisfies  $\|\nu\|_\infty \leq 1$  and is almost a one step minimizer.

**Lemma 2.17.** *Under the assumptions of Lemma 2.16, there exists  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\nu\|_\infty \leq 1$  and  $C > 0$  depending only on the dimension, so that*

$$(2.11) \quad W_2(\rho_{\tau,m}, \nu) \leq Cm^{-1/4} \quad \text{and} \quad E_m(\nu; \mu) \leq E_m(\rho_{\tau,m}; \mu) + Cm^{-1/2}.$$

*Proof.* Define  $a := \int (\rho_{\tau,m} - 1)_+$ . Since  $\rho_{\tau,m}$  is a probability measure,  $a < 1$ , and by Lemma 2.16, we also have  $a \leq \sqrt{(2 + C_d^2)/m}$ .

To construct  $\nu$ , we decompose  $\rho_{\tau,m}$  as  $\rho_{\tau,m} = \rho_{\tau,m}^1 + \rho_{\tau,m}^2$ , where  $\rho_{\tau,m}^1 = \min\{\rho_{\tau,m}, 1 - a\}$  and  $\rho_{\tau,m}^2 = (\rho_{\tau,m} - (1 - a))_+$ . First, note that

$$1 \geq \int \rho_m \geq (1 - a)|\{\rho_m > 1 - a\}| + \int (\rho_m - 1)_+ = (1 - a)|\{\rho_m > 1 - a\}| + a,$$

so subtracting  $a$  from both sides and dividing by  $1 - a$  ensures  $|\{\rho_{\tau,m} > 1 - a\}| \leq 1$ . Thus,

$$(2.12) \quad \int \rho_{\tau,m}^2 \leq \int (\rho_{\tau,m} - 1)_+ + a|\{\rho_{\tau,m} > 1 - a\}| \leq a + a \cdot 1 = 2a.$$

Now, choose  $R_d$  so  $g := \frac{1}{2}\chi_{B_{R_d}(0)} \in \mathcal{P}_2(\mathbb{R}^d)$  and define  $\nu = \rho_{\tau,m}^1 + g * \rho_{\tau,m}^2 \in \mathcal{P}_2(\mathbb{R}^d)$ . By Young's inequality, the definition of  $\rho_{\tau,m}^1$ , and inequality (2.12),

$$\|\nu\|_\infty \leq \|\rho_{\tau,m}^1\|_\infty + \|g\|_\infty \|\rho_{\tau,m}^2\|_1 \leq (1 - a) + \frac{1}{2} \cdot 2a \leq 1.$$

It remains to show that  $\nu$  satisfies (2.11). To show the first inequality, we construct a transport plan between  $\rho_{\tau,m}$  and  $\nu$  as follows: keep all the mass of  $\rho_{\tau,m}^1$  at its original location and distribute the mass of  $\rho_{\tau,m}^2(x)$  uniformly over the disk  $B_{R_d}(x)$ . Since  $\int \rho_{\tau,m}^2 \leq 2a$ , the total cost of this plan is bounded by  $2aR_d^2$ , which gives  $W_2(\rho_{\tau,m}, \nu) \leq \sqrt{2a}R_d \leq R_d(4(2 + C_d^2)/m)^{1/4}$ .

To show the second inequality in (2.11), we abbreviate  $\Phi_{1/m} = \psi_{1/m} * \mathbf{N}\mu$ . Then,

$$\begin{aligned} E_m(\nu; \mu) - E_m(\rho_{\tau,m}; \mu) &= \frac{1}{m-1} \int \nu^m - \frac{1}{m-1} \int \rho_{\tau,m}^m + \int \Phi_{1/m} d\nu - \int \Phi_{1/m} d\rho_{\tau,m} \\ &\leq \|\nu\|_m^m / (m-1) + \int (g * \Phi_{1/m} - \Phi_{1/m}) \rho_{\tau,m}^2 \leq (m-1)^{-1} + 2a \|g * \Phi_{1/m} - \Phi_{1/m}\|_\infty \\ &\leq 2m^{-1} + 2a \operatorname{ess\,sup}_x \left| \int_{y \in B_{R_d}(x)} (\Phi_{1/m}(y) - \Phi_{1/m}(x)) g(x-y) dy \right| \\ &\leq 2m^{-1} + 2a \|\nabla \Phi_{1/m}\|_\infty R_d \|g\|_1 \leq 2m^{-1} + 2C_d R_d \sqrt{(2 + C_d)^2 / m}, \end{aligned}$$

where in the last inequality we use Proposition 2.3.  $\square$

**2.3. Distance between discrete time sequences of  $E_\infty$ ,  $\tilde{E}_\infty$ , and  $E_m$ .** In this section, we apply the previous results to show that as  $m \rightarrow +\infty$ ,  $\rho_m(t)$  converges to  $\rho_\infty(t)$ , with quantitative rates of convergence on bounded time intervals. We accomplish this by first estimating the distance between the discrete time sequences of  $E_\infty$  and  $\tilde{E}_\infty$  and then  $\tilde{E}_\infty$  and  $E_m$ . We begin by showing that one step of the discrete gradient flow of  $E_\infty$  is also one step of the discrete time sequence corresponding to  $\tilde{E}_\infty$ .

**Lemma 2.18** (one-step comparison between  $\rho_\tau$  and  $\tilde{\rho}_\tau$ ). *Given  $\tau > 0$  and  $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ , if  $\rho_\tau$  is a one step minimizer of  $E_\infty$ , then it is also a one step minimizer of  $\tilde{E}_\infty(\cdot, \rho_\tau)$ .*

*Proof.* Assume, for the sake of contradiction, that  $\rho_\tau$  is not a one step minimizer of  $\tilde{E}_\infty(\cdot, \rho_\tau)$ . Then there exists  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\nu\|_\infty \leq 1$ , such that

$$(2.13) \quad \frac{1}{2\tau} W_2^2(\rho, \rho_\tau) + \tilde{E}_\infty(\rho_\tau; \rho_\tau) > \frac{1}{2\tau} W_2^2(\rho, \nu) + \tilde{E}_\infty(\nu; \rho_\tau).$$

Define  $\rho^\varepsilon := (1 - \varepsilon)\rho_\tau + \varepsilon\nu \in \mathcal{P}_2(\mathbb{R}^d)$ , so  $\|\rho^\varepsilon\|_\infty \leq 1$ . We will show that for  $\varepsilon > 0$  small,

$$(2.14) \quad \frac{1}{2\tau} W_2^2(\rho, \rho_\tau) + E_\infty(\rho_\tau) > \frac{1}{2\tau} W_2^2(\rho, \rho^\varepsilon) + E_\infty(\rho^\varepsilon),$$

which contradicts the fact that  $\rho_\tau$  is a one step minimizer of  $E_\infty$ .

By inequality (2.2),  $W_2^2$  is convex along linear interpolations of measures, hence

$$(2.15) \quad W_2^2(\rho, \rho^\varepsilon) \leq (1 - \varepsilon)W_2^2(\rho, \rho_\tau) + \varepsilon W_2^2(\rho, \nu) = W_2^2(\rho, \rho_\tau) - \varepsilon (W_2^2(\rho, \rho_\tau) - W_2^2(\rho, \nu)).$$

Likewise, we use that  $2E_\infty(\rho_\tau) = \tilde{E}_\infty(\rho_\tau; \rho_\tau)$  to estimate the behavior of  $E_\infty$  along  $\rho^\varepsilon$  by

$$\begin{aligned}
(2.16) \quad E_\infty(\rho^\varepsilon) &= \frac{1}{2} \int \mathbf{N}((1-\varepsilon)\rho_\tau + \varepsilon\nu) d((1-\varepsilon)\rho_\tau + \varepsilon\nu) \\
&= (1-\varepsilon)^2 E_\infty(\rho_\tau) + \varepsilon(1-\varepsilon)\tilde{E}_\infty(\nu; \rho_\tau) + \varepsilon^2 E_\infty(\nu) \\
&= E_\infty(\rho_\tau) - \varepsilon(\tilde{E}_\infty(\rho_\tau; \rho_\tau) - \tilde{E}_\infty(\nu; \rho_\tau)) + D\varepsilon^2,
\end{aligned}$$

where  $D := E_\infty(\rho_\tau) + E_\infty(\nu) - \tilde{E}_\infty(\nu, \rho_\tau)$  is a constant independent of  $\varepsilon$ . Multiplying (2.15) by  $1/(2\tau)$  and adding to (2.16) yields

$$\begin{aligned}
&\frac{1}{2\tau} W_2^2(\rho, \rho^\varepsilon) + E_\infty(\rho^\varepsilon) \\
&\leq \frac{1}{2\tau} W_2^2(\rho, \rho_\tau) + E_\infty(\rho_\tau) - \varepsilon \left( \frac{1}{2\tau} (W_2^2(\rho, \rho_\tau) - W_2^2(\rho, \nu)) + \tilde{E}_\infty(\rho_\tau; \rho_\tau) - \tilde{E}_\infty(\nu; \rho_\tau) \right) + D\varepsilon^2.
\end{aligned}$$

By (2.13), the quantity within parentheses is strictly positive, hence we obtain (2.14) for  $\varepsilon$  small.  $\square$

Using this lemma and Proposition 2.8, which proves a contraction inequality for one step of the discrete gradient flows of  $\omega$ -convex energies, we can bound the distance between the discrete gradient flow of  $E_\infty$  and the discrete time sequence corresponding to  $\tilde{E}_\infty$ .

**Proposition 2.19** (multi-step comparison between  $\rho_\tau^n$  and  $\tilde{\rho}_\tau^n$ ). *Given  $T > 0$  and initial data  $\rho \in D(E_\infty)$ , there exist positive constants  $C$  and  $N$  depending on the dimension,  $T$ , and  $E_\infty(\rho)$  so that for  $\tau = t/n$ ,  $0 \leq t \leq T$ , and  $n > N$ ,*

$$W_2(\rho_\tau^n, \tilde{\rho}_\tau^n) \leq C(n^{-1/2})^{1/2} e^{2C_d T},$$

*Proof.* By Lemma 2.18,  $\rho_\tau^n$  is also a time varying discrete gradient flow of  $\tilde{E}_\infty$ , in the sense of Definition 2.12 (ii). Hence, by Proposition 5.5, for any  $T > 0$  there exist positive constants  $C$  and  $N$  (which we allow to change from line to line), depending on the dimension,  $T$ , and  $E_\infty(\rho)$  so that for  $\tau = t/n$ ,  $0 \leq t \leq T$ , and  $n > N$ ,

$$f_\tau^{(2n)}(W_2^2(\rho_\tau^n, \tilde{\rho}_\tau^n)) \leq C\omega(t/n) \leq Cn^{-1/2}$$

Furthermore, combining Corollary 2.15 and the triangle inequality provides the following crude bound for the distance between the two sequences:

$$W_2(\rho_\tau^n, \tilde{\rho}_\tau^n) \leq W_2(\rho_\tau^n, \rho) + W_2(\tilde{\rho}_\tau^n, \rho) \leq 4C_d T \leq C.$$

Therefore, by Proposition 2.7 (iii) and the fact that  $F_t(x)$  is decreasing in time,

$$F_{2n\tau}(W_2^2(\rho_\tau^n, \tilde{\rho}_\tau^n)) \leq Cn^{-1/2} + 2C_d\omega(C)T/n \implies F_{2T}(W_2^2(\rho_\tau^n, \tilde{\rho}_\tau^n)) \leq Cn^{-1/2}$$

Since for  $0 \leq x \leq e^{-1-\sqrt{2}}$ ,  $F_t(x) = x^{e^{C_d t}}$ , for  $n$  sufficiently large, we have

$$W_2(\rho_\tau^n, \tilde{\rho}_\tau^n) \leq C(n^{-1/2})^{1/2} e^{2C_d T},$$

which gives the result.  $\square$

Next, we bound the distance between one step of the discrete time sequences corresponding to  $\tilde{E}_\infty$  and  $E_m$ .

**Proposition 2.20** (one-step comparison between  $\tilde{\rho}_\tau$  and  $\rho_{\tau,m}$ ). *Given  $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\rho\|_\infty, \|\mu\|_\infty \leq 1$  and  $C_d > 0$  as in Proposition 2.3, there exists  $C > 0$  depending only on the dimension so that for all  $0 < \tau < 1/6C_d$  and  $m \geq 2$ ,*

$$W_2(\rho_{\tau,m}, \tilde{\rho}_\tau) \leq Cm^{-1/8} + 2e^{-1/(4C_d\tau)}.$$

*Proof.* Let  $\nu$  be as in Lemma 2.17 and define

$$\eta := \left( \frac{1}{2} \mathbf{t}_\rho^\nu + \frac{1}{2} \mathbf{t}_\rho^{\tilde{\rho}_\tau} \right) \# \rho$$

to be the midpoint on the generalized geodesic from  $\nu$  to  $\tilde{\rho}_\tau$  with base  $\rho$ . Since the  $L^\infty$  norm of a generalized geodesic is bounded by the  $L^\infty$  norm of its endpoints (c.f. [20, inequality (60)]), we have  $\|\nu\|_\infty \leq 1$ . Furthermore, by the optimality of  $\rho_{\tau,m}$  and  $\tilde{\rho}_\tau$ ,

$$\begin{aligned} \frac{1}{2\tau} W_2^2(\rho, \eta) + E_m(\eta; \mu) &\geq \frac{1}{2\tau} W_2^2(\rho, \rho_{\tau,m}) + E_m(\rho_{\tau,m}; \mu), \\ \frac{1}{2\tau} W_2^2(\rho, \eta) + \tilde{E}_\infty(\eta; \mu) &\geq \frac{1}{2\tau} W_2^2(\rho, \tilde{\rho}_\tau) + \tilde{E}_\infty(\tilde{\rho}_\tau; \mu). \end{aligned}$$

Adding these inequalities together and collecting the distance and energy terms gives

$$(2.17) \quad T_W + T_E \geq 0,$$

for

$$\begin{aligned} T_W &:= \frac{1}{\tau} W_2^2(\rho, \eta) - \frac{1}{2\tau} W_2^2(\rho, \rho_{\tau,m}) - \frac{1}{2\tau} W_2^2(\rho, \tilde{\rho}_\tau), \\ T_E &:= E_m(\eta; \mu) + \tilde{E}_\infty(\eta; \mu) - E_m(\rho_{\tau,m}; \mu) - \tilde{E}_\infty(\tilde{\rho}_\tau; \mu). \end{aligned}$$

Next, we find upper bounds on  $T_W$  and  $T_E$ . Define  $A := \|\mathbf{t}_\rho^\nu - \mathbf{t}_\rho^{\tilde{\rho}_\tau}\|_{L^2(\rho)}$ . Since  $W_2^2(\rho, \cdot)$  is 2-convex along generalized geodesics with base  $\rho$  [2, Lemma 9.2.1],

$$W_2^2(\rho, \eta) \leq W_2^2(\rho, \nu)/2 + W_2^2(\rho, \tilde{\rho}_\tau)/2 - A^2/4.$$

Substituting this in the definition of  $T_W$ ,

$$\begin{aligned} T_W &\leq \frac{1}{2\tau} (W_2^2(\rho, \nu) - W_2^2(\rho, \rho_{\tau,m})) - \frac{A^2}{4\tau} \leq \frac{1}{2\tau} W_2(\rho_{\tau,m}, \nu) (W_2(\rho, \nu) + W_2(\rho, \rho_{\tau,m})) - \frac{A^2}{4\tau} \\ &\leq \frac{1}{2\tau} W_2(\rho_{\tau,m}, \nu) (W_2(\rho_{\tau,m}, \nu) + 2W_2(\rho, \rho_{\tau,m})) - \frac{A^2}{4\tau} \leq \frac{C}{2\tau} m^{-1/4} - \frac{A^2}{4\tau}, \end{aligned}$$

where in the last inequality we apply  $W_2(\rho_{\tau,m}, \nu) \leq Cm^{-1/4}$  from Lemma 2.17,  $W_2(\rho_{\tau,m}, \rho) \leq \sqrt{2\tau} + 2C_d\tau$  from Lemma 2.14, and the facts that  $m \geq 2$  and  $\tau < 1$ . We also allow  $C > 0$ , depending only on the dimension, to change from line to line.

In order to bound  $T_E$  from above, we first estimate the difference between  $E_m(\tilde{\mu}, \mu)$  and  $\tilde{E}_\infty(\tilde{\mu}, \mu)$  for any  $\tilde{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$  with  $\|\tilde{\mu}\|_\infty \leq 1$ . As usual, we abbreviate  $\Phi := \mathbf{N}\mu$  and  $\Phi_{1/m} := \psi_{1/m} * \mathbf{N}\mu$ . Given  $R_\psi > 0$  so that  $\text{supp } \psi \subseteq B_{R_\psi}(0)$ , for any  $x \in \mathbb{R}^d$ , Proposition 2.3 ensures

$$|\Phi_{1/m}(x) - \Phi(x)| = \left| \int_{\mathbb{R}^d} (\Phi(x-y) - \Phi(x)) \psi_{1/m}(y) dy \right| \leq R_\psi \|\nabla \Phi\|_\infty m^{-1} \leq C_d m^{-1}.$$

Consequently,

$$\left| E_m(\tilde{\mu}; \mu) - \tilde{E}_\infty(\tilde{\mu}; \mu) \right| \leq \frac{1}{m-1} \int \tilde{\mu}^m + \int |\Phi_{1/m} - \Phi| d\tilde{\mu} \leq (m-1)^{-1} + \|\Phi_{1/m} - \Phi\|_\infty \leq (2+C_d)m^{-1}.$$

Therefore, first applying Lemma 2.17 to the definition of  $T_E$  and then using the above inequality,

$$(2.18) \quad \begin{aligned} T_E &\leq E_m(\eta; \mu) + \tilde{E}_\infty(\eta; \mu) - E_m(\nu, \mu) - \tilde{E}_\infty(\tilde{\rho}_\tau; \mu) + Cm^{-1/2} \\ &\leq 2\tilde{E}_\infty(\eta; \mu) - \tilde{E}_\infty(\nu, \mu) - \tilde{E}_\infty(\tilde{\rho}_\tau; \mu) + Cm^{-1/2}. \end{aligned}$$

Since  $\tilde{E}_\infty(\cdot, \mu)$  is  $\omega$ -convex along generalized geodesics and  $\eta$  is the midpoint along the generalized geodesic from  $\nu$  and  $\tilde{\rho}_\tau$  with base  $\rho$ ,

$$2\tilde{E}_\infty(\eta; \mu) - \tilde{E}_\infty(\nu; \mu) - \tilde{E}_\infty(\tilde{\rho}_\tau; \mu) \leq \frac{C_d}{2}\omega\left(\frac{A^2}{4}\right).$$

Substituting this into inequality (2.18) gives

$$T_E \leq Cm^{-1/2} + C_d\omega(A^2/4).$$

Finally, combining our upper bounds on  $T_W$  and  $T_E$  with inequality (2.17), we obtain

$$(2.19) \quad A^2 \leq Cm^{-1/4} + 4C_d\tau\omega(A^2/4).$$

We now claim that

$$(2.20) \quad A \leq \sqrt{2C}m^{-1/8} + 2e^{-1/(4C_d\tau)}$$

If  $A^2/4 > e^{-1-\sqrt{2}}$ , then combining inequality (2.19) and  $\tau < 1/(6C_d)$  implies

$$A^2 \leq Cm^{-1/4} + 4\tau C_d\omega(A^2/4) \leq Cm^{-1/4} + 3\tau C_d A^2 \implies A \leq \sqrt{2C}m^{-1/8},$$

hence (2.20) holds. Alternatively, if  $A^2/4 \leq e^{-1-\sqrt{2}}$ ,

$$(2.21) \quad A^2 \leq Cm^{-1/4} + 4\tau C_d\omega(A^2/4) = Cm^{-1/4} - C_d\tau A^2 \log(A^2/4).$$

If (2.20) is violated, we have  $A > \sqrt{2C}m^{-1/8}$  and  $A > 2e^{-1/(4C_d\tau)}$ , so  $Cm^{-1/4} < A^2/2$  and  $-C_d\tau A^2 \log(A^2/4) < \frac{A^2}{2}$ . Adding these together would contradict (2.19), so again (2.20) holds.

Since  $A = \|\mathbf{t}_\rho^\nu - \mathbf{t}_{\tilde{\rho}_\tau}^{\tilde{\rho}_\tau}\|_{L^2(\rho)} = \|\mathbf{t}_\rho^\nu \circ \mathbf{t}_{\tilde{\rho}_\tau}^\rho - \mathbf{id}\|_{L^2(\tilde{\rho}_\tau)}$  and  $\mathbf{t}_\rho^\nu \circ \mathbf{t}_{\tilde{\rho}_\tau}^\rho \# \tilde{\rho}_\tau = \nu$ , we have  $W_2(\nu, \tilde{\rho}_\tau) \leq A$ . Therefore, using (2.20) and Lemma 2.17, we may conclude the result,

$$\begin{aligned} W_2(\rho_{\tau,m}, \tilde{\rho}_\tau) &\leq W_2(\rho_{\tau,m}, \nu) + W_2(\nu, \tilde{\rho}_\tau) \leq Cm^{-1/4} + A \\ &\leq Cm^{-1/4} + \sqrt{2C}m^{-1/8} + 2e^{-1/(4C_d\tau)} \leq Cm^{-1/8} + 2e^{-1/(4C_d\tau)}. \end{aligned}$$

□

**Proposition 2.21** (multi-step comparison between  $\tilde{\rho}_\tau^n$  and  $\rho_{\tau,m}^n$ ). *Given  $T > 0$  and initial data  $\rho \in D(E_\infty)$ , there exist positive constants  $C$ ,  $N$ , and  $M$  depending on the dimension,  $T$ ,  $E_\infty(\rho)$ , and  $\psi$  so that for  $\tau = t/n$ ,  $0 \leq t \leq T$ ,  $n > N$ ,  $m > M$ , and  $n = o(m^{1/8})$ ,*

$$W_2(\rho_{\tau,m}^n, \tilde{\rho}_\tau^n) \leq C(n^{-1/4} + nm^{-1/8})1/2e^{2C_dT}$$

*Proof.* Define  $d_i := W_2(\rho_{\tau,m}^i, \tilde{\rho}_\tau^i)$  for any  $i = 1, \dots, n$ . Using Corollary 2.15 and  $\|\rho\|_m^m \leq 1$ , we have the crude bound

$$(2.22) \quad d_i \leq W_2(\rho_{\tau,m}^i, \rho) + W_2(\tilde{\rho}_\tau^i, \rho) \leq \sqrt{4T + 8C_dT^2} + 2C_dT.$$

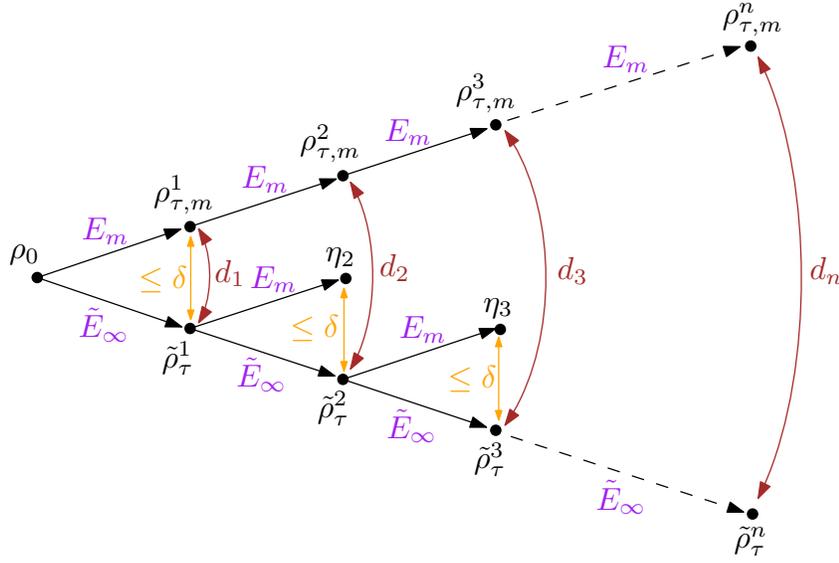


FIGURE 1. An illustration of the tree structure used in the multi-step comparison between  $\tilde{\rho}_\tau^n$  and  $\rho_{\tau,m}^n$

The one step estimates from Proposition 2.20 allow us to control the distance between one-step minimizers of  $E_m$  and  $\tilde{E}_\infty$  when they have the *same* initial data. In particular, for

$$(2.23) \quad \delta := Cm^{-1/8} + 2e^{-1/(4C_d\tau)},$$

we have  $d_1 \leq \delta$ . In order to apply Proposition 2.20 to control  $d_i$  for  $i = 2, \dots, n$ , we use a sequence of densities  $\eta^i$  to serve as a bridge between  $\rho_{\tau,m}^i$  and  $\tilde{\rho}_\tau^i$ , following the tree structure in Figure 1. Specifically, we choose  $\eta^i \in \mathcal{P}_2(\mathbb{R}^d)$  so that, by Proposition 2.20,

$$(2.24) \quad \eta^i \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\tilde{\rho}_\tau^{i-1}, \nu) + E_m(\nu; \rho_\tau^i) \right\} \implies W_2(\tilde{\rho}_\tau^i, \eta^i) \leq \delta.$$

Since  $\eta^i$  and  $\rho_{\tau,m}^i$  are one-step minimizers of the same energy  $E_m(\cdot, \rho_\tau^i)$  with different initial data ( $\tilde{\rho}_\tau^{i-1}$  and  $\rho_{\tau,m}^{i-1}$  respectively), we may control their distance using Proposition 2.8.

First, we obtain a few elementary bounds on how the energy changes along the discrete time sequence. Combining Lemma 2.2, Proposition 2.3, Lemma 2.14, and the definition of  $\rho_{\tau,m}^{i-1}$  as a minimizer,

$$\begin{aligned} E_m(\rho_{\tau,m}^{i-1}; \rho_\tau^i) &= E_m(\rho_{\tau,m}^{i-1}; \rho_{\tau,m}^{i-1}) + E_m(\rho_{\tau,m}^{i-1}; \rho_\tau^i) - E_m(\rho_{\tau,m}^{i-1}; \rho_{\tau,m}^{i-1}) \\ &= E_m(\rho_{\tau,m}^{i-1}; \rho_{\tau,m}^{i-1}) + \int \psi_{1/m} * \mathbf{N}\rho_{\tau,m}^{i-1} d(\rho_\tau^i - \rho_{\tau,m}^{i-1}) \leq E_m(\rho_{\tau,m}^{i-2}; \rho_{\tau,m}^{i-1}) + C_d W_2(\rho_\tau^i, \rho_{\tau,m}^{i-1}) \\ &\leq E_m(\rho_{\tau,m}^{i-2}; \rho_{\tau,m}^{i-1}) + 2C_d^2 \tau \leq \dots \leq E_m(\rho; \rho_\tau^1) + 2C_d^2 T. \end{aligned}$$

Likewise, we may control the first term on the right hand side of the last inequality by

$$\begin{aligned} E_m(\rho; \rho_\tau^1) &= \|\rho\|_m^m / (m-1) + \int \psi_{1/m} * \mathbf{N}\rho_\tau^1 d\rho \leq 1 + 2E_\infty(\rho) + \int \mathbf{N}\rho d(\psi_{1/m} * \rho_\tau^1 - \rho) \\ &\leq 1 + 2E_\infty(\rho) + C_d W_2(\psi_{1/m} * \rho_\tau^1, \rho) \leq 1 + 2E_\infty(\rho) + C_d (W_2(\psi_{1/m} * \rho_\tau^1, \rho_\tau^1) + W_2(\rho_\tau^1, \rho)) \\ &\leq 1 + 2E_\infty(\rho) + C_d((1/m)M_\psi + 2C_d\tau) \end{aligned}$$

where, in the last step, we apply [2, Lemma 7.1.10], which ensures

$$(2.25) \quad W_2(\mu * \psi_{1/m}, \mu) \leq \frac{1}{m} \left( \int |x|^2 \psi(x) dx \right)^{1/2} =: \frac{1}{m} M_\psi.$$

Combining the above two inequalities, we conclude that there exists constant  $C > 0$  (which we allow to change from line to line) depending on the dimension,  $T$ ,  $E_\infty(\rho)$ , and  $\psi$  so that

$$E_m(\rho_{\tau,m}^{i-1}; \rho_\tau^i) \leq C \text{ for all } i = 1, \dots, n.$$

Furthermore, by Proposition 2.3, we also have that  $E_m(\cdot; \cdot)$  is uniformly bounded below.

Using these estimates on the energy, we may now apply Proposition 2.8 to conclude that there exist positive constants  $C$  and  $N$  depending on the dimension,  $T$ ,  $E_\infty(\rho)$ , and  $\psi$ , which we allow to change from line to line, so that for  $\tau = t/n$ ,  $0 \leq t \leq T$ , and  $n > N$ ,

$$f_\tau^{(2)}(W_2^2(\eta^i, \rho_{\tau,m}^i)) \leq d_{i-1}^2 + C_d \tau \omega(CW_2(\eta^i, \tilde{\rho}_\tau^{i-1})) + 2\tau(E_m(\rho_{\tau,m}^{i-1}; \rho_\tau^i) - E_m(\rho_{\tau,m}^i; \rho_\tau^i)) + C\tau^2.$$

By Lemma 2.14 (iii), we have the following bounds for two quantities on the right hand side:

$$\begin{aligned} W_2(\eta^i, \tilde{\rho}_\tau^{i-1}) &\leq \sqrt{\frac{2\tau}{m-1} \|\tilde{\rho}_\tau^{i-1}\|_m^m} + 2C_d \tau \leq C\sqrt{\tau}, \\ E_m(\rho_{\tau,m}^{i-1}; \rho_\tau^i) - E_m(\rho_{\tau,m}^i; \rho_\tau^i) &\leq (\|\rho_{\tau,m}^{i-1}\|_m^m - \|\rho_{\tau,m}^i\|_m^m) + C\sqrt{\tau}. \end{aligned}$$

Therefore,

$$(2.26) \quad f_\tau^{(2)}(W_2^2(\eta^i, \rho_{\tau,m}^i)) \leq d_{i-1}^2 + 2\tau (\|\rho_{\tau,m}^{i-1}\|_m^m - \|\rho_{\tau,m}^i\|_m^m) + C\tau^{5/4}$$

We now use this estimate to bound  $d_i = W_2(\rho_{\tau,m}^i, \tilde{\rho}_\tau^i)$ . By the triangle inequality and (2.24),

$$(2.27) \quad d_i^2 \leq (W_2(\eta^i, \rho_{\tau,m}^i) + W_2(\tilde{\rho}_\tau^i, \eta^i))^2 \leq W_2^2(\eta^i, \rho_{\tau,m}^i) + (2W_2(\eta^i, \rho_{\tau,m}^i) + \delta) \delta$$

Furthermore, by Lemma 2.14 (iii), inequality (2.22), and equation (2.23)

$$W_2(\eta^i, \rho_{\tau,m}^i) \leq W_2(\eta^i, \tilde{\rho}_\tau^{i-1}) + d_{i-1} + \delta \leq C,$$

Thus, by Proposition 2.7, we may apply  $f_\tau^{(2)}$  to both sides of the (2.27) to obtain

$$f_\tau^{(2)}(d_i^2) \leq f_\tau^{(2)}(W_2^2(\eta^i, \rho_{\tau,m}^i)) + C\delta + C\tau^2.$$

Combining this with (2.26) gives, for all  $i = 1, \dots, n$ ,

$$(2.28) \quad f_\tau^{(2)}(d_i^2) \leq d_{i-1}^2 + 2\tau (\|\rho_{\tau,m}^{i-1}\|_m^m - \|\rho_{\tau,m}^i\|_m^m) + C\tau^{5/4} + C\delta.$$

We claim that the result will follow if we can show that, for all  $j = 1, \dots, n$ ,

$$(2.29) \quad f_\tau^{(2j)}(d_n^2) \leq d_{n-j}^2 + 2\tau (\|\rho_{\tau,m}^{n-j}\|_m^m - \|\rho_{\tau,m}^n\|_m^m) + 2C\tau^{5/4}j + C\delta j.$$

In particular, if this holds, then taking  $j = n$  and using that  $e^{-1/4C_d\tau} = O(\tau^{5/4})$  gives

$$f_\tau^{(2n)}(d_n^2) \leq 2\tau \|\rho\|_m^m + 2CT\tau^{1/4} + C\delta n \leq C(n^{-1/4} + nm^{-1/8}).$$

By Proposition 2.7 (iii) and the fact that  $F_t(x)$  is decreasing in  $t$ ,

$$F_{2n\tau}(d_n^2) \leq C(n^{-1/4} + nm^{-1/8}) + 2C_d\omega(C)T/n \implies F_{2T}(d_n^2) \leq C(n^{-1/4} + nm^{-1/8}).$$

For  $0 \leq x \leq e^{-1-\sqrt{2}}$ , we have  $F_t(x) = x^{e^{C_d t}}$ . Thus, for  $n$  and  $m$  sufficiently large, depending on the dimension,  $T$ ,  $E_\infty(\rho)$ , and  $\psi$ , and with  $n = o(m^{1/8})$ , we have

$$d_n \leq C(n^{-1/4} + nm^{-1/8})^{1/2e^{2C_d T}},$$

which gives the result.

It remains to show (2.29). We proceed by induction. The base case for  $j = 1$  follows from (2.28), so we assume the result holds for  $j - 1$ ,

$$f_\tau^{(2(j-1))}(d_n^2) \leq d_{n-j+1}^2 + 2\tau(\|\rho_{\tau,m}^{n-j+1}\|_m^m - \|\rho_{\tau,m}^n\|_m^m) + 2C\tau^{5/4}(j-1) + C\delta(j-1).$$

For any  $j = 1, \dots, n$ , the right hand side is bounded by a constant depending on the dimension,  $T$ ,  $E_\infty(\rho)$ , and  $\psi$ . Thus, by Proposition 2.7, we may apply  $f_\tau^{(2)}$  to both sides to conclude

$$\begin{aligned} f_\tau^{(2j)}(d_n^2) &\leq f_\tau^{(2)}(d_{n-j+1}^2) + 2\tau(\|\rho_{\tau,m}^{n-j+1}\|_m^m - \|\rho_{\tau,m}^n\|_m^m) + 2C\tau^{5/4}(j-1) + C\delta(j-1) + C^2\tau^2 \\ &\leq d_{n-j}^2 + 2\tau(\|\rho_{\tau,m}^{n-j}\|_m^m - \|\rho_{\tau,m}^n\|_m^m) + 2C\tau^{5/4}j + C\delta j \end{aligned}$$

where, in the second inequality, we apply (2.28) and the fact that  $C^2\tau^2 \leq C\tau\sqrt{\tau}$ .  $\square$

Combining the previous propositions, we obtain our main result.

**Theorem 2.22** (convergence of  $\rho_m(t)$  to  $\rho_\infty(t)$ ). *Given  $T > 0$  and initial data  $\rho \in D(E_\infty)$ , there exist positive constants  $C$  and  $M$  depending on  $d$ ,  $T$ ,  $E_\infty(\rho)$ , and  $\psi$  so that for all  $0 \leq t \leq T$  and  $m \geq M$ ,*

$$W_2(\rho_m(t), \rho_\infty(t)) \leq Cm^{-1/144e^{4C_d T}}.$$

*Proof.* Combining Proposition 2.19, Proposition 2.21, [20, Theorem 3.8], and Proposition 5.6, there exist positive constants  $C$  and  $N$  depending on  $d, T, E_\infty(\rho)$ , and  $\psi$  so that for  $\tau = t/n$  and all  $n \geq N$ ,  $m \geq d + 1$ ,  $0 \leq t \leq T$ , and  $n = o(m^{1/8})$ ,

$$\begin{aligned} W_2(\rho_\tau^n, \tilde{\rho}_\tau^n) &\leq Cn^{-1/4e^{2C_d T}}, & W_2(\rho_{\tau,m}^n, \tilde{\rho}_\tau^n) &\leq C(n^{-1/4} + nm^{-1/8})^{1/2e^{2C_d T}} \\ W_2(\rho_\tau^n, \rho_\infty(t)) &\leq Cn^{-1/16e^{2C_d T}}, & W_2(\rho_{t/n,m}^n, \rho_m(t)) &\leq Cn^{-1/16e^{4C_d T}}. \end{aligned}$$

Hence,

$$W_2(\rho_m(t), \rho_\infty(t)) \leq C(n^{-1/16e^{4C_d T}} + n^{1/2e^{2C_d T}} m^{-1/16e^{2C_d T}})$$

Taking  $n = m^{1/9}$  gives the result.  $\square$

### 3. CONVERGENCE OF VISCOSITY SOLUTIONS: DRIFT DIFFUSION PRESSURE TO FREE BOUNDARY PROBLEM

In the previous section, we showed that the gradient flow of the height constrained interaction energy  $E_\infty$ , which is merely a curve in the space of measures, may be approximated by solutions of the nonlinear diffusion equations (PME-D) $_m$  as  $m \rightarrow +\infty$ . This approximation provides the bridge by which we are able to unite the energy methods approach with viscosity solution approach. In the present section, we use this approximation to characterize the dynamics of patch solutions in terms of a Hele-Shaw type free boundary problem. We accomplish this by

considering the nonlinear diffusion equations in terms of their pressure variables: given  $\rho_m$  a weak solution of  $(\text{PME-D})_m$ , the pressure variable  $p_m := \frac{m}{m-1}(\rho_m)^{m-1}$  uniquely solves

$$(P)_m \quad (p_m)_t - (m-1)p_m(\Delta p_m + \Delta \Phi_{1/m}) - \nabla p_m \cdot (\nabla p_m + \nabla \Phi_{1/m}) = 0.$$

For initial data given by (1.9), we show that as  $m \rightarrow +\infty$  the half-relaxed limits of viscosity solutions of  $(P)_m$  satisfy sub- and supersolution properties of  $(P)_\infty$ . The comparison principle of  $(P)_\infty$  then yields that these half-relaxed limits are ordered with respect to the viscosity solution  $p$  of  $(P)_\infty$  with the same initial data. In terms of the density variable, we show that  $\rho_m$  uniformly converges to  $\chi_{\Omega(t)}$  away from  $\partial\Omega(t)$ , where  $\Omega(t) = \{p(\cdot, t) > 0\}$ . It follows that  $\rho_\infty = \chi_{\Omega(t)}$  almost everywhere, and thus  $(P)_\infty$  identifies with  $(P)$ . Due to the fact that the link between  $(P)_\infty$  and  $(P)$  lacks a priori stability estimates as the initial data varies, we must introduce additional perturbations and approximations into our proof of this final result.

*Remark 3.1.* The lack of the comparison principle for the original problem  $(P)$  is not the main reason we consider  $(\text{PME-D})_m$ . We could have considered the drift term given by  $\Phi := \mathcal{N} * \rho_m$ , and thus proved the convergence of the Keller-Segel equation to our problem, if we had known that the corresponding solutions  $\rho_m$  converged to  $\rho_\infty$  as  $m \rightarrow +\infty$ . Obtaining such convergence seems to require a uniform  $L^\infty$  bound on the gradient flow solutions of  $(\text{PME-D})_m$ , which is an open question at the moment.

**3.1. Basic properties of viscosity solutions of  $(P)_m$  and  $(P)_\infty$ .** We refer the reader to Alexander, Kim, and Yao [1, Section 3] and Kim and Lei [33, Section 2.1] for the definitions of classical and viscosity solutions of  $(P)_m$ , and we refer the reader to appendix section 5.1 for the definition of viscosity solutions of  $(P)_\infty$ . To clarify our notion of weak solutions for the original free boundary problem  $(P)$ , we make the following definition:

**Definition 3.2.**  *$p$  is a weak solution of  $(P)$  if it is a viscosity solution of  $(P)_\infty$  with initial data  $p_0$  and  $\rho_\infty = \chi_{\{p>0\}}$  almost everywhere.*

We now recall the several results on well-posedness of viscosity solutions of  $(P)_m$  and the  $L^1$  contraction of the corresponding density variable.

**Lemma 3.3.** *Consider the porous medium equation with drift and source terms,*

$$(3.1) \quad \rho_t = \nabla \cdot (\rho \nabla \Phi_{1/m}) + \Delta \rho^m + \rho f,$$

with  $f \in L^1$  and bounded initial data.

(a) *If  $\rho_1$  and  $\rho_2$  are weak solutions of (3.1) with source terms  $f_1$  and  $f_2$ , then for all  $t \geq 0$ ,*

$$\|\rho_1(\cdot, t) - \rho_2(\cdot, t)\|_{L^1(\mathbb{R}^d)} \leq \|\rho_1(\cdot, 0) - \rho_2(\cdot, 0)\|_{L^1(\mathbb{R}^d)} + \int_0^t \int_{\mathbb{R}^d} |\rho_1 f_1 - \rho_2 f_2|.$$

(b) *Let  $\rho$  be a weak solution of (3.1) for any continuous, compactly supported initial data  $\rho_0$  and continuous function  $f$ . Then the pressure variable  $p_m := \frac{m-1}{m} \rho^{m-1}$  is a viscosity solution to*

$$(p_m)_t - (m-1)p_m(\Delta p_m + \Delta \Phi_{1/m} + f) - \nabla p_m \cdot (\nabla p_m + \nabla \Phi_{1/m}) = 0.$$

*Proof.* (a) is due to [53, Section 3.2.2], and (b) follows from [33, Corollary 2.11].  $\square$

We now turn to the following estimates on the size and support of solutions to  $(P)_m$ , which are uniform in  $m$ . The first ensures that if the initial data is bounded uniformly in  $m$ , it

remains so on bounded time intervals. The second ensures that if the support of the initial data is bounded uniformly in  $m$ , it likewise remains so on bounded time intervals.

**Lemma 3.4** (Estimates on size and support of solutions to  $(P)_m$ ). *Let  $p_m$  be a viscosity solution of  $(P)_m$  with continuous, compactly supported initial data  $p_m(\cdot, 0)$ . Suppose that there exists  $R_0 \geq 1$  sufficiently large so that  $\{p_m(\cdot, 0) > 0\} \subseteq B_{R_0/2}(0)$  and  $p_m(\cdot, 0) \leq R_0^2/4d$ . Define  $R(t) := (R_0 + \frac{C_d}{d})e^{t/d} - \frac{C_d}{d}$ , with  $C_d > 0$  as in (2.8). Then,*

- (a)  $\{p_m(\cdot, t) > 0\} \subseteq B_{R(t)}(0)$  for all  $t \in [0, T]$ ;
- (b)  $p_m(x, t) \leq R(t)^2/2d$  for all  $t > 0$ .

*Proof.* We prove the result by comparison with a classical supersolution of  $(P)_m$ . Define

$$h(x) = \begin{cases} \frac{1-|x|^2}{2d} & \text{for } |x| < 1, \\ 0 & \text{for } |x| \geq 1, \end{cases}$$

so that  $h(x)$  satisfies  $-\Delta h = 1$  in  $|x| < 1$  and  $h = 0$  in  $|x| \geq 1$ . Let  $\phi(x, t) := R(t)^2 h(x/R(t))$ , where  $R(t)$  solves  $R'(t) = \frac{R(t)}{d} + C_d$  with  $r(0) = R_0$ , and  $C_d$  is the upper bound of  $\|\nabla \Phi_{1/m}\|_\infty$  given by (2.8). We claim that such  $\phi$  is a classical supersolution of  $(P)_m$  for all  $m$ . To check this, direct computation gives that in the support of  $\phi$ ,

$$(3.2) \quad \phi_t(x, t) = 2R(t)R'(t) \frac{1 - \frac{|x|^2}{R(t)^2}}{2d} + R(t)^2 \left( -\frac{x}{dR(t)} \right) \cdot \left( -\frac{xR'(t)}{R(t)^2} \right) = \frac{R(t)R'(t)}{d}.$$

and  $\|\nabla \phi(\cdot, t)\|_\infty = R(t)\|\nabla h\|_\infty = \frac{R(t)}{d}$ . In addition, since  $\Delta \phi = -1$  in its support and  $\Delta \Phi_{1/m} \leq 1$  for all  $m$ , we have

$$(3.3) \quad (m-1) \underbrace{\phi(\Delta \phi + \Delta \Phi_{1/m})}_{\leq 0} - \nabla \phi \cdot (\nabla \phi + \nabla \Phi_{1/m}) \leq \frac{R(t)}{d} \left( \frac{R(t)}{d} + C_d \right).$$

Comparing (3.2) with (3.3) gives that  $\phi$  is a classical supersolution if  $R'(t) = \frac{R(t)}{d} + C_d$ . With  $R(0) = R_0$ , we have  $p_m(\cdot, 0) \leq \phi(\cdot, 0)$  for all  $m$ , so comparison principle yields that  $\{p_m(\cdot, t) > 0\} \subseteq B_{R(t)}(0)$  for all  $t$ , and  $p_m(x, t) \leq R(t)^2/2d$  for all  $x, t$ .  $\square$

*Remark 3.5.* Lemma 3.4 (b) and the fact that  $\rho_m = (\frac{m-1}{m} p_m)^{1/(m-1)}$  directly lead to the bound

$$(3.4) \quad \limsup_{m \rightarrow \infty} \|\rho_m(\cdot, t)\|_\infty \leq 1 \quad \text{for all } t \geq 0,$$

which we will make use of in what follows.

A key property of viscosity solutions of  $(P)_m$  is that they satisfy a comparison principle, which we now recall. We say two functions  $f, g : \mathbb{R}^d \rightarrow [0, \infty)$  are *strictly separated*, denoted by  $f \prec g$ , if  $f < g$  in  $\{f > 0\}$ , and  $\overline{\{f > 0\}}$  is a compact subset of  $\{g > 0\}$ .

**Theorem 3.6** (comparison theorem for  $(P)_m$ ). *Suppose  $u$  and  $v$  are viscosity sub- and supersolutions of  $(P)_m$ . If the initial data are strictly ordered, i.e.*

$$u(\cdot, 0) < v(\cdot, 0) \text{ in } \overline{\{u(\cdot, 0) > 0\}} \text{ and } \overline{\{u(\cdot, 0) > 0\}} \text{ is a compact subset of } \{v(\cdot, 0) > 0\},$$

*then  $u(\cdot, t) \prec v(\cdot, t)$  for all  $t > 0$ .*

*Proof.* The result follows from [33, Theorem 2.25].  $\square$

We also have the following comparison theorem for solutions to  $(P)_\infty$ , which we prove at the end of this section.

**Theorem 3.7** (comparison theorem for  $(P)_\infty$ ). *Suppose  $(u, \Sigma)$  and  $v$  are respectively viscosity sub- and supersolutions of  $(P)_\infty$ . If the initial data are strictly ordered, i.e.*

$$(3.5) \quad u(\cdot, 0) < v(\cdot, 0) \text{ in } \overline{\Sigma \cap \{t = 0\}} \text{ and } \overline{\Sigma \cap \{t = 0\}} \text{ is a compact subset of } \{v(\cdot, 0) > 0\},$$

then  $u(\cdot, t) < v(\cdot, t)$  and  $\overline{\Sigma} \cap \{t\} \subset \{v(\cdot, t) > 0\}$  for all  $t > 0$ .

While the above theorem almost provides uniqueness of  $(P)_\infty$ , the requirement that the initial data be strictly ordered prevents us from concluding this result. However, combining the comparison principle with Perron's method yields the following:

**Theorem 3.8.** *For any bounded open set  $\Omega_0 \subseteq \mathbb{R}^d$  with Lipschitz boundary, there exists minimal and maximal viscosity solutions of  $(P)_\infty$ .*

*Proof.* The result follows from [32]. □

We will use this comparison theorem, as well as the  $L^1$  contraction theorem for  $\rho_m$ , to obtain our first main result: we identify  $\rho_\infty$  with the characteristic function on the support of the minimal viscosity solution of  $(P)_\infty$ , when the initial data  $p_0$  is given by (1.9).

3.1.1. *Comparison theorem for  $(P)_\infty$ .* To conclude this section on basic properties of viscosity solutions of  $(P)_m$  and  $(P)_\infty$ , we sketch the proof of the comparison principle for  $(P)_\infty$ , Theorem 3.7. Our approach is to consider the first contact time for regularizations of the sub- and supersolutions, obtained by considering their sup and inf convolutions over space-time smooth sets. Such regularizations are often used to prove comparison principles for free boundary problems (c.f. [1, 13, 14, 32]), as they ensure that, when the free boundaries intersect for the first time, the free boundaries have both the interior and exterior ball property at the contact point. This provides sufficient regularity to consider a first-order asymptotic expansion of the free boundary graph at the contact point. In many ways, our proof parallels previous work by Alexander, Kim, and Yao [1, Theorem 2.7], the main differences being that our drift term  $\nabla \Phi_{1/m}$  has less regularity uniformly in  $m$ . This makes  $(P)_\infty$  more susceptible to perturbations, so we must carefully choose our regularization procedure so that the regularized solutions remain sub- and supersolutions of the original problem.

We now describe the details of these regularizations of the sub- and supersolutions. Fix  $r_0 \in [0, e^{(-1-\sqrt{2})/2})$ . Let  $r(t)$  be the unique solution to

$$(3.6) \quad \begin{cases} r'(t) &= -2C_d \sigma(r(t)), \\ r(0) &= r_0, \end{cases}$$

with  $C_d$  and  $\sigma(x)$  as defined in Proposition 2.6. Given  $(u, \Sigma)$  and  $v$  as in Theorem 3.7, define the spatial sup and inf convolutions

$$(3.7) \quad u^r(x, t) := \sup_{B_{r(t)}(x)} u(y, t), \quad v^r(x, t) := \inf_{B_{r(t)}(x)} v(y, t), \quad \Sigma^r := \cup_{t>0} \overline{\Omega^{r(t)}(t)} \times \{t\},$$

where  $\overline{\Omega^{r(t)}(t)} := \{x : d(x, \Omega(t)) \leq r(t)\}$ . Next we define the spacetime sup and inf convolutions

$$(3.8) \quad \tilde{u}^r(x, t) := \sup_{B_{r^*}(x, t)} u^r(y, s), \quad \tilde{v}^r(x, t) := \inf_{B_{r^*}(x, t)} v^r(y, s), \quad \tilde{\Sigma}^r := \{(x, t) : d((x, t), \Sigma^r) < r^*\}.$$

for fixed  $r_* > 0$ .

**Lemma 3.9.** *Let  $r^* := [\sigma(r(T))/11]^{2d}$ . Then  $(\tilde{u}^r, \tilde{\Sigma}^r)$  and  $\tilde{v}^r$  are viscosity sub- and supersolutions of  $(P)_\infty$  in  $\mathbb{R}^n \times (r_0, T - r^*)$ .*

*Proof.* We will show that  $(\tilde{u}^r, \tilde{\Sigma}^r)$  is a subsolution of  $(P)_\infty$  according to Definition 5.1. Parallel arguments apply to prove that  $\tilde{v}^r$  is a supersolution.

We begin by showing that  $(u^r, \Sigma^r)$  is a subsolution of  $(P)_\infty$ . It follows quickly from the definition that  $u^r$  is upper semicontinuous and  $(u^r, \Sigma^r)$  satisfies (a), so we devote our attention to (b).

Suppose  $u^r - \varphi$  has a local maximum zero at  $(x_0, t_0)$  with respect to  $\overline{\Sigma^r} \cap \{t \leq t_0\}$ . In case (i), if either  $x_0 \in \Omega^r(t_0)$  or  $u^r(x_0, t_0) > 0$ , fix  $y_0 \in \overline{B_{r(t_0)}}(x_0)$  so that  $u(y_0, t_0) = u^r(x_0, t_0)$  and either  $y_0 \in \Omega(t_0)$  or  $u(y_0, t_0) > 0$ . Let  $\tilde{\varphi}(x, t) = \varphi(x + x_0 - y_0, t)$ , so  $u - \tilde{\varphi}$  has a local max zero in  $\overline{\Sigma} \cap \{t \leq t_0\}$  at  $(y_0, t_0)$ . Since  $(u, \Sigma)$  is a subsolution, we have  $-\Delta\varphi(x_0, t_0) = -\Delta\tilde{\varphi}(y_0, t_0) \leq 1$ .

Now consider case (ii), so  $u^r(x_0, t_0) = 0$ , with  $(x_0, t_0) \in \partial\Sigma^r$  and  $|\nabla\varphi|(x_0, t_0) \neq 0$ . Choose  $y_0 \in \partial B_{r(t_0)}(x_0)$  so that  $(y_0, t_0) \in \partial\Sigma$  and  $u(y_0, t_0) = 0$ . Since  $|\nabla\varphi|(x_0, t_0) \neq 0$  and  $u^r - \varphi$  has a local maximum at  $(x_0, t_0)$ ,  $y_0 - x_0$  is parallel to  $v := \nabla\varphi/|\nabla\varphi|(x_0, t_0)$ .

Define  $\tilde{\varphi}(x, t) := \varphi(x - r(t)v, t)$ . Then  $\tilde{\varphi}(y_0, t_0) = \varphi(x_0, t_0)$  and  $u - \tilde{\varphi}$  has a local maximum zero at  $(y_0, t_0)$  with respect to  $\overline{\{u > 0\}} \cap \{t \leq t_0\}$ . Consequently,

$$\min(-\Delta\tilde{\varphi} - 1, \tilde{\varphi}_t - |\nabla\tilde{\varphi}|^2 - \nabla\tilde{\varphi} \cdot \nabla\Phi)(y_0, t_0) \leq 0.$$

By Proposition 2.6,

$$|\nabla\Phi(y_0, t_0) - \nabla\Phi(x_0, t_0)| \leq C_d\sigma(r(t_0)).$$

Since  $\tilde{\varphi}_t(y_0, t_0) = \varphi_t(x_0, t_0) - r'(t_0)|\nabla\varphi|(x_0, t_0)$ , we have

$$\min(-\Delta\varphi - 1, \varphi_t - |\nabla\varphi|^2 - \nabla\varphi \cdot \nabla\Phi - r'|\nabla\varphi| - C_d\sigma(r)|\nabla\varphi|)(x_0, t_0) \leq 0.$$

Hence, since  $r'(t) = -2C_d\sigma(r(t))$ ,

$$(3.9) \quad \min(-\Delta\varphi - 1, \varphi_t - |\nabla\varphi|^2 - \nabla\varphi \cdot \nabla\Phi + C_d\sigma(r)|\nabla\varphi|)(x_0, t_0) \leq 0,$$

which, in particular, implies (ii). Therefore,  $(u^r, \Sigma^r)$  is a subsolution of  $(P)_\infty$ .

Now we use this fact to show that  $(\tilde{u}^r, \tilde{\Sigma}^r)$  is a subsolution of  $(P)_\infty$ . Again,  $\tilde{u}^r$  is upper semicontinuous and  $(\tilde{u}^r, \tilde{\Sigma}^r)$  satisfies (a), so we devote our attention to (b). Property (i) follows as above, so we only discuss (ii), the barrier property on the free boundary. Suppose  $\tilde{u}^r - \varphi$  has a local maximum zero in  $\overline{\tilde{\Sigma}^r} \cap \{t \leq t_0\}$  at  $(x_0, t_0) \in \partial\tilde{\Sigma}^r$  with  $\tilde{u}^r(x_0, t_0) = 0$  and  $|\nabla\varphi|(x_0, t_0) \neq 0$ . Choose  $(y_0, s_0) \in \partial B_{r^*}(x_0, t_0)$  so  $(y_0, s_0) \in \partial\Sigma^r$  and  $u^r(y_0, s_0) = 0$ . Since  $|\nabla\varphi|(x_0, t_0) \neq 0$  and  $\tilde{u}^r - \varphi$  has a local maximum zero at  $(x_0, t_0)$ ,  $(y_0 - x_0, s_0 - t_0)$  is parallel to  $w := \nabla_{x,t}\varphi/|\nabla_{x,t}\varphi|(x_0, t_0)$ .

Define  $\tilde{\varphi}(x, t) := \varphi((x, t) - r^*w)$ . Then  $\tilde{\varphi}(y_0, s_0) = \varphi(x_0, t_0)$  and  $u^r - \tilde{\varphi}$  has a local maximum zero at  $(y_0, s_0)$  with respect to  $\overline{\{u^r > 0\}} \cap \{t \leq t_0\}$ . Consequently, by inequality (3.9) above,

$$\min(-\Delta\tilde{\varphi} - 1, \tilde{\varphi}_t - |\nabla\tilde{\varphi}|^2 - \nabla\tilde{\varphi} \cdot \nabla\Phi + C_d\sigma(r)|\nabla\tilde{\varphi}|)(y_0, s_0) \leq 0.$$

For  $0 \leq x \leq e^{(-1-\sqrt{2})/2}$ , we have  $-\log(x) \leq x^{-1/2} \leq x^{-1+1/2d}$ , hence  $\sigma(x) \leq x^{1/2d}$ . Therefore, by Proposition 2.6,

$$|\nabla\Phi(y_0, s_0) - \nabla\Phi(x_0, t_0)| \leq 11C_d(r^*)^{1/2d} = C_d\sigma(r(T)) \leq C_d\sigma(r(s_0)),$$

and thus it follows that

$$\min(-\Delta\varphi - 1, \varphi_t - |\nabla\varphi|^2 - \nabla\varphi \cdot \nabla\Phi)(x_0, t_0) \leq 0,$$

which concludes the proof.  $\square$

The rest of the proof of Theorem 3.7 parallels [1, Theorem 2.7], so we omit the proof.

**3.2. Convergence of  $(\mathbf{P})_m$  to  $(\mathbf{P})_\infty$ .** In this section, we show that, as  $m \rightarrow +\infty$ , viscosity solutions  $p_m$  of  $(\mathbf{P})_m$  approach a solution  $p$  of  $(\mathbf{P})_\infty$  and use this to show that patch solutions to the congested aggregation equation satisfy  $\rho_\infty = \chi_{\Omega(t)}$  almost everywhere, where  $\Omega(t) = \{p(\cdot, t) > 0\}$ .

We begin with the following lemma which states that  $\rho_m$  converges to  $\rho_\infty$  weakly even if  $\rho_m$  has initial data  $(\frac{m}{m-1}p_0)^{1/(m-1)}$ , instead of requiring the initial data of  $\rho_m$  to coincide with the initial data of  $\rho_\infty$ , as proved in Theorem 2.22.

**Lemma 3.10.** *Let  $\Omega_0 \subseteq \mathbb{R}^d$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty(\cdot, t)$  be the gradient flow of  $E_\infty$  with initial data  $\rho_0 = \chi_{\Omega_0}$ . Let  $\rho_m$  be the weak solution of  $(\text{PME-D})_m$  with initial data  $(\frac{m}{m-1}p_0)^{1/(m-1)}$ , where  $p_0$  is as in (1.9). Then for any  $t \geq 0$  and any  $f \in C(\mathbb{R}^d)$ , we have*

$$(3.10) \quad \lim_{m \rightarrow \infty} \int_{\mathbb{R}^d} \rho_m(x, t) f(x) dx = \int_{\mathbb{R}^d} \rho_\infty(x, t) f(x) dx \quad \text{for all } t \geq 0.$$

*Proof.* We will first prove (3.10) for all  $f \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ , and at the end of the proof we will extend it to all (possibly unbounded) continuous functions.

Let  $\tilde{\rho}_m$  be the weak solution of  $(\text{PME-D})_m$  with initial data  $\chi_{\Omega_0}$ . Theorem 2.22 then yields that  $\lim_{m \rightarrow \infty} W_2(\tilde{\rho}_m(t), \rho_\infty(t)) = 0$  for any  $t > 0$ . By [2, Remark 7.1.11 and Remark 5.1.2], convergence in  $W_2$  distance implies that

$$(3.11) \quad \lim_{m \rightarrow \infty} \int_{\mathbb{R}^d} \tilde{\rho}_m(x, t) f(x) dx = \int_{\mathbb{R}^d} \rho_\infty(x, t) f(x) dx \quad \text{for all } f \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d).$$

To relate  $\tilde{\rho}_m$  with  $\rho_m$ , note that they are both weak solutions to  $(\text{PME-D})_m$ , with different initial data  $\chi_{\Omega_0}$  and  $(\frac{m}{m-1}p_0)^{1/(m-1)}$  respectively. Since  $(\frac{m}{m-1}p_0)^{1/(m-1)} \rightarrow \chi_{\Omega_0}$  pointwise as  $m \rightarrow +\infty$ , we have

$$\lim_{m \rightarrow \infty} \left\| \left( \frac{m}{m-1} p_0 \right)^{1/(m-1)} - \chi_{\Omega_0} \right\|_{L^1(\mathbb{R}^d)} = 0$$

by dominated convergence theorem. Also, recall that for any  $m > 1$  and  $t \geq 0$ , the  $L^1$  contraction result in Lemma 3.4 gives

$$\|\tilde{\rho}_m(\cdot, t) - \rho_m(\cdot, t)\|_{L^1(\mathbb{R}^d)} \leq \left\| \left( \frac{m}{m-1} p_0 \right)^{1/(m-1)} - \chi_{\Omega_0} \right\|_{L^1(\mathbb{R}^d)}.$$

Combining the above two equations yields  $\lim_{m \rightarrow \infty} \|\tilde{\rho}_m(\cdot, t) - \rho_m(\cdot, t)\|_{L^1(\mathbb{R}^d)} = 0$ , hence

$$(3.12) \quad \lim_{m \rightarrow \infty} \int_{\mathbb{R}^d} (\tilde{\rho}_m(x, t) - \rho_m(x, t)) f(x) dx = 0 \quad \text{for all } f \in L^\infty(\mathbb{R}^d).$$

Putting (3.11) and (3.12) together gives us (3.10) for all  $f \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ . To remove the requirement  $f \in L^\infty(\mathbb{R}^d)$ , recall that Lemma 3.4 gives that  $\tilde{\rho}_m(\cdot, t)$  is supported in some bounded set  $B(0, R(t))$  for all  $m$ , and as a result  $\rho_\infty(\cdot, t)$  is supported in it too. If  $f \in C(\mathbb{R}^d)$  is unbounded, we can simply set  $\tilde{f} = f\eta$ , where  $\eta$  is a smooth cut-off function that is 1 in  $B(0, R(t))$  and 0 outside of  $B(0, R(t) + 1)$ . We then have (3.10) holds for  $\tilde{f}$ . Since changing  $f$  to  $\tilde{f}$  will not change the integrals in (3.10), we know (3.10) holds for  $f$  too.  $\square$

We begin our study of the limit of solutions of  $(P)_m$  with the following result, which shows that the half relaxed “limit infimum” of solutions of  $(P)_m$  is a supersolution of  $(P)_\infty$ .

**Proposition 3.11.** *Suppose  $p_m(x, t)$  is a viscosity solution of  $(P)_m$  with initial data  $p_m(\cdot, 0) = p_0$  as given in (1.9). Then the half relaxed limit*

$$(3.13) \quad u_2(x, t) := \liminf_* p_m(x, t) = \lim_{n \rightarrow +\infty} \inf_{\substack{m > n \\ |(x,t)-(y,s)| < 1/n}} p_m(y, s)$$

is a viscosity supersolution of  $(P)_\infty$ .

*Proof.* First, note that  $u_2$  is lower semicontinuous. Next, suppose that  $u_2 - \varphi$  has a local minimum zero at  $(x_0, t_0)$  with respect to  $\mathbb{R}^d \cap \{t \leq t_0\}$ . By subtracting  $\delta(x - x_0)^2 + \delta(t_0 - t)$  from  $\varphi$  with  $\delta > 0$  sufficiently small, we may assume that  $u_2 - \varphi$  has a strict minimum at  $(x_0, t_0)$  with respect to a parabolic neighborhood  $Q$  of  $(x_0, t_0)$ . Define  $(x_m, t_m) := \operatorname{argmin}_Q p_m - \varphi$ ,  $C_m := p_m(x_m, t_m) - \varphi(x_m, t_m)$ , and  $\varphi_m := \varphi + C_m$ , so  $p_m - \varphi_m$  has a local minimum zero at  $(x_m, t_m)$ . As in [1, Theorem 3.4] (see paragraph A.2), up to a subsequence, we have  $(x_m, t_m) \rightarrow (x_0, t_0)$ ,  $\lim_{m \rightarrow +\infty} p_m(x_m, t_m) = \liminf_* p_m(x_0, t_0) = u_2(x_0, t_0)$ , and  $(x_m, t_m) \in \{p_m > 0\}$ . Since  $p_m$  is a viscosity supersolution of  $(P)_m$ ,

$$(3.14) \quad (\varphi_t - (m-1)p_m(\Delta\varphi + \Delta\Phi_{1/m}) - \nabla\varphi \cdot (\nabla\varphi + \nabla\Phi_{1/m}))(x_m, t_m) \geq 0 \quad \text{for all } m.$$

First, consider the case when  $u_2(x_0, t_0) > 0$ , and assume  $-\Delta\varphi(x_0, t_0) < 1$  for the sake of contradiction. Since  $\varphi \in C^{2,1}(Q)$ , we have  $\varphi_t - \nabla\varphi \cdot (\nabla\varphi + \nabla\Phi_{1/m})|_{(x_m, t_m)}$  is uniformly bounded for all  $m$ . Hence if we can show that

$$(3.15) \quad \limsup_{m \rightarrow \infty} p_m(x_m, t_m) (\Delta\varphi(x_m, t_m) + \Delta\Phi_{1/m}(x_m, t_m)) > c > 0$$

for some positive  $c$ , it would imply that the left hand side of (3.14) goes to  $-\infty$  along a subsequence of  $m \rightarrow \infty$ , contradicting (3.14). Let us take a subsequence of  $m \rightarrow +\infty$  such that  $(x_m, t_m) \rightarrow (x_0, t_0)$  and  $p_m(x_m, t_m) > \frac{1}{2}u_2(x_0, t_0) > 0$  for all terms in this subsequence. (We still denote this subsequence by  $m$  for notational simplicity.) Since  $\lim_{m \rightarrow \infty} \Delta\varphi(x_m, t_m) = \Delta\varphi(x_0, t_0) > -1$ , to show (3.34), it suffices to show that  $\Delta\Phi_{1/m}(x_m, t_m) = 1$  for all sufficiently large  $m$  in this subsequence. Since  $(x_m, t_m)$  minimizes  $p_m - \varphi$  in  $Q$ , combining this with  $p_m(x_m, t_m) > \frac{1}{2}u_2(x_0, t_0)$  yields that  $p_m(x, t) > \frac{1}{4}u_2(x_0, t_0) > 0$  for all  $(x, t) \in Q \cap B_{r_0}(x_m, t_m)$ , where  $r_0 > 0$  is a sufficiently small constant only depending on  $u_2(x_0, t_0)$  and  $\|\varphi\|_{C^{1,1}(Q)}$ . Also, note that we have  $B_{r_0/2}(x_0, t_0) \subseteq B_{r_0}(x_m, t_m)$  for sufficiently large  $m$  due to the fact that  $(x_m, t_m) \rightarrow (x_0, t_0)$ . In other words, for all sufficiently large  $m$ , we have

$$\rho_m(x, t) = \left( \frac{m}{m-1} p_m(x, t) \right)^{\frac{1}{m-1}} \geq \left( \frac{u_2(x_0, t_0)}{8} \right)^{\frac{1}{m-1}} =: c_2^{\frac{1}{m-1}} \quad \text{for all } (x, t) \in Q \cap B_{r_0/2}(x_0, t_0).$$

Since  $\lim_{m \rightarrow \infty} c_2^{\frac{1}{m-1}} = 1$ , combining the above inequality with the weak convergence of  $\rho_m$  towards  $\rho_\infty$  in Lemma 3.10 yields that  $\rho_\infty(x, t) \geq 1$  almost everywhere in  $Q \cap B_{r_0/2}(x_0, t_0)$  (indeed, we must have  $\rho_\infty = 1$  in this set since  $\rho_\infty \leq 1$  by definition), hence

$$\Delta\Phi_{1/m}(x_m, t_m) = \psi_{1/m} * \rho_\infty(x_m, t_m) = 1$$

for all sufficiently large  $m$  in this subsequence, which yields (3.34) and thus contradicts (3.14).

Now, we consider the case when  $(x_0, t_0) \in \partial\{u_2 > 0\}$ ,  $u_2(x_0, t_0) = 0$  and  $\varphi$  satisfies (5.2). For a given small  $a > 0$ , let us consider a (second-order) perturbed version of  $\varphi$ ,

$$\tilde{\varphi}(x, t) := \varphi(x, t) - a(x - x_0) \cdot \nu - aC((x - x_0)^T)^2 + M((x - x_0) \cdot \nu)^2.$$

Here  $\nu = D\varphi(x_0, t_0) \neq 0$ ,  $(x - x_0)^T = (x - x_0) - [(x - x_0) \cdot \nu]\nu$ ,  $C = \max |D^2\varphi|$ . We choose  $M$  sufficiently large so that

$$(3.16) \quad \Delta\tilde{\varphi}(x_0, t_0) > 0.$$

Note that in a small neighborhood of  $(x_0, t_0)$  that depends on  $a$  and  $M$ ,  $\tilde{\varphi} \leq \max[\varphi, 0]$ . Thus  $\tilde{\varphi}$  still touches  $u_2$  from below at  $(x_0, t_0)$ . We will show that, for any choice of  $a > 0$ ,

$$(\tilde{\varphi}_t - |\nabla\tilde{\varphi}|^2 - \nabla\tilde{\varphi} \cdot \nabla\Phi)(x_0, t_0) \geq 0,$$

which yields the desired conclusion for  $\varphi$ .

Assume, for the sake of contradiction, that there is  $c_0 > 0$  so

$$(3.17) \quad \tilde{\varphi}_t - |\nabla\tilde{\varphi}|^2 - \nabla\tilde{\varphi} \cdot \nabla\Phi(x_0, t_0) < -c_0 < 0.$$

Let  $(x_m, t_m)$  chosen as before but with  $\tilde{\varphi}$  instead of  $\varphi$ . First, suppose  $p_m(x_m, t_m) > 0$  for sufficiently large  $m$ . Then (3.14) and (3.16)-(3.17) yields a contradiction.

Lastly suppose  $(x_m, t_m) \notin \{p_m > 0\}$  for  $m$  sufficiently large. Since  $(x_m, t_m) \in \overline{\{p_m > 0\}}$ , up to a subsequence, we have  $(x_m, t_m) \in \partial\{p_m > 0\}$ . This contradicts [1, Lemma 3.3]. (Though the proof of this lemma in [1] only considers the time independent case, it continues to hold in our setting.) □

Next we proceed to show that taking a ‘‘limit supremum’’ of  $p_m$  yields a subsolution of (P). Here we need to be a bit careful, due to the fact that subsolution property is based on maximum points only in the support of the subsolution. (See Definition 5.1.) Indeed, due to the nature of one-phase problem it is not possible to perturb a smooth test function  $\varphi$  to create a strict maximum of  $u^1 - \varphi$  without restricting the domain to  $\overline{\{u^1 > 0\}}$ . This can create technical difficulties with arguments along the lines of above proof to ensure that the local maximum points are stable under the limit  $m \rightarrow \infty$ , especially when the support of  $p_m$  degenerates as  $m \rightarrow \infty$ . To overcome this obstacle, we work with a modified notion of viscosity subsolutions, which are comprised of a pair  $(u, \Sigma)$ . This allows the set evolution  $\Sigma$  to be larger than the support of  $u$ . (See Definition 5.1 for details.)

**Proposition 3.12.** *Suppose  $p_m(x, t)$  is a viscosity solution of  $(P)_m$ . Define*

$$(3.18) \quad S(t) := \overline{\bigcap_{M \geq 1} (\bigcup_{m > M} \{p_m(\cdot, t) > 0\})}, \quad \Sigma_1 := \bigcup_{t > 0} (\mathring{S}(t) \times \{t\})$$

*Then if  $u_1$  is the half-relaxed limit of  $p_m$ ,*

$$u_1(x, t) := \limsup^* p_m(x, t) = \lim_{n \rightarrow +\infty} \sup_{\substack{m > n \\ |(x,t) - (y,s)| < 1/n}} p_m(y, s),$$

*$(u_1, \Sigma_1)$  is a viscosity subsolution of  $(P)_\infty$ .*

*Proof.* Since  $u_1$  is upper semicontinuous and  $\mathring{S}(t)$  is open, it remains to check properties (a)-(b) of Definition 5.1. By definition,  $S(t)^c \subseteq \{u_1(\cdot, t) = 0\}$ , hence  $\{u_1(\cdot, t) > 0\}^\circ \subseteq \mathring{S}(t)$ . By [1, Theorem B.1], for all  $t_0 > 0$ ,

$$(3.19) \quad \text{if } (x_0, t_0) \in \Sigma_1 \text{ then } (x_0, t_0) \in (\Sigma_1 \cap \{t \geq t_0\})^\circ$$

In particular, we have,  $\Sigma \cap \{t \leq t_0\} \subseteq \overline{\Sigma \cap \{t < t_0\}}$  for all  $t_0 > 0$ .

Now we turn to property (b). Let  $u - \varphi$  have a local maximum zero at  $(x_0, t_0)$  in  $\overline{\Sigma_1} \cap \{t \leq t_0\}$ . First we consider part (i), where either  $(x_0, t_0) \in \Sigma_1$  or  $u_1(x_0, t_0) > 0$ . By adding  $\delta(x - x_0)^2 + \delta(t_0 - t)$  to  $\varphi$  with  $\delta > 0$  sufficiently small, we may assume there is a parabolic neighborhood  $Q$  of  $(x_0, t_0)$  so that  $u_1 - \varphi$  has a strict maximum with respect to  $\overline{Q} \cap \overline{\Sigma_1} \cap \{t \leq t_0\} = \overline{Q} \cap \overline{\Sigma_1}$ .

First, note that it suffices to consider the case when  $u_1(x_0, t_0) > 0$ . In particular, if  $(x_0, t_0) \in \Sigma_1$ , then by (3.19), we may assume that  $Q$  is sufficiently small so that  $\overline{Q} \subseteq \overline{\Sigma_1}$ . This implies  $u_1(x_0, t_0) > 0$ , since otherwise  $\varphi(\cdot, t_0)$  has a local minimum zero at  $x_0$ , contradicting the fact that it is superharmonic. Likewise, we may assume that  $u_1 - \varphi$  has a strict maximum zero at  $(x_0, t_0)$  with respect to  $\overline{Q}$ , since  $\varphi(x_0, t_0) = u_1(x_0, t_0) > 0$  and  $u_1 = 0$  in  $(\overline{\Sigma_1})^c$ .

We now show that

$$(3.20) \quad -\Delta\varphi(x_0, t_0) \leq f(x_0) \text{ for any } f \in C(\mathbb{R}^d) \text{ such that } f \geq \rho_\infty(\cdot, t_0) \text{ almost everywhere.}$$

In particular, this implies that  $-\Delta\varphi(x_0, t_0) \leq 1$ , which gives the result. Suppose for the sake of contradiction that  $-\Delta\varphi(x_0, t_0) > f(x_0)$  for some  $f$ .

Let  $(x_m, t_m) := \operatorname{argmax}_{\overline{Q}} p_m - \tilde{\varphi}$ ,  $C_m := p_m(x_m, t_m) - \tilde{\varphi}(x_m, t_m)$ , and  $\tilde{\varphi}_m := \tilde{\varphi} + C_m$ , so  $p_m - \tilde{\varphi}_m$  has a maximum zero at  $(x_m, t_m)$  with respect to  $\overline{Q}$ . As in [1, Theorem 3.4] (see paragraph A.2), up to a subsequence, we have  $(x_m, t_m) \rightarrow (x_0, t_0) \in Q$ ,  $\lim_{m \rightarrow +\infty} p_m(x_m, t_m) = \limsup^* p_m(x_0, t_0) = u_1(x_0, t_0) > 0$ . Since  $p_m$  is a viscosity subsolution of  $(P)_m$ ,

$$(3.21) \quad (\varphi_t - (m-1)p_m(\Delta\varphi + \Delta\Phi_{1/m}) - \nabla\varphi \cdot (\nabla\varphi + \nabla\Phi_{1/m}))(x_m, t_m) \leq 0.$$

Because  $\Delta\Phi_{1/m} = \psi_{1/m} * \rho_\infty \leq f + o(1)$  and  $-\Delta\varphi(x_0, t_0) > f(x_0)$ , we have

$$(\Delta\varphi + \Delta\Phi_{1/m})(x_m, t_m) < 0$$

for sufficiently large  $m$ , which is a contradiction.

Now we consider part (ii), where  $(x_0, t_0) \in \partial\Sigma_1$ ,  $u_1(x_0, t_0) = 0$ , and  $|\nabla\varphi|(x_0, t_0) \neq 0$ . Suppose, for the sake of contradiction, that

$$(3.22) \quad -\Delta\varphi(x_0, t_0) > 1 \text{ and } (\varphi_t - |\nabla\varphi|^2 - \nabla\varphi \cdot \nabla\Phi)(x_0, t_0) > 0.$$

We can now apply parallel argument as in the proof of Theorem 3.4 in [AKY] to conclude.  $\square$

We now show that the initial data of the half-relaxed limits coincides with the initial data given in equation (1.9). Specifically, this ensures that the initial data of  $u_1$  and  $u_2$  coincides with the initial data of the sequence  $p_m$ , in spite of the time regularization inherent in the definitions of these half-relaxed limits. In what follows, we make frequent use of the following inner and outer approximations:

$$(3.23) \quad \Omega^{-r} := \{x : d(x, \Omega^c) > r\} \text{ and } \Omega^r := \{x : d(x, \Omega) < r\}.$$

**Lemma 3.13.** *Consider a bounded domain  $\Omega_0 \subseteq \mathbb{R}^d$  with the “no-crack” property  $\mathring{\Omega}_0 = \Omega_0$ . Suppose  $p_m$  are viscosity solutions of  $(P)_m$  with initial data  $p_0$  as given in (1.9). Then, for the half-relaxed limits  $u_1$  and  $u_2$ , we have  $u_1(x, 0) = u_2(x, 0) = p_0(x)$ .*

*Proof.* 1. We begin by proving the following claim on the support of  $p_m$  for given  $\varepsilon > 0$ :

(3.24) There is  $T_\varepsilon > 0$  such that  $\Omega_0^{-\varepsilon} \subseteq \{u_i(\cdot, t) > 0\} \subseteq \Omega_0^\varepsilon$  for all  $t \in [0, T_\varepsilon]$  and  $i = 1, 2$ .

We begin by showing  $\{p_m(\cdot, t) > 0\} \subseteq \Omega_0^\varepsilon$  for  $0 \leq t \leq t_\varepsilon$  for some  $t_\varepsilon \in (0, 1)$  that is independent of  $m$ . Suppose  $x_0 \in (\Omega_0^\varepsilon)^c$ , so that  $p_m = 0$  in  $B_\varepsilon(x_0)$  for all  $m$ .

Let us define

$$\phi(x, t) = \left( \mathcal{N}(x - x_0) - \frac{|x - x_0|^2}{2d} + f(t) \right)_+,$$

where  $f$  is an increasing function which we will determine momentarily. Let  $f(0) = -\mathcal{N}(\varepsilon) + C_1 + 1$ , where  $C_1 > 0$  is such that  $p_m(\cdot, t) \leq C_1$  for all  $m$  and  $t \in [0, 1]$ , given by Lemma 3.4(b). Such choice of  $f$  guarantees that  $\phi(x, 0) \geq C_1 \geq p_m(x, 0)$  in  $\{\varepsilon \leq |x - x_0| \leq 1\}$  (hence  $\phi(\cdot, 0) \geq p_m(\cdot, 0)$  in  $B_1(x_0)$ ), and  $\phi(x, t) \geq C_1 \geq p_m(x, t)$  on  $\partial B_1(x_0)$  for all  $t \in [0, 1]$ .

We claim that if we let  $f(t)$  increase sufficiently fast,  $\phi$  would be a classical supersolution of  $(P)_m$  for all  $m$  in  $B_1(x_0) \times [0, t_\varepsilon]$  for some  $t_\varepsilon > 0$ . Note that at time  $t$ ,  $\phi(\cdot, t)$  has support  $\{r(t) \leq |x - x_0| \leq 1\}$ , where  $r(t) \in (0, 1)$  solves  $\mathcal{N}(r(t)) - \frac{r(t)^2}{2d} + f(t) = 0$ , hence it satisfies  $r(t) > \mathcal{N}^{-1}(-f(t) + 1) > 0$ . (Here  $\mathcal{N}^{-1}$  is the inverse function of  $\mathcal{N}$ ). By definition,  $\Delta\phi = -1$  in its support. Thus in order to make  $\phi$  a classical supersolution of  $(P)_m$ , all we need is  $\phi_t \geq |\nabla\phi|(|\nabla\phi| + |\nabla\Phi_{1/m}|)$  everywhere in its support. In the support of  $\phi$ , we have  $\phi_t = f'(t)$ , and  $|\nabla\phi| \leq \mathcal{N}'(r(t)) + 1 \leq \mathcal{N}'(\mathcal{N}^{-1}(-f(t) + 1)) + 1$ . Finally, let

$$f'(t) = (\mathcal{N}'(\mathcal{N}^{-1}(-f(t) + 1)) + 1 + C_d)^2,$$

where  $C_d$  is the bound for  $|\nabla\Phi_{1/m}|$  as in (2.8). The standard ODE theory guarantees that  $f$  is finite in some  $[0, t_\varepsilon]$  (where  $t_\varepsilon > 0$  depends on  $\varepsilon$ ,  $C_1$  and  $d$ ), hence  $r(t) > 0$  in  $[0, t_\varepsilon]$ . By comparing  $p_m$  with  $\phi$  in the domain  $B_1(x_0) \times [0, t_\varepsilon]$  and using the definition of viscosity solutions, we conclude that  $p_m = 0$  in  $B_{r(t)}(x_0)$  for  $t \in [0, t_\varepsilon]$ . In particular,  $x_0 \in \{p_m(\cdot, t) > 0\}^c$  for all  $m$  and all  $t \in [0, t_\varepsilon]$ , and since  $x_0 \in (\Omega_0^\varepsilon)^c$  was arbitrary, this gives the result.

Similarly we show  $\Omega_0^{-\varepsilon} \subseteq \{p_m(\cdot, t) > 0\}$  for small times by constructing a classical subsolution of  $(P)_m$ . Suppose  $y_0 \in \Omega_0^{-\varepsilon}$ , so that  $B_\varepsilon(y_0) \subseteq \Omega_0$ . Let  $h_m(x, t)$  solve  $-\Delta h_m(\cdot, t) = \frac{1}{m}$  in  $B_{r(t)}(y_0)$ , with  $h_m(\cdot, t) = 0$  on  $\partial B_{r(t)}(y_0)$ . Here  $r(t) := \varepsilon - Mt$ , and  $M$  is a large constant to be determined later. Note that  $h_m$  takes the explicit expression  $h_m(x, t) := \left( \frac{r(t)^2 - |x - y_0|^2}{2dm} \right)_+$ , thus  $|\nabla h_m(\cdot, t)| \leq r(t)/dm$  in its support. So the following holds in the support of  $h_m$ :

$$\begin{aligned} (m-1)h_m(\Delta h_m + \Delta\Phi_{1/m}) + \nabla h_m \cdot (\nabla h_m + \nabla\Phi_{1/m}) &\geq -\frac{m-1}{m}h_m - \frac{r(t)}{dm} \left( \frac{r(t)}{dm} + C_d \right) \\ &\geq -\frac{r(t)^2}{2dm} - \frac{r(t)}{dm} \left( \frac{r(t)}{dm} + C_d \right), \end{aligned}$$

where  $C_d$  is the bound for  $\|\nabla\Phi_{1/m}\|$  by (2.8), and we also used  $\Delta\Phi_{1/m} \geq 0$  in the first inequality. Since  $(h_m)_t = r(t)r'(t)/dm$  in its support, in order for  $h_m$  to be a classical subsolution of  $(P)_m$ , all we need is  $r' \leq -r/2 - (r + \|\nabla\Phi_{1/m}\|_\infty)$ , so we can simply let  $r(t) = \varepsilon - Mt$  with  $M = 1 + C_d$ .

Since  $p_m(\cdot, 0) \geq h_m(\cdot, 0)$  for all  $m > 1$ , comparison principle yields that  $p_m \geq h_m$  for  $0 \leq t \leq \frac{\varepsilon}{2M}$ . It follows that  $p_m \geq h_m \geq \frac{\varepsilon^2}{16dm}$  in  $\Sigma_\varepsilon := B_{\varepsilon/4}(y_0) \times [0, \frac{\varepsilon}{2M}]$ . Even though this lower bound of  $p_m$  is not uniformly positive in  $m$ , we can still conclude that  $\liminf_{m \rightarrow \infty} \rho_m \geq 1$  in  $\Sigma_\varepsilon$  by definition of  $\rho_m = (\frac{m}{m-1}p_m)^{1/(m-1)}$ . Given the weak convergence of  $\rho_m$  to  $\rho_\infty$  in Lemma

3.10, we have that  $\rho_\infty = 1$  almost everywhere in  $\Sigma_\varepsilon$ . This implies that  $\Delta\Phi_{1/m} = \rho_\infty * \psi_{1/m} \equiv 1$  in  $B_{\varepsilon/8}(y_0) \times [0, \frac{\varepsilon}{2M}]$  for all sufficiently large  $m$  (more precisely, for all  $m > 8/\varepsilon$ ).

With this information on  $\Delta\Phi_{1/m}$ , we can now define a new subsolution  $\varphi(x, t)$  that solves  $-\Delta\varphi = 1$  in  $B_{\tilde{r}(t)}(y_0)$ , with  $\varphi(\cdot, t) = 0$  on  $\partial B_{\tilde{r}(t)}(y_0)$ , where  $\tilde{r}(t) = \varepsilon/8 - Mt$ , and  $M = 1 + C_d$ . One can check that  $\varphi$  is a classical subsolution of  $(P)_m$ , hence  $p_m \geq \varphi \geq c_\varepsilon$  for some  $c_\varepsilon > 0$  (that is independent of  $m$ ) in  $B_{\varepsilon/32} \times [0, T_\varepsilon]$  for all sufficiently large  $m$ , where  $T_\varepsilon := \frac{\varepsilon}{16M}$ , yielding (3.24).

2. To show that  $u_1(\cdot, 0) = u_2(\cdot, 0) = p_0$ , we construct our first barrier as follows: suppose  $h_\varepsilon(x)$  solves

$$-\Delta h_\varepsilon = 1 + \varepsilon \text{ in } \Omega^\varepsilon, \quad h_\varepsilon = 0 \text{ on } \partial\Omega^\varepsilon.$$

By (3.24),  $S(t) \subseteq \Omega^{\varepsilon/2}$  for  $t \in [0, \tilde{T}_\varepsilon]$  for some  $\tilde{T}_\varepsilon > 0$ . Thus,  $u_1 \leq h_\varepsilon$  in  $S(t)$ . Furthermore, since  $u_1 = 0$  in  $(S(t))^c$ , we conclude that

$$(3.25) \quad u_1(\cdot, t) \leq h_\varepsilon \text{ in } \Omega^\varepsilon \times [0, \tilde{T}_\varepsilon].$$

Next, comparison of  $p_m$  with the classical subsolution  $\varphi$  given above yields that

$$(3.26) \quad p_m \geq c_\varepsilon \text{ in } \Omega^{-\varepsilon} \times [0, T_\varepsilon].$$

Now we construct our second barrier using (3.26). Consider  $g(x)$  solving

$$-\Delta g = 1 - \varepsilon \text{ in } \Omega^{-\varepsilon}, \quad g = c_\varepsilon \text{ on } \partial\Omega^{-\varepsilon}.$$

Since  $u_2$  is a supersolution of  $(P)_\infty$  and (3.24) ensures that  $\Omega^{-\varepsilon} \subseteq \{u_2 > 0\}$  for  $t \in [0, T_\varepsilon]$ , we have  $-\Delta u_2 \geq 1$  in  $\Omega^{-\varepsilon} \times [0, T_\varepsilon]$ . Furthermore, (3.26) ensures that  $g \leq u_2$  on  $\partial\Omega^{-\varepsilon} \times [0, T_\varepsilon]$ . Therefore,

$$(3.27) \quad g \leq u_2 \text{ in } \Omega^{-\varepsilon} \times [0, T_\varepsilon].$$

Combining inequalities (3.25) and (3.27) and sending  $\varepsilon \rightarrow 0$ , we can conclude.  $\square$

We now show our main convergence theorem.

**Theorem 3.14.** *Let  $\Omega_0 \subseteq \mathbb{R}^d$  be a bounded domain with Lipschitz boundary, and let  $p_m$  solve  $(P)_m$  with initial data  $p_0$  as given in (1.9). Let  $\rho_m$  be the density variable corresponding to  $p_m$ , and let  $U$  be the unique maximal solution of  $(P)_\infty$  with initial data  $p_0$ , i.e.*

$$U(x, t) := (\inf\{w : w \text{ is a viscosity supersolution of } (P)_\infty \text{ with } w(\cdot, 0) \geq p_0\})_*.$$

Then the following holds for each  $t > 0$ :

- (a)  $\rho_\infty(\cdot, t) = \chi_{\{u_1(\cdot, t) > 0\}} = \chi_{\{u_2(\cdot, t) > 0\}} = \chi_{\{U(\cdot, t) > 0\}}$  almost everywhere;
- (b) Let  $\Omega(t) := \{u_2(\cdot, t) > 0\}$ , and  $\Omega^1(t) := \{u_1(\cdot, t) > 0\}$ . Then for every  $t \geq 0$ ,  $\Omega(t)$  is an open set with  $|\partial\Omega(t)| = 0$ , and we also have  $|\partial\Omega^1(t)| = 0$ ;
- (c)  $\rho_m$  converges to 1 uniformly in  $\Omega(t)$  away from its boundary—that is, the convergence is uniform in any compact set  $Q \subseteq \{(x, t) : x \in \Omega(t)\}$ . Furthermore, we have  $\lim_{m \rightarrow \infty} \|\rho_m(\cdot, t) - \chi_{\Omega(t)}\|_{L^1(\mathbb{R}^d)} = 0$  for every  $t \geq 0$ .

*Remark 3.15.* The fact that  $U$  is a solution of  $(P)_\infty$  is a consequence of a standard Perron's method argument.

*Proof.* 1. To begin with, let us define two families of functions that are approximations of  $p_m$ . For  $n \in \mathbb{N}$ , let  $p_0^{n,-}(x)$  and  $p_0^{n,+}(x)$  be solutions to (1.9) with  $\Omega_0$  replaced by  $\Omega_0^{-1/n}$  and  $\Omega_0^{1/n}$  (as defined in (3.23)) respectively. Note that  $p_0^{n,-} \prec p_0 \prec p_0^{n,+}$ . We then let  $p_m^{n,-}$  be the viscosity solution to  $(P)_m$  with initial data  $p_0^{n,-}$ , and denote by  $\rho_m^{n,-}$  the corresponding density function. We let  $p_m^{n,+}$  solve a modified version of  $(P)_m$  with an extra source term, namely

$$p_t = (m-1)p(\Delta p + \Delta \Phi_{1/m}) + \nabla p \cdot (\nabla p + \nabla \Phi_{1/m}) + p f_n,$$

where

$$f_n := \chi_{\Omega_n - \Omega}, \quad \Omega := \{u_2 > 0\} \text{ and } \Omega_n := \{(x, t) : d((x, t), \Omega) \leq \frac{1}{n}\},$$

and denote by  $\rho_m^{n,+}$  the corresponding density function. Finally we let  $(u_1^{n,-}, S^{n,-}(t))$  and  $u_2^{n,+}$  denote the corresponding half-relaxed limits for  $p_m^{n,-}$  and  $p_m^{n,+}$ .

The motivation of these two family of functions is as follows: in step 2, we will show that

$$(3.28) \quad u_1^{n,-} \leq U \leq u_2 \leq u_1 \leq u_2^{n,+} \quad \text{for any } n \in \mathbb{N},$$

and it turns out that in order to show the last inequality we have to let  $p_m^{n,+}$  solve the equation with the extra source term, rather than  $(P)_m$ . In step 3, we will use  $L^1$  contraction result between  $\rho_m^{n,-}$  and  $\rho_m^{n,+}$  to show that for any  $t \geq 0$ ,

$$(3.29) \quad A_n(t) := |\text{supp } u_2^{n,+}(\cdot, t) \setminus \text{supp } u_1^{n,-}(\cdot, t)| \rightarrow 0 \text{ as } n \rightarrow \infty,$$

and by combining it with (3.28) we have that  $U$ ,  $u_1$  and  $u_2$  are supported on the same set.

2. In this step we aim to prove (3.28). The second inequality is a direct consequence from the minimality of  $U$  and the fact that  $u_2$  is supersolution of  $(P)_\infty$  with initial data  $p_0$  (which follows from Proposition 3.11 and Lemma 3.13). The third inequality immediately follows from the definition of the half relaxed limits  $u_2$  and  $u_1$ . As for the first inequality, note that by Proposition 3.12,  $(u_1^{n,-}, S^{n,-}(t))$  is a subsolution of  $(P)_m$ . (Proposition 3.12 does not require the initial data be the same as  $p_0$ .) In addition, we have  $u_1^{n,-}(\cdot, 0) = p_0^{n,-}$  via the same argument as in Lemma 3.13. Since  $p_0^{n,-} \prec p_0$ , combining the above discussion on  $u_1^{n,-}$  with Proposition 3.11 and the comparison principle in Theorem 3.7 yields

$$(3.30) \quad u_1^{n,-} \prec u_2 \text{ with } \overline{S_1^{n,-}(t)} \subseteq \{u_2(\cdot, t) > 0\},$$

which gives us the first inequality.

The last inequality of (3.28) is more difficult to obtain. We point out that this is not a direct consequence of the comparison principle for  $(P)_\infty$ , since we do not know that  $u_2^{n,+}$  is a supersolution of  $(P)_\infty$  due to the fact that  $p_m^{n,+}(\cdot, 0) \neq p_0$ . (In order to apply Proposition 3.11, the initial data must be the same as  $p_0$ .)

To overcome this difficulty, we will show that  $u_1$  and  $u_2^{n,+}$  are sub- and supersolutions of another free boundary problem, for which the comparison principle also holds. From the proof (in particular (3.20)) of Proposition 3.12, it follows that in addition to  $(u_1, S(t))$  being a viscosity subsolution of  $(P)_\infty$ ,  $u_1$  satisfies

$$(3.31) \quad -\Delta u_1(\cdot, t) \leq \rho_\infty$$

in the integral sense. On the other hand, parallel arguments as in Proposition 3.11 yield that  $u_2^{n,+}$  satisfies supersolution properties of  $(P)_\infty$  (see Definition 5.2), but with the interior

operator  $-\Delta - 1$  replaced by  $-\Delta - \rho_\infty$ . In particular we have

$$(3.32) \quad -\Delta u_2^{n,+} \geq \rho_\infty \text{ in } \{u_2^{n,+} > 0\}$$

in the integral sense. As a result,  $(u_1, S(t))$  and  $u_2^{n,+}$  are respectively viscosity sub- and supersolutions of

$$(\tilde{P})_\infty \quad \begin{cases} -\Delta p(\cdot, t) = \rho_\infty & \text{in } \{p > 0\}; \\ V = -\nu \cdot (Dp + D\Psi) & \text{on } \partial\{p > 0\}; \\ \Psi = \mathcal{N} * \rho_\infty. \end{cases}$$

Using this fact, one can modify the proof of comparison principle for  $(P)_\infty$  to show that, for any  $n \in \mathbb{N}$ ,

$$(3.33) \quad u_1 \prec u_2^{n,+} \text{ with } \overline{S_1(t)} \subseteq \{u_2^{n,+}(\cdot, t) > 0\}.$$

The proof of (3.33) is parallel to that of Theorem 2.7 in [1] with the only difference lies showing the second inequality in the interior operator which we discuss below. Let us give a heuristic sketch of the proof under the assumption that  $S(t)$  and  $\{u_2^{n,+}(\cdot, t) > 0\}$  have smooth boundaries: the actual proof is carried out with regularizations as given in (3.7)-(3.8) which generate strict subsolution and supersolution of  $(\tilde{P})_\infty$ . As usual in the proof of comparison principle, we begin with the scenario that  $u_1$  crosses  $u_2^{n,+}$  from below at some time and yield a contradiction. More precisely we suppose that the first crossing time is finite, i.e.

$$t_0 := \sup\{t : u_1(\cdot, s) \prec u_2^{n,+}(\cdot, s) \text{ and } S(t) \subseteq \{u_2^{n,+}(\cdot, t) > 0\} \text{ for } s \leq t\} < \infty.$$

Note that  $t_0 > 0$  since  $S(0) = \overline{\Omega_0} = \overline{\{u_1(\cdot, 0) > 0\}}$  due to Lemma 3.13 and  $u_1 \prec u_2^{n,+}$  at  $t = 0$  from the construction. Observe also that (3.31)-(3.32) rules out the possibility that the crossing occurs at an interior point, i.e.,

$$u_1(\cdot, t) < u_2^{n,+}(\cdot, t) \text{ in } \{u_1(\cdot, t) > 0\}$$

as long as  $\{u_1(\cdot, t) > 0\} \subseteq S(t) \subseteq \overline{\{u_2^{n,+} > 0\}}$ .

Hence this means that the set  $S(t)$  touches the boundary of  $\overline{\{u_2^{n,+}(\cdot, t) > 0\}}$  for the first time at some point  $(x_0, t_0)$ . Then the normal velocity law for the sets  $S(t)$  and  $\{u_2^{n,+}(\cdot, t) > 0\}$ , as well as the fact that  $u_1(\cdot, t_0) \leq u_2^{n,+}(\cdot, t_0)$  yields a contradiction.

Note that (3.32) and the definition of  $\Omega_n$  ensures that the source term for  $u_1$  remains smaller than that of  $u_2^{n,+}$  after the regularization process given in (3.7)-(3.8) if  $r(t)$  is sufficiently small. Based on this fact, the rest of the proof is the same as to that of Theorem 2.7 in [1].

3. Next we will show (3.29). First, note that  $\rho_m^{n,-}$  satisfies (3.1) with no source term, while  $\rho_m^{n,+}$  satisfies (3.1) with source term  $\rho_m^{n,+} f_n$  which is non-negative. Since their initial data is also ordered, comparison principle for (3.1) yields that  $\rho_m^{n,-} \leq \rho_m^{n,+}$ . We then define

$$A_m^n(t) := \int (\rho_m^{n,+}(x, t) - \rho_m^{n,-}(x, t)) dx,$$

which is nonnegative. We can apply the  $L^1$  contraction property of (3.1) in Lemma 3.3 to conclude that, for any  $t > 0$  and any  $m > 1$ ,

$$(3.34) \quad A_m^n(t) \leq \int_0^t \int \rho_m^{n,+}(x, s) f_n(x, s) dx ds + A_m^n(0).$$

By (3.30) and the definition of  $S_1^{n,-}$ , for all  $t \geq 0$  and sufficiently large  $m$ , we have

$$(3.35) \quad \text{supp } \rho_m^{n,-}(t) = \text{supp } p_m^{n,-}(t) \subseteq \text{supp } u_2(t) =: \Omega(t).$$

Thus for all sufficiently large  $m$ , the spatial integral in (3.34) can be controlled as

$$\int \rho_m^{n,+}(x, s) f_n(x, s) dx = \int_{\Omega_n(s) \setminus \Omega(s)} \rho_m^{n,+}(x, s) dx \leq A_m^n(s),$$

where in the last step we used that  $\rho_m^{n,-}(x, s) \equiv 0$  in  $\Omega_n(s) \setminus \Omega(s)$  for all large  $m$ , which follows from (3.35). Plugging this into (3.34) yields  $A_m^n(t) \leq \int_0^t A_m^n(s) ds + A_m^n(0)$ , thus Gronwall's inequality immediately yields that  $A_m^n(t) \leq A_m^n(0)e^t$ . It is easy to check that  $\lim_{m \rightarrow \infty} A_m^n(0) = |\Omega_0^{1/n} \setminus \Omega_0^{-1/n}| \leq C/n$ , which yields  $\liminf_{m \rightarrow \infty} A_m^n(t) \leq Ce^t/n$  for all  $n \in \mathbb{N}, t \geq 0$ .

Next we claim

$$\liminf_{m \rightarrow \infty} A_m^n(t) \geq |\text{supp } u_2^{n,+}(\cdot, t) \setminus \text{supp } u_1^{n,-}(\cdot, t)|.$$

To show this, it suffices to show that

(3.36)

$$\liminf_{m \rightarrow \infty} \int \rho_m^{n,+}(x, t) dx \geq |\text{supp } u_2^{n,+}(\cdot, t)| \quad \text{and} \quad \limsup_{m \rightarrow \infty} \int \rho_m^{n,-}(x, t) dx \leq |\text{supp } u_1^{n,-}(\cdot, t)|.$$

For the first inequality, note that by definition of the half-relaxed limit, for any  $x \in \text{supp } u_2^{n,+}(\cdot, t)$ , we have  $\liminf_{m \rightarrow \infty} p_m^{n,+}(x, t) > 0$ , thus  $\liminf_{m \rightarrow \infty} \rho_m^{n,+}(x, t) \geq 1$ . Therefore

$$\int \liminf_{m \rightarrow \infty} \rho_m^{n,+}(x, t) dx \geq |\text{supp } u_2^{n,+}(\cdot, t)|,$$

and applying Fatou's lemma to it yields the first inequality of (3.36). The second inequality follows from the definition of the half-relaxed limit  $u_1^{n,-}$  and the fact that  $\limsup_{m \rightarrow \infty} \|\rho_m^{n,-}\|_\infty \leq 1$ , which is due to (3.4).

We now combine the above claim with  $\liminf_{m \rightarrow \infty} A_m^n(t) \leq Ce^t/n$  to conclude that the  $A_n(t)$  defined in (3.29) satisfies  $A_n(t) \leq Ce^t/n$  for all  $n \in \mathbb{N}, t \geq 0$ , which yields (3.29). Applying this to (3.28) then yields that  $\chi_{\{u_1 > 0\}} = \chi_{\{u_2 > 0\}} = \chi_{\{U > 0\}}$  almost everywhere. So the proof of part (a) would be finished if we can show  $\rho_\infty$  is also equal to these functions almost everywhere, which we postpone to step 4.

At the end of step 3, let us point out part (b) can be easily proved using the above bound on  $A_n(t)$ : note that  $\Omega(t) = \{u_2(\cdot, t) > 0\}$  is open due to the lower-semicontinuity of  $u_2$ , hence

$$\partial\Omega(t) = \overline{\Omega(t)} \setminus \Omega(t) \subseteq \text{supp } u_2^{n,+}(\cdot, t) \setminus \text{supp } u_1^{n,-}(\cdot, t),$$

where we used (3.28), (3.30) and (3.33) in the last inequality. The above bound on  $A_n(t)$  thus gives  $|\partial\Omega(t)| \leq A_n(t) \leq Ce^t/n$  for all  $n \in \mathbb{N}, t \geq 0$ , and by sending  $n \rightarrow \infty$  we obtain part (b) for  $\Omega(t)$ . In addition, the inequalities (3.28), (3.30) and (3.33) also lead to  $|\partial\Omega^1(t)| \leq A_n(t)$ , hence we also have  $|\partial\Omega^1(t)| = 0$ .

4. To finish the proof of part (a), it suffices to relate  $u_2$  and  $\rho_\infty$  and show that

$$(3.37) \quad \rho_\infty = \chi_{\{u_2 > 0\}} \text{ a.e.}$$

The direction  $\rho_\infty \geq \chi_{\{u_2 > 0\}}$  is easier: take any  $(x_0, t_0)$  such that  $a := u_2(x_0, t_0) > 0$ . By definition of the half-relaxed limit  $u_2$ , there exists some positive  $r_0$  and  $N_0$ , such that  $p_m(x, t_0) \geq a/2 > 0$  for all  $x \in B_{r_0}(x_0), m > N_0$ . Hence  $\rho_m(x, t_0) \geq (\frac{m-1}{m}a)^{1/(m-1)}$  for  $x, m$  as above. Combining this lower bound (which approaches 1 as  $m \rightarrow \infty$ ) with the weak convergence

of  $\rho_m$  towards  $\rho_\infty$  in Lemma 3.10, we have  $\rho_\infty(\cdot, t_0) \geq 1$  in  $B_{r_0}(x_0)$ . Since  $x_0 \in \{u_2(\cdot, t_0) > 0\}$  is arbitrary, we have  $\rho_\infty \geq \chi_{\{u_2 > 0\}}$ .

For the other direction, we will use  $p_m^{n,-}$ . Using the definition of  $S^{n,-}(t)$  (see (3.18)) as well as (3.28), we have

$$(3.38) \quad \limsup_{m \rightarrow \infty} \chi_{\{p_m^{n,-}(\cdot, t) > 0\}} = \chi_{S^{n,-}(t)} \leq \chi_{\{u_2(\cdot, t) > 0\}} \quad \text{for any } n \in \mathbb{N}, t \geq 0.$$

In addition, since  $\limsup_{m \rightarrow \infty} \|\rho_m^{n,-}(t)\|_\infty \leq \limsup_{m \rightarrow \infty} \|\rho_m(t)\|_\infty \leq 1$  by (3.4), it implies

$$(3.39) \quad \limsup_{m \rightarrow \infty} \rho_m^{n,-}(\cdot, t) \leq \limsup_{m \rightarrow \infty} \chi_{\{p_m^{n,-}(\cdot, t) > 0\}} = \limsup_{m \rightarrow \infty} \chi_{\{p_m^{n,-}(\cdot, t) > 0\}}.$$

Finally we will relate  $\rho_\infty$  with  $\rho_m^{n,-}$ . Note that for any continuous, bounded  $f \geq 0$ , we have the following (where we omit the  $x$  dependence in integrals due to space limitation):

$$(3.40) \quad \int \rho_\infty(t) f dx = \lim_{m \rightarrow \infty} \int \rho_m(t) f dx = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \int \rho_m^{n,-}(t) f dx \leq \int \limsup_{m, n \rightarrow \infty} \rho_m^{n,-}(t) f dx,$$

where the first equality follows from Lemma 3.10, the second one is due to the  $L^1$  contraction property of Lemma 3.3(a) and the fact that  $\rho_m^{n,-} \leq \rho_m$ , and the last inequality follows from Fatou's lemma. This implies that

$\rho_\infty(\cdot, t) \leq \limsup_{m, n \rightarrow \infty} \rho_m^{n,-}(\cdot, t)$  a.e., and this with (3.38)-(3.39) implies  $\rho_\infty \leq \chi_{\{u_2 > 0\}}$ , which finishes the proof of (3.37), thus yielding part (a).

5. To prove part (c), take any compact set  $Q \subseteq \{(x, t) : x \in \Omega(t)\}$ . By definition of the half-relaxed limit  $u_2$ , for each  $(x_0, t_0) \in Q$  there is some  $r_0 > 0$ , such that  $p_m \geq u(x_0, t_0)/2$  in  $B_{r_0}(x_0, t_0)$  for all sufficiently large  $m$ , which yields that  $\rho_m$  uniformly approaches 1 in  $B_{r_0}(x_0, t_0)$ . The compactness of  $Q$  then allows us to find a finite number of points  $(x_i, t_i)$  such that  $B_{r_i}(x_i, t_i)$  covers  $Q$ , implying that  $\rho_m \rightarrow 1$  uniformly in  $Q$ .

Finally let us prove the  $L^1$  convergence result, where we use the elementary inequality

$$(3.41) \quad \|f - g\|_{L^1} = \int (g - f)_+ dx + \int (f - g)_+ dx \leq 2 \int (g - f)_+ dx + \left| \int (f - g) dx \right|.$$

Let  $f = \rho_m(\cdot, t)$ ,  $g = \chi_{\Omega(t)} = \rho_\infty(\cdot, t)$ . Since the mass of  $\rho_m(\cdot, t)$  and  $\rho_\infty(\cdot, t)$  are both preserved in time, we have

$$\left| \int (f - g) dx \right| \leq \|\rho_m(\cdot, 0) - \chi_{\Omega_0}\|_{L^1} = \left\| \left( \frac{m}{m-1} p_0 \right)^{1/(m-1)} - \chi_{\Omega_0} \right\|_{L^1} \rightarrow 0 \text{ as } m \rightarrow \infty.$$

To control  $\int (g - f)_+ dx$ , note that  $g = 0$  a.e. in  $\Omega(t)^c$ , hence  $\int (g - f)_+ dx = \int_{\Omega(t)} (1 - \rho_m)_+ dx$ . Since  $\Omega(t)$  is open, for any  $\varepsilon > 0$  we can find a compact set  $D \subseteq \Omega(t)$ , such that  $|\Omega(t) \setminus D| \leq \varepsilon$ . We can then apply the uniform convergence result of  $\rho_m$  in  $D$  to conclude that  $\int (g - f)_+ dx \leq 2\varepsilon$  for sufficiently large  $m$ , and since  $\varepsilon > 0$  is arbitrary we have  $\lim_{m \rightarrow \infty} \int (g - f)_+ dx = 0$ . Plugging the above results into (3.41) yields the  $L^1$  convergence result.  $\square$

#### 4. LONG TIME BEHAVIOR

In this section, we investigate the long-time behavior of a patch solution  $\rho_\infty$  in two dimensions, using the pressure variable characterization of the dynamics of  $\rho_\infty$  obtained in section 3. Throughout this section, we consider our spatial domain to be  $\mathbb{R}^2$ . By Theorem 3.14, we know that  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  for some  $\Omega(t) \subseteq \mathbb{R}^2$  for all  $t \geq 0$ . Our goal is to show that, as  $t \rightarrow \infty$ ,

$\Omega(t)$  converges to the unique disk  $B_0$  with the same mass and the center of mass as  $\Omega_0$ . (See Theorem 4.12)

We proceed as follows: in sections 4.1 and 4.2, we show that the second moment of  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  decreases unless  $\Omega(t)$  is a disk, from which we are able to conclude that  $\Omega(t)$  cannot stay uniformly away from a disk for all times, in terms of its *Fraenkel asymmetry*. In section 4.3, we combine this with the gradient flow structure of  $\rho_\infty$  to show that as  $t \rightarrow +\infty$  the energy  $E_\infty(\rho_\infty(t))$  approaches the minimum of  $E_\infty$ , with a quantitative estimate on the rate. Lastly, in section 4.4, we show that  $\rho_\infty(\cdot, t)$  converges to  $\chi_{B_0}$  strongly in  $L^q$  for any  $1 \leq q < \infty$ .

**4.1. Evolution of the second moment.** Let  $M_2[f] := \int_{\mathbb{R}^2} f(x)|x|^2 dx$  denote the second moment of  $f$ . In this subsection, we investigate the evolution of the second moment of  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$ . Before we present the rigorous derivation of the evolution of the second moment, we begin with the following heuristic computation. As described in the introduction,  $\rho_\infty(\cdot, t)$  formally satisfies the transport equation

$$\rho_t = \nabla \cdot (\rho(\nabla \mathbf{N} \rho + \nabla p)),$$

where  $p$  is a solution to (P). (See equation (1.8).) By definition,  $p(\cdot, t)$  solves  $\Delta p = -1$  in  $\Omega(t)$  and  $p = 0$  on  $\partial\Omega(t)$ . Hence, supposing that  $\partial\Omega(t)$  is smooth, the evolution of  $M_2[\rho_\infty(t)]$  is given by

$$\begin{aligned} \frac{d}{dt} M_2[\rho_\infty(t)] &= -2 \int_{\mathbb{R}^2} \rho_\infty \nabla \mathbf{N} \rho_\infty \cdot x dx - 2 \int_{\mathbb{R}^2} \rho_\infty \nabla p \cdot x dx \\ &= -\frac{1}{\pi} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \rho_\infty(x) \rho_\infty(y) \frac{(x-y) \cdot x}{|x-y|^2} dy dx - 2 \int_{\Omega(t)} \nabla p \cdot x dx \\ (4.1) \quad &= -\frac{1}{2\pi} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \rho_\infty(x) \rho_\infty(y) dy dx + 4 \int_{\Omega(t)} p(x) dx \\ &= -\frac{1}{2\pi} |\Omega(t)|^2 + 4 \int_{\Omega(t)} p(x) dx = -\frac{1}{2\pi} |\Omega_0|^2 + 4 \int_{\Omega(t)} p(x) dx. \end{aligned}$$

In the second equality, we use that, in two dimensions,  $\mathbf{N} \rho_\infty = \mathcal{N} * \rho_\infty$  with  $\mathcal{N}(x) = \frac{1}{2\pi} \log|x|$ . In the third equality, we symmetrize  $x$  and  $y$  in the first integral (hence the extra factor of  $\frac{1}{2}$ ), and in the last equality, we use that  $\rho_\infty$  preserves its mass (which is  $|\Omega_0|$ ) for all time.

In the following proposition, we rigorously obtain the time evolution of  $M_2[\rho_\infty(t)]$  by analyzing the evolution of the second moments for each  $\rho_m$  and sending  $m \rightarrow +\infty$ , using our convergence results from the previous sections. We show that the evolution of the second moment indeed satisfies a time-integral form of (4.1), with the exception that we must substitute  $p(x)$  with  $u_1(x)$ , the half-relaxed limit of  $p_m$  defined in Lemma 3.12, to take into account the fact that  $\Omega(t)$  may not have smooth boundary for all time.

**Proposition 4.1.** *Let  $\Omega_0 \subseteq \mathbb{R}^2$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  be the gradient flow of  $E_\infty$  with initial data  $\rho_0 = \chi_{\Omega_0}$ . Then for any  $T > 0$ ,*

$$(4.2) \quad M_2[\rho_\infty(T)] - M_2[\rho_0] \leq -\frac{1}{2\pi} |\Omega_0|^2 T + 4 \int_0^T \int_{\Omega(t)} u_1(x, t) dx dt,$$

where  $u_1$  is the half-relaxed limit of  $p_m$ , defined in Lemma 3.12, and  $\Omega(t) = \{u_2(\cdot, t) > 0\}$ , as defined in Theorem 3.14 (b).

*Proof.* For any  $m > 1$ , let  $\rho_m$  be the weak solution of (PME-D) $_m$  with initial data  $(\frac{m-1}{m}p_0)^{1/(m-1)}$ , where  $p_0$  is given by equation (1.9). Let  $p_m := \frac{m}{m-1}\rho_m^{m-1}$  be the corresponding solution of (P) $_m$ . Taking  $|x|^2$  as our test function, we have for any  $T > 0$ ,

$$(4.3) \quad \underbrace{\int_{\mathbb{R}^2} \rho_m(x, T) |x|^2 dx}_{=: I_1} - \underbrace{\int_{\mathbb{R}^2} \rho_m(x, 0) |x|^2 dx}_{=: I_2} \\ = -2 \underbrace{\int_0^T \int_{\mathbb{R}^2} \rho_m \nabla \Phi_{1/m}(x, t) \cdot x dx dt}_{=: I_3} + 4 \underbrace{\int_0^T \int_{\mathbb{R}^2} \rho_m^m(x, t) dx dt}_{=: I_4}.$$

(Since  $\rho_m$  has compact support in  $[0, T]$ , our test function is not required to have compact support since we can always take a cut-off sufficiently far away.) As  $m \rightarrow +\infty$ , Lemma 3.10 yields that  $I_1$  converges to  $M_2[\rho_\infty(T)]$  and  $I_2$  converges to  $M_2[\rho_\infty(0)]$ .

To show the convergence of  $I_3$ , we decompose the integral into two parts:

$$I_3 = \int_0^T \int_{\mathbb{R}^2} \rho_m \nabla \Phi(x, t) \cdot x dx dt + \int_0^T \int_{\mathbb{R}^2} \rho_m \nabla (\Phi_{1/m}(x, t) - \Phi(x, t)) \cdot x dx dt =: I_{31} + I_{32}.$$

Since  $\nabla \Phi(x, t) \cdot x \in C(\mathbb{R}^d)$  for any  $t$ , Lemma 3.10 again gives that

$$I_{31} \xrightarrow{m \rightarrow +\infty} \int_0^T \int_{\mathbb{R}^2} \rho_\infty \nabla \Phi(x, t) \cdot x dx dt = \frac{1}{4\pi} \int_0^T \left( \int_{\mathbb{R}^2} \rho_\infty dx \right)^2 dt = \frac{1}{4\pi} |\Omega_0|^2 T,$$

where the last two equalities are obtained by symmetrizing  $x$  and  $y$  in the integrand and using conservation of mass, as in equation (4.1).

To control  $I_{32}$ , we first bound  $\|\nabla \Phi_{1/m} - \nabla \Phi\|_{L^2(\mathbb{R}^2)}$ . By Proposition 2.3,

$$(4.4) \quad \|\nabla \Phi_{1/m} - \nabla \Phi\|_{L^2(\mathbb{R}^2)} \leq W_2(\rho_\infty * \psi_{1/m}, \rho_\infty) \leq \frac{1}{m} \int \psi(x) |x|^2 dx,$$

where, in the last step, we apply [2, Lemma 7.1.10]. Hence

$$|I_{32}| \leq \int_0^T \|\rho_m(\cdot, t) |x|\|_{L^2} \|\nabla \Phi_{1/m} - \nabla \Phi\|_{L^2(\mathbb{R}^2)} dt \rightarrow 0 \text{ as } m \rightarrow +\infty,$$

where the fact that  $\sup_{t \in [0, T]} \sup_{m \geq 1} \|\rho_m(\cdot, t) |x|\|_{L^2} < +\infty$  is a consequence of Lemma 3.4, which ensures  $\rho_m$  is uniformly bounded and compactly supported. Combining the estimates on  $I_{31}$  and  $I_{32}$  yields that  $I_3 \rightarrow \frac{1}{4\pi} |\Omega_0|^2 T$  as  $m \rightarrow +\infty$ .

Finally, we consider  $I_4$ . We will show that

$$(4.5) \quad \limsup_{m \rightarrow \infty} \int_0^T \int_{\mathbb{R}^2} \rho_m^m(x, t) dx dt \leq \int_0^T \int_{\Omega(t)} u_1(x, t) dx dt.$$

The proof is then finished by taking  $\limsup_{m \rightarrow +\infty}$  on both sides of (4.3).

To show (4.5), first note that, since  $p_m := \frac{m}{m-1}\rho_m^{m-1}$ , we may write  $\rho_m^m = \frac{m-1}{m}\rho_m p_m$  and apply Remark 3.5 to obtain

$$(4.6) \quad \limsup_{m \rightarrow \infty} \int_0^T \int_{\mathbb{R}^2} \rho_m^m(x, t) dx dt \leq \limsup_{m \rightarrow \infty} \int_0^T \int_{\mathbb{R}^2} p_m(x, t) dx dt.$$

It remains to show that

$$(4.7) \quad \limsup_{m \rightarrow \infty} \int_0^T \int_{\mathbb{R}^2} p_m(x, t) dx dt \leq \int_0^T \int_{\mathbb{R}^2} u_1(x, t) dx dt.$$

For any  $n \in \mathbb{N}$ , define

$$u_{1,n}(x, t) := \sup_{\substack{m > n \\ |(x,t)-(y,s)| < 1/n}} p_m(y, s).$$

Note that  $u_{1,n}$  is decreasing in  $n$  and  $\lim_{n \rightarrow \infty} u_{1,n} = u_1$  by definition of  $u_1$ . For each  $n \in \mathbb{N}$ , we have  $p_m \leq u_{1,n}$  for all  $m > n$ , hence

$$\limsup_{m \rightarrow \infty} \int_0^T \int_{\mathbb{R}^2} p_m(x, t) dx dt \leq \int_0^T \int_{\mathbb{R}^2} u_{1,n}(x, t) dx dt \quad \text{for all } n \in \mathbb{N}.$$

We can then take  $n \rightarrow +\infty$  in the above inequality and apply the monotone convergence theorem. By Theorem 3.14 (a),  $\Omega(t) = \{u_1(\cdot, t) > 0\}$  almost everywhere. Thus, inequality (4.7) holds.  $\square$

**4.2. Some rearrangement inequalities.** In this subsection, we digress a bit to obtain an upper bound for the quantity

$$(4.8) \quad F(\Omega) = -\frac{1}{2\pi} |\Omega|^2 + 4 \int_{\Omega} p(x) dx,$$

where  $\Omega$  is a bounded set with smooth boundary and  $p : \bar{\Omega} \rightarrow \mathbb{R}$  satisfies  $-\Delta p = 1$  in  $\Omega$  and  $p = 0$  on  $\partial\Omega$ . This quantity appears in our heuristic computation for the evolution of the second moment of  $\rho_{\infty}(t)$ , where we show  $\frac{d}{dt} M_2[\rho_{\infty}] \leq F(\Omega(t))$ . Likewise,  $\int_0^T F(\Omega(t)) dt$  would have appeared on the right hand side of our rigorous result, given in equation (4.2), if the boundary of  $\Omega(t)$  were smooth for all time. While in this subsection we only aim to control  $F(\Omega)$  for smooth domains, in the next subsection we discuss how to use this bound to control the right hand side of (4.2), even when the boundary of  $\Omega(t)$  is not smooth.

The following result, due to Talenti [51], shows that  $F(\Omega) \leq 0$ , with equality if and only if  $\Omega$  is a disk. We sketch the proof below for the sake of completeness. In the subsequent proposition, we will modify the proof to get a stronger inequality.

**Proposition 4.2** (c.f. [51, Theorem 1]). *Let  $\Omega \subseteq \mathbb{R}^2$  be a bounded domain with smooth boundary, and let  $F(\Omega)$  be as in (4.8). Then we have*

$$(4.9) \quad F(\Omega) \leq 0,$$

and the equality is achieved if and only if  $\Omega$  is a disk.

*Proof.* First, note that maximum principle yields that  $p \geq 0$  in  $\bar{\Omega}$  and  $p > 0$  in  $\Omega$ . For any  $k \in [0, \sup_{\Omega} p)$ , let us define

$$\Omega_k := \{x \in \Omega : p(x) > k\} \quad \text{and } g(k) := |\Omega_k|.$$

Note that  $g(0) = |\Omega|$ . By definition of  $p$  and the divergence theorem, we have

$$(4.10) \quad g(k) = \int_{\Omega_k} -\Delta p(x) dx = \int_{\partial\Omega_k} -n \cdot \nabla p d\sigma = \int_{\partial\Omega_k} |\nabla p| d\sigma.$$

On the other hand, by the co-area formula (c.f. [22]),

$$(4.11) \quad g(k) = \int_k^{\infty} \int_{\partial\Omega_s} \frac{1}{|\nabla p|} d\sigma ds, \quad g'(k) = - \int_{\partial\Omega_k} \frac{1}{|\nabla p|} d\sigma.$$

Combining (4.10) and (4.11) and applying the Cauchy-Schwarz inequality,

$$(4.12) \quad g(k)g'(k) = \left( \int_{\partial\Omega_k} |\nabla p| d\sigma \right) \left( - \int_{\partial\Omega_k} \frac{1}{|\nabla p|} d\sigma \right) \leq -P(\Omega_k)^2,$$

where  $P(\Omega_k)$  is the perimeter of  $\Omega_k$ . For any bounded domain  $E \subseteq \mathbb{R}^2$ , the isoperimetric inequality yields

$$(4.13) \quad 2\sqrt{\pi}\sqrt{|E|} \leq P(E).$$

Applying inequality (4.13) to  $\Omega_k$  in (4.12) gives

$$g(k)g'(k) \leq - \left( 2\sqrt{\pi}\sqrt{g(k)} \right)^2 = -4\pi g(k),$$

hence  $g(k)$  satisfies the differential inequality

$$(4.14) \quad g'(k) \leq -4\pi \text{ for all } k \in \left( 0, \sup_{\Omega} g \right).$$

Combining this with  $g(0) = |\Omega|$  yields that  $g(k) \leq (|\Omega| - 4\pi k)_+$  for all  $k \geq 0$ . Therefore,

$$\int_{\Omega} p(x) dx = \int_0^{\sup_{\Omega} p} g(k) dk \leq \int_0^{\infty} (|\Omega| - 4\pi k)_+ dk = \frac{1}{8\pi} |\Omega|^2,$$

which gives (4.9). In order to achieve equality,  $\Omega_k$  must be a disk for almost every  $k > 0$ , hence  $\Omega$  must be a disk.  $\square$

We now prove a stronger version of the above inequality by replacing the isoperimetric inequality in the above argument (see (4.13)) by the following quantitative version due to Fusco, Maggi, and Pratelli [26].

**Lemma 4.3** (c.f. [26, Section 1.2]). *Let  $E \subseteq \mathbb{R}^2$  be a bounded domain. We define the Fraenkel asymmetry  $A(E) \in [0, 1]$  as*

$$A(E) := \inf \left\{ \frac{|E \Delta (x_0 + rB)|}{|E|} : x_0 \in \mathbb{R}^2, \pi r^2 = |E| \right\},$$

where  $B$  is the unit disk. Then there is some constant  $c \in (0, 1)$ , depending only on the dimension, such that

$$P(E) \geq 2\sqrt{\pi}\sqrt{|E|} (1 + cA(E)^2),$$

where  $P(E) = \mathcal{H}^1(\partial E)$  denotes the perimeter of  $E$ .

We begin with the following simple observation regarding the Fraenkel asymmetry.

**Lemma 4.4.** *Let  $E \subseteq \mathbb{R}^2$  be a bounded domain. For all  $U \subseteq E$  satisfying  $|U| \geq |E|(1 - \frac{A(E)}{4})$ , we have*

$$A(U) \geq \frac{A(E)}{4}.$$

*Proof.* Assume the statement is not true, so there exists some disk  $B_U$  with the same area as  $U$  so that

$$\frac{|U \Delta B_U|}{|U|} < \frac{A(E)}{4}.$$

Since  $|U| = |B_U|$ , we have  $|U \triangle B_U| = 2(|U| - |U \cap B_U|)$ . Hence the above inequality becomes

$$|U \cap B_U| > |U| \left(1 - \frac{A(E)}{8}\right).$$

Let  $B_E$  be a disk with the same area as  $E$  that contains  $B_U$ . Then since  $|U| \geq |E|(1 - \frac{A(E)}{4})$ ,

$$|E \cap B_E| \geq |U \cap B_U| > |E| \left(1 - \frac{A(E)}{4}\right) \left(1 - \frac{A(E)}{8}\right) \geq |E| \left(1 - \frac{3A(E)}{8}\right)$$

Therefore,

$$\frac{|E \triangle B_E|}{|E|} = \frac{2(|E| - |E \cap B_E|)}{|E|} < \frac{3}{4}A(E),$$

which contradicts the fact that  $A(E) \leq |E \triangle B_E|/|E|$ . This gives the result.  $\square$

With this lemma, we are now able to conclude a stronger upper bound on  $F(\Omega)$  than provided by Proposition 4.2.

**Proposition 4.5.** *Under the same assumptions as Proposition 4.2, there exists a constant  $c_0 \in (0, 1)$ , such that*

$$F(\Omega) \leq -c_0 A(\Omega)^3 |\Omega|^2.$$

*Proof.* We follow the proof of Proposition 4.2, with the following difference: instead of applying the isoperimetric inequality (4.13) to the set  $\Omega_k$  in inequality (4.12), we now apply the quantitative version from Lemma 4.3 to obtain

$$g'(k) \leq -4\pi (1 + cA(\Omega_k)^2)^2 \leq -4\pi (1 + cA(\Omega_k)^2).$$

To relate  $A(\Omega_k)$  with  $A(\Omega)$ , we apply Lemma 4.4 to obtain that  $A(\Omega_k) \geq \frac{A(\Omega)}{4}$  for any  $k$  such that  $g(k) \geq |\Omega|(1 - \frac{A(\Omega)}{4})$ . In other words, we have

$$(4.15) \quad g'(k) \leq -4\pi \left(1 + \frac{cA(\Omega)^2}{16}\right) \text{ for all } k \text{ such that } g(k) \geq |\Omega| \left(1 - \frac{A(\Omega)}{4}\right).$$

We claim that this implies

$$(4.16) \quad g(k) \leq |\Omega| - 4\pi \left(1 + \frac{cA(\Omega)^2}{16}\right) k \quad \text{for all } k \in \left(0, \frac{A(\Omega)|\Omega|}{32\pi}\right).$$

To see this, note that for all  $k \in (0, \frac{A(\Omega)|\Omega|}{32\pi})$ , the right hand side of (4.16) is greater than  $|\Omega|(1 - \frac{A(\Omega)}{4})$  since  $1 + cA(\Omega)^2/16 \leq 2$ . As a result, if (4.16) is violated at some  $k_0 \in (0, \frac{A(\Omega)|\Omega|}{32\pi})$ , then we must have  $g(k) \geq |\Omega|(1 - \frac{A(\Omega)}{4})$  for all  $k \in (0, k_0)$ , since  $g$  is a decreasing function. We can then integrate (4.15) in  $(0, k_0)$  to conclude that (4.16) actually holds at  $k_0$ , a contradiction.

Let  $h(k) = (|\Omega| - 4\pi k)_+$ . Inequality (4.16) implies that  $g(k) \leq h(k) - \frac{cA(\Omega)^3|\Omega|}{128}$  at  $k = \frac{A(\Omega)|\Omega|}{32\pi}$ . For  $k > \frac{A(\Omega)|\Omega|}{32\pi}$ , recall that by inequality (4.14), we have  $g'(k) \leq -4\pi$  for  $k \in (0, \sup_{\Omega} g)$ , and by definition of  $h$ , we have  $h'(k) = -4\pi$  in  $(0, |\Omega|/(4\pi))$ . This gives that  $g(k) \leq h(k) - \frac{cA(\Omega)^3|\Omega|}{128}$  for  $A(\Omega)|\Omega|/32\pi \leq k \leq |\Omega|/4\pi$ . Since  $A(\Omega) \leq 1$  this range of  $k$  is larger than  $|\Omega|/8\pi$ , and since  $g(k) \leq h(k)$  for all  $k$ , we have

$$\int_0^\infty g(k) dk \leq \int_0^\infty h(k) dk - \frac{cA(\Omega)^3|\Omega|^2}{2000} = \frac{|\Omega|^2}{8\pi} - \frac{cA(\Omega)^3|\Omega|^2}{2000}.$$

Finally, this gives

$$F(\Omega) = \int_0^\infty g(k) - \frac{|\Omega|^2}{8\pi} \leq \frac{cA(\Omega)^3|\Omega|^2}{2000},$$

hence the result holds with  $c_0 := \frac{c}{2000}$ .  $\square$

**4.3. Convergence of energy functional as  $t \rightarrow \infty$ .** In this section, we aim to show that, along the solution  $\rho_\infty(\cdot, t)$ , the energy functional  $E_\infty$  converges to its global minimizer as  $t \rightarrow +\infty$ . We begin by estimating the rate of change of the second moment along  $\rho_\infty$ . Combining Proposition 4.5 with our heuristic computation (4.1) suggests that

$$\frac{d}{dt} M_2[\rho_\infty(t)] \leq -c_0 A(\Omega(t))^3 |\Omega(t)|^2.$$

We now show that this inequality is indeed true in the time-integral sense, even if  $\Omega(t)$  does not have smooth boundary.

**Proposition 4.6.** *Let  $\Omega_0 \subseteq \mathbb{R}^2$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  be the gradient flow of  $E_\infty$  with initial data  $\rho_0 = \chi_{\Omega_0}$ . Then we have*

$$(4.17) \quad M_2[\rho_\infty(T)] - M_2[\rho_0] \leq -c_0 |\Omega_0|^2 \int_0^T A(\Omega(t))^3 dx dt,$$

where  $c_0 \in (0, 1)$  is the constant given in Proposition 4.5.

*Proof.* Since the evolution of the second moment is already given by Proposition 4.1, it remains to show

$$(4.18) \quad -\frac{1}{2\pi} |\Omega_0|^2 T + 4 \int_0^T \int_{\Omega(t)} u_1(x, t) dx dt \leq -c_0 |\Omega_0|^2 \int_0^T A(\Omega(t))^3 dt,$$

where  $u_1$  is the half-relaxed limit of  $p_m$  defined in Lemma 3.12.

Let  $\Omega^1(t) = \{u_1(\cdot, t) > 0\}$ . By Theorem 3.14 (a) and (b), we have  $\Omega^1(t) = \Omega(t)$  almost everywhere, so  $A(\Omega(t)) = A(\Omega^1(t))$ , and  $|\partial\Omega^1(t)| = 0$  for all  $t \in [0, T]$ . Hence for any  $\varepsilon > 0$  and  $t \in [0, T]$ , we can find a set  $D(t) \subseteq \mathbb{R}^2$  with smooth boundary such that  $\overline{\Omega^1(t)} \subseteq D(t)$ , and  $|D(t) \setminus \Omega^1(t)| \leq \varepsilon$ . For any  $t \geq 0$ , we then have a classical solution  $p(\cdot, t)$  such that  $-\Delta p(\cdot, t) = 1$  in  $D(t)$ , and  $p(\cdot, t) = 0$  on  $\partial D(t)$  and  $D(t)^c$ . In addition, we may choose  $D(t)$  so that  $\partial D(t)$  is continuous in time with respect of Hausdorff distance of sets, which ensures that  $p$  is continuous in time.

We first aim to show that

$$(4.19) \quad u_1(x, t) \leq p(x, t).$$

It suffices to show that  $u_1(x, t) \leq ap(x, t)$  for any  $a > 1$ . Towards a contradiction, assume that there exists some  $a > 0$ , such that  $\sup_{x \in \mathbb{R}^2, t \in [0, T]} (u_1 - ap) > 0$ . Since  $p$  is continuous in both space and time, and  $u_1$  is upper semicontinuous by definition as the half-relaxed limit,  $u_1 - ap$  achieves a strictly positive maximum at some  $(x_0, t_0)$ . Furthermore, since  $p \geq 0$ , we have  $u_1(x_0, t_0) > 0$ . Again using that  $(u_1, \Sigma_1)$  is a subsolution of  $(P)_\infty$ , we have that  $-a\Delta p(x_0, t_0) \leq 1$ , which implies that  $-\Delta p(x_0, t_0) < 1$ . However, since  $x_0 \in \Omega^1(t_0) \subseteq D(t_0)$ , we must have  $-\Delta p(x_0, t_0) = 1$ , which gives the contradiction.

We now show inequality (4.18). Since  $|D(t) \setminus \Omega^1(t)| \leq \varepsilon$ , there exists  $C$  depending on  $|\Omega_0|$  so that  $A(\Omega^1(t)) = A(\Omega(t)) \leq A(D(t)) + C\varepsilon$ . Combining this observation with (4.19)

and Proposition 4.5, we obtain the following bound for the left hand side of (4.18), where  $C$  depends on  $\Omega_0$  and  $T$ :

$$\begin{aligned} \int_0^T \left( -\frac{1}{2\pi} |\Omega_0|^2 + \int_{\Omega(t)} u_1(x, t) dx \right) dt &\leq \int_0^T \left( -\frac{1}{2\pi} |D(t)|^2 + \int_{D(t)} p(x, t) dx \right) dt + C\varepsilon \\ &\leq -c_0 \int_0^T A(D(t))^3 |D(t)|^2 dt + C\varepsilon \leq -c_0 |\Omega_0|^2 \int_0^T A(\Omega(t))^3 dt + C\varepsilon. \end{aligned}$$

Sending  $\varepsilon \rightarrow 0$  gives the result.  $\square$

**Corollary 4.7.** *Under the assumptions of Proposition 4.6, for any  $T > 0$ , there exists some  $t_0 \in (0, T)$ , such that*

$$(4.20) \quad A(\Omega(t_0)) \leq C(\Omega_0)T^{-1/3},$$

where  $C(\Omega_0) := (M_2[\chi_{\Omega_0}]/c_0|\Omega_0|^2)^{1/3}$ , for  $c_0$  as in Proposition 4.5.

*Proof.* Fix  $T > 0$ . Towards a contradiction, assume  $A(\Omega(t_0)) > C(\Omega_0)T^{-1/3}$  for all  $t_0 \in (0, T)$ . By Proposition 4.6 and the definition of  $C(\Omega_0)$ ,

$$M_2[\rho_\infty(T)] \leq M_2[\rho_0] - c_0 |\Omega_0|^2 \int_0^T A(\Omega(t))^3 dx dt \leq M_2[\rho_0] - c_0 |\Omega_0|^2 T (C(\Omega_0)T^{-1/3})^3 = 0,$$

which contradicts with the fact that  $M_2[\rho_\infty(t)]$  must be positive for all time.  $\square$

The above corollary does not directly yield that  $\lim_{t \rightarrow \infty} A(\Omega(t)) = 0$ . To show this and conclude that  $\Omega(t)$  converges to a disk, we will use the fact that the energy  $E_\infty$  is decreasing in time along  $\rho_\infty(\cdot, t)$ . In the next lemma, we show that if  $A(\Omega)$  is small, then the energy is close to its minimum.

**Lemma 4.8.** *Let  $\Omega \subseteq \mathbb{R}^2$  be a bounded domain, and let  $B_\Omega \subseteq \mathbb{R}^2$  be a disk with  $|B_\Omega| = |\Omega|$ . Then,*

$$0 \leq E_\infty(\chi_\Omega) - E_\infty(\chi_{B_\Omega}) \leq 40|\Omega|(1 + |\Omega| + M_2[\chi_\Omega])\sqrt{A(\Omega)}$$

*Proof.* The first inequality is a direct consequence of Riesz's rearrangement inequality. To prove the second one, let us first rewrite  $E_\infty(\chi_\Omega) - E_\infty(\chi_{B_\Omega})$  as

$$\begin{aligned} E_\infty(\chi_\Omega) - E_\infty(\chi_{B_\Omega}) &= \frac{1}{2\pi} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} (\chi_\Omega(x) - \chi_{B_\Omega}(x))(\chi_\Omega(y) + \chi_{B_\Omega}(y)) \log|x - y| dx dy \\ &=: \frac{1}{2\pi} I_1 + \frac{1}{2\pi} I_2, \end{aligned}$$

where  $I_1$  and  $I_2$  denote the integral in the domains  $|x - y| \leq 1$  and  $|x - y| > 1$ , respectively.

First, we consider  $I_1$ . Note that for any  $x \in \mathbb{R}^2$ , we have

$$\left| \int_{y \in B(x, 1)} (\chi_\Omega(y) + \chi_{B_\Omega}(y)) \log|x - y| dy \right| \leq 2 \left| \int_{|x - y| \leq 1} \log|x - y| dy \right| = \pi,$$

hence

$$I_1 \leq \|\chi_\Omega - \chi_{B_\Omega}\|_1 \left\| \int_{y \in B(x, 1)} (\chi_\Omega(y) + \chi_{B_\Omega}(y)) \log|x - y| dy \right\|_\infty \leq \pi \|\chi_\Omega - \chi_{B_\Omega}\|_1.$$

Now, we consider  $I_2$ . For  $|x - y| > 1$ ,  $\log|x - y| \leq |x - y| \leq |x| + |y| \leq (1 + |x|)(1 + |y|)$ , so

$$\begin{aligned} I_2 &\leq \left( \int_{\mathbb{R}^2} |\chi_\Omega(x) - \chi_{B_\Omega}(x)|(1 + |x|)dx \right) \left( \int_{\mathbb{R}^2} (\chi_\Omega(y) + \chi_{B_\Omega}(y))(1 + |y|)dy \right) \\ &\leq \|\chi_\Omega - \chi_{B_\Omega}\|_1^{1/2} \left( \int_{\mathbb{R}^2} |\chi_\Omega(x) - \chi_{B_\Omega}(x)|(1 + |x|)^2 dx \right)^{1/2} \\ &\quad \cdot (2|\Omega|)^{1/2} \left( \int_{\mathbb{R}^2} |\chi_\Omega(y) + \chi_{B_\Omega}(y)|(1 + |y|)^2 dy \right)^{1/2} \\ &\leq 2\sqrt{2}|\Omega|^{1/2} \|\chi_\Omega - \chi_{B_\Omega}\|_1^{1/2} (M_2[\chi_\Omega] + M_2[\chi_{B_\Omega}] + 2|\Omega|). \end{aligned}$$

Combining the above estimates on  $I_1$  and  $I_2$  with the facts that  $\|\chi_\Omega - \chi_{B_\Omega}\|_1 \leq 2|\Omega|$  and  $1/\pi \leq 1$ , we have

$$(4.21) \quad E_\infty(\chi_\Omega) - E_\infty(\chi_{B_\Omega}) \leq \frac{1}{2\pi} I_1 + \frac{1}{2\pi} I_2 \leq |\Omega|^{1/2} \|\chi_\Omega - \chi_{B_\Omega}\|_1^{1/2} (1 + M_2[\chi_\Omega] + M_2[\chi_{B_\Omega}] + 2|\Omega|).$$

The proof is then split into the following two cases:  $A(\Omega) \geq 1/2$  and  $A(\Omega) < 1/2$ .

*Case 1:*  $A(\Omega) \geq 1/2$ . In this case, we have  $\|\chi_\Omega - \chi_{B_\Omega}\|_1 \leq 2|\Omega| \leq 4A(\Omega)|\Omega|$  for any disk  $B_\Omega$  with the same measure as  $\Omega$ . Since  $E_\infty$  is invariant under translations, we can simply choose  $B_\Omega$  to be centered at 0. Such a choice directly yields  $M_2[\chi_{B_\Omega}] \leq M_2[\chi_\Omega]$ , hence (4.21) becomes

$$E_\infty(\chi_\Omega) - E_\infty(\chi_{B_\Omega}) \leq |\Omega|^{1/2} (4A(\Omega)|\Omega|)^{1/2} (1 + 2|\Omega| + 2M_2[\chi_\Omega]) \leq 4|\Omega|(1 + |\Omega| + M_2[\chi_\Omega])\sqrt{A(\Omega)},$$

which gives the result.

*Case 2:*  $A(\Omega) < 1/2$ . In this case, we choose  $B_\Omega$  to be the disk minimizing  $|\Omega \Delta B_\Omega|$ , which then gives

$$(4.22) \quad \|\chi_\Omega - \chi_{B_\Omega}\|_1 = A(\Omega)|\Omega|.$$

This choice of  $B_\Omega$  no longer directly gives us  $M_2[\chi_{B_\Omega}] \leq M_2[\chi_\Omega]$ , but we claim that we still have  $M_2[\chi_{B_\Omega}] \leq 36M_2[\chi_\Omega]$ . To see this, first note that  $A(\Omega) < 1/2$  implies  $|B_\Omega \setminus \Omega| < |B_\Omega|/2$ . Also, a simple computation yields that for any  $x, y \in B_\Omega$ , we have  $|x|^2 \leq (|y| + |x - y|)^2 \leq 2|y|^2 + 2|x - y|^2 \leq 2|y|^2 + 8|\Omega|/\pi$ . Therefore,

$$\begin{aligned} M_2[\chi_{B_\Omega}] &= \int_{B_\Omega} |x|^2 dx \leq |B_\Omega| \max_{x \in B_\Omega} |x|^2 \leq 2|B_\Omega \cap \Omega| \max_{x \in B_\Omega} |x|^2 \leq 2 \int_{B_\Omega \cap \Omega} \left( 2|y|^2 + \frac{8}{\pi}|\Omega| \right) dy \\ &\leq 4M_2[\chi_\Omega] + \frac{16}{\pi}|\Omega|^2 \leq 4M_2[\chi_\Omega] + 32 \left( \frac{|\Omega|^2}{2\pi} \right) \leq 4M_2[\chi_\Omega] + 32 \left( \int_0^{\sqrt{|\Omega|/\pi}} r^2 \cdot 2\pi r dr \right) \leq 36M_2[\chi_\Omega], \end{aligned}$$

Combining this and equation (4.22) with inequality (4.21) then yields

$$E_\infty(\chi_\Omega) - E_\infty(\chi_{B_\Omega}) \leq |\Omega|(1 + 37M_2[\chi_\Omega] + 2|\Omega|)A(\Omega)^{1/2},$$

which completes the proof.  $\square$

Combining the above results, we are now able to show that, along the solution  $\rho_\infty(t)$ , the energy functional  $E_\infty$  is converging towards its global minimizer with an explicit rate.

**Theorem 4.9.** *Let  $\Omega_0 \subseteq \mathbb{R}^2$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  be the gradient flow of  $E_\infty$  with initial data  $\rho_0 = \chi_{\Omega_0}$ . Suppose  $B_0$  is a disk with the same area as  $\Omega_0$ . Then, for any  $t > 0$ , we have*

$$0 \leq E_\infty(\chi_{\Omega(t)}) - E_\infty(\chi_{B_0}) \leq C_1(|\Omega_0|, M_2[\Omega_0])t^{-1/6},$$

where  $C_1(|\Omega_0|, M_2[\Omega_0]) = C_2|\Omega|^{2/3}(|\Omega_0| + M_2[\Omega_0])^{7/6}$  and  $C_2$  is a universal constant.

*Proof.* By Corollary 4.7, we have that, for any  $t > 0$ , there exists some  $t_0 \in (0, t)$ , so that

$$A(\Omega(t_0)) \leq (M_2[\chi_{\Omega_0}]/c_0|\Omega_0|^2)^{1/3}t^{-1/3}.$$

By definition of the discrete gradient flow and the lower semicontinuity of  $E_\infty$ ,  $E_\infty(\rho_\infty(t))$  is nonincreasing in time. Therefore, at time  $t$ , we may apply Lemma 4.8 to conclude

$$\begin{aligned} E_\infty(\chi_{\Omega(t)}) - E_\infty(\chi_{B_0}) &\leq E_\infty(\chi_{\Omega(t_0)}) - E_\infty(\chi_{B_0}) \leq 40|\Omega(t_0)|(1 + |\Omega_0| + M_2[\chi_{\Omega(t_0)}])\sqrt{A(\Omega(t_0))} \\ &\leq 40|\Omega_0|(1 + |\Omega_0| + M_2[\chi_{\Omega_0}]) (M_2[\chi_{\Omega_0}]/c_0|\Omega_0|^2)^{1/6}t^{-1/6} \\ &\leq C_2|\Omega_0|^{2/3}(1 + |\Omega_0| + M_2[\Omega_0])^{7/6}t^{-1/6}. \end{aligned}$$

□

*Remark 4.10.* While the rate in Theorem 4.9 is probably not optimal, the following example shows that the optimal power cannot go beneath  $-1$ . For  $0 < \varepsilon \ll 1$ , let  $\Omega_0^\varepsilon = B(x_\varepsilon, \varepsilon) \cup B(0, R_\varepsilon)$ , where  $x_\varepsilon := (\varepsilon^{-1}, 0) \in \mathbb{R}^2$ , and  $R_\varepsilon := \sqrt{1 - \varepsilon^2}$  is chosen such that  $|\Omega_0^\varepsilon| = \pi$ . This definition ensures that  $M_2[\Omega_0^\varepsilon]$  is uniformly bounded for all  $\varepsilon < 1$ . Since  $\partial_r(\mathcal{N} * \chi_{B(0,1)})(r) \sim r^{-1}$  for  $r \gg 1$ , the extra  $\pi\varepsilon^2$  amount of mass will stay outside  $B(0, (2\varepsilon)^{-1})$  for all  $t \in [0, c_1\varepsilon^{-2}]$ , where  $c_1 > 0$  is independent of  $\varepsilon$ . During this time interval, the free energy is at least  $c_2\varepsilon^2|\log \varepsilon|$  greater than its global minimizer for some  $c_2 > 0$ . Hence  $E_\infty(\chi_{\Omega^\varepsilon(T_\varepsilon)}) - E_\infty(\chi_{B(0,1)}) \gtrsim T_\varepsilon^{-1}|\log T_\varepsilon|$  for  $T_\varepsilon = c_1\varepsilon^{-2}$ , implying that the optimal power of  $t$  in Theorem 4.9 cannot be less than  $-1$ .

**4.4. Convergence of  $\rho_\infty(t)$  as  $t \rightarrow \infty$ .** We now conclude our study of asymptotic behavior by showing that, as  $t \rightarrow \infty$ ,  $\rho_\infty(t)$  converges to  $\chi_{B_0}$  in  $L^q$  for any  $1 \leq q < \infty$ , where  $B_0$  is the disk with the same area and the center of mass as  $\Omega_0$ . We begin with the following lemma, which ensures that the center of mass of  $\rho_\infty(t)$  is preserved for all time.

**Lemma 4.11.** *Let  $\Omega_0 \subseteq \mathbb{R}^2$  be a bounded domain with Lipschitz boundary, and let  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  be the gradient flow of  $E_\infty$  with initial data  $\rho_0 = \chi_{\Omega_0}$ . Then for any  $T > 0$ , we have  $\int_{\mathbb{R}^2} \rho_\infty(x, T)xdx = \int_{\mathbb{R}^2} \rho_\infty(x, 0)xdx$ .*

*Proof.* We proceed as in the proof of Proposition 4.1. For any  $m > 1$ , let  $\rho_m$  be the weak solution of (PME-D) $_m$  with initial data  $(\frac{m-1}{m}p_0)^{1/(m-1)}$ , where  $p_0$  is as in equation 1.9. For  $i = 1$  or  $2$ , we take our test function to be  $x_i$ , the  $i$ -th component of  $x$ . Then, for any  $T > 0$ ,

$$(4.23) \quad \int_{\mathbb{R}^2} \rho_m(x, T)x_i dx - \int_{\mathbb{R}^2} \rho_m(x, 0)x_i dx = - \int_0^T \int_{\mathbb{R}^2} \rho_m \partial_i \Phi_{1/m}(x, t) dx dt.$$

By Lemma 3.10, the left hand side of (4.23) converges to  $\int_{\mathbb{R}^2} \rho_\infty(x, T)x_i dx - \int_{\mathbb{R}^2} \rho_\infty(x, 0)x_i dx$  as  $m \rightarrow \infty$ . The right hand side can be controlled in the same way as the term  $I_3$  in the proof of Proposition 4.1, which gives

$$\lim_{m \rightarrow \infty} \int_{\mathbb{R}^2} \rho_m \partial_i \Phi_{1/m}(x, t) dx = \int_{\mathbb{R}^2} \rho_\infty \partial_i \Phi(x, t) dx = \frac{1}{2\pi} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \rho_\infty(x) \rho_\infty(y) \frac{x_i - y_i}{|x - y|^2} dx dy = 0.$$

Hence, sending  $m \rightarrow \infty$  in (4.23), we have  $\int_{\mathbb{R}^2} \rho_\infty(x, T)x_i dx = \int_{\mathbb{R}^2} \rho_\infty(x, 0)x_i dx$  for  $i = 1, 2$ , which finishes the proof.  $\square$

With this control on the center of mass of  $\rho_\infty(x, t)$  in hand, we now turn to the proof of the main result.

**Theorem 4.12.** *Let  $\Omega_0 \subseteq \mathbb{R}^2$  be a bounded domain with Lipschitz boundary, and let  $B_0 \subseteq \mathbb{R}^2$  be a disk such that  $|B_0| = |\Omega_0|$  and  $\int_{B_0} x dx = \int_{\Omega_0} x dx$ . Let  $\rho_\infty(\cdot, t) = \chi_{\Omega(t)}$  be the gradient flow of  $E_\infty$  with initial data  $\chi_{\Omega_0}$ . Then for any  $1 \leq q < +\infty$ , we have*

$$\lim_{t \rightarrow \infty} \|\rho_\infty(\cdot, t) - \chi_{B_0}\|_{L^q(\mathbb{R}^2)} = 0.$$

*Proof.* We first show that, for any  $f \in C_b(\mathbb{R}^2)$ , the space of bounded, continuous functions,

$$(4.24) \quad \lim_{t \rightarrow \infty} \int_{\mathbb{R}^2} \rho_\infty(x, t)f(x)dx = \int_{\mathbb{R}^2} \chi_{B_0}f(x)dx.$$

To show this, take any diverging time sequence  $(t_n)_{n=1}^\infty$ . By Proposition 4.6,  $M_2[\rho_\infty(t_n)]$  is uniformly bounded for all  $n$ . Hence by Prokhorov's Theorem [2, Theorem 5.1.3], there exists a subsequence  $(t_{n_k})_{k=1}^\infty$  and  $\mu \in L^1_+(\mathbb{R}^2)$  so that

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^2} \rho_\infty(x, t_{n_k})f(x)dx = \int_{\mathbb{R}^2} \mu(x)f(x)dx$$

for all  $f \in C_b(\mathbb{R}^2)$ . Choosing suitable test functions  $f$ , we have  $\int \mu dx = |\Omega_0|$  and  $\|\mu\|_\infty \leq \sup_{t \geq 0} \|\rho_\infty(\cdot, t)\|_\infty = 1$ . In addition, by letting the test function  $f$  approach  $f(x) = x$ , we have

$$(4.25) \quad \lim_{k \rightarrow \infty} \int_{\mathbb{R}^2} \rho_\infty(x, t_{n_k})x dx = \int_{\mathbb{R}^2} \mu(x)x dx.$$

Since the energy functional  $E_\infty$  is lower-semicontinuous with respect to weak-\* convergence of probability measures [20, Proposition 4.5], by Theorem 4.9,

$$E_\infty(\mu) \leq \liminf_{k \rightarrow \infty} E_\infty(\rho_\infty(t_{n_k})) = E_\infty(\chi_{B_0}).$$

As the only global minimizers of  $E_\infty$  are translations of  $\chi_{B_0}$ ,  $\mu$  must equal some translation of  $\chi_{B_0}$  almost everywhere. Finally, recall that Lemma 4.11 and the definition of  $B_0$  give that  $\int_{\mathbb{R}^2} \rho_\infty(x, t)x dx = \int \chi_{\Omega_0}x dx = \int \chi_{B_0}x dx$  for all time. Combining this with (4.25), we obtain  $\int_{\mathbb{R}^2} \mu(x)x dx = \int \chi_{B_0}x dx$ , leading to  $\mu = \chi_{B_0}$  a.e.. Thus, any diverging time sequence contains a subsequence satisfying (4.24), so we conclude that (4.24) must hold.

We now show that  $\rho_\infty(\cdot, t) \rightarrow \chi_{B_0}$  in  $L^1(\mathbb{R}^2)$ . Since  $0 \leq \rho_\infty \leq 1$ , we have  $\rho_\infty \leq \chi_{B_0}$  a.e. in  $B_0$  and  $\rho_\infty \geq \chi_{B_0}$  a.e. in  $B_0^c$ . Hence

$$\|\rho_\infty(\cdot, t) - \chi_{B_0}\|_1 = 2 \int_{\mathbb{R}^2} (\chi_{B_0} - \rho_\infty(x, t))\chi_{B_0} dx.$$

Thus, by choosing  $f \in C_b(\mathbb{R}^2)$  sufficiently close to  $\chi_{B_0}$  and applying (4.24), we can show that, for any  $\varepsilon > 0$ ,  $\|\rho_\infty(\cdot, t) - \chi_{B_0}\|_1 \leq \varepsilon$  for sufficiently large  $t$ . This shows that  $\rho_\infty(\cdot, t) \rightarrow \chi_{B_0}$  in  $L^1(\mathbb{R}^2)$ . Finally, for  $1 < q < \infty$ , the convergence in  $L^q$  follows directly from the  $L^1$  convergence and the fact that  $\|\rho_\infty(\cdot, t) - \chi_{B_0}\|_\infty \leq 1$ .  $\square$

## 5. APPENDIX

**5.1. Definition of viscosity solutions of  $(P)_\infty$ .** We begin by recalling some notation. For  $Q \subseteq \mathbb{R}^d \times (0, \infty)$ , we write  $f \in C^{2,1}(Q)$  if  $f$  is twice continuously differentiable in  $x$  and once in  $t$ . We say that  $u - \varphi$  has a local maximum (minimum) zero  $(x_0, t_0)$  in  $Q$  if there exists  $\varepsilon > 0$  such that

$$\varphi(x_0, t_0) = u(x_0, t_0) \text{ and } \varphi \geq u \text{ (} \varphi \leq u \text{) in } Q \cap (B_\varepsilon(x_0) \times (t_0 - \varepsilon, t_0 + \varepsilon)).$$

In other words,  $\varphi$  touches  $u$  from above (below) at  $(x_0, t_0)$  with respect to  $Q$ .

Likewise, given an open set  $\Omega \subseteq \mathbb{R}^d$  and a function  $h : \Omega \times [0, +\infty) \rightarrow \mathbb{R}$ , we denote its upper and lower semicontinuous envelopes by

$$(5.1) \quad h^*(x, t) := \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x-y| \leq \varepsilon, \\ |t-s| \leq \varepsilon}} h(y, s), \quad h_*(x, t) := \lim_{\varepsilon \rightarrow 0} \inf_{\substack{|x-y| \leq \varepsilon, \\ |t-s| \leq \varepsilon}} h(y, s).$$

Note that  $h^*$  is the smallest upper semicontinuous function satisfying  $h \leq h^*$ , and  $h_*$  is the largest lower semicontinuous function satisfying  $h \geq h_*$ .

Now, we define the notion of viscosity subsolution, supersolution, and solution of  $(P)_\infty$ . Instead of proceeding as above and defining solutions of  $(P)_\infty$  by comparison with classical sub- and supersolutions, we follow an approach reminiscent of Kim [32] and Alexander, Kim, and Yao [1]. While the former would ease our proof of the comparison theorem, Theorem 3.7, the latter is more natural from the perspective of the convergence theorem, Theorem 3.14. One notable difference in the definition below from those of Alexander, Kim, and Yao [1] is the separation of the solution and the set evolution in our notion of subsolutions.

**Definition 5.1** (subsolution of  $(P)_\infty$ ). *An upper semicontinuous function  $u : \mathbb{R}^d \times (0, +\infty) \rightarrow [0, +\infty)$ , paired with a space-time set  $\Sigma = \cup_{t>0} (\Omega(t) \times \{t\})$ , is a viscosity subsolution of  $(P)_\infty$  if*

- (a)  $\{u(\cdot, t) > 0\} \subseteq \overline{\Omega(t)}$  and  $\Sigma \cap \{t \leq t_0\} \subseteq \overline{\Sigma \cap \{t < t_0\}}$  for every  $t_0 > 0$ ;
- (b) for all  $\varphi \in C^{2,1}(\mathbb{R}^d \times (0, +\infty))$  so that  $u - \varphi$  has a local maximum zero at  $(x_0, t_0)$  in  $\overline{\Sigma} \cap \{t \leq t_0\}$ ,
  - (i) if  $x_0 \in \Omega(t_0)^\circ$  or  $u(x_0, t_0) > 0$ , then  $-\Delta\varphi(x_0, t_0) \leq 1$ ;
  - (ii) if  $x_0 \in \partial\Omega(t_0)$ ,  $u(x_0, t_0) = 0$ , and  $|\nabla\varphi|(x_0, t_0) \neq 0$ , then

$$\min(-\Delta\varphi - 1, \varphi_t - |\nabla\varphi|^2 - \nabla\varphi \cdot \nabla\Phi)(x_0, t_0) \leq 0.$$

We will say that  $u : \mathbb{R}^d \times [0, +\infty) \rightarrow [0, +\infty)$  has compactly supported initial data  $u_0$  if, in addition,

- (c)  $u(\cdot, 0) = u_0(\cdot)$  and  $\overline{\{u_0 > 0\}} = \overline{\Sigma} \cap \{t = 0\}$ .

We introduce the set  $\Sigma$  for technical reasons, to allow the possibility that  $u$  becomes zero in the evolving set  $\Omega(t)$ . Condition (a) ensures that a subsolution does not jump up from zero. Condition (b)(ii) ensures that limits of viscosity solutions are viscosity solutions, since it is possible that the boundary collapses in a limit and boundary points of the limiting functions become interior points of the limit.

**Definition 5.2** (supersolution of  $(P)_\infty$ ). *A lower semicontinuous function  $v : \mathbb{R}^d \times (0, +\infty) \rightarrow (0, +\infty)$  is a viscosity supersolution of  $(P)_\infty$  with initial data  $v_0$  if for all  $\varphi \in C^{2,1}(\mathbb{R}^d \times (0, +\infty))$  so that  $v - \varphi$  has a local minimum zero at  $(x_0, t_0)$  with respect to  $\mathbb{R}^d \cap \{t \leq t_0\}$ ,*

- (i) if  $(x_0, t_0) \in \{v > 0\}$ ,  $-\Delta\varphi(x_0, t_0) \geq 1$ ;

(ii) if  $(x_0, t_0) \in \partial\{v > 0\}$ ,  $v(x_0, t_0) = 0$ ,  
 (5.2)  $|\nabla\varphi|(x_0, t_0) \neq 0$ , and  $\{\varphi > 0\} \cap \{v > 0\} \cap B \neq \emptyset$  for some ball  $B$  centered at  $(x_0, t_0)$

then  $\max(-\Delta\varphi - 1, \varphi_t - |\nabla\varphi|^2 - \nabla\varphi \cdot \nabla\Phi)(x_0, t_0) \geq 0$ .

Will we say that  $v : \mathbb{R}^d \times [0, +\infty) \rightarrow (0, +\infty)$  has initial data  $v_0$  if  $v(\cdot, 0) = v_0(\cdot)$

Condition (5.2) ensures that  $\varphi$  touches  $v$  from below in a non-degenerate way.

**Definition 5.3.** A lower semi-continuous function  $u$  is a viscosity solution of  $(P)_\infty$  in  $\mathbb{R}^d \times (0, \infty)$  with compactly supported initial data  $u_0$  if  $(u^*, \{u > 0\})$  and  $u$  are respectively viscosity sub- and supersolutions of  $(P)_\infty$  with initial data  $u_0$ .

The following lemma illustrates the fact that the solution of  $(P)_\infty$  is entirely characterized by its support.

**Lemma 5.4.** Suppose  $u$  is a viscosity solution of  $(P)_\infty$  in  $\mathbb{R}^d \times (0, \infty)$  and  $\overline{\{u^* > 0\}} = \overline{\{u > 0\}}$ . Then, for each  $t > 0$ ,  $u(\cdot, t) = (h_t)_*$ , where

$$h_t(x) = \inf\{\alpha(x) : -\Delta\alpha \geq 1 \text{ in an open set } E \text{ containing } \overline{\{u(\cdot, t) > 0\}}; \alpha \geq 0 \text{ on } \overline{E}\}$$

*Proof.* By the definition of a viscosity supersolution,  $-\Delta u(\cdot, t) \geq 1$  in  $\{u(\cdot, t) > 0\}$ , so  $(h_t)_* \leq u(\cdot, t)$ . On the other hand, by the definition of a viscosity subsolution,  $-\Delta u^*(\cdot, t) \leq 1$  in  $\mathbb{R}^d$  and  $u^*(\cdot, t)$  is supported in  $\overline{\{u(\cdot, t) > 0\}}$ . Therefore  $u^*(\cdot, t) \leq \alpha$  for any candidate function  $\alpha(x)$  in the definition of  $h_t$ , so  $u^*(\cdot, t) \leq (h_t)_*$ . Consequently, we conclude that  $u(\cdot, t) = (h_t)_*$ .  $\square$

**5.2. Further properties of gradient flows of  $E_\infty$ ,  $\tilde{E}_\infty$ , and  $E_m$ .** In this section, we collect several results on the gradient flows of  $E_\infty$ ,  $\tilde{E}_\infty$ , and  $E_m$ . We begin by proving Proposition 2.3, which provides elementary estimates on the Newtonian potential of a bounded, integrable function. We use these estimates to conclude that  $E_\infty$  is  $\omega$ -convex along generalized geodesics. (See [20, Theorem 4.3, Proposition 4.4].)

*Proof of Proposition 2.3.* The fourth inequality is a classical potential theory result (c.f. [17, Proposition 2.1], [28, Lemma 2.1]), and the fifth inequality is due to Loeper [39, Theorem 2.7]. (While Loeper only considers the case  $d \geq 3$ , the same argument applies in  $d = 2$ .)

For the bounds on  $\nabla\mathbf{N}\rho$  and  $\Delta\mathbf{N}\rho$ , note that if  $B = B_1(0)$ ,

$$\|\nabla\mathbf{N}\rho\|_\infty \leq \|\nabla\mathcal{N}\|_{L^1(B)}\|\rho\|_\infty + \|\nabla\mathcal{N}\|_{L^\infty(B^c)}\|\rho\|_1 \leq C_d \quad \text{and} \quad \|\Delta\mathbf{N}\rho\|_\infty = \|\rho\|_{L^\infty} \leq 1.$$

Likewise, for the lower bound on  $\int \mathbf{N}\rho d\rho$ , if we let  $\mathcal{N}^-(x)$  denote the negative part of  $\mathcal{N}(x)$ ,

$$\int \mathbf{N}\rho d\mu \geq \int \mathcal{N}^- * \rho(x) d\mu(x) \geq -\|\mathcal{N}^- * \rho\|_\infty \geq -\|\mathcal{N}^-\|_{L^1(B)}\|\rho\|_\infty - \|\mathcal{N}^-\|_{L^\infty(B^c)}\|\rho\|_1 \geq -C_d.$$

$\square$

Next, we prove Proposition 2.5, which ensures that  $\rho_\infty$  is Lipschitz in time, with respect to the Wasserstein metric.

*Proof of Proposition 2.5.* By [20, Theorem 3.11], the function  $S(t) : D(E_\infty) \rightarrow D(E_\infty) : \rho_\infty(\cdot, 0) \mapsto \rho_\infty(\cdot, t)$  is a semigroup, i.e.  $S(t+s) = S(t)S(s)\mu$  for  $t, s \geq 0$ . Therefore, it suffices to show that  $W_2(\rho_\infty(t), \rho_\infty(0)) \leq 2C_d t$  for all  $t \geq 0$ .

Let  $\rho_\tau^n$  be the discrete gradient flow of  $E_\infty$  with initial data  $\rho = \rho_\infty(0)$  and time step  $\tau > 0$ , as defined by equation (i). By [20, Theorem 3.8], if we take  $\tau = t/n$  for any  $t \geq 0$ , then  $\lim_{n \rightarrow +\infty} W_2(\rho_{t/n}^n, \rho_\infty(t)) = 0$ . Therefore,

$$W_2(\rho_\infty(t), \rho_\infty(0)) = \lim_{n \rightarrow +\infty} W_2(\rho_{t/n}^n, \rho) \leq \lim_{n \rightarrow +\infty} \sum_{i=1}^n W_2(\rho_{t/n}^i, \rho_{t/n}^{i-1}) \leq 2C_d t,$$

where the last inequality follows from Lemma 2.14, which ensures  $W_2(\rho_{t/n}^i, \rho_{t/n}^{i-1}) \leq 2C_d(t/n)$ .  $\square$

We now turn to the proof of Proposition 2.6, which concerns the regularity of  $\nabla \mathbf{N}\rho_\infty(x, t)$  in space and time.

*Proof of Proposition 2.6.* The fact that  $\nabla \mathbf{N}\rho_\infty(x, t)$  is log-Lipschitz in space is an immediate consequence of Proposition 2.3. We now consider the continuity with respect to time. By Proposition 2.5,  $\rho_\infty$  is Lipschitz in time with respect to the Wasserstein metric, so it suffices to translate this into continuity in time with respect to the Euclidean norm.

Fix  $\psi \in C_c^\infty(\mathbb{R}^d)$  so that  $\text{supp } \psi \subseteq B_1(0)$  and  $\|\psi\|_\infty \leq 1$ , and let  $\Phi(x, t) = \mathbf{N}\rho_\infty(x, t)$  and  $\Phi_{1/m} := \Phi * \psi_{1/m}$ . Combining the fifth inequality in Proposition 2.3 with Proposition 2.5,

$$\begin{aligned} |\nabla \Phi_{1/m}(x, t) - \nabla \Phi_{1/m}(x, s)| &= |\psi_{1/m} * (\nabla \mathbf{N}\rho_\infty(x, t) - \nabla \mathbf{N}\rho_\infty(x, s))| \\ &\leq \|\psi_{1/m}\|_{L^2(\mathbb{R}^d)} \|\nabla \mathbf{N}\rho_\infty(t) - \nabla \mathbf{N}\rho_\infty(s)\|_{L^2(\mathbb{R}^d)} \leq m^{d/2} W_2(\rho_\infty(t), \rho_\infty(s)) \leq 2C_d m^{d/2} |t - s|. \end{aligned}$$

We now use this inequality controlling the continuity in time of  $\nabla \Phi_{1/m}(x, t)$  to estimate the continuity in time of  $\nabla \Phi(x, t)$ . By Proposition 2.3,

$$\begin{aligned} |\nabla \Phi(x, t) - \nabla \Phi_{1/m}(x, t)| &= \left| \int (\nabla \Phi(x, t) - \nabla \Phi(x - y, t)) \psi_{1/m}(y) dy \right| \leq C_d \int \sigma(|y|) \psi_{1/m}(y) dy \\ &\leq C_d \sigma(1/m) \int \psi_{1/m}(y) dy = C_d \sigma(1/m). \end{aligned}$$

Therefore,

$$\begin{aligned} |\nabla \Phi(x, t) - \nabla \Phi(x, s)| &\leq |\nabla \Phi(x, t) - \nabla \Phi_{1/m}(x, t)| + |\nabla \Phi_{1/m}(x, t) - \nabla \Phi_{1/m}(y, t)| + |\nabla \Phi_{1/m}(y, t) - \nabla \Phi(y, t)| \\ &\leq 2C_d \sigma(1/m) + 2C_d m^{d/2} |t - s|. \end{aligned}$$

Let  $p = 1/2d$ . Since  $|t - s| < e^{(-1-\sqrt{2})/2}$ , if we choose  $m = |t - s|^{(-2/d)(1-p)} \geq 1$ , we have  $m^{d/2} |t - s| = |t - s|^p$ , which takes care of the second term in the above inequality. Furthermore,  $q = 1/(2(2 - 1/d)) < 1/2$  ensures  $|\log(x)| \leq x^{-1/2} \leq x^{q-1}$  for  $0 \leq x \leq 1$ . Therefore,

$$\sigma(1/m) \leq \begin{cases} (1/m)^q & \text{if } 1/m < e^{(-1-\sqrt{2})/2} \\ 3/m & \text{if } 1/m \geq e^{(-1-\sqrt{2})/2} \end{cases} \leq 3(1/m)^q = 3|t - s|^{(2q/d)(1-p)} = 3|t - s|^p.$$

Therefore,  $|\nabla \Phi(x, t) - \nabla \Phi(x, s)| \leq 10C_d |t - s|^{1/2d}$ , which gives the result.  $\square$

In the next proposition, we show that, while the discrete time sequence corresponding to  $\tilde{E}_\infty$  may not be unique, the distance between any two such sequences converges to zero as the time step  $\tau \rightarrow 0$ .

**Proposition 5.5.** *Fix  $T > 0$  and initial data  $\rho \in D(E_\infty)$  and let  $\tilde{\rho}_\tau^n$  and  $\tilde{\mu}_\tau^n$  be two choices for the time discrete time sequence corresponding to  $\tilde{E}_\infty$ , as defined in Definition 2.12 (ii). Then there exist positive constants  $N$  and  $C$ , depending on the dimension,  $T$ , and  $E_\infty(\rho)$ , so that for  $\tau = t/n$  and all  $0 \leq t \leq T$  and  $n > N$ ,*

$$f_\tau^{(2n)}(W_2^2(\tilde{\rho}_\tau^n, \tilde{\mu}_\tau^n)) \leq C\omega(\tau).$$

*Proof.* By Corollary 2.15, we have the following crude bound for all  $i = 1, \dots, n$ ,

$$W_2(\tilde{\rho}_\tau^i, \tilde{\mu}_\tau^i) \leq W_2(\tilde{\rho}_\tau^i, \rho) + W_2(\tilde{\mu}_\tau^i, \rho) \leq 4C_d T.$$

To obtain a more refined bound, we use Proposition 2.8. First, we estimate the behavior of the energy  $\tilde{E}_\infty$  along the discrete time sequence. By Proposition 2.3, Lemma 2.14, and the definition of  $\tilde{\rho}_\tau^i$  as a minimizer

$$\begin{aligned} & \tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^i) \\ &= \tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^{i-1}) + \tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^i) - \tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^{i-1}) = \tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^{i-1}) + \int \mathbf{N} \tilde{\rho}_\tau^{i-1} d(\rho_\tau^i - \rho_\tau^{i-1}) \\ &\leq \tilde{E}_\infty(\tilde{\rho}_\tau^{i-2}; \rho_\tau^{i-1}) + C_d W_2(\rho_\tau^i, \rho_\tau^{i-1}) \leq \tilde{E}_\infty(\tilde{\rho}_\tau^{i-2}; \rho_\tau^{i-1}) + 2C_d^2 \tau \leq \dots \leq \tilde{E}_\infty(\rho; \rho_\tau^1) + 2C_d^2 T. \end{aligned}$$

Likewise, we may control the first term on the right hand side by

$$\tilde{E}_\infty(\rho; \rho_\tau^1) = 2E_\infty(\rho) + \tilde{E}_\infty(\rho; \rho_\tau^1) - \tilde{E}_\infty(\rho; \rho) = 2E_\infty(\rho) + \int \mathbf{N} \rho d(\rho_\tau^1 - \rho) \leq 2E_\infty(\rho) + 2C_d^2 \tau.$$

Thus, there exists  $C > 0$  (which we allow to change from line to line) depending only on the dimension,  $T$ , and  $E_\infty(\rho)$  so that

$$\tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^i) \leq C.$$

Likewise, by Proposition 2.3,  $\tilde{E}_\infty(\cdot; \cdot)$  is uniformly bounded below by  $-C_d$ .

Due to these estimates, we may apply Proposition 2.8 to conclude that there exist positive constants  $C$  and  $N$  depending on the dimension,  $T$ , and  $E_\infty(\rho)$  so that for  $\tau = t/n$ ,  $0 \leq t \leq T$ , and  $n > N$ ,

$$\begin{aligned} & f_\tau^{(2)}(W_2^2(\tilde{\rho}_\tau^i, \tilde{\mu}_\tau^i)) \\ &\leq W_2^2(\tilde{\rho}_\tau^{i-1}, \tilde{\mu}_\tau^{i-1}) + C_d \tau \omega(CW_2(\tilde{\mu}_\tau^i, \tilde{\mu}_\tau^{i-1})) + 2\tau(\tilde{E}_\infty(\tilde{\rho}_\tau^{i-1}; \rho_\tau^i) - \tilde{E}_\infty(\tilde{\rho}_\tau^i; \rho_\tau^i)) + C\tau^2. \end{aligned}$$

By Lemma 2.14 (ii), we may bound the second term by  $C_d \tau \omega(C\tau)$  and the third term by  $4C_d^2 \tau^2$ . Therefore, for all  $i = 1, \dots, n$ ,

$$(5.3) \quad f_\tau^{(2)}(W_2^2(\tilde{\rho}_\tau^i, \tilde{\mu}_\tau^i)) \leq W_2^2(\tilde{\rho}_\tau^{i-1}, \tilde{\mu}_\tau^{i-1}) + C\tau\omega(\tau).$$

We now show that, for all  $j = 1, \dots, n$ ,

$$(5.4) \quad f_\tau^{(2j)}(W_2^2(\tilde{\rho}_\tau^n, \tilde{\mu}_\tau^n)) \leq W_2^2(\tilde{\rho}_\tau^{n-j}, \tilde{\mu}_\tau^{n-j}) + 2C\tau\omega(\tau)j.$$

Once we have this, taking  $j = n$  gives the result. We prove (5.4) by induction. The base case, when  $j = 1$ , is a consequence of (5.3). Suppose that the result holds for  $j - 1$ ,

$$f_\tau^{(2(j-1))}(W_2^2(\tilde{\rho}_\tau^n, \tilde{\mu}_\tau^n)) \leq W_2^2(\tilde{\rho}_\tau^{n-j+1}, \tilde{\mu}_\tau^{n-j+1}) + 2C\tau\omega(\tau)(j-1).$$

By Proposition 2.7, applying  $f_\tau^{(2)}$  to both sides,

$$\begin{aligned} f_\tau^{(2j)}(W_2^2(\tilde{\rho}_\tau^n, \tilde{\mu}_\tau^n)) &\leq f_\tau^{(2)}(W_2^2(\tilde{\rho}_\tau^{n-j+1}, \tilde{\mu}_\tau^{n-j+1})) + 2C\tau\omega(\tau)(j-1) + C\tau^2 \\ &\leq W_2^2(\tilde{\rho}^{n-j}, \tilde{\mu}^{n-j}) + 2C\tau\omega(\tau)j \end{aligned}$$

where the second inequality is a consequence of (5.3) and the fact that  $C\tau^2 \leq C\tau\omega(\tau)$ . This gives the result.  $\square$

Now, we turn to the proof that the discrete time sequence  $\rho_{\tau,m}^n$  corresponding to  $E_m$  converges to the solution of (PME-D) $_m$  as the time step goes to zero.

**Proposition 5.6.** *Given initial data  $\rho \in D(E_\infty)$ , let  $\rho_{\tau,m}^n$  be the discrete time sequence given in Definition 2.12 (iii). Then, for any  $t \geq 0$ ,  $\rho_{t/n,m}^n$  converges as  $n \rightarrow +\infty$  to a limit  $\rho_m(t)$ , and there exist positive constants  $C$  and  $N$  depending on the dimension,  $E_\infty(\rho)$ , and  $T$  so that for all  $n \geq N$ ,  $m \geq d+1$ , and  $0 \leq t \leq T$ ,*

$$W_2(\rho_{t/n,m}^n, \rho_m(t)) \leq Cn^{-1/16}e^{4C_d T}.$$

Furthermore,  $\rho_m(t)$  is the unique weak solution of (PME-D) $_m$ .

*Proof.* Given initial data  $\rho \in D(E_\infty)$ , let  $\rho_\tau^n$  be the discrete gradient flow of  $E_\infty$ , as in Definition 2.12 (i). Using this sequence, we define a time dependent energy  $E_{\tau,m}^n$  by

$$E_{\tau,m}^n(\nu) := E_m(\nu; \rho_\tau^n) = \begin{cases} \frac{1}{m-1} \int_{\mathbb{R}^d} \nu(x)^m dx + \int_{\mathbb{R}^d} \psi_{1/m} * \mathbf{N}\rho_\tau^n(x) d\nu(x) & \text{if } \nu \ll \mathcal{L}^d, \\ +\infty & \text{otherwise.} \end{cases}$$

Then  $\rho_{\tau,m}^n$  given in Definition 2.12 (iii) is the time varying discrete gradient flow of this energy in the sense that

$$(5.5) \quad \rho_{\tau,m}^n \in \operatorname{argmin}_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ \frac{1}{2\tau} W_2^2(\rho_{\tau,m}^{n-1}, \nu) + E_{\tau,m}^n(\nu) \right\} \text{ and } \rho_{\tau,m}^0 := \rho.$$

Consequently, we may apply the first author's results on convergence of the discrete gradient flow of time dependent energies [20, Theorem A.3], provided that we can show  $E_m^n$  satisfies [20, Assumption A.2].

First, by [20, Theorem 4.3, Proposition 4.4],  $E_\tau^n$  satisfies [20, Assumption 2.18] uniformly for  $n \in \mathbb{N}$ ,  $m > 1$ , and  $\tau > 0$ . In particular, there exists a solution to the minimization problem (5.5) and  $E_\tau^n$  is  $\omega$ -convex along generalized geodesics, for  $\lambda_\omega = -C_d$  as in Proposition 2.3 and  $\omega(x)$  as in equation (2.4).

Next, we estimate the behavior of the energies and Wasserstein distance along the discrete gradient flow. By Lemma 2.14 (iii), for all  $1 \leq i \leq n$ ,

$$W_2(\rho_{\tau,m}^i, \rho_{\tau,m}^{i-1}) \leq \sqrt{\frac{2\tau}{m-1} (\|\rho_{\tau,m}^{i-1}\|_m^m - \|\rho_{\tau,m}^i\|_m^m) + 2C_d\tau} \leq \sqrt{2\tau (1 + n\tau C_d^2/2)} + 2C_d\tau.$$

where, in the second inequality, we use that  $\|\rho^0\|_m^m \leq 1$ . Likewise, by Corollary 2.15,

$$W_2(\rho_{m,\tau}^n, \rho) \leq \sqrt{4n\tau(1 + 8C_d^2 n\tau)}.$$

Finally, since Proposition 2.3 ensures  $E_{\tau,m}^n$  is uniformly bounded below by  $-C_d$ , there exists a constant  $\tilde{C}_d > 0$ , depending only on the dimension, so that

$$E_{\tau,m}^0(\rho) - E_{\tau,m}^n(\rho_{\tau,m}^n) \leq E_m(\rho, \rho) + C_d \leq 1 + \int \mathbf{N}\rho(x)\psi_{1/m} * \rho(x) dx + C_d \leq \tilde{C}_d + E_\infty(\rho),$$

where in the last inequality we use that  $\mathbf{N}\rho(x)$  is a continuous function with at most quadratic growth and  $\psi_{1/m} * \rho \xrightarrow[W_2]{m \rightarrow +\infty} \rho$ , so  $\int \mathbf{N}\rho(x)\psi_{1/m} * \rho(x)dx \xrightarrow{m \rightarrow +\infty} \int \mathbf{N}\rho(x)\rho(x)dx$ .

It remains to show that  $E_{\tau,m}^n$  possesses sufficient continuity in  $n\tau$ . To do this, we first estimate the continuity of  $\rho_\tau^n$  in  $n\tau$ . By Lemma 2.14, we have the following crude bound

$$W_2^2(\rho_\tau^n, \rho_h^k) \leq (2C_d(n\tau + kh))^2 \leq 16C_d^2 T^2.$$

Combining this with Proposition 2.7 (iii) and [20, Theorem 3.6], we obtain that for any  $T > 0$ , there exists  $\bar{\tau} = \bar{\tau}(T, d)$  and  $\bar{C} = \bar{C}(T, d)$  so that for all  $0 \leq h < \tau < \bar{\tau}$  and  $k, n \in \mathbb{N}$  with  $kh, n\tau \leq T$ ,

$$\begin{aligned} F_{2kh}(W_2^2(\rho_\tau^n, \rho_h^k)) &\leq \bar{C} \left[ \sqrt{(n\tau - kh)^2 + \tau^2 n} + hk\tilde{\omega}(\sqrt{\tau}) + h^2 k + \tilde{\omega}(h^2)k \right] \\ &\quad + 2h(E_\infty(\rho) - \inf E_\infty) + C_d \omega(16C_d^2 T^2) T/n. \end{aligned}$$

Since  $F_t(x)$  is decreasing in  $t$ , this implies there exists  $\tilde{C} = \tilde{C}(T, d, E_\infty(\rho))$  so that for  $0 < \tau < \bar{\tau}$ ,

$$(5.6) \quad F_{2T}(W_2^2(\rho_\tau^n, \rho_h^k)) \leq \tilde{C} \left[ \sqrt{(n\tau - kh)^2 + \tau^2 n} + \sqrt{\tau} |\log \tau| \right].$$

Since  $F_{2T}(x)$  is strictly increasing and convex in  $x$ ,  $F_{2T}^{-1}(x)$  is strictly increasing and concave. Therefore,

$$\sigma(x) := \sqrt{F_{2T}^{-1}(x)}$$

is a continuous, nondecreasing, concave function that vanishes only at zero. In particular,  $\sigma(x)$  is also subadditive, so (5.6) implies that, for some  $C' = C'(T, d, E_\infty(\rho))$ ,

$$(5.7) \quad W_2(\rho_\tau^n, \rho_h^k) \leq C' \left[ \sigma((n\tau - kh)^2) + \sigma(\tau |\log \tau|^2) \right].$$

We use this estimate to show that  $E_{\tau,m}^n$  is continuous in  $n\tau$ , up to an error that decreases with  $\tau$ . Since  $f := \mathbf{N}(\psi_{1/m} * \rho_{\tau,m}^i) \in C^1$ , by Lemma 2.2,

$$\begin{aligned} |E_{\tau,m}^n(\rho_{\tau,m}^i) - E_{h,m}^k(\rho_{\tau,m}^i)| &= |E_m(\rho_{\tau,m}^i; \rho_\tau^n) - E_m(\rho_{\tau,m}^i; \rho_h^k)| = \left| \int_{\mathbb{R}^d} \mathbf{N}(\psi_{1/m} * \rho_{\tau,m}^i) d(\rho_\tau^n - \rho_h^k) \right| \\ &\leq C' \|\nabla f\|_\infty \left[ \sigma((n\tau - kh)^2) + \sigma(\tau |\log \tau|^2) \right]. \end{aligned}$$

Finally,  $\|\nabla f\|_\infty$  is bounded uniformly in  $m, i$ , and  $\tau$ , since for  $B = B_1(0)$ , there exists  $c$  depending only on the dimension (and which we allow to change from line to line) so that, for all  $m \geq d + 1$ ,

$$\begin{aligned} \|\nabla f\|_\infty &\leq \|\nabla \mathcal{N}\|_{L^\infty(\mathbb{R}^d \setminus B)} + \|\nabla \mathcal{N}\|_{L^{m'}(B)} \|\rho_{\tau,m}^i\|_{L^m(\mathbb{R}^d)} \leq c + (1/\alpha_d)^{(m-1)/m} \|\rho_{\tau,m}^i\|_{L^m(\mathbb{R}^d)} \\ &\leq c \left( 1 + \|\rho\|_m + ((m-1)TC_d^2/2)^{1/m} \right) \leq c \end{aligned}$$

where the fourth inequality uses Lemma 2.14.

Thus, [20, Assumption A.2] is satisfied, so by [20, Theorem A.3], we conclude that for all  $0 \leq t \leq T$ , there exists  $C = C(E_\infty(\rho), T, d)$  (which we allow to change from line to line) so

$$F_{2t} \left( W_2^2(\rho_{t/n,m}^n, \rho_m(t)) \right) \leq C \left[ t/\sqrt{n} + t\omega \left( \sqrt{t/n} \right) + \sigma(t^2/n) + \sigma(t/n |\log(t/n)|^2) \right].$$

Hence, using again that  $F_t(x)$  is decreasing in  $t$ ,

$$F_{2T} \left( W_2^2(\rho_{t/n,m}^n, \rho_m(t)) \right) \leq C \left[ n^{-1/2} \log n + \sqrt{F_{2T}^{-1}(t/\sqrt{n})} + \sqrt{F_{2T}^{-1}(\sqrt{t/n} |\log(t/n)|)} \right].$$

For  $0 \leq x \leq e^{-1-\sqrt{2}}$ ,  $F_t(x) = x^{e^{C_d t}}$  and  $n^{-1/2} \log n = O(n^{-1/4})$ , so for  $n$  sufficiently large,

$$\left( W_2(\rho_{t/n,m}^n, \rho_m(t)) \right)^{2e^{2C_d T}} \leq C(n^{-1/8})^{1/e^{2C_d T}} \implies W_2(\rho_{t/n,m}^n, \rho_m(t)) \leq Cn^{-1/16e^{4C_d T}}.$$

Finally, it remains to show that the limit  $\rho_m$  is the unique solution of (PME-D) $_m$ . Following a parallel argument as in Jordan, Kinderlehrer, and Otto's original work on the convergence of the discrete gradient flow to solutions of the Fokker-Planck equation [27], one can show that for all  $\zeta \in C_0^\infty(\mathbb{R}^d \times [0, +\infty))$ ,

$$(5.8) \quad 0 = \int_{\mathbb{R}^d} \rho_m(x, 0) \zeta(x, 0) dx + \int_0^{+\infty} \int_{\mathbb{R}^d} \rho_m(x, s) (\partial_s \zeta(x, s) - \nabla \Phi_{1/m}(x, s) \nabla \zeta(x, s)) dx ds \\ + \int_0^{+\infty} \int_{\mathbb{R}^d} \rho_m(x, s)^m \Delta \zeta(x, s) dx ds.$$

To conclude that  $\rho_m(x, t)$  is the unique weak solution of (PME-D) $_m$ , it remains to show that for all  $0 < t < +\infty$ ,

$$(5.9) \quad \int_0^t \int_{\mathbb{R}^d} |\rho_m(x, s)|^m dx ds < +\infty,$$

$$(5.10) \quad \int_0^t \left( \int_{\mathbb{R}^d} \left| \frac{\nabla \rho_m(x, s)^m}{\rho_m(x, s)} + \nabla \Phi_{1/m}(x, s) \right|^2 \rho_m(x, s) dx \right)^{1/2} ds < +\infty.$$

(See, for example, [15, Theorem 6.1] and [34, Theorem 7.1]. While these references do not consider the case of a time-dependent drift, an identical argument applies to the present case.)

To show (5.9) and (5.10), we define the following piecewise constant interpolations:

$$\bar{\rho}_{t/n,m}(x, s) := \rho_{t/n,m}^i(x) \quad \text{and} \quad \bar{\rho}_{t/n}(x, s) := \rho_{t/n}^i \quad \text{for } s \in ((i-1)t/n, it/n].$$

Using [20, Theorem 3.6] (see [20, Appendix A.3] for the adaptation to ‘‘time dependent’’ gradient flows), one can show that

$$\bar{\rho}_{t/n,m}(x, s) \xrightarrow[W_2]{n \rightarrow +\infty} \rho_m(x, s) \quad \text{and} \quad \bar{\rho}_{t/n}(x, s) \xrightarrow[W_2]{n \rightarrow +\infty} \rho_\infty(x, s).$$

We begin with (5.9). By the lower semicontinuity of  $\|\cdot\|_m^m$  with respect to Wasserstein convergence [43, Lemma 3.4], Fatou's Lemma, and Lemma 2.14, which bounds  $\|\rho_{t/n,m}^i\|_m^m$  uniformly in  $i$  and  $n$ ,

$$\int_0^t \int_{\mathbb{R}^d} |\rho_m(x, s)|^m dx ds \leq \liminf_{n \rightarrow +\infty} \int_0^t \int_{\mathbb{R}^d} |\bar{\rho}_{t/n,m}(x, s)|^m dx ds = \liminf_{n \rightarrow +\infty} \sum_{i=1}^n \frac{t}{n} \int_{\mathbb{R}^d} |\rho_{t/n,m}^i(x)|^m dx < +\infty.$$

We now turn to (5.10). To ease notation, we recall the definition of the Rényi entropy and its metric slope [2, Theorem 10.4.6],

$$S_m(\mu) := \frac{1}{m-1} \int \mu(x)^m dx, \quad |\partial S_m|(\mu) = \left( \int_{\mathbb{R}^d} \left| \frac{\nabla \mu(x)^m}{\mu(x)} \right|^2 \mu(x) dx \right)^{1/2}.$$

By inequality (2.8),

$$\int_0^t \left( \int_{\mathbb{R}^d} |\nabla \Phi_{1/m}(x, s)|^2 \rho_m(x, s) dx \right)^{1/2} ds \leq C_d t < +\infty.$$

Thus, by the triangle inequality and Jensen's inequality, it suffices to show

$$(5.11) \quad \left( \int_0^t |\partial S_m|^2(\rho_m(s)) ds \right)^{1/2} < +\infty.$$

By [2, Theorem 3.1.6, Theorem 10.4.13] and the fact that  $E_m$  is  $\lambda$ -convex for  $\lambda = \lambda(m)$ ,

$$\frac{t}{n} \int \left| \frac{\nabla(\rho_{t/n,m}^i)^m}{\rho_{t/n,m}^i} + \nabla\psi_{1/m} * \mathbf{N}\rho_{t/n}^i \right|^2 d\rho_{t/n,m}^i \leq \frac{1}{1 + \lambda t/n} \left[ E_m(\rho_{t/n,m}^{i-1}; \rho_{t/n}^i) - E_m(\rho_{t/n,m}^i; \rho_{t/n}^i) \right].$$

Summing both sides from  $i = 1, \dots, n$  and using the definition of  $\bar{\rho}_{t/n,m}$ ,  $\bar{\rho}_{t/n}$ , and  $E_m$ ,

$$\begin{aligned} & \int_0^t \int_{\mathbb{R}^d} \left| \frac{\nabla(\bar{\rho}_{t/n,m}(x,s))^m}{\bar{\rho}_{t/n,m}(x,s)} + \nabla\psi_{1/m} * \mathbf{N}\bar{\rho}_{t/n}(x,s) \right|^2 \bar{\rho}_{t/n,m}(x,s) dx ds \\ & \leq \frac{1}{1 + \lambda t/n} \left[ E_m(\rho; \rho_{t/n}^1) - E_m(\rho_{t/n,m}^n; \rho_{t/n}^n) + \sum_{i=1}^n \int_{\mathbb{R}^d} \mathbf{N}\psi_{1/m} * \rho_{t/n,m}^i(x) \left[ \rho_{t/n}^{i+1}(x) - \rho_{t/n}^i(x) \right] dx \right], \\ & \leq \frac{1}{1 + \lambda t/n} \left[ E_m(\rho; \rho_{t/n}^1) + C_d + 2C_m C_d t \right], \end{aligned}$$

where in the last inequality we use Propositions 2.2, 2.3, and 2.14, for  $C_m$  chosen so that  $\|\nabla\mathbf{N}\psi_{1/m} * \rho_{t/n,m}^i\|_\infty < C_m$ . Taking the square root of both sides and applying the reverse triangle inequality for the  $L^2(d\bar{\rho}_{t/n,m}(\cdot, s))$  norm and Proposition 2.3,

$$\left( \int_0^t |\partial S_m|^2(\bar{\rho}_{t/n,m}(s)) ds \right)^{1/2} \leq \frac{1}{\sqrt{1 + \lambda t/n}} \left[ E_m(\rho; \rho_{t/n}^1) + C_d + 2C_m C_d t \right]^{1/2} + C_d.$$

Taking the  $\liminf_{n \rightarrow +\infty}$  and using that  $|\partial S_m|$  is lower semicontinuous with respect to Wasserstein convergence [2, Corollary 2.4.10], we conclude the result. □

**Acknowledgements.** The authors would like to thank José Antonio Carrillo, Alessio Figalli, and Alpár Mészáros for their helpful comments.

#### REFERENCES

- [1] Alexander, D., Kim, I., Yao, Y.: Quasi-static evolution and congested crowd transport. *Nonlinearity* **27**(4), 823–858 (2014). DOI 10.1088/0951-7715/27/4/823. URL <http://dx.doi.org/10.1088/0951-7715/27/4/823>
- [2] Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the space of probability measures, second edn. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2008)
- [3] Ambrosio, L., Serfaty, S.: A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.* **61**(11), 1495–1539 (2008). DOI 10.1002/cpa.20223. URL <http://dx.doi.org/10.1002/cpa.20223>
- [4] Balagué, D., Carrillo, J., Laurent, T., Raoul, G.: Nonlocal interactions by repulsive-attractive potentials: Radial ins/stability. *Phys. D.* **260**, 5–25 (2013)
- [5] Balagué, D., Carrillo, J.A., Yao, Y.: Confinement for repulsive-attractive kernels. *Discrete Contin. Dyn. Syst. Ser. B* **19**(5), 1227–1248 (2014). DOI 10.3934/dcdsb.2014.19.1227. URL <http://dx.doi.org/10.3934/dcdsb.2014.19.1227>
- [6] Benedetto, D., Caglioti, E., Carrillo, J.A., Pulvirenti, M.: A non-Maxwellian steady distribution for one-dimensional granular media. *J. Statist. Phys.* **91**(5-6), 979–990 (1998). DOI 10.1023/A:1023032000560. URL <http://dx.doi.org/10.1023/A:1023032000560>

- [7] Bertozzi, A.L., Carrillo, J.A., Laurent, T.: Blow-up in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity* **22**(3), 683–710 (2009). DOI 10.1088/0951-7715/22/3/009. URL <http://dx.doi.org/10.1088/0951-7715/22/3/009>
- [8] Bertozzi, A.L., Kolokolnikov, T., Sun, H., Uminsky, D., von Brecht, J.: Ring patterns and their bifurcations in a nonlocal model of biological swarms. *Commun. Math. Sci.* **13**(4), 955–985 (2015). DOI 10.4310/CMS.2015.v13.n4.a6. URL <http://dx.doi.org/10.4310/CMS.2015.v13.n4.a6>
- [9] Bertozzi, A.L., Laurent, T., Léger, F.: Aggregation and spreading via the Newtonian potential: the dynamics of patch solutions. *Math. Models Methods Appl. Sci.* **22**(suppl. 1), 1140,005, 39 (2012). DOI 10.1142/S0218202511400057. URL <http://dx.doi.org/10.1142/S0218202511400057>
- [10] Blanchet, A.: A gradient flow approach to the Keller-Segel systems. to appear in RIMS Kokyuroku’s lecture notes, preprint at <http://publications.ut-capitole.fr/16518/>
- [11] Blanchet, A., Carlen, E.A., Carrillo, J.A.: Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model. *J. Funct. Anal.* **262**(5), 2142–2230 (2012). DOI 10.1016/j.jfa.2011.12.012. URL <http://dx.doi.org/10.1016/j.jfa.2011.12.012>
- [12] Burger, M., Fetecau, R., Huang, Y.: Stationary states and asymptotic behavior of aggregation models with nonlinear local repulsion. *SIAM J. Appl. Dyn. Syst.* **13**(1), 397–424 (2014). DOI 10.1137/130923786. URL <http://dx.doi.org/10.1137/130923786>
- [13] Caffarelli, L., Salsa, S.: A geometric approach to free boundary problems, *Graduate Studies in Mathematics*, vol. 68. American Mathematical Society, Providence, RI (2005). DOI 10.1090/gsm/068. URL <http://dx.doi.org/10.1090/gsm/068>
- [14] Caffarelli, L., Vazquez, J.L.: Viscosity solutions for the porous medium equation. In: *Differential equations: La Pietra 1996 (Florence)*, *Proc. Sympos. Pure Math.*, vol. 65, pp. 13–26. Amer. Math. Soc., Providence, RI (1999). DOI 10.1090/pspum/065/1662747. URL <http://dx.doi.org/10.1090/pspum/065/1662747>
- [15] Carlier, G., Laborde, M.: On systems of continuity equations with nonlinear diffusion and nonlocal drifts. preprint at <http://arxiv.org/abs/1505.01304>
- [16] Carrillo, J.A., Hittmeir, S., Volzone, B., Yao, Y.: Nonlinear aggregation-diffusion equations: Radial symmetry and long time asymptotics. in preparation
- [17] Carrillo, J.A., Lisini, S., Mainini, E.: Uniqueness for Keller-Segel-type chemotaxis models. *Discrete Contin. Dyn. Syst.* **34**(4), 1319–1338 (2014). DOI 10.3934/dcds.2014.34.1319. URL <http://dx.doi.org/10.3934/dcds.2014.34.1319>
- [18] Carrillo, J.A., McCann, R.J., Villani, C.: Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.* **179**(2), 217–263 (2006). DOI 10.1007/s00205-005-0386-1. URL <http://dx.doi.org/10.1007/s00205-005-0386-1>
- [19] Chuang, Y.L., Huang, Y., D’Orsogna, M., Bertozzi, A.: Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials. *IEEE International Conference on Robotics and Automation* pp. 2292–2299 (2007)
- [20] Craig, K.: Nonconvex gradient flow in the Wasserstein metric and applications to constrained nonlocal interactions. preprint at <http://arxiv.org/abs/1512.07255>
- [21] Doye, J.P.K., Wales, D.J., Berry, R.S.: The effect of the range of the potential on the structures of clusters. *J. Chem. Phys.* **103**, 4234–4249 (1995)
- [22] Federer, H.: Curvature measures. *Trans. Amer. Math. Soc.* **93**, 418–491 (1959)
- [23] Fellner, K., Raoul, G.: Stable stationary states of non-local interaction equations. *Math. Models Methods Appl. Sci.* **20**(12), 2267–2291 (2010). DOI 10.1142/S0218202510004921. URL <http://dx.doi.org/10.1142/S0218202510004921>
- [24] Fetecau, R.C., Huang, Y.: Equilibria of biological aggregations with nonlocal repulsive-attractive interactions. *Phys. D* **260**, 49–64 (2013). DOI 10.1016/j.physd.2012.11.004. URL <http://dx.doi.org/10.1016/j.physd.2012.11.004>
- [25] Fetecau, R.C., Huang, Y., Kolokolnikov, T.: Swarm dynamics and equilibria for a nonlocal aggregation model. *Nonlinearity* **24**(10), 2681–2716 (2011). DOI 10.1088/0951-7715/24/10/002. URL <http://dx.doi.org/10.1088/0951-7715/24/10/002>
- [26] Fusco, N., Maggi, F., Pratelli, A.: Stability estimates for certain Faber-Krahn, isocapacitary and Cheeger inequalities. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **8**(5), 51–71 (2009)

- [27] Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17 (1998). DOI 10.1137/S0036141096303359. URL <http://dx.doi.org/10.1137/S0036141096303359>
- [28] Judovič, V.I.: Non-stationary flows of an ideal incompressible fluid. *Ž. Vyčisl. Mat. i Mat. Fiz.* **3**, 1032–1066 (1963)
- [29] Keller, E., Segel, L.: Initiation of slide mold aggregation viewed as an instability. *J. Theoret. Biol.* **26** (1970)
- [30] Kim, I., Pozar, N.: Porous medium equation to Hele-Shaw flow with general initial density. preprint at <http://arxiv.org/abs/1509.06287>
- [31] Kim, I., Yao, Y.: The patlak-keller-segel model and its variations: properties of solutions via maximum principle. *SIAM Journal on Mathematical Analysis* **44**(2), 568–602 (2012)
- [32] Kim, I.C.: Uniqueness and existence results on the Hele-Shaw and the Stefan problems. *Arch. Ration. Mech. Anal.* **168**(4), 299–328 (2003). DOI 10.1007/s00205-003-0251-z. URL <http://dx.doi.org/10.1007/s00205-003-0251-z>
- [33] Kim, I.C., Lei, H.K.: Degenerate diffusion with a drift potential: a viscosity solutions approach. *Discrete Contin. Dyn. Syst.* **27**(2), 767–786 (2010). DOI 10.3934/dcds.2010.27.767. URL <http://dx.doi.org/10.3934/dcds.2010.27.767>
- [34] Laborde, M.: On some non linear evolution systems which are perturbations of Wasserstein gradient flows. preprint at <http://arxiv.org/abs/1506.00126>
- [35] Lieb, E.H., Loss, M.: Analysis, *Graduate Studies in Mathematics*, vol. 14. American Mathematical Society, Providence, RI (1997)
- [36] Lieb, E.H., Yau, H.T.: The Chandrasekhar theory of stellar collapse as the limit of quantum mechanics. *Comm. Math. Phys.* **112**(1), 147–174 (1987). URL <http://projecteuclid.org/euclid.cmp/1104159813>
- [37] Lin, F., Zhang, P.: On the hydrodynamic limit of Ginzburg-Landau vortices. *Discrete Contin. Dynam. Systems* **6**(1), 121–142 (2000). DOI 10.3934/dcds.2000.6.121. URL <http://dx.doi.org/10.3934/dcds.2000.6.121>
- [38] Lions, P.L.: The concentration-compactness principle in the calculus of variations. The locally compact case. I. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **1**(2), 109–145 (1984). URL [http://www.numdam.org/item?id=AIHPC\\_1984\\_\\_1\\_2\\_109\\_0](http://www.numdam.org/item?id=AIHPC_1984__1_2_109_0)
- [39] Loeper, G.: Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *J. Math. Pures Appl.* (9) **86**(1), 68–79 (2006). DOI 10.1016/j.matpur.2006.01.005. URL <http://dx.doi.org/10.1016/j.matpur.2006.01.005>
- [40] Masmoudi, N., Zhang, P.: Global solutions to vortex density equations arising from sup-conductivity. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **22**(4), 441–458 (2005). DOI 10.1016/j.anihpc.2004.07.002. URL <http://dx.doi.org/10.1016/j.anihpc.2004.07.002>
- [41] Maury, B., Roudneff-Chupin, A., Santambrogio, F.: A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.* **20**(10), 1787–1821 (2010). DOI 10.1142/S0218202510004799. URL <http://dx.doi.org/10.1142/S0218202510004799>
- [42] Maury, B., Roudneff-Chupin, A., Santambrogio, F., Venel, J.: Handling congestion in crowd motion modeling. *Netw. Heterog. Media* **6**(3), 485–519 (2011). DOI 10.3934/nhm.2011.6.485. URL <http://dx.doi.org/10.3934/nhm.2011.6.485>
- [43] McCann, R.J.: A convexity principle for interacting gases. *Adv. Math.* **128**(1), 153–179 (1997). DOI 10.1006/aima.1997.1634. URL <http://dx.doi.org/10.1006/aima.1997.1634>
- [44] Mellet, A., Perthame, B., Quiros, F.: A Hele-Shaw problem for tumor growth. preprint at <http://arxiv.org/abs/1512.069957>
- [45] Perea, L., Gómez, G., Elosegui, P.: Extension of the Cucker–Smale control law to space flight formations. *AIAA J. of Guidance, Control, and Dynamics* **32**, 527–537 (2009)
- [46] Poupaud, F.: Diagonal defect measures, adhesion dynamics and Euler equation. *Methods Appl. Anal.* **9**(4), 533–561 (2002). DOI 10.4310/MAA.2002.v9.n4.a4. URL <http://dx.doi.org/10.4310/MAA.2002.v9.n4.a4>
- [47] Rechtsman, M., Stillinger, F., Torquato, S.: Optimized interactions for targeted self-assembly: application to a honeycomb lattice. *Phys. Rev. Lett.* **95**(22) (2005)
- [48] Santambrogio, F.: Optimal transport for applied mathematicians. *Progress in Nonlinear Differential Equations and their Applications*, 87. Birkhäuser/Springer, Cham (2015). DOI 10.1007/978-3-319-20828-2. URL <http://dx.doi.org/10.1007/978-3-319-20828-2>. Calculus of variations, PDEs, and modeling

- [49] Sugiyama, Y.: Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller-Segel systems. *Differential Integral Equations* **19**(8), 841–876 (2006)
- [50] Sun, H., Uminsky, D., Bertozzi, A.L.: Stability and clustering of self-similar solutions of aggregation equations. *J. Math. Phys.* **53**(11), 115,610, 18 (2012). DOI 10.1063/1.4745180. URL <http://dx.doi.org/10.1063/1.4745180>
- [51] Talenti, G.: Elliptic equations and rearrangements. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **3**(4), 697–718 (1976)
- [52] Topaz, C.M., Bertozzi, A.L., Lewis, M.A.: A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.* **68**(7), 1601–1623 (2006). DOI 10.1007/s11538-006-9088-6. URL <http://dx.doi.org/10.1007/s11538-006-9088-6>
- [53] Vázquez, J.L.: The porous medium equation. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford (2007). Mathematical theory
- [54] Villani, C.: Topics in optimal transportation, *Graduate Studies in Mathematics*, vol. 58. American Mathematical Society, Providence, RI (2003)
- [55] Wales, D.: Energy landscapes of clusters bound by short-ranged potentials. *Chem. Eur. J. Chem. Phys.* **11**, 2491–2494 (2010)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES

DEPARTMENT OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY