

# CONVERGENCE OF A ONE-DIMENSIONAL CAHN–HILLIARD EQUATION WITH DEGENERATE MOBILITY\*

MATIAS G. DELGADINO†

**Abstract.** We consider a one-dimensional periodic forward-backward parabolic equation, regularized by a degenerate nonlinear fourth-order term of order  $\varepsilon^2 \ll 1$ . This equation is known in the literature as Cahn–Hilliard equation with degenerate mobility. Under the hypothesis of the initial data being well prepared, we prove that as  $\varepsilon \rightarrow 0$ , the solution converges to the solution of a well-posed degenerate parabolic equation. The proof exploits the gradient flow nature of the equation in  $\mathcal{W}^2(\mathbb{T})$  and utilizes the framework of convergence of gradient flows developed by Sandier and Serfaty (*Comm. Pure Appl. Math.*, 57 (2004), pp. 1627–1672; *Discrete Contin. Dyn. Syst.*, 31 (2011), pp. 1427–1451). As an incidental, we study fine energetic properties of solutions to the thin-film equation  $\partial_t \nu = -(\nu \nu_{xxx})_x$ .

**Key words.** gradient flow, Cahn–Hilliard, Wasserstein, gamma convergence, fourth-order parabolic equations, thin-film equation

**AMS subject classifications.** 35K25, 47J35, 28A33, 49J40

**DOI.** 10.1137/15M1045429

**1. Introduction.** Given a smooth nonconvex potential  $W : \mathbb{R}_+ \rightarrow \mathbb{R}$ , we are interested in the properties of solutions  $\nu^\varepsilon$  to

$$(1) \quad \begin{cases} \partial_t \nu &= (\nu(W'(\nu) - \varepsilon^2 \nu_{xx})_x)_x & \text{in } (0, \infty) \times \mathbb{T}, \\ \nu(0) &= \nu_i^\varepsilon & \text{on } \{0\} \times \mathbb{T} \end{cases}$$

and more specifically, in their behavior as  $\varepsilon \rightarrow 0^+$ , where  $\mathbb{T}$  denotes the one-dimensional flat torus  $\mathbb{R}/\mathbb{Z}$ .

Equation (1) is known in the literature as the Cahn–Hilliard equation [8, 22], [17]. The function  $\nu^\varepsilon$  models the concentration of one of two phases in a system undergoing phase separation. Mathematically, this equation could be considered as a fourth-order regularization of a forward-backward parabolic equation, by the fourth-order term  $-\varepsilon^2(\nu \nu_{xxx})_x$ . In the case where  $W$  vanishes identically, we are left with a fourth-order parabolic equation

$$\partial_t \nu = -\varepsilon^2(m(\nu) \nu_{xxx})_x$$

known as the thin-film equation, with mobility  $m(\nu) = \nu$ , which is interesting on its own (see, for instance, [14], [20], [7]). Note that the Dirichlet energy is a Lyapunov functional and that when  $m(\nu) = \nu$ , the equation is formally the gradient flow of this energy under the  $\mathcal{W}^2(\mathbb{T})$  metric. This observation was made in the seminal paper by Otto [23] and has been exploited for some generalizations in [12], [18, 17].

The main result of this paper is the fact that, under some assumptions (see Theorem 3.1),  $\nu^\varepsilon$  converges, as  $\varepsilon \rightarrow 0$ , to the unique solution  $\nu_0$  of the following well-posed degenerate parabolic equation:

\*Received by the editors October 26, 2015; accepted for publication (in revised form) April 20, 2018; published electronically August 21, 2018.

<http://www.siam.org/journals/sima/50-4/M104542.html>

**Funding:** This work was partially supported by NSF grant DMS-1501067.

†Department of Mathematics, University of Maryland, College Park, MD 20742 (matias.delgadino@gmail.com).

$$(2) \quad \begin{cases} \partial_t \nu &= (\nu(W^{**'}(\nu))_x)_x, \\ \nu(0) &= \nu_i, \end{cases}$$

where  $W^{**}$  denotes the convex envelope of  $W$ .

The mathematical intuition behind this convergence comes from the fact that, formally at least, we know that  $\nu^\varepsilon$  is the gradient flow of

$$(3) \quad E^\varepsilon[\mu] = \int_{\mathbb{T}} \frac{\varepsilon^2}{2} |\mu_x|^2 + W(\mu) \, dx,$$

while  $\nu_0$  is the gradient flow of

$$E^{**}[\mu] = \int_{\mathbb{T}} W^{**}(\mu) \, dx,$$

with respect to  $\mathcal{W}^2(\mathbb{T})$ , and it is somewhat classical that the energy  $E^\varepsilon$   $\Gamma$ -converges to  $E^{**}$  in  $\mathcal{W}^2(\mathbb{T})$ . Unfortunately, it is well known that the  $\Gamma$ -convergence of the energy is not enough to prove the convergence of the gradient flows.

Indeed, to be able to prove the convergence of the gradient flows we need an additional condition on the gradient of the energy. A sufficient condition for Hilbert spaces was given in the paper by Sandier and Serfaty [24], which was later extended to metric spaces by Serfaty in [25]. This additional condition is usually written as follows:

$$(4) \quad \Gamma - \liminf_{\varepsilon \rightarrow 0^+} |\nabla E^\varepsilon| \geq |\nabla E^{**}| \text{ (see section 2 for definitions),}$$

and proving this inequality is always the hard part of the Sandier–Serfaty approach. However, in our case  $|\nabla E^\varepsilon|$  is not well understood, so we need to introduce a relaxation for which we prove a condition similar to (4) (see Theorem 3.2).

The framework of Sandier–Serfaty has been applied to an array of diverse problems. To name a few we have Allen–Cahn [21], Cahn–Hilliard [15], [4], nonlocal interactions energies [11], TV flow [10], and Fokker Plank [3]. The most relevant reference for this paper was written by Belletini et al. [4], where they consider the convergence of the one-dimensional Cahn–Hilliard equation on the Torus with mobility equal to one:

$$(5) \quad \partial_t \nu = (W'(\nu) - \varepsilon^2 \nu_{xx})_{xx}.$$

We actually borrow some of the notations and the ideas on how to track the oscillations of the solution. The main difference between [4] and our work is that (5) is a gradient flow of (3) in the Hilbert space  $H^{-1}(\mathbb{T})$ , instead of the metric space  $\mathcal{W}^2(\mathbb{T})$ . Besides bringing some nontrivial technical issues, working with a degenerate mobility coefficient also means that the estimates degenerate when the solution is near zero; this actually turns out to be a major issue that keeps showing up in the thin-film equation literature as well.

In the spirit of being self-contained, we give a brief introduction into the framework developed by Ambrosio, Gigli, and Savare for gradient flows in  $\mathcal{W}^2(\mathbb{T})$  (see Appendix A). We try to outline all of the tools and terminologies used in the paper, but it is in no way complete and the interested reader should definitely take the time to read [2].

A word of caution is that the framework developed in [2] cannot be applied to the functional  $E^\varepsilon$  as it is neither  $\lambda$ -convex in the sense given at [2], nor regular. For a proof see [9] or [13]. In fact, the subdifferential of  $E^\varepsilon$  is not really well understood.

Indeed, no matter how regular the measure is, if it vanishes at some point, it has not been proven that the natural candidate is a strong upper gradient.

We deal with this setback by considering Otto's approach in [23], which constructs solutions to the equation as the limit of Minimizing Movements, an idea that was originated by De Giorgi. In the case of  $E^\varepsilon$ , this has been made rigorous in [17], where the authors are even able to prove a uniform  $L_t^2(H_x^2)$  estimate for the constant interpolant of the discrete approximations, by using a discrete version of the entropy dissipation inequality (for the continuum case see [7]). In this paper we go a bit further and obtain an energy dissipation inequality (see (26)) by defining  $\mathcal{G}^\varepsilon$ , a weak lower semicontinuous relaxation of  $|\partial E^\varepsilon|$  (see section 2). To our knowledge, this is a novel result in the literature and gives a starting point to understand the  $\mathcal{W}^2(\mathbb{T})$  gradient flows of energies involving derivatives. It should be noted that there is no available well-posedness theory for solutions to the one-dimensional thin-film equation.

Once we are able to prove the existence of an appropriate solution to our equation, the main obstacle we encounter, when we try to prove the convergence, is oscillatory behavior, known as the wrinkling phenomenon. Numerical simulations show that the functions  $\nu^\varepsilon$  tend to oscillate quickly in the whole of the unstable set

$$\Sigma = cl(\{W > W^{**}\}).$$

However, in this paper we only prove that the wrinkling phenomenon occurs at least in a subset of  $\Sigma$  and we do not explore further if it can be proven analytically that when oscillations occur, they actually encompass the whole of  $\Sigma$ .

We prove that oscillations only occur inside of  $\Sigma$  by proving that  $d(\nu^\varepsilon, \Sigma)$  is uniformly lower semicontinuous in  $\varepsilon$  (see Corollary 4.4), which allows us to derive a uniform  $H_{loc}^1$  estimate away from the unstable set (see Proposition 5.2). The degenerate diffusion at  $\{\nu^\varepsilon = 0\}$  makes the control near zero very subtle. Only a careful study of the behavior of the solution near zero can rule out uncontrolled jumps (see proof of Theorem 4.3).

It is the intention of this paper that the proofs make a clear connection between where the oscillations can occur and the tangent lines of the graph of  $W$ . In short, in the regions where the tangent lines do not cross the graph of  $W$ , the function cannot have large oscillations (see (43)). In this way, the function  $W^{**}$  appears naturally.

As usual with the framework of [24], [25], we have to make an assumption on the initial data being well prepared with respect to the energy, meaning that

$$\lim_{\varepsilon \rightarrow 0^+} E^\varepsilon[\nu_i^\varepsilon] = E^{**}[\nu_i].$$

In our case, the well preparedness can be interpreted as the fact that the approximations  $\nu_i^\varepsilon$  stay away from  $\Sigma$ , so the convergence we prove only tells us that asymptotically the dynamic keeps it that way. With this assumption, we are missing how the wrinkling phenomenon is actually affecting the dynamic in the limit, which is a really interesting question on its own, but needs to be analyzed more carefully with other types of techniques.

The paper is organized as follows. The rest of this section deals with motivation. Section 2 provides the definitions and hypothesis of the objects we work with and introduces a suitable notion of solution to (1). Section 3 contains the statements of the main result of the convergence (see Theorem 3.1) and the main auxiliary result of the lower semicontinuity of the size of the gradients (see Theorem 3.2). Section 4 presents and proves the result on where can oscillations occur (see Theorem 4.3). Section 5 proves that away from  $\Sigma$  the functions are in  $H^1$  (see Proposition 5.2).

Section 6 proves Theorem 3.2. Section 7 proves Theorem 3.1. Appendix A gives the necessary background of gradient flows in  $\mathcal{W}^2(\mathbb{T})$ .

**1.1. Motivation.** Our original motivation for studying (1) came from a model for biological aggregation introduced in [26] which we describe now.

We consider  $\nu$  a population density that moves with velocity  $v = \nabla K * \nu - \nu \nabla \nu$ . If  $\nu_0$  is the initial population, then  $\nu$  satisfies

$$(6) \quad \begin{cases} \partial_t \nu + \nabla \cdot (\nu(\nabla K * \nu - \nu \nabla \nu)) = 0, \\ \nu(x, 0) = \nu_0. \end{cases}$$

The velocity  $v = v_a + v_d$  is decomposed in an attractive and a dispersal part. The attractive part of the velocity  $v_a = \nabla K * \nu$  is nonlocal and appears as individuals try to climb the gradients of a sensing function

$$s(x) = \int_{\mathbb{R}^n} K(x - y) \nu(y) dy.$$

The kernel  $K$  is typically radially symmetric, compactly supported, and of unit mass. The dispersal part of the velocity  $v_d = \nu \nabla \nu$  is local and arises as an anticrowding mechanism, which operates over a much shorter length scale.

Now, by rescaling (6), we want to consider what happens to a large population as we zoom out, over a large period of time. We thus set

$$\int_{\mathbb{R}^n} \nu_0 dx = \varepsilon^{-n}$$

for some  $\varepsilon \ll 1$  and we rescale time and space as follows:

$$(7) \quad \nu^\varepsilon(x, t) = \nu\left(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^2}\right);$$

the scaling in  $x$  is chosen such that  $\int \nu_0^\varepsilon = 1$ . Using (6), we obtain the following equation for  $\nu^\varepsilon$ :

$$(8) \quad \begin{cases} \partial_t \nu^\varepsilon + \nabla \cdot (\nu^\varepsilon(\nabla K^\varepsilon * \nu^\varepsilon - \nu^\varepsilon \nabla \nu^\varepsilon)) = 0, \\ \nu^\varepsilon(x, 0) = \nu_0^\varepsilon, \end{cases}$$

where  $K^\varepsilon = \frac{1}{\varepsilon^n} K(\frac{x}{\varepsilon})$  is an approximation of the  $\delta$  measure.

Adding and subtracting  $\nabla \cdot (\nu^\varepsilon \nabla \nu^\varepsilon)$ , we can rewrite (8) as

$$(9) \quad \partial_t \nu^\varepsilon + \nabla \cdot (\nu^\varepsilon(\nabla \nu^\varepsilon - \nu^\varepsilon \nabla \nu^\varepsilon + (\nabla K^\varepsilon * \nu^\varepsilon - \nabla \nu^\varepsilon))) = 0.$$

Assuming  $\nu^\varepsilon$  to be smooth and taking a Taylor expansion of  $\nu^\varepsilon$ , we get that

$$(10) \quad K^\varepsilon * \nu^\varepsilon(x) - \nu^\varepsilon(x) = \varepsilon^2 k_0 \Delta \nu^\varepsilon(x) + \mathcal{O}(\varepsilon^4),$$

where

$$k_0 = \int |x|^2 K(x) dx.$$

Replacing (10) in (9), disregarding the  $\mathcal{O}(\varepsilon^4)$  term, we finally obtain (1):

$$\begin{cases} \partial_t \nu^\varepsilon + \nabla \cdot (\nu^\varepsilon(-\nabla W'(\nu^\varepsilon) + \varepsilon^2 k_0 \nabla \Delta \nu^\varepsilon)) = 0, \\ \nu^\varepsilon(x, 0) = \nu_0^\varepsilon, \end{cases}$$

where  $W'(x) = \frac{x^2}{2} - x$ .

The Cahn–Hilliard equation we are studying in this paper is thus an approximation of the nonlocal equation (8). Unfortunately, the techniques used in this paper to control the oscillations of solutions to (1) could not be generalized to deal with solutions of (8). The main issue being that the nonlocality does not allow us to integrate exactly against the derivative of the solution. We are thus unable, at the present time, to fully describe the behavior of the solutions of (8) as  $\varepsilon \rightarrow 0$ . The only result that carries through is a uniform in  $\varepsilon$ ,  $L^\infty$  estimate for the solutions of (9), which follows almost exactly as Lemma 4.1.

*Remark 1.1.* It is worth noticing that (9) is the gradient flow of

$$F^\varepsilon[\nu] = \int \frac{\nu^3(x)}{6} - \frac{1}{2} K^\varepsilon * \nu(x) \nu(x) \, dx$$

with respect to the metric induced by the  $\mathcal{W}^2$  distance. By adding and subtracting  $\frac{\nu^2(x)}{2}$ , in the expression, we obtain, after some calculations,

$$(11) \quad F^\varepsilon[\nu] = \int W(\nu) \, dx + \frac{1}{4} \int \int K^\varepsilon(x-y) (\nu(x) - \nu(y))^2 \, dx dy$$

with  $W(x) = \frac{x^3}{6} - \frac{x^2}{2}$ .

The seminorm

$$\frac{1}{4} \int \int K^\varepsilon(x-y) (\nu(x) - \nu(y))^2 \, dx dy$$

is, up to a constant, a smooth nonlocal approximation of

$$\frac{\varepsilon^2}{2} \int |\nabla \nu|^2;$$

therefore (11) can be considered as a smooth nonlocal approximation of (3).

The gradient flow of  $F^\varepsilon$  is more difficult to track than  $E^\varepsilon$  due to the lack of exact computations involving derivatives.

*Remark 1.2.* Different scalings of time in (7) can be considered. The case of  $\frac{t}{\varepsilon}$  is related, in the limit  $\varepsilon \rightarrow 0$ , to motion by mean curvature (see [15]).

*Remark 1.3.* A similar heuristic relationship between (1) and the nonlocal model in [26] has been drawn independently in [6].

**2. Notation and assumptions.** Throughout the paper, we always consider measures  $\mu \in \mathcal{P}(\mathbb{T})$  that are absolutely continuous with respect to the Lebesgue measure, and we do not make any distinction between the measure and its density. Also, we use the term sequence loosely: it may denote family of measures labeled by the continuous parameter  $\varepsilon$ .

**2.1. Assumptions on  $W$ .** We assume that  $W$  is in  $C^2([0, \infty), \mathbb{R})$ ; we denote by  $W^{**}$  its convex envelope. We define the auxiliary function  $Q$  such that

$$(12) \quad Q'(y) = yW'(y) - W(y);$$

we use the notation with a prime, because its derivative is related to the second derivative of  $W$ , namely

$$(13) \quad Q''(y) = yW''(y).$$

Moreover, we assume that  $W$  has the following properties

- (H1) There exists a constant  $C > 0$  such that for every  $y \in \mathbb{R}$

$$(14) \quad |Q'(y)| \leq C(1 + W(y))$$

and

$$(15) \quad |W'(y)| \leq C(1 + W(y)).$$

- (H2)  $\lim_{y \rightarrow +\infty} Q'(y) = +\infty$ .
- (H3) The unstable set is a finite union of closed intervals. That is to say,  $\Sigma = cl(\{W > W^{**}\} \cup \{0\}) = \cup_{i=1}^p \Sigma_i$ , where  $p \in \mathbb{N}$  and  $\Sigma_i = [a_i, b_i]$  with  $a_{i+1} > b_i$  and  $a_1 = 0$ .

The first interval could be degenerate in the sense of  $a_1 = b_1 = 0$ . As the dynamics near zero will be special, we will distinguish a value

$$(16) \quad m_0 = \begin{cases} b_1 + 1 & \text{if } p = 1, \\ \frac{b_1 + a_2}{2} & \text{if } p \geq 2. \end{cases}$$

- (H4) Given  $K \subset \Sigma^c$  compact,  $\inf_{A \in K} W''(A) > 0$ . Equivalently,  $W''(p) > 0$  for any  $p \in \Sigma^c$ .

*Remark 2.1.* Equations (1) and (2) are not affected by adding an affine function to  $W$ , so without loss of generality we will consider the case  $W(0) = 0$  and  $W'(0) = 0$ .

*Remark 2.2.* The potential that arises from section 1.1

$$(17) \quad W(y) = \frac{|y|^3}{6} - \frac{|y|^2}{2}$$

satisfies all of the hypotheses above.

**2.2. Functionals  $E^\varepsilon$ ,  $E^{**}$ ,  $\mathcal{G}^\varepsilon$ ,  $|\nabla E^{**}|$ .** For any  $\varepsilon > 0$ , we define

$$E^\varepsilon : \mathcal{P}(\mathbb{T}) \rightarrow [0, +\infty],$$

the functional

$$E^\varepsilon[\nu] = \begin{cases} \int_{\mathbb{T}} \frac{\varepsilon^2}{2} |\nu_x|^2 + W(\nu) \, dx & \text{if } \nu \in H^1(\mathbb{T}), \\ +\infty & \text{elsewhere.} \end{cases}$$

Formally, the subdifferential of  $E^\varepsilon$  at  $\mu$  with respect to  $\mathcal{W}^2$  is given by

$$\partial_{\mathcal{W}^2} E^\varepsilon[\mu] = (W'(\mu) - \varepsilon^2 \mu_{xx})_x$$

and we have

$$(18) \quad |\partial_{\mathcal{W}^2} E^\varepsilon|(\mu) = \left( \int_{\mathbb{T}} \mu |W'(\mu) - \varepsilon^2 \mu_{xx}|^2 \, dx \right)^{\frac{1}{2}}.$$

However, to our knowledge, unless  $\mu$  is assumed to be strictly positive, nobody has proven that  $E^\varepsilon$  is actually subdifferentiable at  $\mu$ , no matter how regular  $\mu$  is. For this reason, we introduce a relaxation of (18),

$$\mathcal{G}^\varepsilon(\cdot) : \mathcal{P}(\mathbb{T}) \cap H^1(\mathbb{T}) \rightarrow [0, +\infty],$$

defined by

$$(19) \quad |\mathcal{G}^\varepsilon(\mu)|^2 = \sup_{\varphi \in H^2(\mathbb{T})} 2 \int_{\mathbb{T}} \left( \mu W'(\mu) - W(\mu) + \frac{3\varepsilon^2}{2} |\mu_x|^2 \right) \varphi_x + \varepsilon^2 \mu \mu_x \varphi_{xx} \, dx - \int_{\mathbb{T}} \mu \varphi^2 \, dx.$$

*Remark 2.3.* The functional  $|\mathcal{G}^\varepsilon(\cdot)|^2$  is lower semicontinuous under strong  $H^1(\mathbb{T})$  convergence.

In this paper we mostly deal with  $\mu \in H^2(\mathbb{T})$ , where the previous definition simplifies a little bit using the auxiliary map

$$(20) \quad G^\varepsilon(\mu) = \mu W'(\mu) - W(\mu) + \frac{\varepsilon^2}{2} |\mu_x|^2 - \varepsilon^2 \mu \mu_{xx}.$$

We notice that for  $\mu \in H^2(\mathbb{T})$ , we have

$$(21) \quad |\mathcal{G}^\varepsilon(\mu)|^2 = \sup_{\varphi \in H^1(\mathbb{T})} 2 \int_{\mathbb{T}} G^\varepsilon(\mu) \varphi_x \, dx - \int_{\mathbb{T}} \mu \varphi^2 \, dx.$$

The map  $G^\varepsilon$  keeps appearing in the paper due to its relation with the dissipation  $G^\varepsilon(\mu)_x = \mu(W'(\mu) - \varepsilon^2 \mu_{xx})_x$ .

We should note that  $\mathcal{G}^\varepsilon$  is a reasonable relaxation for  $|\partial_{\mathcal{W}^2} E^\varepsilon|$ . Indeed, using our formal definition we would like to show

$$(22) \quad |\partial_{\mathcal{W}^2} E^\varepsilon|^2 = \int_{\mathbb{T}} \mu |W'(\mu) - \varepsilon^2 \mu_{xx}|^2 \, dx \geq |\mathcal{G}^\varepsilon(\mu)|^2.$$

If the left-hand side is infinite, there is nothing to prove. If the left-hand side is finite, then

$$G^\varepsilon(\mu)_x = \mu(W'(\mu) - \varepsilon^2 \mu_{xx})_x$$

belongs, at least, to  $L^2(\mathbb{T})$ . So, (22) follows by integrating by parts and applying Young's inequality.

*Remark 2.4.* The idea behind this cumbersome definition is that we have some control over  $G^\varepsilon$ . In particular, we can control the  $BV$  seminorm of  $G^\varepsilon$  in the following way:

$$2|G^\varepsilon(\mu)|_{BV} - 1 \leq \sup_{\varphi \in H^1(\mathbb{T}) : \|\varphi\|_{L^\infty} \leq 1} 2\langle G^\varepsilon(\mu), \varphi_x \rangle - \int_{\mathbb{T}} \mu \varphi^2 \, dx \leq |\mathcal{G}^\varepsilon(\mu)|^2.$$

Moreover, when  $\mu$  is regular in  $\{\mu > 0\}$ , then

$$\int_{\mu > 0} \mu |W'(\mu) - \varepsilon^2 \mu_{xx}|^2 \, dx = \sup_{\delta > 0} 2\langle G^\varepsilon(\mu)_x, \varphi_\delta \rangle - \int_{\mathbb{T}} \mu \varphi_\delta^2 \, dx \leq |\mathcal{G}(\mu)|^2,$$

where  $\varphi_\delta = (W'(u) - \varepsilon^2 u_{xx})_x \eta_\delta$  and  $\eta_\delta$  is a smooth bump function such that  $\mathbb{1}_{\{u > \delta\}} \leq \eta_\delta \leq \mathbb{1}_{\{u > 0\}}$ . The fact that the integral is only on the set  $\{\mu > 0\}$  is a standard inconvenience in the thin-film equation literature and is the source of some technical difficulties.

We also define

$$E^{**} : \mathcal{P}(\mathbb{T}) \rightarrow [0, +\infty],$$

the functional

$$E^{**}[\nu] = \begin{cases} \int_{\mathbb{T}} W^{**}(\nu) dx & \text{if } \nu \in L^1(\mathbb{T}), \\ +\infty & \text{elsewhere,} \end{cases}$$

where  $E^{**}$  is convex (see [19]). The subdifferential  $E^{**}$  is given by

$$\partial_{\mathcal{W}^2} E^{**}[\mu] = (W^{**'}(\mu))_x.$$

Therefore, the metric upper gradient

$$(23) \quad |\partial_{\mathcal{W}^2} E^{**}| : \mathcal{P}(\mathbb{T}) \rightarrow [0, +\infty]$$

is given by

$$(24) \quad |\partial_{\mathcal{W}^2} E^{**}(\mu)| = \begin{cases} \left( \int_{\mathbb{T}} \mu |(W^{**'}(\mu))_x|^2 dx \right)^{\frac{1}{2}} & \text{if } (Q^{**'}(\mu)) \in W^{1,1}(\mathbb{T}), \\ +\infty & \text{elsewhere,} \end{cases}$$

where  $Q^{**'}(z) = zW^{**'}(z) - W^{**}(z)$ . For more details, see section 10.4.3 in [2].

*Remark 2.5.* Theorem 10.4.6 in [2] has a doubling assumption on the potential. Namely, there exists  $C > 0$  such that  $W^{**}(z+w) \leq C(1 + W^{**}(z) + W^{**}(w))$  for any  $z$  and  $w$ . We omit the subtlety of the doubling condition because we deal with measures that are uniformly bounded in  $L^\infty$ .

*Remark 2.6.* Because  $E^{**}$  is convex, we have that  $|\nabla E^{**}|$  is a strong upper gradient. (See Definitions A.5 and A.3)

**2.3. Existence of  $\nu^\varepsilon$  and  $\nu_0$ .** Given  $\varepsilon > 0$  and an initial condition  $\nu_i^\varepsilon$ , such that

$$E^\varepsilon[\nu_i^\varepsilon] < +\infty,$$

we consider  $\nu^\varepsilon(x, t)$  a solution of (1) given by the following proposition.

**PROPOSITION 2.7.** *Given  $\nu_i^\varepsilon \in \mathcal{P}(\mathbb{T})$ , such that  $E^\varepsilon[\nu_i^\varepsilon] < \infty$ , then there exist  $\nu^\varepsilon \in L^\infty((0, \infty); H^1(\mathbb{T})) \cap L^2_{loc}((0, \infty); H^2(\mathbb{T})) \cap C^{1,4}_{loc}(\{\nu^\varepsilon > 0\})$  such that*

$$(25) \quad \int_0^\infty \int_{\mathbb{T}} \nu^\varepsilon \phi_t dx dt - \int_0^\infty \int_{\mathbb{T}} (\varepsilon^2 \nu^\varepsilon_{xx} - W'(\nu^\varepsilon)) (\nu^\varepsilon \phi_x)_x dx dt = 0$$

for every  $\phi \in C_c^\infty((0, \infty) \times \mathbb{T})$ .

Moreover,

$$(26) \quad E^\varepsilon[\nu^\varepsilon(t)] + \frac{1}{2} \int_0^t |\mathcal{G}(\nu^\varepsilon)|^2 ds + \frac{1}{2} \int_0^t |\nu^{\varepsilon'}|^2 ds \leq E^\varepsilon[\nu_i^\varepsilon] \quad \forall t > 0,$$

where  $|\nu^{\varepsilon'}|$  is the size of the metric derivative of  $\nu^\varepsilon$  with respect to  $\mathcal{W}^2(\mathbb{T})$  (see Definition A.2).

*Remark 2.8.* We cannot claim that  $\nu^\varepsilon$  is a curve of maximal slope, as defined in [2], since we do not prove that  $\mathcal{G}^\varepsilon$  is an upper gradient of  $E^\varepsilon$  (see Definition A.3).

*Remark 2.9.* From the inclusion  $H^2 \subset C^{1, \frac{1}{2}}$ , we get that for almost every  $t$ ,  $\nu^\varepsilon(t) \in C^{1, \frac{1}{2}}$ .



*Proof.* To lighten the notation, we prove the proposition in the case  $\varepsilon = 1$  and drop the dependence.

First, we notice that  $\nu \in C_{loc}^{1,4}(\{\nu > 0\})$  follows from Schauder estimates (see [5]). The existence of  $\nu \in L^\infty((0, \infty); H^1(\mathbb{T})) \cap L_{loc}^2((0, \infty); H^2(\mathbb{T}))$ , which satisfies (25), is a particular case of Theorem 1 in [17]. More precisely,  $\nu$  is constructed as any accumulation point of the discrete interpolation of the solutions of the appropriate JKO scheme. We are interested in showing that  $\nu$  also holds the energy dissipation inequality (26).

The JKO scheme defines inductively

$$(27) \quad \mu_\tau^0 = \nu_i^1, \quad \mu_\tau^{n+1} = \operatorname{argmin}_{\rho \in \mathcal{P}(\mathbb{T})} \{d_2^2(\mu_\tau^n, \rho) + 2\tau E(\rho)\}.$$

Note that  $\nu$  may not be unique, so we fix such a  $\nu$  and a corresponding sequence of  $\tau_n \rightarrow 0$ , for which the constant interpolation of  $\{\mu_{\tau_n}\}_{n=0}^\infty$  converges to  $\nu$  in  $L^2((0, T); H^1(\mathbb{T}))$ .

Subsequently, we define the De Giorgi variational interpolation by

$$\bar{\mu}_\tau(t) = \operatorname{argmin}_{\rho \in \mathcal{P}(\mathbb{T})} \{d_2^2(\mu_\tau^n, \rho) + 2(t - (n-1)\tau)E(\rho)\} \quad \text{when } t \in ((n-1)\tau, n\tau).$$

By Lemmas 3.1.3 and 3.2.2 in [2], we know that  $E$  is strongly subdifferentiable at  $\bar{\mu}_{\tau_n}(t)$  for every  $t > 0$  (see Definition A.16), and that  $\bar{\mu}_{\tau_n}(t) \rightarrow \nu(t)$  for all  $t \geq 0$ . Moreover, for every  $n \in \mathbb{N}$ ,

$$E(\mu_\tau^n) + \frac{1}{2\tau} \sum_{k=1}^n d_2^2(\mu_\tau^n, \mu_\tau^{n-1}) + \frac{1}{2} \int_0^{n\tau} |\partial E(\bar{\mu}_\tau(t))|^2 dt \leq E(\mu_\tau^0) = E(\nu_i).$$

Therefore, because of the strong subdifferentiability we know that by Lemma A.17 and the definition of  $\mathcal{G}$

$$|\mathcal{G}(\bar{\mu}_\tau(t))|^2 \leq |\partial E(\bar{\mu}_\tau(t))|^2 = \int_{\mathbb{T}} |(W'(\bar{\mu}_\tau(t)) - \bar{\mu}_{\tau xx}(t))_x|^2 \bar{\mu}_\tau(t) dx \quad \forall t > 0.$$

We deduce

$$(28) \quad E(\mu_\tau^n) + \frac{1}{2\tau} \sum_{k=1}^n d_2^2(\mu_\tau^n, \mu_\tau^{n-1}) + \frac{1}{2} \int_0^{n\tau} |\mathcal{G}(\bar{\mu}_\tau(t))|^2 dt \leq E(\nu_i).$$

Due to the  $L^\infty((0, T); H^1(\mathbb{T}))$  bound we have that  $\bar{\mu}_{\tau_n}(t) \rightarrow \nu(t)$  weakly in  $H^1(\mathbb{T})$  and strongly in  $L^2(\mathbb{T})$  for all  $t > 0$ . This implies that  $\bar{\mu}_{\tau_n} \rightarrow \nu$  strongly in  $L^2((0, T); L^2(\mathbb{T}))$ . Interpolating the bounds  $L^\infty((0, T); H^1(\mathbb{T}))$  and  $L^2((0, T); H^2(\mathbb{T}))$  (Proposition 4.1 in [17]), one can show that, up to subsequence,  $\bar{\mu}_{\tau_n} \rightarrow \nu$  in  $L^2((0, T); H^1(\mathbb{T}))$ . Therefore,  $\bar{\mu}_{\tau_n}(t) \rightarrow \nu(t)$  strongly in  $H^1(\mathbb{T})$  almost everywhere in  $(0, T)$ .

The energy inequality (26) follows by taking the limit  $\tau_n \rightarrow 0$  in (28). More precisely, the metric derivative term in the energy inequality (26) follows exactly as in the proof of Theorem 2.3.3 in [2]. The term involving  $\mathcal{G}$  follows from Fatou's lemma and from the lower semicontinuity  $\mathcal{G}$  under the strong  $H^1(\mathbb{T})$  convergence (see Remark 2.3).  $\square$

As the functional  $E^{**}$  is convex, then, using Theorem A.19, we denote by  $\nu_0$  the unique gradient flow of  $E^{**}$  emanating from  $\nu_i$ . Moreover,  $\nu_0$  is also the unique distributional solution to

$$(29) \quad \begin{cases} \partial_t \nu &= ((W^{**'}(\nu))_x \nu)_x, \\ \nu(0) &= \nu_i. \end{cases}$$

It can be characterized by either the energy inequality, also known as the maximal slope condition

$$(30) \quad E^{**}[\nu(t)] + \frac{1}{2} \int_0^t |\nabla E^{**}(\nu)|^2 ds + \frac{1}{2} \int_0^t |\nu'|^2 ds \leq E^{**}[\nu_i] \quad \forall t > 0,$$

or the energy equality

$$(31) \quad E^{**}[\nu(t)] + \frac{1}{2} \int_0^t |\nabla E^{**}(\nu)|^2 ds + \frac{1}{2} \int_0^t |\nu'|^2 ds = E^{**}[\nu_i] \quad \forall t > 0.$$

**3. Statement of the result.** The main result of this paper is the following.

**THEOREM 3.1.** *Let  $\{\nu_i^\varepsilon\}_\varepsilon \subset \mathcal{P}(\mathbb{T})$  and  $\nu_i \in \mathcal{P}(\mathbb{T})$  be such that*

$$E^\varepsilon[\nu_i^\varepsilon] < +\infty \quad \text{and} \quad E^{**}[\nu_i] < +\infty.$$

*Suppose that*

$$(32) \quad \lim_{\varepsilon \rightarrow 0^+} \nu_i^\varepsilon = \nu_i \quad \text{in } \mathcal{W}^2(\mathbb{T})$$

*and*

$$(33) \quad \lim_{\varepsilon \rightarrow 0^+} E^\varepsilon[\nu_i^\varepsilon] = E^{**}[\nu_i].$$

*Then, for any  $T > 0$ ,*

$$\lim_{\varepsilon \rightarrow 0^+} \nu^\varepsilon = \nu_0 \quad \text{in } C^0([0, T]; \mathcal{W}^2(\mathbb{T})),$$

$$\lim_{\varepsilon \rightarrow 0^+} \int_0^T (\mathcal{G}^\varepsilon(\nu^\varepsilon(t)) - |\nabla E^{**}(\nu_0(t))|)^2 dt = 0,$$

*and*

$$\lim_{\varepsilon \rightarrow 0^+} E^\varepsilon[\nu^\varepsilon(t)] = E^{**}[\nu_0(t)] \quad \forall t \geq 0,$$

where  $\nu^\varepsilon$  is the solution of (1) given by Proposition 2.7 with initial condition  $\nu_i^\varepsilon$  and  $\nu_0$  is the unique gradient flow of  $E^{**}$  (solution of (29)) with initial condition  $\nu_i$ , with respect to the metric  $\mathcal{W}^2(\mathbb{T})$ .

As in [21], [15], [4], [11], [10], [3] the key step in the proof of Theorem 3.1 is to prove the  $\Gamma$ -liminf of  $\mathcal{G}^\varepsilon$  to  $|\nabla E^{**}|$ ; more specifically, we need to prove the following.

**THEOREM 3.2.** *Let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T})$  such that  $\rho^\varepsilon \in C^1(\mathbb{T}) \cap C_{loc}^4\{\rho^\varepsilon > 0\}$ ,  $\rho^\varepsilon \rightarrow \rho_0$  in  $\mathcal{W}^2(\mathbb{T})$ , and  $\sup_\varepsilon E^\varepsilon(\rho^\varepsilon) < \infty$ ; then*

$$(34) \quad \liminf_{\varepsilon \rightarrow 0^+} \mathcal{G}^\varepsilon(\rho^\varepsilon) \geq |\nabla E^{**}|(\rho_0).$$

The next two sections are devoted to some preliminary compactness results, which are used in the proof of Theorem 3.2 that can be found in section 6. The proof of Theorem 3.1 can be found in section 7.

#### 4. Preliminary to the proof of Theorem 3.2.

**4.1. Uniform  $L^\infty$  estimate.** The first step is to prove a uniform  $L^\infty$  estimate.

**PROPOSITION 4.1.** *Given  $C > 0$ , let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2(\mathbb{T})$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$ , then there exists  $M(C) > 0$ , such that*

$$\sup_\varepsilon \|\rho^\varepsilon\|_\infty \leq M.$$

Moreover, there exists  $\rho_0 \in L^\infty(\mathbb{T})$  such that up to a subsequence,

$$\rho^\varepsilon \rightharpoonup \rho^0 \quad \text{weak-}^* L^\infty.$$

*Proof.* Consider  $G^\varepsilon(\rho^\varepsilon)$  as in (20):

$$G^\varepsilon(\rho^\varepsilon) = -\varepsilon^2 \rho^\varepsilon \rho_{xx}^\varepsilon + \frac{\varepsilon^2}{2} (\rho_x^\varepsilon)^2 + Q'(\rho^\varepsilon)$$

with  $Q'$  defined by (12). By Remark 2.4, we know that

$$|G^\varepsilon(\rho^\varepsilon)|_{BV} \leq C.$$

Moreover,

$$\int_{\mathbb{T}} G^\varepsilon(\rho^\varepsilon) dx = \int_{\mathbb{T}} \frac{3}{2} \varepsilon^2 (\rho_x^\varepsilon)^2 + Q'(\rho^\varepsilon) dx \leq 3E^\varepsilon[\rho^\varepsilon] + D \int_{\mathbb{T}} (W(\rho^\varepsilon) + 1) dx$$

by (H1). Therefore,  $G^\varepsilon(\rho^\varepsilon)$  is uniformly in  $BV(\mathbb{T})$ , which implies

$$\sup_\varepsilon \|G^\varepsilon(\rho^\varepsilon)\|_\infty < \infty.$$

Now, let us prove that  $\rho^\varepsilon$  is uniformly in  $L^\infty$ : take  $x_0$ , such that  $\rho^\varepsilon(x_0) = \|\rho^\varepsilon\|_\infty$ ; then  $\rho_x(x_0) = 0$  and  $\rho_{xx}(x_0) \leq 0$ . We should note that because  $\mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$ , then  $\rho^\varepsilon \in C_{loc}^{2, \frac{1}{2}}(\{\rho^\varepsilon > 0\})$ , and  $\rho_{xx}(x_0)$  has a well-defined value.

Hence, we have the bound

$$\|G^\varepsilon(\rho^\varepsilon)\|_\infty \geq G^\varepsilon(\rho^\varepsilon)(x_0) \geq Q'(\rho^\varepsilon(x_0)),$$

which, by assumption (H2), gives a bound for  $\sup_\varepsilon \|\rho^\varepsilon\|_\infty$ .  $\square$

**COROLLARY 4.2.** *Under the assumptions of Proposition 4.1,  $G^\varepsilon(\rho^\varepsilon)$  is bounded in  $H^1(\mathbb{T})$  uniformly in  $\varepsilon$ . More precisely, we have the bound*

$$\|G^\varepsilon(\rho^\varepsilon)_x\|_{L^2} \leq \frac{|\mathcal{G}^\varepsilon(\rho^\varepsilon)|^2 + \|\rho^\varepsilon\|_\infty}{2} \leq C.$$

Therefore,  $G^\varepsilon(\rho^\varepsilon) \in C^{\frac{1}{2}}(\mathbb{T})$  uniformly in  $\varepsilon$ .

*Proof.* Considering  $\varphi \in H^1(\mathbb{T})$  such that  $\|\varphi\|_{L^2(\mathbb{T})} = 1$ , we have the inequality

$$(35) \quad \int_{\mathbb{T}} G^\varepsilon(\rho^\varepsilon) \phi_x dx - \|\rho^\varepsilon\|_\infty \leq 2 \int_{\mathbb{T}} G^\varepsilon(\rho^\varepsilon) \phi_x dx - \int_{\mathbb{T}} \rho^\varepsilon \varphi^2 dx \leq |\mathcal{G}^\varepsilon(\rho^\varepsilon)|^2$$

by the definition of  $\mathcal{G}^\varepsilon$ . The result now follows.  $\square$

**4.2. Control of the oscillations in the good set.** The key in the proof of Theorem 3.2 is to control the size of the oscillations of  $\rho^\varepsilon$  in the good sets. This will be given by the following theorem.

**THEOREM 4.3.** *Given  $C > 0$ , let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2(\mathbb{T})$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$ ; then, for any  $L \geq 0$  and  $\eta > 0$ , there exists  $\delta(\eta, C) > 0$ , independent of  $\varepsilon$ , such that for any  $\varepsilon < \varepsilon_0(\eta, C, L)$  and any pair of sequences  $\{x_\varepsilon\}_\varepsilon, \{y_\varepsilon\}_\varepsilon \subset \mathbb{T}$  satisfying*

- $0 < y_\varepsilon - x_\varepsilon < \delta$  and
- $|\rho_x^\varepsilon(x_\varepsilon)| < L$  and  $|\rho_x^\varepsilon(y_\varepsilon)| < L$ ,

*we have either*

$$d(\rho^\varepsilon(z), \Sigma) < \eta \quad \forall z \in [x_\varepsilon, y_\varepsilon]$$

*or*

$$|\rho^\varepsilon(x_\varepsilon) - \rho^\varepsilon(y_\varepsilon)| < \eta.$$

Theorem 4.3 is similar to Lemma 5.5 in the paper by Belletini et al. [4]. The main difference in the proof is that in [4] they have control of the  $H^1$  norm of

$$(36) \quad e^\varepsilon(\rho^\varepsilon) = W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon,$$

while we only have control on

$$(37) \quad \int_{\mathbb{T}} |e^\varepsilon(\rho^\varepsilon)_x|^2 \rho^\varepsilon dx,$$

which is degenerate near  $\{\rho^\varepsilon = 0\}$ .

Theorem 4.3 can be interpreted as a uniform lower semicontinuity for  $d(\rho^\varepsilon, \Sigma)$ :

**COROLLARY 4.4.** *Let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2(\mathbb{T})$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$  and that  $\rho^\varepsilon \rightarrow \rho_0$  in  $\mathcal{W}^2(\mathbb{T})$ ; then for  $x$ , any Lebesgue point of  $\rho_0$ , there exists  $\varepsilon_x$  and  $\delta' = \delta'(d(\rho_0(x), \Sigma))$  such that for every  $\varepsilon < \varepsilon_x$  and  $y \in (x - \delta', x + \delta')$ , we have*

$$d(\rho_0(y), \Sigma) > \frac{d(\rho_0(x), \Sigma)}{2}.$$

Moreover,  $\Omega := \{\rho_0 \notin \Sigma\}$  has an open representative.

*Proof of Corollary 4.4.* We start with the following claim.

**CLAIM.** *For any  $\beta > 0$ , we define  $\delta(\beta) = \frac{\delta(\beta, C)}{4}$  and  $\varepsilon(\beta) = \varepsilon(C, \beta, \frac{4M}{\delta})$  (given by Theorem 4.3). If for some  $\varepsilon \in (0, \varepsilon_\beta)$ , we have that  $d(\rho^\varepsilon(x), \Sigma) > 2\beta$ ; then  $d(\rho^\varepsilon(y), \Sigma) > \beta$  for all  $y \in (x - \delta_\beta, x + \delta_\beta)$ .*

*Proof of the Claim.* We take  $\delta = \delta(\eta, C)$  given by Theorem 4.3. Because we know that  $\sup_\varepsilon |\rho^\varepsilon| \leq M$ , we have  $\text{osc}_{(x+\frac{\delta}{4}, x+\frac{\delta}{2})} \rho^\varepsilon \leq M$ . Therefore, there exists  $x_1 \in (x + \frac{\delta}{4}, x + \frac{\delta}{2})$  such that

$$|\rho_x^\varepsilon(x_1)| \leq \frac{4M}{\delta}.$$

Similarly, there exists  $x_2 \in (x - \frac{\delta}{2}, x - \frac{\delta}{4})$  such that  $|\rho_x^\varepsilon(x_2)| \leq \frac{4M}{\delta}$ .

If  $\varepsilon < \varepsilon(C, \eta, \frac{4M}{\delta})$  given by Theorem 4.3, then we can estimate the difference between the maximum and the minimum in  $[x_2, x_1]$  (they are either a critical point or a boundary point). Therefore, we know that

$$\text{osc}_{(x_2, x_1)} \rho^\varepsilon < \eta,$$

by taking  $\eta = \beta$  the Claim follows.

Because  $x$  is a Lebesgue point, for all  $r$  small enough, we know that

$$(38) \quad \left| \frac{1}{2r} \int_{x-r}^{x+r} \rho_0(y) dy - \rho_0(x) \right| < \frac{d(\rho_0(x), \Sigma)}{6}.$$

We will fix  $r_x < \frac{1}{2} \delta(\frac{d(\rho_0(x), \Sigma)}{3})$  such that (38) holds.

By Proposition 4.1, we know that  $\rho^\varepsilon \rightarrow \rho_0$  weak-\*  $L^\infty$ ; therefore, there exists  $\varepsilon_x$  such that for all  $\varepsilon < \varepsilon_x$

$$\left| \frac{1}{2r_x} \int_{x-r_x}^{x+r_x} \rho^\varepsilon(y) dy - \frac{1}{2r_x} \int_{x-r_x}^{x+r_x} \rho_0(y) dy \right| < \frac{d(\rho_0(x), \Sigma)}{6}.$$

So, if  $\varepsilon < \varepsilon_x$ , there exists  $x_\varepsilon \in (x - r_x, x + r_x)$ , such that

$$|\rho^\varepsilon(x_\varepsilon) - \rho_0(x)| < \frac{d(\rho_0(x), \Sigma)}{3};$$

hence

$$d(\rho^\varepsilon(x_\varepsilon), \Sigma) > \frac{2}{3} d(\rho_0(x), \Sigma).$$

By the Claim, if  $\varepsilon$  is small enough, it follows that  $d(\rho^\varepsilon(y), \Sigma) > \frac{d(\rho_0(x), \Sigma)}{3}$  for all

$$y \in \left( x_\varepsilon - \delta \left( \frac{d(\rho_0(x), \Sigma)}{3} \right), x_\varepsilon + \delta \left( \frac{d(\rho_0(x), \Sigma)}{3} \right) \right).$$

The result follows because

$$\begin{aligned} & \left( x - \frac{1}{2} \delta \left( \frac{d(\rho_0(x), \Sigma)}{3} \right), x + \frac{1}{2} \delta \left( \frac{d(\rho_0(x), \Sigma)}{3} \right) \right) \\ & \subset \left( x_\varepsilon - \delta \left( \frac{d(\rho_0(x), \Sigma)}{3} \right), x_\varepsilon + \delta \left( \frac{d(\rho_0(x), \Sigma)}{3} \right) \right). \end{aligned} \quad \square$$

To prove Theorem 4.3, we start by looking at the behavior of  $\rho^\varepsilon$  on the set  $\{\rho^\varepsilon > h\}$  with  $h > 0$ . This case follows exactly as Lemma 5.5 in [4]; our main contribution here is to give a different proof in a simple case that makes the set  $\Sigma = cl\{W > W^{**}\}$  appear more naturally.

**LEMMA 4.5.** *Let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2(\mathbb{T})$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$ ; then for any  $h > 0$  and  $L \geq 0$  there exists  $\delta(\eta, C, h) > 0$ , independent of  $\varepsilon$ , such that for any  $\varepsilon < \varepsilon_0(\eta, C, h, L)$  and any pair of sequences  $\{x_\varepsilon\}_\varepsilon, \{y_\varepsilon\}_\varepsilon \subset \mathbb{T}$  satisfying:*

- $0 < y_\varepsilon - x_\varepsilon < \delta$ ,
- $|\rho^\varepsilon_x(x_\varepsilon)| < L$  and  $|\rho^\varepsilon_x(y_\varepsilon)| < L$ ,
- $\rho^\varepsilon(z) > h \ \forall z \in [x_\varepsilon, y_\varepsilon]$ .

*Then we have either*

$$d(\rho^\varepsilon(z), \Sigma) < \eta \quad \forall z \in [x_\varepsilon, y_\varepsilon]$$

*or*

$$|\rho^\varepsilon(x_\varepsilon) - \rho^\varepsilon(y_\varepsilon)| < \eta.$$

*Proof.* We only give a sketch of the proof, which shows why the set  $\Sigma$  appears naturally. For a complete proof see Lemma 5.5 in [4].

Since  $\rho^\varepsilon(z) > h$  for every  $z \in (x_\varepsilon, y_\varepsilon)$  and  $\varepsilon$ , then

$$(e^\varepsilon(\rho^\varepsilon))_x = (W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x$$

is bounded in  $L^2(x_\varepsilon, y_\varepsilon)$  uniformly in  $\varepsilon$  (see (36), (37), and Remark 2.4). Moreover, we have

$$(39) \quad \int_{x_\varepsilon}^{y_\varepsilon} e^\varepsilon(\rho^\varepsilon)(z) dz = \int_{x_\varepsilon}^{y_\varepsilon} W'(\rho^\varepsilon(z)) - \varepsilon^2 \rho_{xx}^\varepsilon(z) dz \leq C + 2\varepsilon^2 L.$$

Sobolev's embedding theorem implies that  $e^\varepsilon(\rho^\varepsilon)$  is also uniformly in bounded  $C^{\frac{1}{2}}$ .

Without loss of generality, we will assume that  $\rho^\varepsilon(x_\varepsilon) \leq \rho^\varepsilon(y_\varepsilon)$ , and that  $\rho_x^\varepsilon(x_\varepsilon) \geq 0$ ,  $\rho_x^\varepsilon(y_\varepsilon) \leq 0$ , if not we work with the closest minimum to  $x_\varepsilon$  and the closest maximum to  $y_\varepsilon$ , inside the interval. We will also assume  $\rho_{xx}^\varepsilon(x_\varepsilon) \geq 0$  and  $\rho_{xx}^\varepsilon(y_\varepsilon) \leq 0$ ; if this condition is not satisfied, we can take

$$\tilde{x}_\varepsilon = \inf\{z : z \in (x_\varepsilon, y_\varepsilon) \cap \rho_{xx}^\varepsilon(z) < 0 \cap \rho_x^\varepsilon \geq 0\}.$$

Then, we obtain

$$|\rho^\varepsilon(x_\varepsilon) - \rho^\varepsilon(\tilde{x}_\varepsilon)| \leq L\delta.$$

If  $\rho^\varepsilon(\tilde{x}_\varepsilon)_x > 0$ , then  $\rho^\varepsilon(\tilde{x}_\varepsilon)_{xx} > 0$ . If  $\rho^\varepsilon(\tilde{x}_\varepsilon)_x = 0$  and  $\rho^\varepsilon(\tilde{x}_\varepsilon)_{xx} < 0$ , then  $\rho^\varepsilon(\tilde{x}_\varepsilon)$  is a maximum. If this happens, we consider  $\tilde{z}_\varepsilon$  the closest minimum to  $\tilde{x}_\varepsilon$ , so that we get  $\rho^\varepsilon(\tilde{z}_\varepsilon)_{xx} \geq 0$ . We split the interval in three,  $(x_\varepsilon, \tilde{x}_\varepsilon)$ ,  $(\tilde{x}_\varepsilon, \tilde{z}_\varepsilon)$ , and  $(\tilde{z}_\varepsilon, y_\varepsilon)$ , and we control each of the pieces separately.

If  $\rho_{xx}^\varepsilon(y_\varepsilon) > 0$ , we can repeat the same arguments.

Multiplying  $e^\varepsilon(\rho^\varepsilon)$  by  $\rho_x^\varepsilon(z)$  and integrating between  $x_\varepsilon$  and  $y_\varepsilon$  we get the following:

$$\begin{aligned} \int_{x_\varepsilon}^{y_\varepsilon} e^\varepsilon(\rho^\varepsilon) \rho_x^\varepsilon dz &= \frac{\varepsilon^2}{2} (|\rho_x^\varepsilon(y_\varepsilon)|^2 - |\rho_x^\varepsilon(x_\varepsilon)|^2) + W(\rho^\varepsilon(y_\varepsilon)) - W(\rho^\varepsilon(x_\varepsilon)) \\ &\leq \frac{\varepsilon^2}{2} L^2 + W(\rho^\varepsilon(y_\varepsilon)) - W(\rho^\varepsilon(x_\varepsilon)). \end{aligned}$$

On the other hand, integrating by parts we also find

$$\int_{x_\varepsilon}^{y_\varepsilon} e^\varepsilon(\rho^\varepsilon) \rho_x^\varepsilon dz = - \int_{x_\varepsilon}^{y_\varepsilon} e_x^\varepsilon(\rho^\varepsilon) \rho^\varepsilon dz + [e^\varepsilon(\rho^\varepsilon) \rho^\varepsilon]_{x_\varepsilon}^{y_\varepsilon}.$$

Because  $e_x^\varepsilon$  is uniformly in  $L^2$  and  $\rho^\varepsilon$  uniformly in  $L^\infty$ , we have

$$\left| \int_{x_\varepsilon}^{y_\varepsilon} e_x^\varepsilon(\rho^\varepsilon) \rho^\varepsilon dz \right| \leq C(y_\varepsilon - x_\varepsilon)^{\frac{1}{2}} \leq C\delta^{\frac{1}{2}}.$$

We decompose

$$(40) \quad [e^\varepsilon(\rho^\varepsilon) \rho^\varepsilon]_{x_\varepsilon}^{y_\varepsilon} = e^\varepsilon(\rho^\varepsilon)(x_\varepsilon) [\rho^\varepsilon(y_\varepsilon) - \rho^\varepsilon(x_\varepsilon)] + [e^\varepsilon(\rho^\varepsilon)]_{x_\varepsilon}^{y_\varepsilon} \rho^\varepsilon(y_\varepsilon);$$

using that  $e^\varepsilon$  is uniformly in  $C^{\frac{1}{2}}$ , we see that

$$|[e^\varepsilon(\rho^\varepsilon)]_{x_\varepsilon}^{y_\varepsilon} \rho^\varepsilon| \leq C(y_\varepsilon - x_\varepsilon)^{\frac{1}{2}} \leq C\delta^{\frac{1}{2}}.$$

Combining the five equations above, we see that given any  $\lambda > 0$ , we can choose  $\varepsilon$  and  $\delta$  small enough, such that

$$W(\rho^\varepsilon(x_\varepsilon)) + e^\varepsilon(\rho^\varepsilon)(x_\varepsilon)(\rho^\varepsilon(y_\varepsilon) - \rho^\varepsilon(x_\varepsilon)) + \lambda \geq W(\rho^\varepsilon(y_\varepsilon)).$$

Using the assumption  $\rho^\varepsilon(y_\varepsilon) \geq \rho^\varepsilon(x_\varepsilon)$ , the definition of  $e^\varepsilon(\rho^\varepsilon)$  and the fact that  $\rho_{xx}^\varepsilon(x_\varepsilon) \geq 0$ , we obtain

$$(41) \quad W(\rho^\varepsilon(x_\varepsilon)) + W'(\rho^\varepsilon(x_\varepsilon))(\rho^\varepsilon(y_\varepsilon) - \rho^\varepsilon(x_\varepsilon)) + \lambda \geq W(\rho^\varepsilon(y_\varepsilon)).$$

Exchanging the roles of  $x_\varepsilon$  and  $y_\varepsilon$  in (40) and using the fact that  $\rho_{xx}^\varepsilon(y_\varepsilon) \leq 0$ , we obtain similarly

$$(42) \quad W(\rho^\varepsilon(y_\varepsilon)) - W'(\rho^\varepsilon(y_\varepsilon))(\rho^\varepsilon(y_\varepsilon) - \rho^\varepsilon(x_\varepsilon)) + \lambda \geq W(\rho^\varepsilon(x_\varepsilon)).$$

To use these conditions analytically, we define the sets

$$U_\lambda^W(A) = \{B \in \mathbb{R}_+ : W(A) + W'(A)(B - A) + \lambda \geq W(B)\},$$

so conditions (41) and (42) can be reformulated as

$$(43) \quad \rho^\varepsilon(y_\varepsilon) \in U_\lambda^W(\rho^\varepsilon(x_\varepsilon)) \text{ and } \rho^\varepsilon(x_\varepsilon) \in U_\lambda^W(\rho^\varepsilon(y_\varepsilon)).$$

We now finish the proof under the extra assumption that  $\Sigma \cap (h, \infty)$  contains only one interval.

By (H4), we know that for any fixed  $\eta > 0$ ,  $W$  is uniformly convex in  $\{p \in \mathbb{R}^+ : d(p, \Sigma) \geq \eta\}$ ; therefore we can choose  $\lambda_0$  such that for all  $\lambda < \lambda_0$ , we have

$$(44) \quad U_\lambda^W(A) \subset (A - \eta, A + \eta) \quad \forall A \in \{p \in \mathbb{R}^+ : p > h \cap d(p, \Sigma) \geq \eta\}.$$

If  $d(\rho^\varepsilon(x_\varepsilon), \Sigma) > \eta$ , then, with  $A = \rho^\varepsilon(x_\varepsilon)$ , (43) and (44) imply

$$|\rho^\varepsilon(x_\varepsilon) - \rho^\varepsilon(y_\varepsilon)| < \eta.$$

The same holds for  $d(\rho^\varepsilon(y_\varepsilon), \Sigma) > \eta$ .

On the other hand, if  $d(\rho^\varepsilon(x_\varepsilon), \Sigma) < \eta$  and  $d(\rho^\varepsilon(y_\varepsilon), \Sigma) < \eta$ , we take

$$z = \operatorname{argmax}_{t \in [x_\varepsilon, y_\varepsilon]} d(\rho^\varepsilon(t), \Sigma);$$

if  $d(\rho^\varepsilon(z), \Sigma) < \eta$ , we are done. If  $d(\rho^\varepsilon(z), \Sigma) > \eta$ , because we assume that  $\Sigma \cap (h, \infty)$  is an interval, we know that  $\rho_x^\varepsilon(z) = 0$ . Therefore, the intervals  $(x_\varepsilon, z)$  and  $(z, y_\varepsilon)$  satisfy the hypotheses of the Lemma. Then, arguing as before, because  $d(\rho^\varepsilon(z), \Sigma) > \eta$ , we have

$$|\rho^\varepsilon(x_\varepsilon) - \rho^\varepsilon(z)| < \eta \quad \text{and} \quad |\rho^\varepsilon(y_\varepsilon) - \rho^\varepsilon(z)| < \eta.$$

By our definition of  $z$ , we can conclude

$$|\rho^\varepsilon(x_\varepsilon) - \rho^\varepsilon(y_\varepsilon)| < \eta,$$

which proves the Lemma with the extra assumption of  $\Sigma \cap (h, \infty)$  contains one interval.

A more convoluted argument than the one in the proof of Theorem 4.3 (below) can be made for the cases when  $\Sigma \cap (h, \infty)$  contains more than one interval. It is not included here, as a proof of this Lemma can already be found in [4] and the ideas of the argument can be found in the proof below.  $\square$

We now turn to the proof of Theorem 4.3.

*Proof of Theorem 4.3.* We prove Theorem 4.3 by contradiction. Due to Lemma 4.5, we know that the theorem would be proven if we can prove that there is no  $\eta_0 > 0$ , such that there exist a sequence of  $\varepsilon_i \rightarrow 0$  and sequences of points  $\{x_i\}$ ,  $\{y_i\}$  that satisfy for all  $i$ :

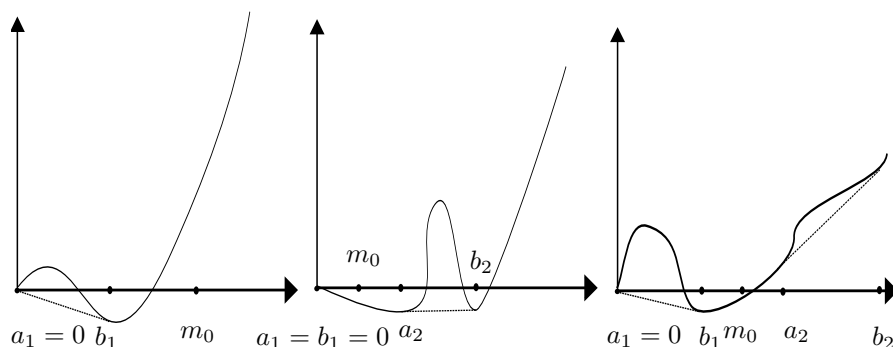


FIG. 1. From left to right: Case  $p = 1$  with  $a_1 \neq b_1$ , case  $p = 2$  with  $a_1 = b_1 = 0$ , and case  $p = 2$  with  $a_1 \neq b_1$ .

- $|y_i - x_i| < \frac{1}{i}$ ,
- $\max(|\rho_x^{\varepsilon_i}(x_i)|, |\rho_x^{\varepsilon_i}(y_i)|) < L$ ,
- $\rho^{\varepsilon_i}(x_i) \rightarrow 0$ ,
- $\rho^{\varepsilon_i}(y_i) > b_1 + \eta_0$  (see (H3)).

So, let us assume that such  $\eta_0$  exists and derive a contradiction.

Without loss of generality, we assume, as in the proof of Lemma 4.5, that  $\rho_x^{\varepsilon_i}(x_i) \leq 0$ ,  $\rho_x^{\varepsilon_i}(y_i) \geq 0$ ,  $\rho_{xx}^{\varepsilon_i}(x_i) \geq 0$ , and  $\rho_{xx}^{\varepsilon_i}(y_i) \leq 0$ .

From the proof of Proposition 4.1, we know that the function

$$G^\varepsilon(\rho^\varepsilon) = -\varepsilon^2 \rho^\varepsilon \rho_{xx}^\varepsilon + \varepsilon^2 \rho_x^{\varepsilon^2} + Q'(\rho^\varepsilon)$$

is uniformly in  $C^{\frac{1}{2}}$ . Using the fact that  $\rho_{xx}^{\varepsilon_i}(x_i) \geq 0$  and  $\rho_{xx}^{\varepsilon_i}(y_i) \leq 0$ , we can conclude that

$$Q'(\rho^{\varepsilon_i}(y_i)) \leq Q'(\rho^{\varepsilon_i}(x_i)) + C \frac{1}{i^{\frac{1}{2}}} + \frac{\varepsilon_i^2}{2} L^2.$$

Moreover, since  $\rho^\varepsilon(x_i) \rightarrow 0$ , for every  $\kappa > 0$ , there exists  $i_0$  such that

$$(45) \quad Q'(\rho^{\varepsilon_i}(x_i)) + C \frac{1}{i^{\frac{1}{2}}} + \frac{\varepsilon_i^2}{2} L^2 < \kappa$$

for all  $i > i_0$ . This implies that  $Q'(\rho^{\varepsilon_i}(y_i)) < \kappa$  for all  $i > i_0$ .

The first observation is that  $Q'(b_1) = b_1 W'(b_1) - W(b_1) = 0$ . This is just saying that the tangent of  $W$  at  $b_1$  intersects the origin, which is satisfied because  $b_1 = \inf_{t>0} \{W(t) = W^{**}(t)\}$  (for a picture see Figure 1).

The second observation is that by (H4), we have  $\int_{s_1}^{s_2} t W''(t) = \int_{s_1}^{s_2} Q''(t) > 0$  for  $s_1, s_2 \in (b_1, m_0)$ , where  $m_0$  is defined in (16). We deduce that, for every  $\eta > 0$ , there exists  $\kappa_0$ , such that if  $A \in (b_1, m_0)$  and  $Q'(A) < \kappa_0$ , then  $A - b_1 < \eta$ . Therefore, we will get a contradiction, if we show that  $\rho^{\varepsilon_i}(y_i) < m_0$ .

CLAIM I. *If  $i$  is large enough, then  $\rho^{\varepsilon_i}(y_i) < m_0$ .*

Again, we will prove Claim I by contradiction; if  $\rho^{\varepsilon_i}(y_i) \geq m_0$ , then there exists  $z_0^i \in [x_i, y_i]$  such that  $\rho^i(z_0^i) = m_0$ . If we assume also that  $\rho_{xx}^{\varepsilon_i}(z_0^i) \leq 0$ , then proceeding as above, we get

$$0 < Q'(m_0) + \frac{\varepsilon_i |\rho_x(z)|^2}{2} \leq Q'(\rho^{\varepsilon_i}(x_i)) + C \frac{1}{i^{\frac{1}{2}}} + \frac{\varepsilon_i^2}{2} L^2 \rightarrow_{i \rightarrow \infty} 0$$



and taking  $i$  large enough it yields our desired contradiction. Therefore, we want to prove that we can indeed assure that  $\rho_{xx}^{\varepsilon_i}(z_0^i) \leq 0$  for  $i$  large enough.

CLAIM II. *Let*

$$(46) \quad z_0 = \sup\{t \in (x_i, y_i) : \rho^{\varepsilon_i}(t) = m_0\}.$$

*If*

$$(47) \quad W(m_0) + W'(m_0)(\rho^{\varepsilon_i}(y_i) - m_0) < W(\rho^{\varepsilon_i}(y_i)) - C(y_i - z_0)^{\frac{1}{2}}(\rho^{\varepsilon_i}(y_i) - m_0),$$

*for some  $C$  independent of  $i$ , then for all  $i$  big enough*

$$\rho_{xx}^{\varepsilon_i}(z_0) \leq 0.$$

*Proof of Claim II.* Due to the assumption on  $z_0$ , we know that  $\rho^{\varepsilon_i} > m_0$  in  $(z_0, y_i)$ . Therefore, we know that  $e^{\varepsilon_i}(t) = W'(\rho^{\varepsilon_i}(t)) - \varepsilon_i^2 \rho_{xx}^{\varepsilon_i}(t)$  is uniformly in  $H^1(z_0, y_i)$  (see (39)). We perform the following calculation:

$$\begin{aligned} \int_{z_0}^{y_i} e^{\varepsilon_i}(t) \rho_x^{\varepsilon_i}(t) dt &= W(\rho^{\varepsilon_i}(y_i)) - W(\rho^{\varepsilon_i}(z_0)) - \frac{\varepsilon_i^2}{2} |\rho_x^{\varepsilon_i}(y_i)|^2 + \frac{\varepsilon_i^2}{2} |\rho_x^{\varepsilon_i}(z_0)|^2 \\ &\geq W(\rho^{\varepsilon_i}(y_i)) - W(m_0) - \frac{\varepsilon_i^2}{2} L^2. \end{aligned}$$

Using the same arguments used to derive (41) in the proof of Lemma 4.5, we get

$$\begin{aligned} e^{\varepsilon_i}(z_0)(\rho^{\varepsilon_i}(y_i) - \rho^{\varepsilon_i}(z_0)) + C(y_i - z_0)^{\frac{1}{2}}(\rho^{\varepsilon_i}(y_i) - \rho^{\varepsilon_i}(z_0)) \\ \geq W(\rho^{\varepsilon_i}(y_i)) - W(\rho^{\varepsilon_i}(z_0)) - \frac{\varepsilon_i^2}{2} L^2. \end{aligned}$$

If  $\rho_{xx}^{\varepsilon_i}(z_0) \geq 0$ , then  $W'(\rho^{\varepsilon_i}(z_0)) \geq e^{\varepsilon_i}(z_0)$ , and so

$$\begin{aligned} W(\rho^{\varepsilon_i}(z_0)) + W'(\rho^{\varepsilon_i}(z_0))(\rho^{\varepsilon_i}(y_i) - \rho^{\varepsilon_i}(z_0)) &\geq W(\rho^{\varepsilon_i}(y_i)) - C(y_i - z_0)^{\frac{1}{2}}(\rho^{\varepsilon_i}(y_i) - \rho^{\varepsilon_i}(z_0)) \\ &\quad - \frac{\varepsilon_i^2}{2} L^2. \end{aligned}$$

Since  $\frac{\varepsilon_i^2}{2} L^2 \rightarrow 0$ , if  $i$  is large enough this contradicts (47), and thus proves Claim II.

To finish the proof of Claim I, we have to show that if  $z_0$  defined by (46) exists (in particular, if  $\rho^{\varepsilon_i}(y_i) \geq m_0$ ), then (47) holds. First, we note that

$$W(m_0) + W'(m_0)(t - m_0) < W(t) \quad \forall t \neq m_0,$$

due to (H4). Therefore, there exists  $\kappa_0 > 0$  such that

$$W(m_0) + W'(m_0)(t - m_0) < W(t) - \kappa_0 \quad \text{for every } t \text{ s.t. } Q'(t) < \frac{Q'(m_0)}{2}$$

(the choice of  $\frac{Q'(m_0)}{2}$  is arbitrary).

Using (45)

$$\limsup_{i \rightarrow \infty} Q'(\rho^{\varepsilon_i}(y_i)) \leq 0,$$

we can take  $i$  large enough such that

$$Q'(\rho^{\varepsilon_i}(y_i)) < \frac{Q'(m_0)}{2}$$

and

$$\frac{C}{i^{\frac{1}{2}}}(\rho^{\varepsilon_i}(y_i) - \rho^{\varepsilon_i}(z_0)) < \kappa_0;$$

therefore, we can conclude that (47) holds and Claim II yields

$$\rho_{xx}^{\varepsilon_i}(z_0^i) \leq 0.$$

This completes the proof of Claim I and of the theorem.  $\square$

**5.  $H^1$  estimate in the good set  $\Omega$ .** We want to show that  $\rho^\varepsilon$  is bounded in  $H_{loc}^1(\Omega)$  uniformly in  $\varepsilon$ , with  $\Omega = \{\rho_0 \notin \Sigma\}$ ; in other words, that  $\rho^\varepsilon$  does not oscillate in the “good” limiting set. We start with the following proposition.

**PROPOSITION 5.1.** *Let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2(\mathbb{T})$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$  and  $\rho^\varepsilon \rightarrow \rho_0$  in  $\mathcal{W}^2(\mathbb{T})$ , given  $\phi \in \mathcal{D}(\Omega)$ , there exists  $\varepsilon_0 > 0$  such that for every  $\varepsilon < \varepsilon_0$ , we have  $W''(\rho^\varepsilon) \geq \lambda_\phi$  and  $\rho^\varepsilon \geq \lambda_\phi$  in the support of  $\phi$  for some constant  $\lambda_\phi > 0$  independent of  $\varepsilon$ .*

*Proof.* By assumption, if  $x \in \Omega$ , then  $d(\rho_0(x), \Sigma(W)) > 0$ . By Corollary 4.4, for any Lebesgue point  $x \in \Omega$  there exists  $\varepsilon_x$  and  $\delta_x$  such that for every  $\varepsilon < \varepsilon_x$  and every  $z \in (x - \delta_x, x + \delta_x)$

$$d(\rho^\varepsilon(z), \Sigma) > \frac{d(\rho_0(x), \Sigma)}{3}.$$

Now, the family of intervals  $\{(x - \delta_x, x + \delta_x)\}_{x \in \hat{\Omega}}$ , where  $\hat{\Omega}$  is the Lebesgue points of  $\Omega$ , is an open covering of the support of  $\phi$ , therefore by compactness there exists a finite subcovering, which proves the proposition.  $\square$

Using Proposition 5.1 we can now prove the following.

**PROPOSITION 5.2.** *Let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2(\mathbb{T})$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$  and  $\rho^\varepsilon \rightarrow \rho_0$  in  $\mathcal{W}^2(\mathbb{T})$ , for any  $K \subset \Omega$  compact, there exists  $C$  and  $\varepsilon_K > 0$  such that*

$$\int_K |\rho_x^\varepsilon|^2 dx < C \quad \forall \varepsilon < \varepsilon_K.$$

Therefore, up to a subsequence,  $\rho^\varepsilon$  in  $C_{loc}^0(\Omega)$  to  $\rho_0$ .

*Proof.* Take  $\phi \in \mathcal{D}(\Omega)$  with  $\phi \geq \chi_K$ . We start with the following computation:

$$\int_{\mathbb{T}} \rho_x^\varepsilon \partial_x [W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon] \phi dx = \int_{\mathbb{T}} W''(\rho^\varepsilon) |\rho_x^\varepsilon|^2 \phi dx + \varepsilon^2 \int_{\mathbb{T}} |\rho_{xx}^\varepsilon|^2 \phi dx + \varepsilon^2 \int_{\mathbb{T}} \rho_x^\varepsilon \rho_{xx}^\varepsilon \phi_x dx,$$

from which we deduce

$$\begin{aligned} \int_{\mathbb{T}} W''(\rho^\varepsilon) |\rho_x^\varepsilon|^2 \phi + \varepsilon^2 \int_{\mathbb{T}} |\rho_{xx}^\varepsilon|^2 \phi dx &= -\varepsilon^2 \int_{\mathbb{T}} \rho_x^\varepsilon \rho_{xx}^\varepsilon \phi_x dx + \int_{\mathbb{T}} \rho_x^\varepsilon \partial_x [W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon] \phi dx \\ &= \frac{\varepsilon^2}{2} \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi_{xx} dx + \int_{\mathbb{T}} \rho_x^\varepsilon \partial_x [W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon] \phi dx \\ &\leq C(\phi_{xx}) \varepsilon^2 \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 dx + \left( \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi dx \right)^{\frac{1}{2}} \left( \int_{\mathbb{T}} \phi |\partial_x [W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon]|^2 dx \right)^{\frac{1}{2}} \\ &\leq C \varepsilon^2 \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 dx + \frac{\lambda_\phi}{2} \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi dx + C(\lambda_\phi) \int_{\mathbb{T}} \phi |\partial_x [W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon]|^2 dx \end{aligned}$$

with the constant  $\lambda_\phi$  given by Proposition 5.1. Therefore, we get

$$\left(\inf_{x \in \text{supp}\{\phi\}} W''(\rho^\varepsilon(x)) - \frac{\lambda_\phi}{2}\right) \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi \, dx \leq C\varepsilon^2 \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \, dx \\ + C(\lambda_\phi) \int_{\text{supp}\{\phi\}} |\partial_x [W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon]|^2 \, dx.$$

Using Proposition 5.1 we can conclude that in the support of  $\phi$  we have that  $W''(\rho^\varepsilon) > \lambda_\phi$  and that  $\rho^\varepsilon > \lambda_\phi$ , for  $\varepsilon < \varepsilon_0$ , so we deduce

$$\int_K |\rho_x^\varepsilon|^2 \, dx \leq \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi \, dx \leq C(\phi) E^\varepsilon[\rho^\varepsilon] + \frac{C(\lambda_\phi)}{\lambda_\phi} \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C \quad \forall \varepsilon < \varepsilon_0. \quad \square$$

**PROPOSITION 5.3.** *Let  $\{\rho^\varepsilon\}_{\varepsilon>0}$  be a sequence of functions in  $\mathcal{P}(\mathbb{T}) \cap H^2$  such that  $\sup_\varepsilon E^\varepsilon[\rho^\varepsilon] + \mathcal{G}^\varepsilon(\rho^\varepsilon) \leq C$  and  $\rho^\varepsilon \rightarrow \rho_0$  in  $\mathcal{W}^2(\mathbb{T})$ ; then*

$$(W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x \rho^\varepsilon \rightarrow Q'(\rho_0)_x \quad \text{in } D'(\Omega).$$

*Proof.* Fix  $\phi \in \mathcal{D}(\Omega)$ ; then using that  $\rho W''(\rho) = Q''(\rho)$  with an integration by parts, we get

$$(48) \quad \int_{\mathbb{T}} (W'(\rho^\varepsilon)_x - \varepsilon^2 \rho_{xxx}^\varepsilon) \rho^\varepsilon \phi \, dx = - \int_{\mathbb{T}} Q'(\rho^\varepsilon) \phi_x \, dx + \varepsilon^2 \int_{\mathbb{T}} \rho_{xx}^\varepsilon \rho_x^\varepsilon \phi \, dx + \varepsilon^2 \int_{\mathbb{T}} \rho_{xx}^\varepsilon \rho^\varepsilon \phi_x \, dx.$$

The first term converges to what we are looking for,

$$\lim_{\varepsilon \rightarrow 0} - \int_{\mathbb{T}} Q'(\rho^\varepsilon) \phi_x \, dx = - \int_{\mathbb{T}} Q'(\rho_0) \phi_x \, dx,$$

by Lebesgue's dominated convergence and Proposition 5.2.

It remains to show that the last two terms in (48) go to zero. Integrating by parts again, we get

$$\varepsilon^2 \int_{\mathbb{T}} \rho_{xx}^\varepsilon \rho_x^\varepsilon \phi \, dx + \varepsilon^2 \int_{\mathbb{T}} \rho_{xx}^\varepsilon \rho^\varepsilon \phi_x \, dx = - \frac{3}{2} \varepsilon^2 \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi_x \, dx - \varepsilon^2 \int_{\mathbb{T}} \rho_x^\varepsilon \rho^\varepsilon \phi_{xx} \, dx.$$

The first term goes to zero, by applying Proposition 5.2. The second term can be rewritten as

$$\frac{1}{2} \varepsilon^2 \int_{\mathbb{T}} |\rho_x^\varepsilon|^2 \phi_{xxx} \, dx,$$

which goes to zero, because  $\rho^\varepsilon$  is in  $L^\infty$  uniformly.  $\square$

## 6. Proof of Theorem 3.2.

*Proof.* To begin with, we assume that  $\liminf \mathcal{G}^\varepsilon(\rho^\varepsilon) < \infty$ ; otherwise, there is nothing to prove. Therefore, up to relabeling, we can consider  $\rho^\varepsilon$  such that

$$\sup_\varepsilon E^\varepsilon(\rho^\varepsilon) + \mathcal{G}^\varepsilon(\rho^\varepsilon) < \infty.$$

By Proposition 4.1, we know that, up to subsequence,  $\rho^\varepsilon \rightharpoonup \rho_0$  weak- $*$   $L^\infty$ ; we define  $\Omega = \{\rho_0 \notin \Sigma\}$ . We start with the following bound: using Proposition 5.1, for any  $K \subset \Omega$  compact, we have, for all  $\varepsilon$  small enough,

$$\mathcal{G}^\varepsilon(\rho^\varepsilon)^2 \geq \int_{\{\rho^\varepsilon > 0\}} |(W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x|^2 \rho^\varepsilon \, dx \geq \int_K |(W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x|^2 \rho^\varepsilon \, dx.$$

Furthermore, we have

$$\int_K |(W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x|^2 \rho^\varepsilon dx \geq 2 \int_{\mathbb{T}} (W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x \phi \rho^\varepsilon dx - \int_{\mathbb{T}} \phi^2 \rho^\varepsilon dx \quad \forall \phi \in \mathcal{D}(K),$$

which implies

$$\liminf_{\varepsilon \rightarrow 0} \mathcal{G}^\varepsilon(\rho^\varepsilon)^2 \geq \liminf_{\varepsilon \rightarrow 0} \left[ 2 \int_{\mathbb{T}} (W'(\rho^\varepsilon) - \varepsilon^2 \rho_{xx}^\varepsilon)_x \phi \rho^\varepsilon dx - \int_{\mathbb{T}} \phi^2 \rho^\varepsilon dx \right] \quad \forall \phi \in \mathcal{D}(K).$$

By Proposition 5.3, we deduce

$$\liminf_{\varepsilon \rightarrow 0} \mathcal{G}^\varepsilon(\rho^\varepsilon)^2 \geq \sup_{\phi \in \mathcal{D}(K)} -2 \int_{\mathbb{T}} Q'(\rho_0) \phi_x dx - \int_{\mathbb{T}} \phi^2 \rho_0 dx.$$

By Proposition 5.2 and the lower semicontinuity of the  $H^1$  seminorm, we also know that  $\rho_0$  is in  $H_{loc}^1(\Omega)$ , so we can integrate by parts:

$$\begin{aligned} \lim_{\varepsilon} \mathcal{G}^\varepsilon(\rho^\varepsilon)^2 &\geq \sup_{\phi \in \mathcal{D}(K)} 2 \int_{\mathbb{T}} Q'(\rho_0)_x \phi dx - \int_{\mathbb{T}} \phi^2 d\rho_0 dx \\ &= \sup_{\phi \in \mathcal{D}(K)} 2 \int_{\mathbb{T}} W'(\rho_0)_x \rho_0 \phi dx - \int_{\mathbb{T}} \phi^2 d\rho_0 dx \\ &= \|W'(\rho_0)_x\|_{L_{\rho_0}^2(K)}^2. \end{aligned}$$

Taking  $K \rightarrow \Omega$  we obtain

$$\lim_{\varepsilon} |\mathcal{G}^\varepsilon|^2(\rho^\varepsilon) \geq \|W'(\rho_0)_x\|_{L_{\rho_0}^2(\Omega)}^2.$$

The rest of the proof is devoted to proving

$$\|W'(\rho_0)_x\|_{L_{\rho_0}^2(\Omega)}^2 = |\nabla E^{**}|(\rho_0).$$

First, to have  $\|W^{**'}(\rho_0)_x\|_{L_{\rho_0}^2(\mathbb{T})}^2 = |\nabla E^{**}|(\rho_0)$ , we need to prove that  $Q^{**'} \in W^{1,1}$  (see (24)). We prove this by proving that

$$Q^{**'}(\rho_0)_x = Q^{**'}(\rho_0)_x \mathbb{1}_\Omega \quad \text{in } \mathcal{D}'(\mathbb{T}).$$

Since  $\rho^\varepsilon$  is continuous, if  $\rho^\varepsilon(x) \in \Sigma_i$  and  $\rho^\varepsilon(y) \in \Sigma_j$ , there exists  $z \in (x, y)$  such that  $d(\rho^\varepsilon(z), \Sigma) \geq \inf_{i \neq j} \frac{d(\Sigma_i, \Sigma_j)}{2}$ . By Corollary 4.4, we know that  $d(\rho^\varepsilon, \Sigma)$  is uniformly lower semicontinuous. Therefore there exists  $\delta_0$ , independent of  $\varepsilon$ , such that  $d(\rho^\varepsilon(t), \Sigma) > 0$ , for any  $t \in (z_0 - \delta_0, z_0 + \delta_0)$ ; then  $|x - y| > 2\delta_0$ . This implies that the sets  $C_i = \{\rho_0 \in \Sigma_i\}$  are at a nonzero distance from each other.

We define, as an auxiliary function,  $w$  in  $\Sigma$  by

$$w(x) = Q^{**'}(\Sigma_i) \quad \text{if } x \in C_i,$$

and we extend it to the whole of  $\mathbb{T}$  by linear interpolation. Since the sets  $C_i$  are separated, the function  $w$  is Lipschitz. Moreover,  $Q^{**'}(\rho_0) = w$  in  $\Omega^c$ ; then for every  $\phi \in \mathcal{D}(\mathbb{T})$ ,

$$\int_{\mathbb{T}} (Q^{**'}(\rho_0) - w) \phi_x dx = \int_{\Omega} (Q^{**'}(\rho_0) - w) \phi_x dx.$$

Integrating by parts we have no boundary term, and so

$$\int_{\mathbb{T}} (Q^{**'}(\rho_0) - w) \phi_x dx = - \int_{\Omega} (Q^{**'}(\rho_0) - w)_x \phi dx.$$

Because  $w$  is Lipschitz and  $w_x = 0$  in  $\Sigma$ , then

$$\int_{\Omega} w_x \phi \, dx = \int_{\mathbb{T}} w_x \phi \, dx = - \int_{\mathbb{T}} w \phi_x \, dx.$$

Therefore, we obtain that

$$\int_{\mathbb{T}} Q^{**'}(\rho_0) \phi_x \, dx = - \int_{\Omega} Q^{**'}(\rho_0)_x \phi \, dx$$

for every  $\phi \in \mathcal{D}(\mathbb{T})$ .

Similarly, we can prove that  $W^{**'}(\rho_0)_x = W'(\rho_0)_x \mathbb{1}_{\Omega}$ ; so we obtain the desired equality:

$$\|W'(\rho_0)_x\|_{L^2_{\rho_0}(\Omega)}^2 = |\nabla E^{**}|(\rho_0) = \|W^{**'}(\rho_0)_x\|_{L^2_{\rho_0}(\mathbb{T})}^2. \quad \square$$

## 7. Proof of Theorem 3.1.

*Proof.* To be able to apply the framework developed by Sandier and Serfaty [24], we have to prove the compactness of the family  $\nu^\varepsilon$  with respect to time.

Because the diameter of  $\mathbb{T}$  is finite, we know that the diameter of  $\mathcal{W}^2(\mathbb{T})$  is also finite; then

$$\nu^\varepsilon \in L^\infty([0, T]; \mathcal{W}^2(\mathbb{T})).$$

By the energy inequality (26), we know that

$$\int_0^T |\nu^{\varepsilon'}(t)|^2 \, dt$$

is uniformly bounded and therefore we know that

$$\nu^\varepsilon \text{ is uniformly bounded in } H^1((0, T); \mathcal{W}^2(\mathbb{T})).$$

By [16], we deduce that  $\nu^\varepsilon$  is precompact in  $L^2([0, T]; \mathcal{W}^2(\mathbb{T}))$ , so up to a subsequence

$$\nu^\varepsilon \rightarrow \mu \quad \text{in } L^2([0, T]; \mathcal{W}^2(\mathbb{T})).$$

Also,

$$\int_0^T |\nu^{\varepsilon'}(t)|^2 \, dt = \sup_{h \in (0, T)} \int_0^{T-h} \frac{d_2(\nu^\varepsilon(t), \nu^\varepsilon(t+h))}{h} \, dt$$

is lower semicontinuous with respect to the convergence in  $L^2([0, T]; \mathcal{W}^2(\mathbb{T}))$ , hence

$$(49) \quad \liminf_{\varepsilon \rightarrow 0} \int_0^T |\nu^{\varepsilon'}(s)|^2 \, ds \geq \int_0^T |\mu'(s)|^2 \, ds.$$

Furthermore, by Arselà–Ascoli, we also know that up to a further subsequence,

$$\nu^\varepsilon \rightarrow \mu \quad \text{in } C^0([0, T]; \mathcal{W}^2(\mathbb{T})).$$

In particular,

$$\nu_i^\varepsilon \rightarrow \nu_i = \mu(0).$$

Now, we only have to follow the proof in [24] and obtain that  $\mu$  is the gradient flow of  $E^{**}$  with initial condition  $\nu_i$ .

By (26), we know that

$$E^\varepsilon[\nu_i^\varepsilon] - E^\varepsilon[\nu^\varepsilon(t)] \geq \frac{1}{2} \int_0^t \mathcal{G}^\varepsilon(\nu^\varepsilon)^2 ds + \frac{1}{2} \int_0^t |\nu^{\varepsilon'}|^2 ds.$$

Taking the limit  $\varepsilon \rightarrow 0$ , using Fatou's Lemma, Theorem 3.2, and (49), we get that

$$(50) \quad \liminf_{\varepsilon \rightarrow 0^+} (E^\varepsilon[\nu_i^\varepsilon] - E^\varepsilon[\nu^\varepsilon(t)]) \geq \frac{1}{2} \int_0^t |\nabla E^{**}(\mu)|^2 ds + \frac{1}{2} \int_0^t |\mu'|^2 ds.$$

By Young's inequality, we know that

$$(51) \quad \frac{1}{2} \int_0^t |\nabla E^{**}(\mu)|^2 ds + \frac{1}{2} \int_0^t |\mu'|^2 ds \geq \int_0^t |\nabla E^{**}(\mu)| |\mu'| ds.$$

Because  $E^{**}$  is convex with respect to the geodesics in  $\mathcal{W}^2(\mathbb{T})$ , we can apply Theorem A.6 to obtain

$$(52) \quad \int_0^t |\nabla E^{**}(\mu)| |\mu'| ds \geq E^{**}[\mu(0)] - E^{**}[\mu(t)].$$

Since  $\lim_\varepsilon E^\varepsilon[\nu_i^\varepsilon] = E^{**}[\nu_i] = E^{**}[\mu(0)]$  by the well preparedness assumption, (50), (51), and (52) imply

$$\limsup E^\varepsilon[\nu^\varepsilon(t)] \leq E^{**}[\mu(t)].$$

The reverse inequality comes from the  $\Gamma$ -convergence of  $E^\varepsilon$  to  $E^{**}$ , so we have proven that

$$\lim E^\varepsilon[\nu^\varepsilon(t)] = E^{**}[\mu(t)],$$

and the inequalities (50), (51), and (52) are in fact equalities. In particular, (50) yields

$$E^{**}[\nu_i] - E^{**}[\mu(t)] = \frac{1}{2} \int_0^t |\nabla E^{**}(\mu)|^2 ds + \frac{1}{2} \int_0^t |\mu'|^2 ds$$

for all  $t > 0$ . Finally, by Theorem A.19, we deduce that  $\nu_0 = \mu$ .  $\square$

## Appendix A. Gradient flows in $\mathcal{W}^2(\mathbb{T})$ .

**A.1. General metric spaces.** We briefly review some important definitions and theorems from *Gradient Flows: In Metric Spaces and in the Space of Probability Measures* by Ambrosio, Gigli, and Savare [2]. We start with some notions defined for a general complete metric space  $(X, d)$ , which we later analyze in the  $\mathcal{W}^2(\mathbb{T}) = (\mathcal{P}(\mathbb{T}), d_2)$  case. We begin with the notion of an absolutely continuous curve.

**DEFINITION A.1.** Let  $v : (0, 1) \rightarrow X$  be a curve. We say that  $v \in AC^p(a, b; X)$  with  $p \in [1, \infty)$ , if there exists  $m \in L^p(a, b)$  such that

$$(53) \quad d(v(s), v(t)) \leq \int_s^t m(r) dr \quad \forall a < s < t < b.$$

If  $p = 1$ , we suppress the superscript and just denote it by  $AC$ .

Absolute continuity is enough to define the size of a derivative at almost every point; this is the subject of the next theorem.

THEOREM A.2. Let  $v \in AC^p(a, b; X)$ ; then the limit

$$|v'(t)| := \lim_{h \rightarrow 0} \frac{d(v(t+h), v(t))}{h}$$

exists a.e. in  $(a, b)$  and  $|v'| \in L^p(a, b)$ . Moreover, it is minimal in the sense that it holds (53), and if

$$d(v(s), v(t)) \leq \int_s^t m(r) \, dr \quad \forall a < s < t < b,$$

then  $|v'(t)| \leq m(t)$  a.e. in  $(a, b)$ .

Now that we have the concept of the size of the derivative of a curve, we can give a notion the size of gradients for functionals defined in  $X$ . From now on,  $\phi$  is a lower semicontinuous real-valued function on  $X$ .

DEFINITION A.3. A function  $g : X \rightarrow [0, +\infty]$  is a strong upper gradient for  $\phi$  if for any  $v \in AC(a, b; X)$ , the function  $g \circ v$  is borel and

$$|\phi(v(t)) - \phi(v(s))| \leq \int_s^t g \circ v(r) |v'(r)| \, dr \quad \forall a < s < t < b.$$

In particular, if  $g \circ v(r) |v'(r)| \in L^1(a, b)$ , then  $\phi \circ v$  is absolutely continuous and

$$|(\phi \circ v)'(t)| \leq g \circ v(r) |v'(r)| \quad \text{a.e. in } (a, b).$$

The most natural candidate to satisfy the definition above is the slope of  $\phi$ .

DEFINITION A.4. The slope of  $\phi$  at  $v$  is defined by

$$|\partial\phi(v)| := \limsup_{w \rightarrow v} \frac{(\phi(w) - \phi(v))^+}{d(w, v)}.$$

To be able to relate the two definitions, we need to consider a more restrictive set of functionals, for instance  $\lambda$ -convex functionals.

DEFINITION A.5. Given  $\lambda \in \mathbb{R}$ , we say that  $\phi$  is  $\lambda$ -convex with respect to the geodesics, if for every  $\gamma_t : [0, 1] \rightarrow X$  constant speed geodesic, we have that

$$\phi(\gamma_t) \leq (1-t)\phi(\gamma_0) + t\phi(\gamma_1) - \frac{1}{2}\lambda t(1-t)(d(\gamma_0, \gamma_1))^2.$$

With this definition we can write the following theorem.

THEOREM A.6. Suppose that  $\phi$  is  $\lambda$  convex with respect to the geodesics; then  $|\partial\phi|$  is a strong upper gradient.

*Proof.* See Corollary 2.4.10 in [2]. □

Now, we are ready to define the curves of maximal slope for  $\lambda$ -convex functionals.

DEFINITION A.7. We say that the locally absolutely continuous map  $u : (a, b) \rightarrow X$  is a curve of maximal slope of  $\phi$  with respect to its upper gradient  $|\partial\phi|$  if

$$(54) \quad \phi(u(t)) - \phi(u(s)) \geq \int_s^t \frac{|u'(r)|^2}{2} + \frac{|\partial\phi(u(r))|^2}{2} \, dr.$$

Remark A.8. If  $(X, d)$  is a Hilbert space, and  $\phi$  is  $\lambda$ -convex, then  $|\partial\phi(v)|$  is actually the norm of the minimal selection in the subdifferential at  $v$ . Moreover,  $u(\cdot)$  is a curve of maximal slope, if and only if,  $u(\cdot)$  is a gradient flow. This follows from an application of the Cauchy–Schwarz and Young inequality.

**A.2. Wasserstein metric.** We start with some auxiliary definitions to be able to define the  $L^2$ -Wasserstein distance,  $d_2$ .

DEFINITION A.9. Given  $\mu, \nu \in \mathcal{P}(\mathbb{T})$ , we call  $\pi \in \mathcal{P}(\mathbb{T} \times \mathbb{T})$  a transference plan if

$$\pi(A \times \mathbb{T}) = \mu(A) \quad \text{and} \quad \pi(\mathbb{T} \times A) = \nu(A)$$

for every Borel set  $A$ .

We denote the set of transference plans from  $\mu$  to  $\nu$  as  $\Pi(\mu, \nu)$ .

Remark A.10.  $\mu \times \nu \in \Pi(\mu, \nu)$ .

DEFINITION A.11. Given  $\mu, \nu \in \mathcal{P}(\mathbb{T})$ , we define their  $L^2$ -Wasserstein distance as

$$d_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left\{ \int_{\mathbb{T} \times \mathbb{T}} |x - y|^2 d\pi(x, y) \right\}.$$

This distance has been extensively studied in the literature, and we recommend [27] and [1], which also contain a pedagogical introduction to the gradient flow theory. In this work, we are mostly interested on its differential structure.

THEOREM A.12. Let the curve  $\mu_t : I \rightarrow \mathcal{P}(\mathbb{T})$  be absolutely continuous with respect to  $d_2$  and let  $|\mu'| \in L^1(I)$  be its metric derivative, then there exists a Borel vector field  $v$  such that

$$\|v(\cdot, t)\|_{L_{\mu_t}^2} \leq |\mu'(t)| \quad \text{a.e. } t \in I$$

and the continuity equation

$$(55) \quad \partial_t \mu_t + \nabla \cdot (v(\cdot, t) \mu_t) = 0$$

is solved in the sense of distributions.

Conversely, if  $\mu_t : I \rightarrow \mathcal{P}(\mathbb{T})$  is continuous with respect to  $d_2$  and satisfies the continuity equation (55) for some Borel velocity field  $v$  with  $\|v(\cdot, t)\|_{L_{\mu_t}^2} \in L^1(I)$ , then  $\mu_t$  is absolutely continuous and  $|\mu'(t)| \leq \|v(\cdot, t)\|_{L_{\mu_t}^2}$  a.e.  $t \in I$ .

*Proof.* See Theorem 8.3.1 in [2].  $\square$

Remark A.13. We are missing an extra condition to uniquely determine the vector field  $v$ , as we could always perturb it by a field  $w$  such that  $\nabla \cdot (w \mu_t) = 0$ , without changing the continuity equation (55).

DEFINITION A.14. Let  $\mu \in \mathcal{P}(\mathbb{T})$ . We define

$$\text{Tan}_\mu \mathcal{P}(\mathbb{T}) = \text{cl}(\{\nabla \phi : \phi \in C^\infty(\mathbb{T})\}),$$

where  $\text{cl}$  denotes the closure with respect to the  $L_\mu^2$  topology.

THEOREM A.15. Let  $\mu_t : I \rightarrow \mathcal{P}(\mathbb{T})$  be an absolutely continuous curve and let  $v$  be such that the continuity equation (55) is satisfied. Then,  $|\mu'(t)| = \|v(\cdot, t)\|_{L_{\mu_t}^2}$  a.e.  $t \in I$ , if and only if,  $v \in \text{Tan}_{\mu_t} \mathcal{P}(\mathbb{T})$  a.e.  $t \in I$ . Moreover, the vector field  $v$  is a.e. uniquely determined by  $|\mu'(t)| = \|v(\cdot, t)\|_{L_{\mu_t}^2}$ .

*Proof.* See Theorem 8.3.1 in [2].  $\square$

Using the inner product structure in  $L_\mu^2$ , we are able to define the subdifferential of a  $\lambda$ -convex functional.



DEFINITION A.16. We say that  $\zeta \in L^2_\mu(\mathbb{T})$  is a strong subdifferential of  $\phi$  at  $\mu$ , denoted by  $\partial\phi(\mu)$ , if

$$\phi(H\#\mu) - \phi(\mu) \leq \int_{\mathbb{T}} \langle \zeta(x), H(x) - x \rangle d\mu(x) + o(\|H - I\|_{L^2_\mu(\mathbb{T})}),$$

where  $H$  is a Borel vector field and the push-forward  $H\#\mu$  is defined by the condition  $H\#\mu(A) = \mu(H^{-1}(A))$  for every Borel set  $A$ .

We would like to characterize the strong subdifferentials of functionals like  $E^\varepsilon$ ; therefore we consider

$$\mathcal{F}[\mu] = \begin{cases} \int_{\mathbb{T}} F(x, \rho(x), \nabla \rho(x)) dx & \text{if } \mu = \rho d\mathcal{L} \text{ and } \rho \in C^1(\mathbb{T}), \\ +\infty & \text{otherwise,} \end{cases}$$

where  $d\mathcal{L}$  is the Lebesgue measure in  $\mathbb{T}$ . We denote  $(x, z, p) \in \mathbb{T} \times \mathbb{R} \times \mathbb{R}$  the variables of  $F$ . To simplify, we ask that  $F \in C^2$  and that  $F(x, 0, p) = 0$  for every  $x$  and  $p$ .

LEMMA A.17. If  $\mu = \rho d\mathcal{L} \in \mathcal{P}(\mathbb{T})$ , with  $\rho \in C^1$ , satisfies  $\mathcal{F}[\mu] < \infty$ , then any strong subdifferential of  $\mathcal{F}$  at  $\mu$  is  $\mu$ -a.e. equal to

$$(56) \quad \nabla \frac{\delta \mathcal{F}}{\delta \rho} = \nabla(F_z(x, \rho(x), \nabla \rho(x)) - \nabla \cdot F_p(x, \rho(x), \nabla \rho(x))).$$

*Proof.* See Lemma 10.4.1 in [2]. □

Now we can define the notion of gradient flow for a functional  $\phi$ .

DEFINITION A.18. We say that a map  $\mu_t \in AC^2((0, \infty), \mathcal{P}(\mathbb{T}))$  is a solution to the gradient flow equation, if the vector field  $v$  from Theorem A.15 satisfies

$$v(\cdot, t) \in \partial\phi(\mu_t) \quad \forall t > 0.$$

Now, in the  $\lambda$ -convex case, we can make the connection between gradient flows and curves of maximal slope.

THEOREM A.19. If  $\phi$  is  $\lambda$ -convex, then  $\mu_t$  is a curve of maximal slope with respect to  $|\partial\phi|$ , if and only if,  $\mu_t$  is a gradient flow and  $\phi(\mu_t)$  is equal a.e. to a function of bounded variation.

Moreover, given two gradient flows  $\mu_t^1$  and  $\mu_t^2$ , such that  $\mu_t^1 \rightarrow \mu_1$  and  $\mu_t^2 \rightarrow \mu_2$  as  $t \rightarrow 0$ , then

$$d_2(\mu_t^1, \mu_t^2) \leq e^{-\lambda t} d_2(\mu_1, \mu_2).$$

In particular, there is a unique gradient flow  $\mu_t$  with initial condition  $\mu_0$  and it satisfies the maximal slope condition (54) with equality.

*Proof.* See Theorems 11.1.3 and 11.1.4 in [2]. □

**Acknowledgments.** The author thanks Professor Antoine Mellet for his infinite patience. The author also thanks the anonymous referee for suggesting a major improvement in the definition of the relaxation  $\mathcal{G}^\varepsilon$ .

## REFERENCES

- [1] L. AMBROSIO AND N. GIGLI, *A user's guide to optimal transport*, in Modelling and Optimisation of Flows on Networks, Springer, Heidelberg, 2013, pp. 1–155.
- [2] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, Birkhäuser, Basel, 2008.

- [3] S. ARNRICH, A. MIELKE, M. A. PELETIER, G. SAVARÉ, AND M. VENERONI, *Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction*, Calc. Var. Partial Differential Equations, 44 (2012), pp. 419–454.
- [4] G. BELLETTINI, L. BERTINI, M. MARIANI, AND M. NOVAGA, *Convergence of the one-dimensional Cahn–Hilliard equation*, SIAM J. Math. Anal., 44 (2012), pp. 3458–3480.
- [5] F. BERNIS AND A. FRIEDMAN, *Higher order nonlinear degenerate parabolic equations*, J. Differential Equations, 83 (1990), pp. 179–206.
- [6] A. J. BERNOFF AND C. M. TOPAZ, *Biological aggregation driven by social and environmental factors: A nonlocal model and its degenerate Cahn–Hilliard approximation*, SIAM J. Appl. Dyn. Syst., 15 (2016), pp. 1528–1562.
- [7] A. BERTOZZI AND M. PUGH, *The lubrication approximation for thin viscous films: Regularity and long-time behavior of weak solutions*, Comm. Pure Appl. Math., 49 (1996), pp. 85–123.
- [8] J. W. CAHN AND J. E. HILLIARD, *Free energy of a nonuniform system. I. Interfacial free energy*, J. Chem. Phys., 28 (1958), pp. 258–267.
- [9] J. A. CARRILLO AND D. SLEPČEV, *Example of a displacement convex functional of first order*, Calc. Var. Partial Differential Equations, 36 (2009), pp. 547–564.
- [10] M. COLOMBO AND M. GOBBINO, *Passing to the limit in maximal slope curves: From a regularized Perona–Malik equation to the total variation flow*, Math. Models Methods Appl. Sci., 22 (2012), 1250017.
- [11] K. CRAIG AND I. TOPALOGLU, *Convergence of regularized nonlocal interaction energies*, SIAM J. Math. Anal., 48 (2016), pp. 34–60.
- [12] J. DOLBEAULT, B. NAZARET, AND G. SAVARÉ, *A new class of transport distances between measures*, Calc. Var. Partial Differential Equations, 34 (2009), pp. 193–231.
- [13] E. KAMALINEJAD, *An Optimal Transport Approach to Nonlinear Evolution Equations*, Ph.D. thesis, University of Toronto, 2012.
- [14] D. KINDERLEHRER AND G. STAMPACCHIA, *An Introduction to Variational Inequalities and Their Applications*, Classics Appl. Math. 31, SIAM, Philadelphia, PA, 1980.
- [15] N. Q. LE, *A Gamma-Convergence Approach to the Cahn–Hilliard equation*, Calc. Var. Partial Differential Equations, 32 (2008), pp. 499–522.
- [16] S. LISINI, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations, 28 (2007), pp. 85–120.
- [17] S. LISINI, D. MATTHES, AND G. SAVARÉ, *Cahn–Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics*, J. Differential Equations, 253 (2012), pp. 814–850.
- [18] D. MATTHES, R. J. MCCANN, AND G. SAVARÉ, *A family of nonlinear fourth order equations of gradient flow type*, Comm. Partial Differential Equations, 34 (2009), pp. 1352–1397.
- [19] R. J. MCCANN, *A convexity principle for interacting gases*, Adv. Math., 128 (1997), pp. 153–179.
- [20] A. MELLET, *The thin film equation with non-zero contact angle: A singular perturbation approach*, Comm. Partial Differential Equations, 40 (2015), pp. 1–39.
- [21] L. MUGNAI AND M. RÖGER, *The Allen–Cahn action functional in higher dimensions*, Interfaces Free Bound., 10 (2008), pp. 45–78.
- [22] A. NOVICK-COHEN AND L. A. SEGEL, *Nonlinear aspects of the Cahn–Hilliard equation*, Phys. D, 10 (1984), pp. 277–298.
- [23] F. OTTO, *Lubrication approximation with prescribed nonzero contact angle*, Comm. Partial Differential Equations, 23 (1998), pp. 2077–2164.
- [24] E. SANDIER AND S. SERFATY, *Gamma-convergence of gradient flows with applications to Ginzburg–Landau*, Comm. Pure Appl. Math., 57 (2004), pp. 1627–1672.
- [25] S. SERFATY, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete Contin. Dyn. Syst., 31 (2011), pp. 1427–1451.
- [26] C. M. TOPAZ, A. L. BERTOZZI, AND M. A. LEWIS, *A nonlocal continuum model for biological aggregation*, Bull. Math. Biol., 68 (2006), pp. 1601–1623.
- [27] C. VILLANI, *Optimal Transport: Old and New*, Grundlehren Math. Wiss. 338, Springer Science & Business Media, Berlin, Heidelberg, 2008.