Sticky-reflecting diffusion as a Wasserstein gradient flow

Jean-Baptiste Casteras, Léonard Monsaingeon, Filippo Santambrogio

Abstract

In this paper we identify the Fokker-Planck equation for (reflected) Sticky Brownian Motion as a Wasserstein gradient flow in the space of probability measures. The driving functional is the relative entropy with respect to a non-standard reference measure, the sum of an absolutely continuous interior part plus a singular part supported on the boundary. Taking the small time-step limit in a minimizing movement (JKO scheme) we prove existence of weak solutions for the coupled system of PDEs satisfying in addition an Energy Dissipation Inequality.

1 Introduction

Since the work of R. Jordan, D. Kinderlehrer, and F. Otto [22] it is now well understood that the classical Fokker-Planck equation

$$\begin{cases} \partial_t \rho = \Delta\rho + \operatorname{div}(\rho\nabla V) & \text{in } \Omega \\ (\nabla\rho + \rho\nabla V)\cdot\nu = 0 & \text{on } \partial\Omega \end{cases}$$

can be reinterpreted as the Wasserstein gradient flow of the relative entropy in the space of probability measures

$$\frac{d\rho}{dt} = -\operatorname{grad}_{\mathcal{W}} \mathcal{H}(\rho|\mu),$$

where $\mu = \frac{1}{\mathcal{Z}}e^{-V} \in \mathcal{P}(\Omega)$ is the stationary Gibbs measure associated with the background potential $V : \Omega \to \mathbb{R}$. Here $\mathcal{W}$ is the quadratic Wasserstein distance over the smooth domain $\Omega \subset \mathbb{R}^d$, and $\mathcal{H}(\rho|\mu) = \int_\Omega \frac{\rho}{\mu}\log\left(\frac{\rho}{\mu}\right)\mathrm{d}\mu$ is the relative entropy (Kullback-Leibler divergence) of ρ w.r.t. the reference measure μ. This has shed a whole new light on variational evolution of probability measures as gradient flows and the theory now covers advection-diffusion-aggregation equations [9, 5], Porous Medium Equation [31] and doubly nonlinear parabolic equations [1], fourth-order quantum drift-diffusion [27], reaction-diffusion equations with mass variations [24, 17, 12], and many more. We refer to the by-now classical textbooks [38, 37, 34] as well as to the survey [35] for further discussions, references, and applications of this steadily growing topics.

A common feature of all the aforementioned models and equations is that they deal with measures $\rho = \rho(x)\cdot\mathrm{d}x$ which are absolutely continuous with respect to either the Lebesgue measure in Euclidean domains, or its equivalent - the volume form - in Riemannian manifolds. For the standard Fokker-Planck equation this is clearly legitimate due to strong regularizing properties of the Laplacian, the generator of the standard (reflected) Brownian Motion with stationary measure precisely given by the Lebesgue measure. Recently however, the so-called (reflected) *Sticky Brownian Motion* [15] (SBM in short) started attracting renewed interest [33, 13, 7, 23]. Roughly speaking, SBM is a $\mathbb{R}^d$-valued stochastic process which behaves as a standard diffusion as long as it remains in the interior of a prescribed domain $\Omega \subset \mathbb{R}^d$. When it hits the boundary $\Gamma = \partial\Omega$ it sticks there for a random positive amount of time, while following an intrinsic tangential diffusion (generated by the Laplace-Beltrami operator Δ_Γ thereon). Eventually SBM almost surely reenters the domain and resumes its standard Brownian behaviour, until hitting the boundary again, and so on. We refer to [20, 14] for the

construction and regularity properties of SBM via Dirichlet-forms. In functional analytic terms, SBM is best described by its generator

$$\mathcal{L}f(x) = \lim_{t\downarrow 0}\frac{\mathbb{E}f(X_t) - f(x)}{t} = \begin{cases} \frac{1}{2}\Delta f(x) & \text{if } x \in \Omega \\ \frac{a}{2}\Delta_\Gamma f(x) - b\partial_\nu f(x) & \text{if } x \in \Gamma = \partial\Omega \end{cases}$$

were the parameters $a, b > 0$ are the tangential diffusivity and stickiness of the SBM process, respectively. The stationary measure accordingly comprises an interior $\mathcal{L}_\Omega^d$ and boundary $\mathcal{L}_\Gamma^{d-1}$ Lebesgue measures, and the laws of motion couple the interior domain and its boundary through an exchange term. The relevant probability distributions thus cannot simply be absolutely continuous, and we write throughout

$$\rho = \omega + \gamma \in \mathcal{P}(\overline{\Omega}) \qquad \text{with} \qquad \begin{cases} \omega = \rho \llcorner \Omega \in \mathcal{M}^+(\Omega) \\ \text{and} \\ \gamma = \rho \llcorner \Gamma \in \mathcal{M}^+(\Gamma) \end{cases}$$

for the associated interior/boundary decomposition $\overline{\Omega} = \Omega \cup \Gamma$. Discarding the probabilistic $\frac{1}{2}$ factors for convenience, the abstract Fokker-Planck equation $\partial_t\rho = \mathcal{L}^*\rho$ for pure SBM reads here

$$\begin{cases} \partial_t\omega = \Delta\omega & \text{in } \Omega \\ \omega = b\gamma & \text{on } \partial\Omega \\ \partial_t\gamma = a\Delta_\Gamma\gamma - \partial_\nu\omega & \text{in } \Gamma \end{cases} \tag{1.1}$$

In terms of PDEs this is a coupled system of bulk/interface diffusions, and the $\partial_\nu\omega$ exchange term corresponds at the stochastic level to the jump rate between the two interior/boundary behaviours. Here we focus on the case $a = b = 1$, and in this paper we shall make a case that this is again a gradient flow

$$\begin{cases} \partial_t\omega = \Delta\omega & \text{in } \Omega \\ \omega = \gamma & \text{on } \partial\Omega \\ \partial_t\gamma = \Delta_\Gamma\gamma - \partial_\nu\omega & \text{in } \Gamma \end{cases} \quad \Leftrightarrow \quad \frac{d\rho}{dt} = -\operatorname{grad}_\mathcal{W} \mathcal{H}(\rho\,|\,\mu) \quad \text{for} \quad \mu := \mathcal{L}_\Omega^d + \mathcal{L}_\Gamma^{d-1}. \tag{1.2}$$

Perhaps surprisingly, the transportation distance involved here is still the classical, quadratic Wasserstein distance $\mathcal{W}$ over $\overline{\Omega}$. This is due to our crucial assumption that $a = 1$, for which the interior Δ and boundary Δ_Γ diffusions are tangentially matched. For $a \neq 1$ the transportation distance must be adapted and the analysis is quite different. This will be investigated in a future work [10]. The case of coefficients $b \neq 1$ can however be treated by simply adapting the reference measure $\mu_b = b\mathcal{L}_\Omega^d + \mathcal{L}_\Gamma^{d-1}$ in the driving functional $\rho \mapsto \mathcal{H}(\rho\,|\,\mu_b)$, and we chose to focus on $b = 1$ only to clarify the exposition. Similarly, background potentials V_Ω, V_Γ could very well be included in the total free energy $\mathcal{F} = \mathcal{H} + \int_\Omega V_\Omega d\omega + \int_\Gamma V_\Gamma d\gamma$ in order to account for drift in the Fokker-Planck equation, but we simply ignore this possibility in order not to overburden the analysis. Note that a rather complete mathematical analysis (existence, uniqueness, regularity) of (1.1) has been carried out in [36] within a $H^1(\Omega) \times H^1(\Gamma)$ functional framework. For the smooth, positive solutions constructed therein, integration by parts gives

$$\begin{aligned}
\frac{d}{dt}\mathcal{H}(\rho_t\,|\,\mu) &= \frac{d}{dt}\left(\int_\Omega \omega_t \log\omega_t + \int_\Gamma \gamma_t \log\gamma_t\right) \\
&= \int_\Omega (1 + \log\omega_t)\Delta\omega_t + \int_\Gamma (1 + \log\gamma_t)\left[\Delta_\Gamma\gamma_t - \partial_\nu\omega_t\right] \\
&= \left[-\int_\Omega |\nabla\log\omega_t|^2\omega_t + \int_{\partial\Omega}(1 + \log\omega_t)\partial_\nu\omega_t\right] \\
&\quad + \left[-\int_\Gamma |\nabla_\Gamma\log\gamma_t|^2\gamma_t - \int_\Gamma (1 + \log\gamma_t)\partial_\nu\omega_t\right] \\
&= -\int_\Omega |\nabla\log\omega_t|^2\omega_t - \int_\Gamma |\nabla_\Gamma\log\gamma_t|^2\gamma_t =: -\mathcal{I}(\omega_t) - \mathcal{I}(\gamma_t). \tag{1.3}
\end{aligned}$$

2

In the last equality we used the Dirichlet boundary condition $\omega_t|_{\partial\Omega} = \gamma_t$ to cancel out the boundary terms (more on this in a moment). This shows that the relative entropy $\mathcal{H}(\rho \,|\, \mu) = \int_\Omega \omega \log \omega + \int_\Gamma \gamma \log \gamma$ of $\rho = \omega + \gamma$ is dissipated by the sum of the full Fisher information $\mathcal{I}(\omega)$ inside Ω and the tangential Fisher information $\mathcal{I}(\gamma)$ along the boundary $\Gamma = \partial\Omega$. Here our focus is not really on the well-posedness (although our analysis will provide as a byproduct existence of weak solutions for initial data with merely finite entropy), and we rather take interest in the variational gradient flow structure underlying this dissipation relation.

Although one can certainly come up with models naturally taking into account bulk/interface interactions directly into the transportation distance [30, 19], we find surprising that the coupled system (1.2) still fits within a completely standard optimal transport framework, as far as the metric is concerned. Another striking aspect of our analysis is the following: For standard Fokker-Planck equations the usual no-flux boundary condition encodes local conservation of mass, while in (1.2) one may wonder where the boundary condition $\omega|_{\partial\Omega} = \gamma$ stems from (this trace compatibility was crucial in order to get to (1.3)). This is actually unrelated to mass conservation: Our two evolution equations for ω, γ put together are already mass conservative regardless of the extra boundary condition $\omega|_{\partial\Omega} = \gamma$, since $\partial_\nu \omega$ in the second PDE is precisely the outflux of the momentum $\nabla \omega$ appearing in the first one so whatever "comes out of Ω" is just transferred to the boundary (the two PDEs together can indeed be reinterpreted as a single continuity equation $\partial_t \rho + \operatorname{div} m = 0$ with no-flux condition for the total density $\rho = \omega + \gamma$, but with a singular part γ living on the boundary in addition to a usual absolutely continuous part ω in the interior). The Dirichlet boundary condition $\omega|_{\partial\Omega} = \gamma$ is also not merely encoded as a constraint in the entropy functional itself, since finiteness of $\mathcal{H}(\rho \,|\, \mu)$ certainly does not require or guarantee the trace compatibility. For the sake of completeness let us mention that in [16] a transportation distance was constructed on $\mathcal{M}^+(\Omega)$, using the boundary $\Gamma = \partial\Omega$ as an infinite reservoir allowing to store and release arbitrary amounts of mass. It was then showed that the gradient flow of the standard entropy (with $\mathrm{d}x$ as a reference measure) with respect to this new distance corresponds to the heat equation with Dirichlet boundary condition $\rho|_{\partial\Omega} = 1$. Although similar in spirit this is actually unrelated to our approach here, since our model is really mass conservative in all regards. As we shall see later on, our boundary condition rather arises from the energy-dissipation mechanism, and more precisely from the *metric slope* of the driving functional (see Theorem 1). To the best of our knowledge this is the first example of an optimal transport model in which entropy dissipation (rather than entropy itself) gives rise to Dirichlet boundary conditions. In order to illustrate this even further, one could consider different energies, for example of the form

$$\mathcal{F}(\rho) = \int_\Omega F_\Omega(\omega) \mathrm{d}x + \int_\Gamma F_\Gamma(\gamma) \mathrm{d}x, \qquad \rho = \omega + \gamma$$

for some functions $F_\Omega, F_\Gamma : \mathbb{R}^+ \to \mathbb{R}$ satisfying suitable structural conditions. The gradient flow would then read

$$\frac{d\rho}{dt} = -\operatorname{grad}_{\mathcal{W}} \mathcal{F}(\rho) \qquad \Leftrightarrow \qquad \begin{cases} \partial_t \omega = \operatorname{div}(\omega \nabla F_\Omega'(\omega)) & \text{in } \Omega \\ F_\Omega'(\omega)|_{\partial\Omega} = F_\Gamma'(\gamma) & \text{on } \partial\Omega \\ \partial_t \gamma = \operatorname{div}_\Gamma(\gamma \nabla_\Gamma F_\Gamma'(\gamma)) - \omega \partial_\nu F_\Omega'(\omega) & \text{in } \Gamma. \end{cases}$$

The scalar fields $F_\Omega'(\omega), F_\Gamma'(\gamma)$ are known in classical optimal transport as *pressure variables* (in this respect the two evolution equations are nothing but Darcy's law). Very heuristically, a pressure difference along Γ would create somehow an infinite force. From a variational standpoint this should be prohibited for dissipative systems. Thus the Dirichlet boundary condition can be reinterpreted as a dissipative, pressure matching condition.

On a slightly different note, the theory of gradient flows in abstract metric spaces and *curves of maximal slope* was initiated by De Giorgi [11], and more recently developed in [2]. One possible way of formalizing the notion of such abstract gradient flows is to prove convergence of De Giorgi's *minimizing movement* in the small time-step limit $\tau \to 0$, of which the original JKO

scheme [22] is a particular instanciation in the Wasserstein space $(\mathcal{P}(\Omega), \mathcal{W})$. In order to back-up our claim (1.2) that sticky diffusion is indeed a gradient flow we will pursue this by now classical approach, with however a few twists. First, there are two classical ways of proving that the limiting curve is a "solution". The first one consists in exploiting purely metric tools to retrieve in the limit an *Energy Dissipation Inequality* (EDI in short), where a dissipation functional $\mathcal{D}(\rho_t) = -\frac{d}{dt}\mathcal{E}(\rho_t)$ plays a key role together with the *metric speed* $|\dot\rho_t|^2$ (both computed with respect to the Wasserstein distance). Usually the dissipation is related to the *metric slope* $\mathcal{D}(\rho) = |\partial\mathcal{E}|^2(\rho)$, typically a Fisher information functional $\mathcal{I}(\rho)$, and this guarantees that the limit is a curve of maximal slope. Moreover an upper chain rule argument shows that the energy dissipation forces equality in a Cauchy-Schwarz inequality, thus relating the driving momentum m in the continuity equation $\partial_t\rho + \operatorname{div} m = 0$ to spatial gradients of the energy $m = -\rho\nabla\frac{\delta\mathcal{E}}{\delta\rho}$. This gives the PDE in the end, but really requires a chain rule (computing the derivative in time of the energy $\mathcal{E}$ along a curve). This is well understood for classical functionals [2, chapter 10], less so here, and we did not fully succeed in this respect (see Theorem 1 and our Conjecture 2).

The second classical approach is directly PDE-oriented and less related to dissipation: Leveraging ad-hoc tools from optimal transportation theory, one usually writes down the discrete Euler-Lagrange optimality condition for each step $\rho^n \rightsquigarrow \rho^{n+1}$ of the JKO scheme, and tries passing to the limit $\tau \to 0$ to retrieve the continuous PDE directly from the one-step optimality. In the classical Fokker-Planck case this strongly relies on the Brenier-McCann theorem [8, 29], allowing to relate the optimal map in $\rho^{n+1} = T^n{}_\#\rho^n$ with the energy. More precisely, the Euler-Lagrange equation $\frac{T^n(x)-x}{\tau} = -\nabla\frac{\delta\mathcal{E}}{\delta\rho}(\rho^{n+1})$ typically gives the discrete velocity field driving particles around in the end (displacement $T^n(x) - x$ divided by time τ). This technical tool is unfortunately not available here, since our entropy functional always forces $\rho^n = \omega^n + \gamma^n$ to have a singular part γ^n supported on the boundary and therefore systematically prevents any application of the Brenier-McCann theorem (which precisely requires ρ not to give mass to small $\mathcal{H}^{d-1}$ sets in dimension d!). In some sense the bulk-interface interaction, corresponding to the $\partial_\nu\omega$ exchange term in (1.2), forces mass splitting already at the discrete level. In order to circumvent this technical obstacle and still obtain a useful Euler-Lagrange equation we perform instead an ad-hoc ε-regularization of the entropy functional, and show that the barycentric momentum conveys enough information in order to first pass to the limit $\varepsilon \to 0$ and then take $\tau \to 0$.

Let us now fix some notations. We always work in a smooth, bounded, convex domain $\Omega \subset \mathbb{R}^d$, $g \geq 2$. (We assume convexity for technical convenience only, but the result should hold for general domains.) We set

$$H(z) := z\log z - z + 1 \geq 0, \qquad z \geq 0,$$

and the reference measure will always be

$$\mu := \mathcal{L}_\Omega^d + \mathcal{L}_\Gamma^{d-1} \in \mathcal{M}^+(\overline\Omega).$$

The entropy

$$\mathcal{E}(\rho) := \mathcal{H}(\rho\,|\,\mu) = \begin{cases} \int_{\overline\Omega} H\left(\frac{\mathrm{d}\rho}{\mathrm{d}\mu}(x)\right)\mu(\mathrm{d}x) & \text{if } \rho \ll \mu \\ +\infty & \text{else} \end{cases} \tag{1.4}$$

is then nonnegative, strictly convex, and lower semi-continuous. As suggested by (1.3) we define the dissipation

$$\mathcal{D}(\rho) := \begin{cases} \mathcal{I}(\omega) + \mathcal{I}(\gamma) & \text{if } \rho = \omega + \gamma \text{ and } \omega|_{\partial\Omega} = \gamma \\ +\infty & \text{else} \end{cases} \tag{1.5}$$

(see later on for a rigorous definition of the Fisher information functional $\mathcal{I}$ as well as the precise meaning of the boundary trace $\omega|_{\partial\Omega}$). This penalization of the constraint $\omega|_{\partial\Omega} = \gamma$ is not technically artificial and arises intrinsically when trying to compute the metric slope $|\partial\mathcal{E}|$, see Theorem 1 below.

4

Given an initial datum $\rho_0 \in \mathcal{P}(\overline{\Omega})$, the JKO scheme consists in initializing $\rho^0 = \rho_0$ and solving recursively

$$\rho^{n+1} \in \underset{\rho \in \mathcal{P}(\overline{\Omega})}{\text{Argmin}} \left\{ \frac{1}{2\tau} \mathcal{W}^2(\rho, \rho^n) + \mathcal{E}(\rho) \right\}. \tag{1.6}$$

With such a discrete sequence one can define the piecewise constant interpolation

$$\overline{\rho}_t^\tau := \rho^{n+1} \qquad \text{for} \qquad t \in (n\tau, (n+1)\tau],$$

and we will establish

Main result. *Fix any $\rho_0 \in \mathcal{P}(\overline{\Omega})$ with $\mathcal{E}(\rho_0) < \infty$. For any small $\tau > 0$ the JKO scheme (1.6) is uniquely well-posed, and there is a discrete sequence $\tau \to 0$ such that the interpolant $\overline{\rho}^\tau$ converges to a continuous curve $\rho : [0, \infty) \to \mathcal{P}(\overline{\Omega})$*

$$\mathcal{W}\left(\overline{\rho}_t^\tau, \rho_t\right) \xrightarrow[\tau \to 0]{} 0 \qquad \text{for all } t \geq 0.$$

The limit satisfies the Energy Dissipation Inequality

$$\mathcal{E}(\rho_T) + \int_0^T \left(\frac{1}{2} |\dot{\rho}_t|^2 + \frac{1}{2} \mathcal{D}(\rho_t) \right) \mathrm{d}t \leq \mathcal{E}(\rho_0), \qquad \forall T > 0$$

and is a weak solution of the PDE (1.2). Moreover, there is $\lambda > 0$ only depending on Ω such that

$$\mathcal{H}(\rho_t \,|\, \bar{\mu}) \leq e^{-2\lambda t} \mathcal{H}(\rho_0 \,|\, \bar{\mu}) \quad and \quad |\rho_t - \bar{\mu}|_{TV} \leq e^{-\lambda t} \sqrt{\frac{1}{2} \mathcal{H}(\rho_0 \,|\, \bar{\mu})}, \qquad \forall t \geq 0 \tag{1.7}$$

where $\bar{\mu} := \frac{1}{\mu(\overline{\Omega})} \mu$ it the renormalized stationary measure. Finally, if $\rho_0 \leq \bar{c}\mu$ for some constant $\bar{c} > 0$ (resp. $\rho_0 \geq \underline{c}\mu$ for some $\underline{c}$) then $\rho_t \leq \bar{c}\mu$ for all $t \geq 0$ (resp. $\rho_t \geq \underline{c}\mu$.)

It is worth stressing that something is still missing in order to obtain a rigorous metric gradient flow. Indeed, in order for EDI to fully characterize curves of maximal slope one should really prove that $g(\rho) = \sqrt{\mathcal{D}(\rho)}$ is an *upper gradient* [2, chapter 1]. The local slope $|\partial \mathcal{E}|$ is always a (weak) upper gradient, and we conjecture that $\mathcal{D} = |\partial \mathcal{E}|^2$ thus it is plausible that $\sqrt{\mathcal{D}}$ should be an upper gradient. However we only managed to prove that $\mathcal{D} \leq |\partial \mathcal{E}|^2$, and we were also not able to prove directly that $\sqrt{\mathcal{D}}$ is an upper gradient.

Another key concept for the abstract theory of metric gradient flows is that of geodesic convexity, or, in the specific optimal transport framework, McCann's displacement convexity [28]. In smooth, complete Riemannian manifolds the λ-convexity of $\mathcal{H}$ (relatively to the reference volume measure) is equivalent to λ-Ricci lower bounds. In Euclidean domains, it is known to hold with $\lambda = 0$ if and only if Ω is convex, which can also be reinterpreted as the fact that the subspace $\mathcal{P}_{\mathcal{L}_\Omega^d}^{ac}(\overline{\Omega}) \subset \mathcal{P}(\overline{\Omega})$ of absolutely continuous measures (w.r.t. to the reference Lebesgue measure $\mathcal{L}_\Omega^d$) is geodesically convex. It is worth stressing that this whole picture collapses here: Due to our choice of reference measure $\mu = \mathcal{L}_\Omega^d + \mathcal{L}_\Gamma^{d-1}$ the entropy $\mathcal{E}(\rho) = \mathcal{H}(\rho \,|\, \mu)$ cannot be λ-convex for any $\lambda \in \mathbb{R}$, even if the underlying domain Ω is strongly uniformly convex. This is due to the fact that $\mathcal{P}_\mu^{ac}(\overline{\Omega})$ itself fails to be geodesically convex. For a counterexample, take Ω the unit ball, and consider two measures $\rho_0 = 0 + \gamma_0, \rho_1 = 0 + \gamma_1$ smoothly supported on the boundary as in Figure 1, with finite entropies $\mathcal{E}(\rho_i) = \mathcal{H}(\gamma_i \,|\, \mathcal{L}_\Gamma^{d-1}) < \infty$. It is not difficult in this configuration to check that the unique geodesic ρ_t remains singular (and absolutely continuous w.r.t. $\mathcal{H}^{d-1}$) and supported strictly inside Ω for all intermediate times $t \in (0, 1)$. As a consequence $\rho_t = \omega_t + 0$ is not absolutely continuous w.r.t. μ and thus $\mathcal{E}(\rho_t) = +\infty$ for all $t \in (0, 1)$, while $\mathcal{E}(\rho_0), \mathcal{E}(\rho_1) < +\infty$.

It is also known that geodesic convexity usually leads to logarithmic Sobolev and entropy-entropy production inequalities, which in turn should yield exponential convergence as $t \to \infty$ towards the unique entropy minimizer, here $\overline{\mu} = \frac{1}{|\Omega| + |\Gamma|} \mu \in \mathcal{P}(\overline{\Omega})$. Although convexity completely fails here as just discussed, it was proved nonetheless in [6] that $\overline{\mu}$ always satisfies a
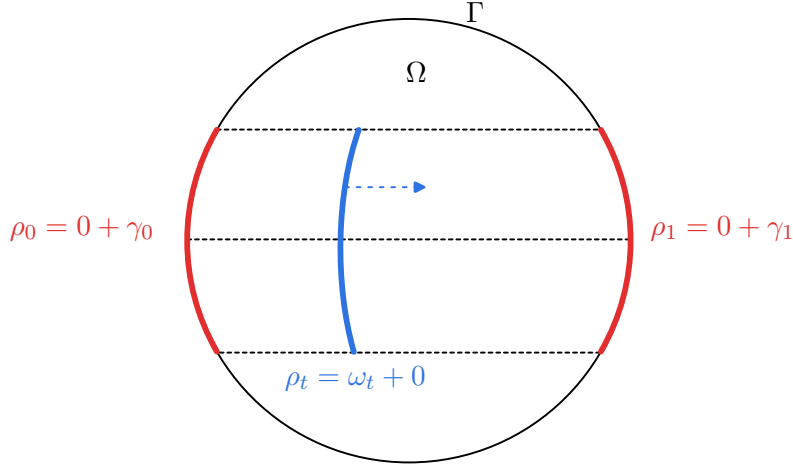
Figure 1: Counterexample to displacement convexity

boundary-trace logarithmic Sobolev inequality. The usual entropy-entropy production argument will thus apply in our framework as well and will allow us to establish the exponential convergence (1.7).

The paper is organized as follows: In Section 1 we fix some notations and define a few relevant objects. We also establish a (partial) connection between the abstract metric slope $|\partial\mathcal{E}|(\rho)$ of our relative entropy $\mathcal{E}(\rho) = \mathcal{H}(\rho\,|\,\mu)$ and its dissipation functional $\mathcal{D}(\rho)$, Theorem 1. Section 2 gathers several a-priori estimates for the discrete JKO solutions, including the Euler-Lagrange optimality condition and a discrete dissipation inequality based on De Giorgi's *variational interpolant*. In Section 3 we leverage the previous work to pass to the limit $\tau \to 0$, and prove the main result. For convenience we split this into Proposition 3.1 (convergence to a limiting curve satisfying EDI) and Proposition 3.3 (the limit is a weak solution of the PDE). Finally, we defer to Appendix A some technical statements and lemmas needed along the way. Let us first fix once and for all the framework as well as a few notations.

- $\Omega \subset \mathbb{R}^d$ ($d \geq 2$) is a smooth, bounded, convex domain with boundary $\Gamma = \partial\Omega$. We tend to write $\partial\Omega$ for quantities arising from the interior, such as $\omega|_{\partial\Omega}$, while we prefer the notation Γ for quantities intrinsically defined thereon, such as γ. The Euclidean distance on $\overline{\Omega}$ is denoted by $\mathsf{d}_{\overline{\Omega}}$, while the intrinsic Riemannian distance on the boundary is denoted d_Γ. The full gradient, divergence and Laplacian in Ω are denoted $\nabla, \mathrm{div}, \Delta$, while their intrinsic Riemannian counterpart along Γ are denoted $\nabla_\Gamma, \mathrm{div}_\Gamma, \Delta_\Gamma$.

- We write $\mathcal{X}$ for smooth Riemannian manifolds, possibly with boundaries, and $\mathsf{d}_\mathcal{X}$ the Riemannian distance. Typically we shall take $\mathcal{X} = \overline{\Omega}$, a flat manifold with boundary $\partial\Omega$, or $\mathcal{X} = \Gamma$, a curved manifold of its own right (with induced Riemannian metric).

- With a clear abuse of notations we call Lebesgue measure the volume form on a manifold $\mathcal{X}$ and we write indistinctly $\mathrm{d}x$ for the Lebesgue measures $\mathcal{L}^d_\Omega$ or $\mathcal{L}^{d-1}_\Gamma$. Unless otherwise specified, and with a slight abuse of notations, the Boltzmann entropy $\mathcal{H}(\nu)$ always denotes the entropy $\mathcal{H}(\nu\,|\,\mathrm{d}x)$ of a nonnegative measure $\nu \in \mathcal{M}^+(\mathcal{X})$ with respect to this $\mathrm{d}x$ measure.

- For probability measures $\rho \in \mathcal{P}(\overline{\Omega})$ we always decompose $\rho = \rho\llcorner\Omega + \rho\llcorner\Gamma =: \omega + \gamma$ and identify $\omega \in \mathcal{M}^+(\Omega), \gamma \in \mathcal{M}^+(\Gamma)$ with their densities with respect to the corresponding

6

Lebesgue measures, i-e we $\omega = \omega(x) \cdot \mathcal{L}_\Omega^d$ and $\gamma = \gamma(x) \cdot \mathcal{L}_\Gamma^{d-1}$. Typically this notation allows to simply write

$$\mathcal{E}(\rho) = \mathcal{H}\left(\omega + \gamma \mid \mathcal{L}_\Omega^d + \mathcal{L}_\Gamma^{d-1}\right) = \mathcal{H}(\omega) + \mathcal{H}(\gamma).$$

- The narrow convergence of measures is defined in duality with bounded, continuous test functions

$$\rho^n \xrightarrow{*} \rho \qquad \text{iff} \qquad \int_\mathcal{X} \varphi(x)\rho^n(\mathrm{d}x) \to \int_\mathcal{X} \varphi(x)\rho(\mathrm{d}x), \quad \forall \varphi \in C_b(\mathcal{X})$$

as $n \to \infty$

- For $\mu, \nu \in \mathcal{M}^+(\mathcal{X})$ with compatible masses $M = \int \mu = \int \nu$ the (squared) Wasserstein distance is

$$\mathcal{W}_\mathcal{X}^2(\mu, \nu) = \min_{\pi \in \Pi(\mu,\nu)} \iint_{\mathcal{X} \times \mathcal{X}} \mathsf{d}_\mathcal{X}^2(x, y)\pi(\mathrm{d}x, \mathrm{d}y),$$

where $\Pi(\mu, \nu)$ the set of admissible plans $\pi \in \mathcal{M}^+(\mathcal{X} \times \mathcal{X})$ with mass M and marginals $\pi_x = \mu, \pi_y = \nu$. We refer to [34, 37, 38] for more material on optimal transport theory. Unless otherwise specified $\mathcal{W}$ will simply denote $\mathcal{W}_{\overline{\Omega}}$, the Wasserstein distance over $\overline{\Omega}$ for the quadratic cost $\mathsf{d}_{\overline{\Omega}}^2(x, y) = |y - x|^2$.

- Following [2], in a metric space (X, d) a curve $\{x_t\}_{t \in I}$ is said to be absolutely continuous if

$$\mathsf{d}(x_{t_0}, x_{t_1}) \leq \int_{t_0}^{t_1} \eta(t)\mathrm{d}t$$

for some $\eta \in L^1(I)$ and all $[t_0, t_1] \subseteq I$. In that case the *metric speed*

$$|\dot{x}_t| := \lim_{h \to 0} \frac{\mathsf{d}(x_{t+h}, x_t)}{h}$$

exists for a.e. $t \in I$ and is the smallest function η satisfying the above inequality. The (local) *metric slope* of a functional $\mathsf{E} : \mathsf{X} \to \mathbb{R} \cup \{+\infty\}$ is defined as

$$|\partial \mathsf{E}|(x) = \limsup_{y \to x} \frac{[\mathsf{E}(x) - \mathsf{E}(y)]^+}{\mathsf{d}(x, y)}. \tag{1.8}$$

- We use the t subscript for curves of measures $t \mapsto \rho_t = \omega_t + \gamma_t$, while the time derivative is denoted ∂_t.

Definition 1.1 (Fisher information [2, defs. 10.4.15 and 10.4.16]). *On a smooth Riemannian manifold (possibly with boundary) $\mathcal{X}$ we say that a nonnegative measure $\mu \in \mathcal{M}^+(\mathcal{X})$ has logarithmic derivative if $\mu = f(x) \cdot \mathrm{d}x$ for some $f \in L^1(\mathrm{d}x)$ and if its distributional gradient $\nabla f \in L_{loc}^1(\mathrm{d}x)$ with moreover $|\nabla f| \ll f$. In that case we write with a slight abuse of notations $\nabla \log f = \frac{\nabla f}{f}$ for the Radon-Nikodym derivative. The Fisher information is then*

$$\mathcal{I}(\mu) := \begin{cases} \int_\mathcal{X} |\nabla \log f|^2 \mathrm{d}\mu & \text{if } \mu = f \cdot \mathrm{d}x \text{ has logarithmic derivative } \nabla \log f \\ +\infty & \text{otherwise} \end{cases}$$

We shall often exploit the following well-known facts, which we recall here without proof.

Lemma 1.2 (Properties of the Fisher information). *For any nonnegative, absolutely continuou $f \in L^1(\mathcal{X})$ with mass $M = \int_\mathcal{X} f < \infty$ we write $\mu = f \cdot \mathrm{d}x$. Then*

- *One has the equivalent representations*

$$\mathcal{I}(\mu) = \int_\mathcal{X} |\nabla \log f|^2 f = \int_\mathcal{X} \left|\frac{\nabla f}{f}\right|^2 f = \int_\mathcal{X} \left|\frac{\nabla f}{\sqrt{f}}\right|^2 = 4\left\|\nabla \sqrt{f}\right\|_{L^2}^2$$

7

- *In dimension d there holds*

$$\|f\|_{W^{1,p}(\mathcal{X})} \leq C_{p,\mathcal{X}}[M + \mathcal{I}(\mu)] \qquad for \qquad \begin{cases} p = 2 & if\ d = 1 \\ p = 2^- & if\ d = 2 \\ p = \frac{d}{d-1} & if\ d \geq 3 \end{cases} \tag{1.9}$$

($p = 2^-$ *means as usual "for any $p < 2$"*)

- *$\mathcal{I}$ is convex and lower semi-continuous for the narrow convergence.*

Recall that, on smooth bounded domains $\Omega \subset \mathbb{R}^d$, the trace operator $\mathrm{tr} : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is continuous for $p \geq 1$ and compact for $p > 1$ [25, chapter 18]. Owing to (1.9) we see that a measure $\mu = f \cdot \mathrm{d}x$ with $\mathcal{I}(\mu) < \infty$ always has a well-defined trace $f|_{\partial\Omega} = \mathrm{tr}\, f \in L^p(\partial\Omega)$ for some $p > 1$. This clarifies the rigorous meaning of the dissipation functional $\mathcal{D}$ in (1.5): Either $\mathcal{I}(\omega)$ and $\mathcal{I}(\gamma)$ are finite and therefore the trace $\omega|_{\partial\Omega} = \mathrm{tr}\,\omega$ is well-defined, or simply $\mathcal{D}(\rho) = +\infty$. This will also give a meaning to the boundary condition $\omega|_{\partial\Omega} = \gamma$ in the Fokker-Planck equation (1.2), whose solutions will satisfy by construction $\int_0^T \mathcal{D}(\rho_t)\,\mathrm{d}t < +\infty$ and thus $\mathcal{D}(\rho_t) < +\infty$ for a.e. $t \geq 0$.

The next result is specific to our optimal transport context with $\mathcal{E}(\rho) = \mathcal{H}(\omega) + \mathcal{H}(\gamma)$, and will be key for the subsequent analysis. It will allow to retrieve simultaneously the boundary conditions as well as natural dissipation estimates for the PDE directly from the small-time step limit $\tau \to 0$ in the JKO scheme, and more importantly justifies our claim in the introduction that the trace compatibility arises from the dissipation only.

Theorem 1. *Let $\mathcal{D}$ be the dissipation functional in (1.5). Then*

$$|\partial\mathcal{E}|^2(\rho) \geq \mathcal{D}(\rho) \tag{1.10}$$

and $\mathcal{D}$ is (sequentially) narrowly lower semi-continuous.

We were not able to prove the reverse inequality, and for future reference we record here

Conjecture 2. *Equality holds in (1.10).*

Proof of Theorem 1. Let us first show that $|\partial\mathcal{E}|^2(\rho) \geq \mathcal{I}(\omega) + \mathcal{I}(\gamma)$. Take $\rho = \omega + \gamma \in \mathrm{Dom}(\mathcal{E})$ with masses $M_\Omega = \int_\Omega \omega \geq 0$ and $M_\Gamma = \int_\Gamma \gamma \geq 0$, $M_\Omega + M_\Gamma = 1$. For any arbitrary $\tilde{\omega} \in \mathcal{M}^+(\overline{\Omega}), \tilde{\gamma} \in \mathcal{M}^+(\Gamma)$ with same masses M_Ω, M_Γ let $\tilde{\rho} = \tilde{\omega} + \tilde{\gamma} \in \mathcal{P}(\overline{\Omega})$. Let $\pi_\Omega \in \Pi(\omega, \tilde{\omega})$ and $\pi_\Gamma \in \Pi(\gamma, \tilde{\gamma})$ be optimal plans for the transportation problems set on $(\overline{\Omega}, \mathsf{d}_{\overline{\Omega}})$ and $(\Gamma, \mathsf{d}_\Gamma)$, respectively. The plan $\pi := \pi_\Omega + \pi_\Gamma$ is easily seen to be admissible for the transportation problem from $\rho = \omega + \gamma$ to $\tilde{\rho} = \tilde{\omega} + \tilde{\gamma}$, set on $(\overline{\Omega}, \mathsf{d}_{\overline{\Omega}})$. Since the induced distance $\mathsf{d}_\Gamma(x,y) \geq \mathsf{d}_{\overline{\Omega}}(x,y)$ along the boundary we have in particular

$$\mathcal{W}^2(\rho, \tilde{\rho}) \leq \iint_{\overline{\Omega}^2} \mathsf{d}_{\overline{\Omega}}^2(x,y)\pi(\mathrm{d}x, \mathrm{d}y) = \iint_{\Omega^2} \mathsf{d}_{\overline{\Omega}}^2(x,y)\pi_\Omega(\mathrm{d}x, \mathrm{d}y) + \iint_{\Gamma^2} \mathsf{d}_{\overline{\Omega}}^2(x,y)\pi_\Gamma(\mathrm{d}x, \mathrm{d}y)$$

$$\leq \iint_{\Omega^2} \mathsf{d}_{\overline{\Omega}}^2(x,y)\pi_\Omega(\mathrm{d}x, \mathrm{d}y) + \iint_{\Gamma^2} \mathsf{d}_\Gamma^2(x,y)\pi_\Gamma(\mathrm{d}x, \mathrm{d}y) = \mathcal{W}_{\overline{\Omega}}^2(\omega, \tilde{\omega}) + \mathcal{W}_\Gamma^2(\gamma, \tilde{\gamma}). \tag{1.11}$$

Let now $(\omega_t, \gamma_t)_{t \geq 0}$ be the heat flows running on Ω, Γ, respectively, started from ω, γ (with of course no-flux boundary condition on $\partial\Omega$ for the first flow). Exploiting the *duality formula for*

8

the local slope [2, lemma 3.1.5] and writing $\rho_t = \omega_t + \gamma_t$ we can bound

$$\frac{1}{2}|\partial\mathcal{E}|^2(\rho) = \limsup_{t\downarrow 0}\frac{1}{t}\sup_{\tilde{\rho}}\left\{\mathcal{E}(\rho) - \mathcal{E}(\tilde{\rho}) - \frac{1}{2t}\mathcal{W}^2(\rho,\tilde{\rho})\right\}$$

$$= \limsup_{t\downarrow 0}\frac{1}{t}\sup_{\tilde{\rho}=\tilde{\omega}+\tilde{\gamma}}\left\{[\mathcal{H}(\omega) + \mathcal{H}(\gamma)] - [\mathcal{H}(\tilde{\omega}) + \mathcal{H}(\tilde{\gamma})] - \frac{1}{2t}\mathcal{W}^2(\rho,\tilde{\rho})\right\}$$

$$\geq \limsup_{t\downarrow 0}\frac{1}{t}\left\{[\mathcal{H}(\omega) + \mathcal{H}(\gamma)] - [\mathcal{H}(\omega_t) + \mathcal{H}(\gamma_t)] - \frac{1}{2t}\mathcal{W}^2(\rho,\rho_t)\right\}$$

$$\overset{(1.11)}{\geq} \limsup_{t\downarrow 0}\frac{1}{t}\left\{\left[\mathcal{H}(\omega) - \mathcal{H}(\omega_t) - \frac{1}{2t}\mathcal{W}_\Omega^2(\omega,\omega_t)\right] + \left[\mathcal{H}(\gamma) - \mathcal{H}(\gamma_t) - \frac{1}{2t}\mathcal{W}_\Gamma^2(\gamma,\gamma_t)\right]\right\}$$

$$\geq \liminf_{t\downarrow 0}\frac{1}{t}\left[\mathcal{H}(\omega) - \mathcal{H}(\omega_t) - \frac{1}{2t}\mathcal{W}_\Omega^2(\omega,\omega_t)\right] + \liminf_{t\downarrow 0}\frac{1}{t}\left[\mathcal{H}(\gamma) - \mathcal{H}(\gamma_t) - \frac{1}{2t}\mathcal{W}_\Gamma^2(\gamma,\gamma_t)\right].$$

By Proposition A.1 the two $\liminf$ are bounded from below by half the respective Fisher informations, and $|\partial\mathcal{E}|^2(\rho) \geq \mathcal{I}(\omega) + \mathcal{I}(\gamma)$ follows.

Let us now prove the the lower semicontinuity. Take a sequence $\rho^n \overset{*}{\rightharpoonup} \rho$. Up to extraction of a subsequence if needed we can assume that $\liminf \mathcal{D}(\rho^n) = \lim \mathcal{D}(\rho^n) < \infty$. By definition of $\mathcal{D}$ we have $\mathcal{I}(\omega^n) + \mathcal{I}(\gamma^n) \leq C$ and $\omega^n|_{\partial\Omega} = \gamma^n$ for all n, and from (1.9) we have moreover that ω^n, γ^n are bounded in $W^{1,p}$ for some $p > 1$. By Rellich-Kondrachov compactness we get that $\omega^n \to \omega$ and $\gamma^n \to \gamma$ at least in $L^1(\Omega), L^1(\Gamma)$, respectively. This implies in particular that $\rho = \lim \rho^n$ has interior/boundary decomposition $\omega + \gamma = \lim \omega^n + \lim \gamma^n$. By compactness of the trace operator we also get $\omega|_{\partial\Omega} = \operatorname{tr}\omega = \lim(\operatorname{tr}\omega^n) = \lim \gamma^n = \gamma$, hence $\mathcal{D}(\rho) = \mathcal{I}(\omega) + \mathcal{I}(\gamma)$. By standard lower semi-continuity of $\mathcal{I}$ we have next $\mathcal{D}(\rho) = \mathcal{I}(\omega) + \mathcal{I}(\gamma) \leq \liminf \mathcal{I}(\omega^n) + \liminf \mathcal{I}(\gamma^n) \leq \liminf\{\mathcal{I}(\omega^n) + \mathcal{I}(\gamma^n)\} = \liminf \mathcal{D}(\rho^n)$ and the claim follows.

It only remains now to address the much more delicate trace compatibility, i.e. we need to prove that if $\mathcal{I}(\omega) + \mathcal{I}(\gamma) < \infty$ but $\omega|_{\partial\Omega} \neq \gamma$ then $|\partial\mathcal{E}|(\rho) = +\infty$. Assume accordingly that there is $x_0 \in \partial\Omega = \Gamma$ such that $\omega|_{\partial\Omega}(x_0) \neq \gamma(x_0)$. (This is meaningful in the sense that both $\omega|_{\partial\Omega}$ and γ are defined $\mathcal{L}_\Gamma^{d-1}$ a.e.) Consider first the case

$$\omega_0 := \omega|_{\partial\Omega}(x_0) > \gamma(x_0) =: \gamma_0.$$

For some small $\varepsilon, \delta, \theta$ to be adjusted later on we will construct below a suitable $(\varepsilon, \delta, \theta)$-perturbation $\tilde{\omega}, \tilde{\gamma}$ of ω, γ in a neighborhood of x_0 by transferring a small fraction θ of mass from the interior to the boundary in order to "diminish the gap $\omega_0 - \gamma_0 > 0$" close to x_0. This will be achieved by paying a transportation cost one order of magnitude smaller that the gain in entropy as $\varepsilon, \delta, \theta \to 0$, hence by definition (1.8) of the slope we will have $|\partial\mathcal{E}|(\rho) \geq \liminf_{\varepsilon,\delta,\theta\to 0}\frac{\mathcal{E}(\rho) - \mathcal{E}(\tilde{\rho})}{\mathcal{W}(\rho,\tilde{\rho})} = +\infty$. Roughly speaking δ will be a small tangential length-scale along the boundary and ε a small normal scale, so that the perturbation will be constructed in a box of size $\delta^{d-1}\varepsilon$. The parameter θ will control the small fraction of mass to be taken from the interior and mapped to the boundary, and in some sense θ will be used to "linearise" $H(\omega) - H(\tilde{\omega}) \sim H'(\tilde{\omega})(\omega - \tilde{\omega})$ and $H(\gamma) - H(\tilde{\gamma}) \sim H'(\tilde{\gamma})(\gamma - \tilde{\gamma})$. The important point is that the transfer of mass from the interior to the boundary should take place on a normal scale $\varepsilon \ll \delta$ much smaller than the tangential scale δ. This reflects the normal derivative $\partial_\nu\omega$ appearing in (1.2), which consistently encodes the bulk/interface interactions.

Since our construction of the perturbation below will be localized, it is enough to work in local charts and we thus work in the following configuration: $\Omega = \mathbb{R}_+^d$ is a half-space, $\Gamma = \partial\Omega = \{0\} \times \mathbb{R}^{d-1}$, and $x_0 = (0, 0_{\mathbb{R}^{d-1}}) \in \Gamma$. We write $x = (x_1, x') \in \mathbb{R}^+ \times \mathbb{R}^{d-1}$ for the natural coordinates, and denote by
$$T(x) = \operatorname{Proj}_{\mathbb{R}^{d-1}}(x) = (0, x')$$
the projection on the boundary. We take $\Gamma_\delta = B_\delta^{d-1}(0)$ a small neighborhood of the origin in Γ, and also define the ε-thickening $\Omega_{\varepsilon,\delta} = (0, \varepsilon) \times \Gamma_\delta$ as in Figure 2. For reasons that will appear

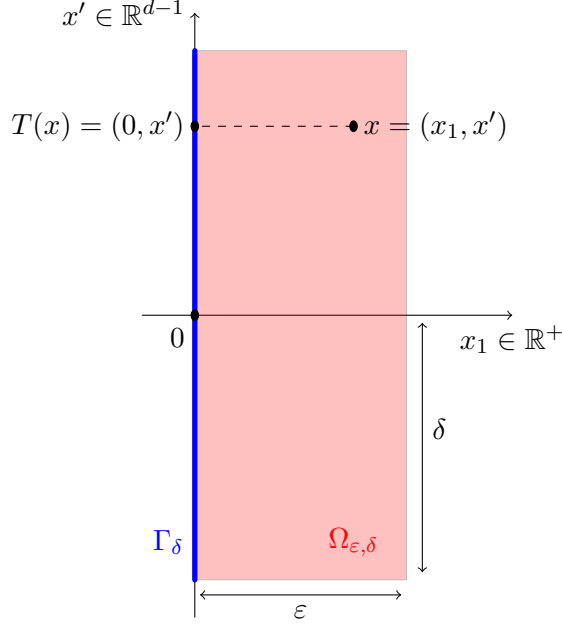Figure 2: The (ε, δ) boundary layer

clear later on we impose

$$\varepsilon^{p-1} = \mathcal{O}\left(\delta^{d-1}\right) \qquad \text{for } p = \frac{d}{d-1} > 1. \tag{1.12}$$

Writing $\hat{\gamma} := T_{\#}(\omega \llcorner \Omega_{\varepsilon, \delta})$ for simplicity, the perturbation is then defined as

$$\tilde{\omega}(x) := \begin{cases} (1-\theta)\omega(x) & \text{if } x \in \Omega_{\varepsilon, \delta} \\ \omega(x) & \text{else} \end{cases} \qquad \text{and} \qquad \tilde{\gamma}(x') := \begin{cases} \gamma(x') + \theta\hat{\gamma}(x') & \text{if } x' \in \Gamma_\delta \\ \gamma(x') & \text{else} \end{cases} \tag{1.13}$$

In other words, we take a small fraction $\theta \ll 1$ of ω inside $\Omega_{\varepsilon, \delta}$ and transport it to the boundary Γ_δ along the x_1 direction, where it is simply added to the boundary density already present there. Meanwhile, ω, γ remain unchanged outside of the small control boxes $\Omega_{\varepsilon, \delta}, \Gamma_\delta$, respectively. Notice that $\tilde{\omega}, \tilde{\gamma}$ do *not* satisfy the compatibility condition. We perturb however "in the good direction" by decreasing the trace gap, since $\tilde{\omega} \leq \omega$ and $\tilde{\gamma} \geq \gamma$ and therefore $[\tilde{\omega} - \tilde{\gamma}](x_0) \leq [\omega - \gamma](x_0)$.

Step 1: variation in entropy. In order to estimate the first variation of $\mathcal{E}$ we first exploit the convexity of H to write

$$\mathcal{E}(\rho) - \mathcal{E}(\tilde{\rho}) = [\mathcal{H}(\omega) - \mathcal{H}(\tilde{\omega})] + [\mathcal{H}(\gamma) - \mathcal{H}(\tilde{\gamma})] \geq \underbrace{\int_{\Omega_{\varepsilon, \delta}} H'(\tilde{\omega})[\omega - \tilde{\omega}]}_{:=A} + \underbrace{\int_{\Gamma_\delta} H'(\tilde{\gamma})[\gamma - \tilde{\gamma}]}_{:=B}$$

10

For the first term we use the convexity of $z \mapsto zH'(z) = z \log z$ to apply Jensen's inequality as

$$
\begin{aligned}
A = \int_{\Omega_{\varepsilon,\delta}} H'[(1-\theta)\omega]\theta\omega &= \frac{\theta}{1-\theta} \int_{\Omega_{\varepsilon,\delta}} H'[(1-\theta)\omega](1-\theta)\omega \\
&= \frac{\theta|\Omega_{\varepsilon,\delta}|}{1-\theta} \fint_{\Omega_{\varepsilon,\delta}} H'[(1-\theta)\omega](1-\theta)\omega \\
&\geq \frac{\theta|\Omega_{\varepsilon,\delta}|}{1-\theta} H'\left((1-\theta)\fint_{\Omega_{\varepsilon,\delta}}\omega\right)\left((1-\theta)\fint_{\Omega_{\varepsilon,\delta}}\omega\right).
\end{aligned}
$$

From Lemma 1.2 we have that $\omega \in W^{1,p}$ with $p = \frac{d}{d-1} > 1$. Because we took care to impose the ε, δ regime (1.12) we can apply Lemma A.3 to conclude that $\fint_{\Omega_{\varepsilon,\delta}}\omega \to \omega_0 > 0$ and thus

$$
\liminf_{\varepsilon,\delta,\theta} \frac{A}{\theta|\Omega_{\varepsilon,\delta}|} \geq H'(\omega_0)\omega_0. \tag{1.14}
$$

(Note that, since we assumed $\omega_0 > \gamma_0$, we have in particular $\omega_0 > 0$ so there is no issue with $\log 0$ here.) The B term is more involved and requires extra work. Observe that $\hat{\gamma} = T_\#(\omega \llcorner \Omega_{\varepsilon,\delta})$ is given more explicitly by

$$
\hat{\gamma}(x') = \int_0^\varepsilon \omega(x_1, x')\mathrm{d}x_1, \qquad x' \in \Gamma_\delta.
$$

This is small as $\varepsilon \to 0$ so in order to obtain something nontrivial one should actually look at the rescaled quantity

$$
\bar{\gamma}(x') := \frac{1}{\varepsilon}\hat{\gamma}(x') = \int_0^\varepsilon \omega(x_1, x')\frac{\mathrm{d}x_1}{\varepsilon}.
$$

For any $p \geq 1$ we have by Jensen's inequality

$$
\begin{aligned}
\fint_{\Gamma_\delta} |\bar{\gamma} - \omega_0|^p = \fint_{\Gamma_\delta} \left|\int_0^\varepsilon [\omega(x_1,x') - \omega_0]\frac{\mathrm{d}x_1}{\varepsilon}\right|^p \mathrm{d}x' \\
\leq \fint_{\Gamma_\delta} \int_0^\varepsilon |\omega(x_1,x') - \omega_0|^p \frac{\mathrm{d}x_1}{\varepsilon}\mathrm{d}x' = \fint_{\Omega_{\varepsilon,\delta}} |\omega - \omega_0|^p \to 0 \quad (1.15)
\end{aligned}
$$

as soon as $p \leq \frac{d}{d-1}$, due to Lemma 1.2 and Lemma A.3 exactly as before. This confirms that $\bar{\gamma}$ remains indeed of order $\omega_0 = \mathcal{O}(1)$ in any $L^p, p \leq \frac{d}{d-1}$. Fixing any $M > \gamma_0$, we define now

$$
\underline{H}'_M(z) := H'(z)\chi_{(0,M]}(z) \qquad \text{and} \qquad \overline{H}'_M(z) := H'(z)\chi_{(M,\infty)}(z),
$$

the lower and upper truncations of $H'(z) = \log z$. We split next, after changing variables $x' = \delta y'$ and writing $\phi^\delta(y') = \phi(\delta y')$ for any relevant function $\phi = \gamma, \tilde{\gamma}, \hat{\gamma}, \bar{\gamma}$,

$$
\begin{aligned}
B = \int_{\Gamma_\delta} H'(\tilde{\gamma})[\gamma - \tilde{\gamma}] &= -|\Gamma_\delta|\fint_{\Gamma_\delta} H'(\gamma + \varepsilon\theta\bar{\gamma})\varepsilon\theta\bar{\gamma} = -\theta\varepsilon|\Gamma_\delta|\fint_{B_1^{d-1}} H'(\gamma^\delta + \varepsilon\theta\bar{\gamma}^\delta)\bar{\gamma}^\delta \\
&= \theta|\Omega_{\varepsilon,\delta}|\left(\underbrace{\fint_{B_1^{d-1}} -\underline{H}'_M(\gamma^\delta + \varepsilon\theta\bar{\gamma}^\delta)\bar{\gamma}^\delta}_{:=\underline{B}} - \underbrace{\fint_{B_1^{d-1}} \overline{H}'_M(\gamma^\delta + \varepsilon\theta\bar{\gamma}^\delta)\bar{\gamma}^\delta}_{:=\overline{B}}\right)
\end{aligned}
$$

Let us first handle the $\underline{B}$ term. By Lebesgue's differentiation theorem we have

$$
\fint_{B_1^{d-1}} |\gamma^\delta - \gamma_0| = \fint_{B_\delta^{d-1}} |\gamma - \gamma_0| \to 0,
$$

11

in particular $\gamma^\delta \to \gamma_0$ a.e. in B_1^{d-1}. Owing to (1.15) we have similarly

$$\fint_{B_1^{d-1}} |\bar\gamma^\delta - \omega_0| = \fint_{B_\delta^{d-1}} |\bar\gamma - \omega_0| \le \fint_{\Omega_{\varepsilon,\delta}} |\omega - \omega_0| \to 0,$$

hence $\bar\gamma^\delta \to \omega_0$ a.e as well. As a consequence $-\underline{H}'_M(\gamma^\delta + \varepsilon\theta\bar\gamma^\delta)\bar\gamma^\delta \to -\underline{H}'_M(\gamma_0)\omega_0 = -H'(\gamma_0)\omega_0$ at least a.e. inside B_1^{d-1} (here it is crucial that the truncation in $\underline{H}'_M$ was chosen at level $M > \gamma_0$ so that $\underline{H}'_M(\gamma_0) = H'(\gamma_0)$). Since $H'(M) - \underline{H}'_M(z) \ge 0$ for fixed M (this is mainly why we had to truncate) we conclude by Fatou's lemma that

$$\liminf_{\varepsilon,\delta,\theta} \underline{B} = \liminf_{\varepsilon,\delta,\theta} \left\{ \fint_{B_1^{d-1}} [H'(M) - \underline{H}'_M(\gamma^\delta + \varepsilon\theta\bar\gamma^\delta)] \bar\gamma^\delta + \fint_{B_1^{d-1}} -H'(M)\bar\gamma^\delta \right\}$$

$$= \liminf_{\varepsilon,\delta,\theta} \left\{ \fint_{B_1^{d-1}} [H'(M) - \underline{H}'_M(\gamma^\delta + \varepsilon\theta\bar\gamma^\delta)] \bar\gamma^\delta \right\} - H'(M) \lim_{\varepsilon,\delta,\theta} \fint_{B_1^{d-1}} \bar\gamma^\delta$$

$$\ge \fint_{B_1^{d-1}} [H'(M) - H'(\gamma_0)]\omega_0 - H'(M)\omega_0 = -H'(\gamma_0)\omega_0.$$

Note that this allows for $\gamma_0 = 0$ with $-(\log\gamma_0)\omega_0 = +\infty$, since in that case $\omega_0 > \gamma_0 = 0$.

For the $\overline{B}$ term we write

$$-\overline{B} = \fint_{B_1^{d-1}} \underbrace{\overline{H}'_M(\gamma^\delta + \varepsilon\theta\bar\gamma^\delta)}_{:=f} \underbrace{\bar\gamma^\delta}_{:=g} .$$

Exactly as before, and again because we truncated at level $M > \gamma_0$, we have convergence $\overline{H}'_M(\gamma^\delta + \varepsilon\theta\bar\gamma^\delta)\bar\gamma^\delta \to \overline{H}'_M(\gamma_0 + 0)\omega_0 = 0$ a.e. In order to apply Vitali's convergence theorem it suffices to show that the product $\{fg\}_{\varepsilon,\delta,\theta}$ is uniformly integrable in B_1^{d-1} as $\varepsilon,\delta,\theta \to 0$. Since $\overline{H}'_M(z) = (\log z)\chi_{(M,\infty)}(z)$ has been truncated one has that the negative part $[\overline{H}'_M]^-(z) \le \max\{0, -\log M\}$, which in turn easily yields

$$\exp\left(\left|\overline{H}'_M(z)\right|\right) \le C_M(1+z), \qquad \forall z \ge 0$$

for some suitable $C_M > 0$. Whence

$$\fint_{B_1^{d-1}} \exp f \le \fint_{B_1^{d-1}} C_M \left(1 + \gamma^\delta + \varepsilon\theta\bar\gamma^\delta)\right) = C_M \fint_{B_\delta^{d-1}} (1 + \gamma + \varepsilon\theta\bar\gamma) \to C_M [1 + \gamma_0 + 0].$$

Moreover by Lemma 1.2 we have $\omega \in W^{1,\frac{d}{d-1}}$. One more application of Lemma A.3 gives

$$\fint_{B_1^{d-1}} |g|^{\frac{d}{d-1}} = \fint_{B_\delta^{d-1}} |\bar\gamma|^{\frac{d}{d-1}} \to \omega_0^{\frac{d}{d-1}}.$$

We use now the elementary inequality $(ab)^p \le C\left[1 + \exp(a) + b^q\right]$ for $a, b \ge 0$ and any $1 \le p < q$ (for some $C = C_{p,r}$). Indeed, choosing any $p \in \left(1, \frac{d}{d-1}\right)$ we conclude from the above estimates that

$$\fint_{B_1^{d-1}} |fg|^p \le \fint_{B_1^{d-1}} C\left(1 + \exp(f) + |g|^{\frac{d}{d-1}}\right) = \mathcal{O}(1)$$

as $\varepsilon,\delta,\theta \to 0$, for some fixed $p > 1$. This gives equi-integrability and we can thus legitimately apply Vitali's convergence theorem to conclude that $\lim_{\varepsilon,\delta,\theta\to 0} \overline{B} = 0$. Whence

$$\liminf_{\varepsilon,\delta,\theta\to 0} \frac{B}{\theta|\Omega_{\varepsilon,\delta}|} \ge \liminf_{\varepsilon,\delta,\theta\to 0} \frac{\underline{B}}{\theta|\Omega_{\varepsilon,\delta}|} \ge -H'(\gamma_0)\omega_0. \tag{1.16}$$

Gathering (1.14)(1.16), and by strict convexity of H with $\omega_0 > \gamma_0$, we obtain

$$\liminf_{\varepsilon,\delta,\theta\to 0} \frac{\mathcal{E}(\rho) - \mathcal{E}(\tilde\rho)}{\theta|\Omega_{\varepsilon,\delta}|} \ge [H'(\omega_0) - H'(\gamma_0)]\omega_0 > 0. \tag{1.17}$$

12

Step 2: transportation cost. Recalling the definition (1.13) of our perturbation $\tilde{\rho} = \tilde{\omega} + \tilde{\gamma}$, it is easy to check that the plan $\pi \in \mathcal{P}\left(\overline{\Omega} \times \overline{\Omega}\right)$ defined by duality as

$$\iint_{\overline{\Omega} \times \overline{\Omega}} \varphi(x,y)\pi(\mathrm{d}x, \mathrm{d}y) = \int_{\overline{\Omega} \setminus \Omega_{\varepsilon,\delta}} \varphi(x,x)\rho(\mathrm{d}x) + (1-\theta)\int_{\Omega_{\varepsilon,\delta}} \varphi(x,x)\rho(\mathrm{d}x) + \theta \int_{\Omega_{\varepsilon,\delta}} \varphi(x,T(x))\rho(\mathrm{d}x)$$

is admissible from ρ to $\tilde{\rho}$. Since $|T(x) - x| \leq \varepsilon$ in $\Omega_{\varepsilon,\delta}$ we obtain

$$\mathcal{W}^2(\rho, \tilde{\rho}) \leq \iint_{\overline{\Omega}^2} |x-y|^2 \pi(\mathrm{d}x, \mathrm{d}y) = \theta \int_{\Omega_{\varepsilon,\delta}} |x - T(x)|^2 \rho(\mathrm{d}x)$$

$$= \theta \int_{\Omega_{\varepsilon,\delta}} |x - T(x)|^2 \omega(\mathrm{d}x) \leq \theta |\Omega_{\varepsilon,\delta}| \varepsilon^2 \fint_{\Omega_{\varepsilon,\delta}} \omega \sim \theta |\Omega_{\varepsilon,\delta}| \varepsilon^2 \omega_0. \quad (1.18)$$

Step 3: conclusion. Take ε small enough so that (1.12) and therefore (1.17)(1.18) hold. Recalling that $|\Omega_{\varepsilon,\delta}| = |B_1^{d-1}|\varepsilon\delta^{d-1}$ we obtain at last

$$|\partial\mathcal{E}(\rho)| = \limsup_{\mu \to \rho} \frac{[\mathcal{E}(\rho) - \mathcal{E}(\mu)]^+}{\mathcal{W}(\rho,\mu)} \geq \liminf_{\varepsilon,\delta,\theta \to 0} \frac{\theta|\Omega_{\varepsilon,\delta}|[H'(\omega_0) - H'(\gamma_0)]\omega_0}{(\theta|\Omega_{\varepsilon,\delta}|\varepsilon^2\omega_0)^{\frac{1}{2}}}$$

$$= \underbrace{\sqrt{\omega_0}[H'(\omega_0) - H'(\gamma_0)]}_{=C_0>0} \liminf_{\varepsilon,\delta,\theta \to 0} \frac{\sqrt{\theta|\Omega_{\varepsilon,\delta}|}}{\varepsilon} = C_0' \liminf_{\varepsilon,\delta,\theta \to 0} \left(\frac{\theta\delta^{d-1}}{\varepsilon}\right)^{\frac{1}{2}} = +\infty$$

provided $\varepsilon \ll \theta\delta^{d-1}$ - which is of course compatible with (1.12). This finally settles the case $\omega_0 > \gamma_0 \geq 0$.

If now $\gamma_0 > \omega_0 \geq 0$ we proceed similarly, by spreading uniformly along horizontal lines the excess of γ-mass from the boundary Γ_δ to the interior $\Omega_{\varepsilon,\delta}$. More explicitly the new perturbation is defined as

$$\tilde{\gamma}(x') := \begin{cases} (1-\theta)\gamma(x') & \text{if } x' \in \Gamma_\delta \\ \gamma(x') & \text{else} \end{cases} \quad \text{and} \quad \tilde{\omega}(x) := \begin{cases} \omega(x) + \theta\frac{1}{\varepsilon}\gamma(x') & \text{if } x = (x_1, x') \in \Omega_{\varepsilon,\delta} \\ \gamma(x) & \text{else} \end{cases}.$$

The rest of the argument is very similar and for the sake of brevity we omit the details. (This case is actually easier because the horizontal oscillations of the infinitesimal variations $\frac{1}{\varepsilon}\gamma(x')$ of ω are trivially under control, while they previously required some $W^{1,p}$ control via the Fisher information in order to apply Lemma A.3). $\qquad \square$

2 Discrete estimates

In this section we lay the groundwork for the convergence $\rho_t^\tau \to \rho_t$ later on. We start with

Lemma 2.1. *Given any $\rho^n \in \mathcal{P}(\overline{\Omega})$ there exists a unique minimizer $\rho^{n+1} = \omega^{n+1} + \gamma^{n+1}$ of (1.6), satisfying*

$$|\partial\mathcal{E}|(\rho^{n+1}) \leq \frac{\mathcal{W}(\rho^n, \rho^{n+1})}{\tau} < +\infty$$

with moreover $\omega^{n+1} \in W^{1,\frac{d}{d-1}}(\Omega), \gamma^{n+1} \in W^{1,\frac{d-1}{d-2}}(\Gamma)$ and the trace compatibility

$$\omega^{n+1}\big|_{\partial\Omega} = \gamma^{n+1}.$$

Proof. Recall first that $\mathcal{E}$ and $\mathcal{W}^2(\cdot, \rho^n)$ are classically nonnegative and lower semi-continuous w.r.t. narrow convergence. The latter is linearly convex, while the former is strictly convex. Hence the direct method in the calculus of variations gives existence and uniqueness of a minimizer ρ^{n+1} for the JKO functional $\mathcal{F}_\tau^n(\rho) := \frac{1}{2\tau}\mathcal{W}^2(\rho, \rho^n) + \mathcal{E}(\rho)$ in (1.6). Moreover this minimizer has finite entropy and therefore we can decompose $\rho^{n+1} = \omega^{n+1} + \gamma^{n+1}$ unambiguously.

13

Note that $\mathcal{F}_\tau^n$ implicitly defines a Moreau-Yosida regularization $\mathcal{E}_\tau$ of $\mathcal{E}$ [2, chapter 3], and more precisely $\mathcal{F}_\tau^n(\rho^{n+1}) = \mathcal{E}_\tau(\rho^n)$. As a consequence the slope estimate $|\partial\mathcal{E}|(\rho^{n+1}) \leq \frac{\mathcal{W}(\rho^n, \rho^{n+1})}{\tau}$ follows from standard properties of the Moreau-Yosida regularization, see e.g. [2, lemma 3.1.3]. The rest of the statement follows by Theorem 1 (the slope controls the Fisher informations and the trace condition) and Lemma 1.2 (properties of the Fisher information). $\square$

For a momentum $m \in \mathcal{M}^d(\overline{\Omega})$ and mass density $\rho \in \mathcal{P}(\overline{\Omega})$ we denote the quadratic kinetic energy functional

$$\mathcal{J}(m, \rho) := \int_{\overline{\Omega}} \frac{|m|^2}{\rho},$$

which is properly defined as the usual extension of to the space of measures of the integral with density $J(M, R) = \frac{|M|^2}{R}$ (by one-homogeneity).

The following lemma gives in some sense the Euler-Lagrange optimality condition for the minimizer ρ^{n+1} in each JKO step.

Lemma 2.2. *There exists an optimal plan $\pi^{n+1} \in \Pi(\rho^n, \rho^{n+1})$ such that the momentum $m^{n+1} \in \mathcal{M}^d(\overline{\Omega})$ defined through*

$$\int_{\overline{\Omega}} \xi(y) \cdot m^{n+1}(\mathrm{d}y) := \iint_{\overline{\Omega} \times \overline{\Omega}} \frac{y - x}{\tau} \cdot \xi(y) \pi^{n+1}(\mathrm{d}x, \mathrm{d}y), \qquad \forall \xi \in C_b(\overline{\Omega})^d \qquad (2.1)$$

can be written as

$$m^{n+1} = -\nabla \omega^{n+1} \mathcal{L}_\Omega^d - \nabla_\Gamma \gamma^{n+1} \mathcal{L}_\Gamma^{d-1} - \mathfrak{n}_\Gamma^{n+1}, \qquad (2.2)$$

for some normal measure $\mathfrak{n}_\Gamma^{n+1} \in \mathcal{M}^d(\Gamma)$. Moreover we have

$$\mathcal{I}(\omega^{n+1}) + \mathcal{I}(\gamma^{n+1}) \leq \mathcal{J}(m^{n+1}, \rho^{n+1}) \leq \frac{1}{\tau^2} \mathcal{W}^2(\rho^n, \rho^{n+1}) \qquad (2.3)$$

Let us stress here that $\pi^{n+1} \in \Pi(\rho^n, \rho^{n+1})$ is a "forward-in time" plan. Accordingly, the term $\frac{y-x}{\tau}$ in (2.1) should be seen as the terminal velocity of the time-τ geodesic from x to y, a tangent vector at y. This convention for the discrete momentum is chosen so that $\partial_t \rho + \operatorname{div} m = 0$ in the end (after taking $\tau \to 0$), and this is consistent with the minus sign in (2.2) ultimately leading to forward-in time diffusion. Let us also anticipate on the fact that the normal measure $\mathfrak{n}_\Gamma = \lim_{\tau \to 0} \mathfrak{n}_\Gamma^{n+1}$ will vanish in the end (we were not able to prove that $\mathfrak{n}_\Gamma^{n+1} = 0$ already at the discrete level) so the key point in (2.2) is to identify the interior contribution $\nabla \omega$ and the tangential component $\nabla_\Gamma \gamma$ of the boundary part.

Proof. Fix $\varepsilon > 0$ and let $\Gamma^\varepsilon = \{x \in \Omega, \mathsf{d}(x, \Gamma \leq \varepsilon)\}$ be an interior ε-neighborhood of $\Gamma = \partial\Omega$. We let $\Omega^\varepsilon = \Omega \setminus \Gamma^\varepsilon$, which is again a smooth domain if $\varepsilon > 0$ is small enough. On Γ^ε one can use a system of normal and tangential coordinates, which we avoid defining explicitly for the sake of clarity. Using this system of coordinates it is easy to define a smooth potential $V^\varepsilon : \overline{\Omega} \to \mathbb{R}$ which is constant inside Ω^ε, only depends on the normal coordinate inside Γ^ε, and most importantly satisfies

$$\mu^\varepsilon := e^{-V^\varepsilon(x)} \mathcal{L}_\Omega^d \overset{*}{\rightharpoonup} \mu = \mathcal{L}_\Omega^d + \mathcal{L}_\Gamma^{d-1}$$

as $\varepsilon \to 0$ (one can simply use a standard ε-mollification of the one-dimensional Dirac-delta in the normal coordinate). We consider the approximate problem

$$\rho^\varepsilon = \underset{\rho \in \mathcal{P}(\overline{\Omega})}{\operatorname{Argmin}} \left\{ \frac{1}{2\tau} \mathcal{W}^2(\rho, \rho^n) + \mathcal{H}(\rho \,|\, \mu^\varepsilon) \right\}.$$

Just as in Lemma 2.1 there exists a unique, absolutely continuous minimizer $\rho^\varepsilon = \rho^\varepsilon(x) \mathcal{L}_\Omega^d$. By standard arguments, and because $\mu^\varepsilon \overset{*}{\rightharpoonup} \mu$, the functional $\mathcal{F}_{\tau,\varepsilon}^n := \frac{1}{2\tau} \mathcal{W}^2(\cdot, \rho^n) + \mathcal{H}(\cdot \,|\, \mu^\varepsilon)$

14

Gamma-converges to the JKO functional $\mathcal{F}_\tau^n$ as $\varepsilon \to 0$. As a consequence ρ^ε converges narrowly to the unique minimizer ρ^{n+1} of $\mathcal{F}_\tau^n$ as $\varepsilon \to 0$.

For $\varepsilon > 0$ the minimizer ρ^ε is absolutely continuous, hence by the Brenier-McCann theorem there exists a unique (forward) optimal plan $\pi^\varepsilon = (T^\varepsilon, \mathrm{id})_{\#}\rho^\varepsilon$ from ρ^n to ρ^ε, induced by the unique (backward) optimal map T^ε from ρ^ε to $\rho^n = T^\varepsilon_{\#}\rho^\varepsilon$. Standard stability of optimal plans [34, thm. 1.52] gives that $\pi^\varepsilon \overset{*}{\rightharpoonup} \pi^{n+1}$ as $\varepsilon \to 0$ for some optimal plan $\pi^{n+1} \in \Pi_{opt}(\rho^n, \rho^{n+1})$. The momentum m^ε, defined by duality as

$$\int_{\overline{\Omega}} \xi(y) \cdot m^\varepsilon(\mathrm{d}y) := \iint_{\overline{\Omega} \times \overline{\Omega}} \frac{y - x}{\tau} \cdot \xi(y) \pi^\varepsilon(\mathrm{d}x, \mathrm{d}y), \qquad \forall \xi \in C_b(\overline{\Omega})^d$$

therefore converges narrowly to m^{n+1} similarly defined by (2.1). For $\varepsilon > 0$, by-now fairly standard arguments [34, chapter 8] give the Euler-Lagrange optimality condition for the ε-problem in the form

$$m^\varepsilon(\mathrm{d}y) = \frac{y - T^\varepsilon(y)}{\tau} \rho^\varepsilon(\mathrm{d}y) = -\rho^\varepsilon \nabla \left(\log \rho^\varepsilon + V^\varepsilon\right) = -\nabla \rho^\varepsilon - \rho^\varepsilon \nabla V^\varepsilon \qquad (2.4)$$

with a slight abuse of notations. We have moreover

$$\mathcal{J}(m^\varepsilon, \rho^\varepsilon) = \int_{\overline{\Omega}} \frac{|m^\varepsilon|^2}{\rho^\varepsilon} = \int_{\overline{\Omega}} \left|\frac{T^\varepsilon(y) - y}{\tau}\right|^2 \rho^\varepsilon(\mathrm{d}y) = \frac{\mathcal{W}^2(\rho^\varepsilon, \rho^n)}{\tau^2} \xrightarrow[\varepsilon \to 0]{} \frac{\mathcal{W}^2(\rho^{n+1}, \rho^n)}{\tau^2}.$$

Since $(\rho^\varepsilon, m^\varepsilon) \overset{*}{\rightharpoonup} (\rho^{n+1}, m^{n+1})$ this gives, by lower semi-continuity of $\mathcal{J}$,

$$\mathcal{J}(m^{n+1}, \rho^{n+1}) \leq \liminf_{\varepsilon \to 0} \mathcal{J}(m^\varepsilon, \rho^\varepsilon) = \frac{\mathcal{W}^2(\rho^{n+1}, \rho^n)}{\tau^2} < \infty.$$

In particular by standard properties of $\mathcal{J}$ we have $|m^{n+1}| \ll \rho^{n+1} = \omega^{n+1} + \gamma^{n+1} \in L^1(\Omega) + L^1(\Gamma)$, and this defines unambiguously the interior and boundary parts

$$m^{n+1} = m^{n+1} \llcorner \Omega + m^{n+1} \llcorner \Gamma =: m_\Omega^{n+1} + m_\Gamma^{n+1}.$$

Recall from Lemma 2.1 that $\omega^{n+1}, \gamma^{n+1}$ have $W^{1,p}$ Sobolev regularity for some $p > 1$, and observe that, owing to $\mathcal{J}(m^{n+1}, \rho^{n+1}) < \infty$ and $\rho^{n+1} = \omega^{n+1} + \gamma^{n+1}$, we have absolute continuity $m_\Omega^{n+1} \ll \mathcal{L}_\Omega^d$ and $m_\Gamma^{n+1} \ll \mathcal{L}_{\partial\Omega}^{d-1}$. We first claim that $m_\Omega^{n+1} = -\nabla \omega^{n+1} \mathcal{L}_\Omega^d$. To this end, note that m^{n+1} has finite mass since

$$|m^{n+1}| = \left|\lim_{\varepsilon \to 0} m^\varepsilon\right| \leq \liminf_{\varepsilon \to 0} |m^\varepsilon| = \liminf_{\varepsilon \to 0} \left(\int_{\overline{\Omega}} \left|\frac{y - T^\varepsilon(y)}{\tau}\right| \rho^\varepsilon\right)$$

$$\leq \liminf_{\varepsilon \to 0} \left(\int_{\overline{\Omega}} \left|\frac{T^\varepsilon(y) - y}{\tau}\right|^2 \rho^\varepsilon\right)^{\frac{1}{2}} = \liminf_{\varepsilon \to 0} \frac{\mathcal{W}(\rho^\varepsilon, \rho^n)}{\tau} = \frac{\mathcal{W}(\rho^{n+1}, \rho^n)}{\tau}.$$

Take any test-function $\xi \in C_c^1(\Omega)^d$ supported away from the boundary. Then

$$\int_\Omega \xi \cdot m_\Omega^{n+1} = \int_{\overline{\Omega}} \xi \cdot m^{n+1} = \lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \xi \cdot m^\varepsilon \overset{(2.4)}{=} -\lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \xi \cdot \left(\nabla \rho^\varepsilon + \rho^\varepsilon \nabla V^\varepsilon\right) = -\lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \xi \cdot \nabla \rho^\varepsilon,$$

where the last equality holds because for ε small enough ∇V^ε is supported in Γ^ε where ξ vanishes (being compactly supported away from the boundary). Integrating by parts gives

$$\int_\Omega \xi \cdot m_\Omega^{n+1} = -\lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \xi \cdot \nabla \rho^\varepsilon = \lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \mathrm{div}(\xi) \rho^\varepsilon = \int_{\overline{\Omega}} \mathrm{div}(\xi) \rho^{n+1} = \int_\Omega \mathrm{div}(\xi) \, \omega^{n+1},$$

where the last equality holds again because ξ is supported away from the boundary. Since m_Ω^{n+1} and ω^{n+1} are absolutely continuous this fully characterizes $m_\Omega^{n+1} = -\nabla \omega^{n+1} \mathcal{L}_\Omega^d$ as claimed.

For the more delicate boundary part let us decompose into tangential and normal components

$$m_\Gamma^{n+1} = \mathfrak{t}_\Gamma^{n+1} + \mathfrak{n}_\Gamma^{n+1}.$$

We need to show

$$\mathfrak{t}_\Gamma^{n+1} = -\nabla_\Gamma \gamma^{n+1} \mathcal{L}_\Gamma^{d-1}.$$

Take a tangential field $\xi \in C^1(\Gamma)^d$ and extend it to $\zeta \in C^1(\overline{\Omega})^d$ such that ζ remains tangential in an interior δ-neighborhood Γ^δ of $\Gamma = \partial\Omega$, for a small $\delta > 0$ fixed. We compute next

$$\int_\Gamma \xi \cdot m_\Gamma^{n+1} + \int_\Omega \zeta \cdot m_\Omega^{n+1} = \int_{\overline{\Omega}} \zeta \cdot m^{n+1} = \lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \zeta \cdot m^\varepsilon = -\lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \zeta \cdot (\nabla\rho^\varepsilon + \rho^\varepsilon \nabla V^\varepsilon).$$

Now, for fixed δ and small enough ε, we have by construction $\zeta \cdot \nabla V^\varepsilon = 0$ inside Γ^ε (where ∇V^ε is normal and ζ is tangential due to $\varepsilon < \delta$), while ∇V^ε vanishes identically outside of Γ^ε. Hence $\zeta \cdot \nabla V^\varepsilon \equiv 0$ in $\overline{\Omega}$, and since on the boundary $\zeta|_{\partial\Omega} = \xi$ is tangential we can integrate by parts the term $\int_\Omega \zeta \cdot \nabla\rho^\varepsilon$ for any fixed $\varepsilon > 0$ to get

$$\int_\Gamma \xi \cdot m_\Gamma^{n+1} + \int_\Omega \zeta \cdot m_\Omega^{n+1} = -\lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \zeta \cdot \nabla\rho^\varepsilon = \lim_{\varepsilon \to 0} \int_{\overline{\Omega}} \operatorname{div}(\zeta)\rho^\varepsilon$$

$$= \int_{\overline{\Omega}} \operatorname{div}(\zeta)\rho^{n+1} = \int_\Gamma \operatorname{div}(\zeta)\gamma^{n+1} + \int_\Omega \operatorname{div}(\zeta)\omega^{n+1}.$$

Because we just proved that $m_\Omega^{n+1} = \nabla\omega^{n+1}$, and since $\zeta|_{\partial\Omega} \cdot \nu = \xi \cdot \nu = 0$ on the boundary, we can cancel out $\int_\Omega \zeta \cdot m_\Omega^{n+1} = -\int_\Omega \zeta \cdot \nabla\omega^{n+1} = \int_\Omega \operatorname{div}(\zeta)\,\omega^{n+1}$ in the previous equality and we are left with

$$\int_\Gamma \xi \cdot m_\Gamma^{n+1} = \int_\Gamma \operatorname{div}(\zeta)\gamma^{n+1}.$$

Since the extension ζ is actually tangential in a neighborhood of $\partial\Omega$ the full divergence $\operatorname{div}(\zeta)|_{\partial\Omega}$ matches the intrinsic tangential divergence $\operatorname{div}_\Gamma(\xi)$, hence for any tangential field $\xi \in C^1(\Gamma)$

$$\int_\Gamma \xi \cdot \mathfrak{t}_\Gamma^{n+1} = \int_\Gamma \xi \cdot m_\Gamma^{n+1} = \int_\Gamma \operatorname{div}_\Gamma(\xi)\,\gamma^{n+1}.$$

Because Γ has no boundary this fully characterizes $\mathfrak{t}_\Gamma^{n+1} = -\nabla_\Gamma \gamma^{n+1}$.

Finally, since we known now that $-m^{n+1} = \nabla\omega^{n+1}\mathcal{L}_\Omega^d + \nabla_\Gamma\gamma^{n+1}\mathcal{L}_\Gamma^{d-1} + \mathfrak{n}_\Gamma^{n+1}$ we get for free that

$$\mathcal{I}(\omega^{n+1}) + \mathcal{I}(\gamma^{n+1}) = \int_\Omega \frac{|\nabla\omega^{n+1}|^2}{\omega^{n+1}} + \int_\Gamma \frac{|\nabla_\Gamma\gamma^{n+1}|^2}{\gamma^{n+1}}$$

$$\leq \int_\Omega \frac{|\nabla\omega^{n+1}|^2}{\omega^{n+1}} + \int_\Gamma \frac{|\mathfrak{n}_\Gamma^{n+1}|^2}{\gamma^{n+1}} + \int_\Gamma \frac{|\nabla_\Gamma\gamma^{n+1}|^2}{\gamma^{n+1}}$$

$$= \int_{\overline{\Omega}} \frac{|m^{n+1}|^2}{\rho^{n+1}} = \mathcal{J}(m^{n+1}, \rho^{n+1})$$

and the proof is complete. $\qquad\square$

Lemma 2.3. *Given any $\rho^n \in \mathcal{P}(\overline{\Omega})$ the unique minimizer ρ^{n+1} defined in Lemma 2.1 satisfies:*

- *if $\rho^n \geq \underline{c}\mu$ for some positive constant $\underline{c}$, then $\rho^{n+1} \geq \underline{c}\mu$;*
- *if $\rho^n \leq \overline{c}\mu$ for some positive constant $\overline{c}$, then $\rho^{n+1} \leq \overline{c}\mu$.*

Proof. We consider an approximation similar to the one presented in Lemma 2.2. Yet, not only we want to regularize μ into μ^ε, but we also need to modify ρ^n into a new measure $\rho^{n,\varepsilon}$ so that we preserve the inequality $\rho^{n,\varepsilon} \geq \underline{c}\mu^\varepsilon$ (resp. $\rho^{n,\varepsilon} \leq \overline{c}\mu^\varepsilon$). To do so, we can simply apply the heat semigroup in the domain Ω to both μ and ρ^n. Note that this construction was not used

16

in Lemma 2.2 because we needed a certain structure in the approximation, and in particular we wanted to guarantee $\mu^\varepsilon = e^{-V_\varepsilon}$ for a potential V_ε which is function of the distance to $\partial\Omega$, constant far enough from $\partial\Omega$, which does not happen for the heat flow. On the other hand, we did not need in Lemma 2.2 to modify ρ^n.

We then observe that we can apply [21, Lemma 2.4] which concerns one step of the JKO scheme for a functional of the form $\rho \mapsto \int F(\frac{\mathrm{d}\rho}{\mathrm{d}\mu})\mathrm{d}\mu$. We apply this to the approximated JKO functional where μ is replaced by μ^ε because [21] did not explicitly mention the case where μ is singular. We obtain the inequalities $\rho^\varepsilon \geq \underline{c}\mu^\varepsilon$ (resp., $\rho^\varepsilon \leq \overline{c}\mu^\varepsilon$) and we can take the limit $\varepsilon \to 0$ (the ε-JKO functional is easily checked to Γ-converge, hence the minimizers converge as well). $\qquad\square$

In order to retrieve later on a weak solution in the limit $\tau \to 0$ we define three very classical time interpolants, based on the discrete solutions $(\rho^n)_{n\geq 0}$. Writing for simplicity $t^n = n\tau$, the piecewise-constant interpolation $\overline{\rho}^\tau : [0,\infty) \to \mathcal{P}(\overline{\Omega})$ is

$$\overline{\rho}_t^\tau := \rho^{n+1}, \qquad \text{for } t \in (t^n, t^{n+1}]. \tag{2.5}$$

The piecewise geodesic interpolation $\rho^\tau : [0,\infty) \to \mathcal{P}(\overline{\Omega})$ is defined as

$$\rho_t^\tau := \left(\frac{t^{n+1} - t}{\tau}x + \frac{t - t^n}{\tau}y \right)_{\#}\pi^{n+1}, \qquad \text{for } t \in [t^n, t^{n+1}], \tag{2.6}$$

where $\pi^{n+1} \in \Pi(\rho^n, \rho^{n+1})$ is the optimal plan from Lemma 2.2. Finally, we also define *De Giorgi's variational interpolant* $\tilde{\rho}^\tau : [0,\infty) \to \mathcal{P}(\overline{\Omega})$ as

$$\tilde{\rho}_t^\tau := \underset{\rho\in\mathcal{P}(\overline{\Omega})}{\text{Argmin}} \left\{ \frac{1}{2r}\mathcal{W}^2(\rho, \rho^n) + \mathcal{E}(\rho) \right\} \qquad \text{for } t = t^n + r \in (t^n, t^{n+1}], \quad r \in (0,\tau]. \tag{2.7}$$

This is well-defined: Arguing exactly as in Lemma 2.1 there is a unique minimizer for all $r > 0$. Note that for $t = t^{n+1}$ we have $r = \tau$, in which case $\tilde{\rho}^\tau(t^{n+1}) = \rho^{n+1}$ is nothing but the next JKO step. Standard properties of the Moreau-Yosida regularization [2, lemma 3.1.2] guarantee that $\tilde{\rho}_t^\tau \to \rho^n$ as $t \searrow t^n \Leftrightarrow r \searrow 0$ in (2.7), thus $\tilde{\rho}_t^\tau$ indeed interpolates continuously between successive steps ρ^n, ρ^{n+1} of the JKO minimizing movement. Note that by definition all three interpolants also match at discrete times, i-e $\rho^\tau(t^n) = \overline{\rho}^\tau(t^n) = \tilde{\rho}^\tau(t^n) = \rho^n$. Moreover, since ρ^τ is piecewise geodesic we have that the metric speed

$$|\dot{\rho}_t^\tau|^2 = cst = \frac{\mathcal{W}^2(\rho^n, \rho^{n+1})}{\tau^2} \qquad \text{for } t \in [t^n, t^{n+1}].$$

Our last a-priori estimate is completely standard for minimizing movements in metric spaces and is just a discrete Energy Dissipation Inequality. Note that this involves both the geodesic interpolation ρ^τ and De Giorgi's interpolant $\tilde{\rho}^\tau$.

Lemma 2.4. *For any $\tau > 0$ and $t^N = N\tau$, $N \in \mathbb{N}$, there holds*

$$\mathcal{E}\left(\rho^\tau(t^N)\right) + \int_0^{t^N} \left(\frac{1}{2}|\dot{\rho}_t^\tau|^2 + \frac{1}{2}|\partial\mathcal{E}|^2(\tilde{\rho}_t^\tau) \right) \mathrm{d}t \leq \mathcal{E}(\rho_0). \tag{2.8}$$

In particular since $\mathcal{E} \geq 0$ one has

$$\sup_{\tau\to 0} \int_0^\infty \left(\frac{1}{2}|\dot{\rho}_t^\tau|^2 + \frac{1}{2}|\partial\mathcal{E}|^2(\tilde{\rho}_t^\tau) \right) \mathrm{d}t \leq \mathcal{E}(\rho_0) < +\infty. \tag{2.9}$$

Proof. For fixed $n \in \mathbb{N}$, standard properties of the Moreau-Yosida approximation give the one-step dissipation estimate

$$\mathcal{H}(\rho^{n+1}) + \frac{\mathcal{W}^2(\rho^n, \rho^{n+1})}{2\tau} + \frac{1}{2}\int_{t^n}^{t^{n+1}} |\partial\mathcal{E}|^2(\tilde{\rho}_t^\tau)\mathrm{d}t \leq \mathcal{E}(\rho^n),$$

17

see e.g. [32, lemma 7.5 and eq. (7.16)] or [2, theorem 3.1.4 with lemma 3.1.3]. Recall that by definition ρ^τ is the geodesic interpolation and therefore has constant speed $|\dot\rho_t^\tau|^2 = cst = \frac{\mathcal{W}^2(\rho^n, \rho^{n+1})}{\tau^2}$. Summing the previous inequality from $n = 0$ to $n = N - 1$, we get a telescopic sum for the $\mathcal{E}\left(\rho^{n+1}\right) - \mathcal{E}\left(\rho^n\right)$ term with by definition $\rho^N = \rho^\tau(t^N)$ and $\rho^0 = \rho_0$, and (2.8) follows. $\qquad\square$

3 Convergence to a dissipative weak solution

In this section we fully exploit the discrete estimates from Section 2 in order to prove our main result. For the ease of exposition we split the proof into separate statements. Our first result is a rather soft and generic convergence statement

Proposition 3.1. *There exists a discrete sequence $\tau \to 0$ (not relabeled for simplicity) and a continuous curve $\rho : [0, \infty) \to \mathcal{P}(\overline{\Omega})$ with $\rho(0) = \rho_0$ such that*

1. *the interpolants $\rho^\tau, \tilde\rho^\tau, \overline\rho^\tau$ converge locally uniformly towards the same limit ρ*

$$\sup_{t \in [0,T]} \left[\mathcal{W}(\rho_t, \overline\rho_t^\tau) + \mathcal{W}(\rho_t, \rho_t^\tau) + \mathcal{W}(\rho_t, \tilde\rho_t^\tau) \right] \xrightarrow[\tau \to 0]{} 0$$

   in any finite time interval $[0, T]$.

2. *we have the Energy Dissipation Inequality*

$$\mathcal{E}(\rho_T) + \int_0^T \left(\frac{1}{2} |\dot\rho_t|^2 + \frac{1}{2} \mathcal{D}(\rho_t) \right) \mathrm{d}t \leq \mathcal{E}(\rho_0), \qquad \forall\, T > 0. \tag{3.1}$$

Proof. First, owing to the Prokhorov theorem, $\mathcal{P}(\overline\Omega)$ is relatively compact for the narrow convergence. Since in bounded domains the Wasserstein distance metrizes the narrow convergence [38, thm 6.9] we have that $\{\rho_t^\tau\}_{\tau>0}$ lies in a fixed $\mathcal{W}$-relatively compact set for any fixed $t \geq 0$. By (2.9) we also have $\sup_{\tau \to 0} \int_0^T |\dot\rho_t^\tau|^2 \mathrm{d}t < +\infty$ for any fixed $T > 0$. In particular $\mathcal{W}(\rho_t^\tau, \rho_s^\tau) \leq C_T |t - s|^{\frac{1}{2}}$ in any finite interval. By the Ascoli-Arzelá theorem there is a discrete sequence such that ρ^τ converges $\mathcal{W}$-uniformly to a limit curve ρ. By diagonal extraction if needed we can assume that the limit does not depend on the finite horizon $T > 0$, hence $\rho^\tau \to \rho$ locally uniformly.

Fix now any $t \geq 0$, and let $n = \lfloor t/\tau \rfloor$ such that $t = t^n + r \in [t^n, t^{n+1})$ for some $r \in [0, \tau)$. Testing $\rho = \rho^n$ as a competitor in (2.7) and recalling that $\mathcal{E}(\tilde\rho_t^\tau) \geq 0$, we get the rough bound

$$\frac{1}{2r} \mathcal{W}^2(\tilde\rho_t^\tau, \rho^n) \leq \mathcal{E}(\tilde\rho_t^\tau) + \frac{1}{2r} \mathcal{W}^2(\tilde\rho_t^\tau, \rho^n) \leq \mathcal{E}(\rho^n),$$

so that

$$\mathcal{W}^2(\tilde\rho_t^\tau, \rho^n) \leq 2r\mathcal{E}(\rho^n) \leq 2\tau\mathcal{E}(\rho_0).$$

Similarly testing $\rho = \rho^n$ in (1.6), and because ρ^τ interpolates geodesically between ρ^n, ρ^{n+1}, one gets

$$\mathcal{W}^2(\rho_t^\tau, \rho^n) \leq \mathcal{W}^2(\rho^{n+1}, \rho^n) \leq 2\tau\mathcal{E}(\rho^n) \leq 2\tau\mathcal{E}(\rho_0).$$

By triangular inequality we get

$$\mathcal{W}(\tilde\rho_t^\tau, \rho_t^\tau) \leq \mathcal{W}(\tilde\rho_t^\tau, \rho^n) + \mathcal{W}(\rho^n, \rho_t^\tau) \leq 2\sqrt{2\tau\mathcal{E}(\rho_0)} \to 0$$

and therefore $\tilde\rho^\tau$ also converges uniformly to the same limit ρ as ρ^τ. For the same reason

$$\mathcal{W}(\overline\rho_t^\tau, \rho_t^\tau) = \mathcal{W}(\rho^{n+1}, \rho_t^\tau) \leq \mathcal{W}(\rho^{n+1}, \rho^n) \leq \sqrt{2\tau\mathcal{E}(\rho_0)} \to 0$$

and thus $\overline\rho^\tau$ also converges uniformly to ρ. This common limit trivially satisfies the initial condition in the sense that $\rho(0) = \lim \rho^\tau(0) = \lim \rho_0 = \rho_0$.

Finally let us turn into the dissipation inequality (3.1). For fixed $T > 0$ let $N = \lfloor T/\tau \rfloor$ and $t^N = N\tau$, so that $T \in [t^N, t^{N+1})$. By continuity of $t \mapsto \rho_t$ and arguing exactly as before we have $\mathcal{W}(\rho_T, \rho_{t^{N+1}}^\tau) \leq \mathcal{W}(\rho_T, \rho_T^\tau) + \mathcal{W}(\rho_T^\tau, \rho_{t^{N+1}}^\tau) \to 0$ and therefore by lower semi-continuity

$$\mathcal{E}(\rho_T) \leq \liminf_{\tau \to 0} \mathcal{E}(\rho_{t^{N+1}}^\tau). \tag{3.2}$$

For the dissipation terms, we note first from $T \leq t^{N+1}$ and $\mathcal{D} \leq |\partial \mathcal{E}|$ in Theorem 1 that

$$\int_0^T \left(\frac{1}{2} |\dot\rho_t^\tau|^2 + \frac{1}{2} \mathcal{D}(\tilde\rho_t^\tau) \right) \mathrm{d}t \leq \int_0^{t^{N+1}} \left(\frac{1}{2} |\dot\rho_t^\tau|^2 + \frac{1}{2} |\partial\mathcal{E}|^2 (\tilde\rho_t^\tau) \right) \mathrm{d}t$$

simply because $T \leq t^{N+1}$. By lower semicontinuity of the H^1 action $\rho \mapsto \int |\dot\rho_t|^2 \mathrm{d}t$ w.r.t. uniform convergence (as a supremum of continuous functions) we have that

$$\int_0^T |\dot\rho_t|^2 \mathrm{d}t \leq \liminf_{\tau \to 0} \int_0^T |\dot\rho_t^\tau|^2 \mathrm{d}t \leq \liminf_{\tau \to 0} \int_0^{t^{N+1}} |\dot\rho_t^\tau|^2 \mathrm{d}t. \tag{3.3}$$

By Fatou's lemma and the lower semicontinuity of $\mathcal{D}$ (Theorem 1) we get

$$\int_0^T \mathcal{D}(\rho_t) \mathrm{d}t \leq \int_0^T \liminf_{\tau \to 0} \mathcal{D}(\rho_t^\tau) \mathrm{d}t \leq \liminf_{\tau \to 0} \int_0^T \mathcal{D}(\rho_t^\tau) \mathrm{d}t \leq \liminf_{\tau \to 0} \int_0^{t^{N+1}} |\partial\mathcal{E}|^2 (\rho_t^\tau) \mathrm{d}t \tag{3.4}$$

Gathering (3.2)(3.3)(3.4), we finally retrieve

$$\mathcal{E}(\rho_T) + \int_0^T \left(\frac{1}{2} |\dot\rho_t|^2 + \frac{1}{2} \mathcal{D}(\rho_t) \right) \mathrm{d}t$$

$$\leq \liminf_{\tau \to 0} \mathcal{E}(\rho_{t^{N+1}}^\tau) + \frac{1}{2} \liminf_{\tau \to 0} \int_0^{t^{N+1}} |\dot\rho_t^\tau|^2 \mathrm{d}t + \frac{1}{2} \liminf_{\tau \to 0} \int_0^{t^{N+1}} |\partial\mathcal{E}|^2 (\rho_t^\tau) \mathrm{d}t$$

$$\leq \liminf_{\tau \to 0} \left\{ \mathcal{E}(\rho_{t^{N+1}}^\tau) + \int_0^{t^{N+1}} \left(\frac{1}{2} |\dot\rho_t^\tau|^2 \mathrm{d}t + \frac{1}{2} |\partial\mathcal{E}|^2 (\rho_t^\tau) \right) \mathrm{d}t \right\} \overset{(2.8)}{\leq} \mathcal{E}(\rho_0)$$

and the proof is complete. $\qquad\square$

So far this proof of convergence was purely abstract and relied solely on soft and fairly general metric considerations. The dissipation estimate (3.1), however, was retrieved in terms of the explicit functional $\mathcal{D}(\rho) = \mathcal{I}(\omega) + \mathcal{I}(\gamma)$ rather than the less tractable slope $|\partial\mathcal{E}|(\rho)$. Since we could not prove that $\mathcal{D} = |\partial\mathcal{E}|^2$ in Theorem 1, and because we could not establish a general upper chain rule, this is where metric arguments become limited. In order to retrieve the PDE we thus have to resort to more direct arguments, mainly based on the Euler-Lagrange optimality condition from Lemma 2.2.

For technical reasons we first need to improve the narrow convergence $\overline\rho_t^\tau \overset{*}{\rightharpoonup} \rho_t$ into a separate convergence of the interior and boundary parts.

Lemma 3.2. *For the same discrete sequence $\tau \to 0$ as in Proposition 3.1 there holds*

$$\overline\omega_t^\tau \rightharpoonup \omega_t \text{ in } L^1(\Omega) \quad \text{and} \quad \overline\gamma_t^\tau \rightharpoonup \gamma_t \text{ in } L^1(\Gamma) \qquad \text{for all } t \geq 0,$$

Proof. Fix any $t \geq 0$. From the entropy bounds $\mathcal{H}(\overline\omega_t^\tau) + \mathcal{H}(\overline\gamma_t^\tau) = \mathcal{E}(\overline\rho_t^\tau) \leq \mathcal{E}(\rho_0)$ in (2.8), for any subsequence τ_k there is a sub-subsequence τ_{k_l} such that $\overline\omega_t^{\tau_{k_l}} \rightharpoonup u$ and $\overline\gamma_t^{\tau_{k_l}} \rightharpoonup v$ in $L^1(\Omega), L^1(\Gamma)$ as $l \to \infty$, for some limits u, v. But since we already know that $\overline\rho_t^\tau \overset{*}{\rightharpoonup} \rho_t$ this shows that any such limits are necessarily $u = \rho_t \llcorner \Omega = \omega_t$ and $v = \rho_t \llcorner \partial\Omega = \gamma_t$. This classically implies that the whole sequence converges and concludes the proof. $\qquad\square$

Proposition 3.3. *The curve $\rho = \omega + \gamma$ from Proposition 3.1 solves*

$$\begin{cases} \partial_t \omega = \Delta \omega & \text{in } \Omega \\ \omega = \gamma & \text{on } \partial\Omega \\ \partial_t \gamma = \Delta_\Gamma \gamma - \partial_\nu \omega & \text{in } \partial\Omega \end{cases}$$

in the sense that $\omega_t|_{\partial\Omega} = \gamma_t$ for a.e. $t \geq 0$, $\omega \in L^1_{loc}([0,\infty); W^{1,p}(\Omega))$ and $\gamma \in L^1_{loc}([0,\infty); W^{1,q}(\Gamma))$ with $p, q > 1$ given by (1.9) (with respectively $d \geq 2$ and $d - 1 \geq 1$), and

$$\int_{\overline{\Omega}} \varphi \rho_{t_1} - \int_{\overline{\Omega}} \varphi \rho_{t_0} + \int_{t_0}^{t_1} \int_\Omega \nabla\varphi \cdot \nabla\omega + \int_{t_0}^{t_1} \int_\Gamma \nabla_\Gamma \varphi \cdot \nabla_\Gamma \gamma = 0 \qquad \forall \varphi \in C^1(\overline{\Omega}), \quad \forall 0 \leq t_0 < t_1 < \infty.$$

Proof. We first observe from the energy dissipation (3.1) that $\int_0^T \mathcal{D}(\rho_t)dt < \infty$. By Theorem 1 this immediately implies $\omega_t|_{\partial\Omega} = \gamma_t$ for a.e. $t \geq 0$ and, combined with (1.9), also entails the claimed Sobolev regularity for ω, γ.

Let us now focus on the weak formulation of the PDE. Recalling that the piecewise geodesic interpolation ρ_t^τ is defined by (2.6), we naturally define the corresponding momentum $m_t^\tau \in \mathcal{M}(\overline{\Omega})^d$ as follows. Let $\pi^{n+1} \in \Pi(\rho^n, \rho^{n+1})$ be the optimal plan from Lemma 2.2. For $x, y \in \overline{\Omega}$ we denote by

$$z_t^{xy} := \frac{t^{n+1} - t}{\tau} x + \frac{t - t^n}{\tau} y, \qquad t \in [t^n, t^{n+1}]$$

the time-τ geodesic from x to y. Setting

$$\int_{\overline{\Omega}} \xi(z) \cdot m_t^\tau(\mathrm{d}z) := \iint_{\overline{\Omega} \times \overline{\Omega}} \frac{y - x}{\tau} \cdot \xi(z_t^{xy}) \, \pi^n(\mathrm{d}x, \mathrm{d}y), \qquad \forall \xi \in C_b(\overline{\Omega})^d$$

and $m^\tau := m_t^\tau \mathrm{d}t \in \mathcal{M}^d([0,\infty) \times \overline{\Omega})$, a classical computation shows that (ρ^τ, m^τ) solve the continuity equation $\partial_t \rho^\tau + \operatorname{div} m^\tau = 0$ on $(0,\infty) \times \overline{\Omega}$ with no flux boundary conditions, i.e.

$$\int_{\overline{\Omega}} \varphi \, \rho_{t_1}^\tau - \int_{\overline{\Omega}} \varphi \, \rho_{t_0}^\tau + \int_{t_0}^{t_1} \int_{\overline{\Omega}} \nabla\varphi \cdot m^\tau = 0 \qquad \forall \varphi \in C^1(\overline{\Omega}), \quad \forall 0 \leq t_0 < t_1. \tag{3.5}$$

Moreover for any fixed $t \in [t^n, t^{n+1}]$ we have

$$\left| \int_{\overline{\Omega}} \xi(z) \cdot m_t^\tau(\mathrm{d}z) \right| \leq \left(\iint_{\overline{\Omega} \times \overline{\Omega}} \left| \frac{y - x}{\tau} \right| \pi^n(\mathrm{d}x, \mathrm{d}y) \right) \|\xi\|_{L^\infty(\overline{\Omega})}$$

$$\leq \left(\iint_{\overline{\Omega} \times \overline{\Omega}} \frac{|y - x|^2}{\tau^2} \pi^n(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{2}} \|\xi\|_{L^\infty(\overline{\Omega})}$$

$$= \frac{\mathcal{W}(\rho^n, \rho^{n+1})}{\tau} \|\xi\|_{L^\infty(\overline{\Omega})} = |\dot\rho_t^\tau| \, \|\xi\|_{L^\infty(\overline{\Omega})}.$$

As a consequence in any finite time interval $Q_T = [0, T] \times \overline{\Omega}$ we have

$$|m^\tau|(Q_T) = \int_0^T |m_t^\tau|(\overline{\Omega}) \, \mathrm{d}t \leq \int_0^T |\dot\rho_t^\tau| \, \mathrm{d}t \leq \sqrt{T} \left(\int_0^T |\dot\rho_t^\tau|^2 \, \mathrm{d}t \right)^{\frac{1}{2}} \leq C_T$$

uniformly in $\tau > 0$, where the last inequality follows from (2.9). This gives compactness $m^\tau \overset{*}{\rightharpoonup} m$ narrowly as $\tau \to 0$ in any finite time interval, for some limit m. Moreover the weak formulation (3.5) for $\partial_t \rho^\tau + \operatorname{div} m^\tau = 0$ immediately passes to the limit and (ρ, m) therefore also solve the same continuity equation.

Similarly, if m^{n+1} is the discrete momentum from Lemma 2.2, it is easy to see that the piecewise constant momentum $\overline{m}^\tau = \overline{m}_t^\tau \mathrm{d}t$ with

$$\overline{m}_t^\tau := m^{n+1} \qquad \text{for } t \in (t^n, t^{n+1}]$$

satisfies $|\overline{m}^\tau|(Q_T) \le C_T$ and therefore also converges to some $\overline{m}$. Exactly as in [34, lemma 8.9] we claim that $\overline{m} = m$. Indeed, fix any Lipschitz function $\xi : \overline{\Omega} \to \mathbb{R}^d$ and $t \in [0, T]$. For $t \in (t^n, t^{n+1}]$ the time-τ geodesic z_t^{xy} between x, y satisfies of course $|z_t^{xy} - y| \le |x - y|$, hence by definition of m and $\overline{m}$

$$\left| \int_{\overline{\Omega}} \xi \cdot [m_t^\tau - \overline{m}_t^\tau] \right| = \left| \iint_{\overline{\Omega} \times \overline{\Omega}} \frac{y - x}{\tau} \cdot [\xi(z_t^{xy}) - \xi(y)] \pi^{n+1}(\mathrm{d}x, \mathrm{d}y) \right|$$

$$\le \mathrm{Lip}(\xi) \iint_{\overline{\Omega} \times \overline{\Omega}} \frac{|y - x|^2}{\tau} \pi^{n+1}(\mathrm{d}x, \mathrm{d}y)$$

$$= \mathrm{Lip}(\xi) \frac{\mathcal{W}^2(\rho^n, \rho^{n+1})}{\tau} = \tau \, \mathrm{Lip}(\xi) |\dot{\rho}_t^\tau|^2.$$

As a consequence for any Lipschitz function $\zeta : [0, T] \times \overline{\Omega} \to \mathbb{R}^d$ we have

$$\left| \int_0^T \int_{\overline{\Omega}} \zeta \cdot [m^\tau - \overline{m}^\tau] \right| \le \tau \, \mathrm{Lip}(\zeta) \int_0^T |\dot{\rho}_t^\tau|^2 \, \mathrm{d}t \le C\tau \to 0,$$

and since $m^\tau \overset{*}{\rightharpoonup} m, \overline{m}^\tau \overset{*}{\rightharpoonup} \overline{m}$ the claim follows.

Finally, let us show that $m = -\nabla\omega \cdot \mathcal{L}_\Omega^d - \nabla_\Gamma\gamma \cdot \mathcal{L}_\Gamma^{d-1}$. For expediency we denote $m_\Omega = m \llcorner [0, \infty) \times \Omega$ and $m_\Gamma = m \llcorner [0, \infty) \times \Omega$. We first deal with the interior part. By Lemma 2.2 we have, for fixed $0 \le t_0 < t_1$ and any $\xi \in C_c^1(\Omega)$ compactly supported away from the boundary,

$$\int_{t_0}^{t_1} \int_\Omega \xi \cdot m_\Omega = \int_{t_0}^{t_1} \int_{\overline{\Omega}} \xi \cdot m = \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_{\overline{\Omega}} \xi \cdot \overline{m}^\tau$$

$$= \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_\Omega \xi \cdot \overline{m}_\Omega^\tau \overset{(2.2)}{=} - \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_\Omega \xi \cdot \nabla\overline{\omega}^\tau$$

$$= \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_\Omega \mathrm{div}(\xi) \overline{\omega}^\tau = \int_{t_0}^{t_1} \int_\Omega \mathrm{div}(\xi) \omega,$$

where the last equality is a straightforward application of Lebesgue's dominated convergence theorem together with the pointwise weak L^1 convergence $\overline{\omega}_t^\tau \rightharpoonup \omega_t$ for all $t \ge 0$ from Lemma 3.2 (and the bound $|\int_\Omega \mathrm{div}(\xi)\overline{\omega}_t^\tau| \le \|\mathrm{div}\,\xi\|_\infty$ uniformly in t). Since $t_0 < t_1$ and $\xi \in C_c^1(\Omega)$ were arbitrary this identifies the interior part $m_\Omega = -\nabla\omega$ at least in the distributional sense.

In order to identify now the boundary contribution $m_\Gamma = -\nabla_\Gamma\gamma$, observe first that $\partial_t\rho + \mathrm{div}\,\overline{m} = \partial_t\rho + \mathrm{div}\,m = 0$ with no-flux boundary condition. One should thus expect m_Γ to be tangential, otherwise mass would either exit or enter the domain through Γ. However the no-flux condition is enforced only in a weak form, similar to (3.5) with ρ, m in place of ρ^τ, m^τ. The argument therefore needs some care and for simplicity we defer it to Proposition A.2. Thus taking for granted that m_Γ is tangential, we only need to show equality $m_\Gamma = -\nabla_\Gamma\gamma$ in duality with tangential vector-fields. To this end take any such tangential $\xi \in C^1(\Gamma)$, and extend it to $\zeta \in C^1(\overline{\Omega})$. We have now

$$\int_{t_0}^{t_1} \int_\Gamma \xi \cdot m_\Gamma = \int_{t_0}^{t_1} \int_\Gamma \zeta|_\Gamma \cdot m_\Gamma = \int_{t_0}^{t_1} \int_{\overline{\Omega}} \zeta \cdot m - \int_{t_0}^{t_1} \int_\Omega \zeta \cdot m_\Omega$$

$$= \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_{\overline{\Omega}} \zeta \cdot \overline{m}^\tau - \int_{t_0}^{t_1} \int_\Omega \zeta \cdot \nabla\omega = \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_{\overline{\Omega}} \zeta \cdot \overline{m}^\tau + \int_{t_0}^{t_1} \int_\Omega \mathrm{div}(\zeta)\omega$$

$$= \lim_{\tau \to 0} \left\{ \int_{t_0}^{t_1} \int_{\overline{\Omega}} \zeta \cdot \overline{m}^\tau + \int_{t_0}^{t_1} \int_\Omega \mathrm{div}(\zeta)\overline{\omega}^\tau \right\}$$

$$\overset{(2.2)}{=} \lim_{\tau \to 0} \left\{ -\int_{t_0}^{t_1} \int_{\overline{\Omega}} \zeta \cdot [\nabla\overline{\omega}^\tau \mathcal{L}_\Omega^d + \nabla_\Gamma\overline{\gamma}^\tau \mathcal{L}_\Gamma^{d-1} + \overline{\mathfrak{n}}_\Gamma^\tau] + \int_{t_0}^{t_1} \int_\Omega \mathrm{div}(\zeta)\overline{\omega}^\tau \right\}$$

$$= -\lim_{\tau \to 0} \int_{t_0}^{t_1} \int_\Gamma \xi \cdot \nabla_\Gamma\overline{\gamma}^\tau = \lim_{\tau \to 0} \int_{t_0}^{t_1} \int_\Gamma \mathrm{div}_\Gamma(\xi)\overline{\gamma}^\tau = \int_{t_0}^{t_1} \int_\Gamma \mathrm{div}_\Gamma(\xi)\gamma.$$

The last equality follows again from the pointwise L^1 convergence $\overline{\gamma}_t^\tau \rightharpoonup \gamma_t$ in $L^1(\Gamma)$ for all $t \geq 0$ from Lemma 3.2 and Lebesgue's dominated convergence. Here we also heavily exploited that $\zeta|_\Gamma = \xi$ is tangential on the boundary in order to suitably integrate by parts in both directions, as well as to disregard the normal part $\zeta \cdot \overline{\mathfrak{n}}_\Gamma^\tau = 0$ of $\overline{m}^\tau$ along the boundary. This identifies $m_\Gamma = -\nabla_\Gamma \gamma$ in the distributional sense and the proof is complete. $\qquad\square$

In order to fully establish our main result we are still missing the long-time convergence (1.7) and the propagation of initial pointwise bounds. Observe first that, since $\bar{\mu} = \frac{1}{\mu(\Omega)}\mu$ is a simple renormalization to unit mass, our driving entropy is just a vertical shift $\mathcal{E}(\rho) = \mathcal{H}(\rho\,|\,\mu) = \mathcal{H}(\rho\,|\,\bar{\mu}) + C_\mu$ and the Energy dissipation (3.1) holds with $\mathcal{H}(\rho_T\,|\,\bar{\mu})$ and $\mathcal{H}(\rho_0\,|\,\bar{\mu})$ in place of $\mathcal{E}(\rho_T), \mathcal{E}(\rho_0)$. Moreover, Proposition 3.3 implies in particular that ρ solves the continuity equation $\partial_t \rho + \operatorname{div} m = 0$ with driving momentum $m_t = \nabla \omega_t \cdot \mathcal{L}_\Omega^d + \nabla_\Gamma \gamma_t \cdot \mathcal{L}_\Gamma^{d-1}$. By standard characterization of $\mathcal{W}_2$-absolutely continuous curves [2] we see that $\frac{1}{2}|\dot{\rho}_t|^2 \leq \mathcal{J}(m_t, \rho_t) = \frac{1}{2}\int_\Omega \frac{|m_t|^2}{\rho_t} = \frac{1}{2}\int_\Omega \frac{|\nabla \omega_t|^2}{\omega_t} + \frac{1}{2}\int_\Gamma \frac{|\nabla_\Gamma \gamma_t|^2}{\gamma_t} = \frac{1}{2}\mathcal{D}(\rho_t)$ for a.e. t. As a consequence we have

$$\mathcal{H}(\rho_T\,|\,\bar{\mu}) + \int_0^T \mathcal{D}(\rho_t)\mathrm{d}t \leq \mathcal{H}(\rho_0\,|\,\bar{\mu}), \qquad \forall T \geq 0.$$

We exploit next a *boundary trace Logarithmic-Sobolev inequality* from [6, section 4], which can be phrased in our precise framework as

$$\mathcal{H}(\rho\,|\,\bar{\mu}) \leq \frac{1}{2\lambda}\mathcal{D}(\rho), \qquad \forall \rho \in \mathcal{P}(\overline{\Omega})$$

for some $\lambda \equiv \lambda(\Omega) > 0$. This shows that

$$\mathcal{H}(\rho_T\,|\,\bar{\mu}) + 2\lambda \int_0^T \mathcal{H}(\rho_t\,|\,\bar{\mu})\mathrm{d}t \leq \mathcal{H}(\rho_0\,|\,\bar{\mu}), \qquad \forall T \geq 0$$

and immediately entails the first exponential decay in (1.7) by Grönwall's lemma. The decay in total variation is as usual a consequence of the general Csiszár-Kullback-Pinsker inequality $|a - b|_{TV} \leq \sqrt{\frac{1}{2}\mathcal{H}(a\,|\,b)}$ for arbitrary probability distributions a, b. Finally, the propagation of pointwise bounds $\rho_0 \leq \bar{c}\mu$ (resp. $\rho_0 \geq \underline{c}\mu$) is an easy consequence of Lemma 2.3, iterated along the JKO scheme.

A Appendix

Proposition A.1. *Let $\mathcal{X}$ be a smooth, compact Riemannian manifold, possibly with (smooth) boundary. Let $\mathcal{H}(\rho) = \int_\mathcal{X} \rho \log \rho$ and $\mathcal{I}(\rho) = \int_\mathcal{X} |\nabla \log \rho|^2 \rho$ be the entropy and Fisher information (relatively to the volume measure $\mathrm{d}x$), and denote $\mathcal{W}_\mathcal{X}$ the Wasserstein distance with quadratic cost $c(x, y) = \mathsf{d}_\mathcal{X}^2(x, y)$. For any $\rho \in \mathcal{P}(\mathcal{X})$ let $(\rho_t)_{t \geq 0}$ be the heat flow started from ρ (with no-flux boundary condition on $\partial \mathcal{X}$ if needed). Then*

$$\frac{1}{2}|\partial \mathcal{H}|^2(\rho) \geq \liminf_{t \to 0} \frac{1}{t}\left[\mathcal{H}(\rho) - \mathcal{H}(\rho_t) - \frac{1}{2t}\mathcal{W}_\mathcal{X}^2(\rho, \rho_t)\right] \geq \frac{1}{2}\mathcal{I}(\rho). \qquad (A.1)$$

Equality is actually expected to hold: At least formally this is clear because the heat flow is nothing but the Wasserstein gradient flow of the entropy [22]

$$\frac{d\rho_t}{dt} = -\operatorname{grad}_{\mathcal{W}_\mathcal{X}} \mathcal{H}(\rho_t) \qquad (A.2)$$

and therefore

$$\frac{1}{t}\left[\mathcal{H}(\rho) - \mathcal{H}(\rho_t) - \frac{1}{2t}\mathcal{W}_{\mathcal{X}}^2(\rho, \rho_t)\right]$$

$$= \frac{\mathcal{H}(\rho) - \mathcal{H}(\rho_t)}{t} - \frac{1}{2}\left(\frac{\mathcal{W}_{\mathcal{X}}(\rho, \rho_t)}{t}\right)^2 \underset{t \to 0}{\sim} - \frac{d}{dt}\mathcal{H}(\rho_t)\bigg|_{t=0} - \frac{1}{2}\left\|\frac{d\rho_t}{dt}\right\|_{t=0}^2$$

$$= -\left\langle \operatorname{grad}_{\mathcal{W}_{\mathcal{X}}} \mathcal{H}(\rho), \frac{d\rho_t}{dt}\bigg|_{t=0}\right\rangle - \frac{1}{2}\left\|\frac{d\rho_t}{dt}\right\|_{t=0}^2 \overset{(A.2)}{=} \left\|\operatorname{grad}_{\mathcal{W}_{\mathcal{X}}} \mathcal{H}(\rho)\right\|^2 - \frac{1}{2}\left\|\operatorname{grad}_{\mathcal{W}_{\mathcal{X}}} \mathcal{H}(\rho)\right\|^2$$

$$= \frac{1}{2}\left\|\operatorname{grad}_{\mathcal{W}_{\mathcal{X}}} \mathcal{H}(\rho)\right\|^2 = \frac{1}{2}|\partial\mathcal{H}|^2(\rho).$$

We wish to stress that nothing is really new or surprising here, but since we could not find anywhere in the literature a precise statement fitting our purpose we opted for giving a self-contained proof. This generic result is well documented in the case of λ-convex functionals, in particular equality in (A.1) is proved in [2, theorem 10.4.17] over the whole space in (possibly infinite-dimensional) Hilbert settings. For our current choice of the entropy $\mathcal{H}$ this convexity is well-known to be related to Ricci curvature lower bounds on the underlying manifold. Note that, owing to our assumption on $\mathcal{X}$, we always have such a uniform bound $\operatorname{Ric}_x \geq \lambda$ for some $\lambda \in \mathbb{R}$. The point is that the statement is independent of the precise value of this lower bound and requires no curvature assumption on the boundary, if any. This should be expected since all quantities in (A.1) are first order in nature, while curvature is a second order notion. However, this is more subtle than meets the eye. For, when $\mathcal{X} = \Omega \subset \mathbb{R}^d$ is a smooth, banana-shaped domain, the geodesic displacement convexity [28] of the entropy may completely fail ($\mathcal{H}$ may not be λ-convex for any $\lambda \in \mathbb{R}$, even very negative). In other words, in the presence of a negatively curved boundary, the connection between Ricci lower bounds and a modulus of displacement convexity is compromised. As a consequence the characterization of the metric slope in [2] does not apply here due to the possible lack of convexity. In fact this should not be a fundamental obstruction: displacement convexity is a global notion, while as already discussed our statement is local in nature (and in fact our proof below will solely rely on local PDE arguments). Another possible interpretation is as follows: In a purely metric context, λ-convexity can be seen as a quantitative regularity for the driving functional, and may therefore prove to be key for the abstract theory. However we are dealing here with the explicit heat flow: the latter of course enjoys many useful properties such as regularizing effects, which in turn can be exploited in this particular context to compensate for the lack of "metric regularity" (the λ-convexity). One may also hope to apply [3, thm. 7.6], but this requires checking a priori the lower semi-continuity of the (relaxed) metric slope, which is hard to achieve in the absence of any explicit representation (such as $|\partial\mathcal{H}|^2 = \mathcal{I}$, which is precisely at stake here.)

Proof. The first inequality in (A.1) is trivial owing to the *duality formula for the local slope* [2, lemma 3.1.5], which holds in arbitrary metric spaces and reads here

$$\frac{1}{2}|\partial\mathcal{H}|^2(\rho) = \limsup_{\tau \downarrow 0} \frac{1}{\tau} \sup_{\nu \in \mathcal{P}(\mathcal{X})} \left\{\mathcal{H}(\rho) - \mathcal{H}(\nu) - \frac{1}{2\tau}\mathcal{W}_{\mathcal{X}}^2(\rho, \nu)\right\}.$$

Thus we only focus on the second inequality.

Note first that, by usual properties of the heat flow, ρ_t is bounded away from zero and C^∞ smooth (up to the boundary if $\partial\mathcal{X} \neq \emptyset$) for any $t > 0$. The following computations and integration by parts are therefore completely legitimate as long as we step away from $t = 0$. Since $\operatorname{Dom}(|\partial\mathcal{H}|) \subset \operatorname{Dom}\mathcal{H}$, we can assume that $\mathcal{H}(\rho) < \infty$ and $\rho = \rho(y)\mathrm{d}y$ has density w.r.t. the volume measure $\mathrm{d}y$. Let $K_t(x, y)$ be the relevant heat kernel. Owing to the integral representation $\rho_t(x) = \int_{\mathcal{X}} K_t(x, y)\rho(y)\mathrm{d}y$ and the convexity of $z \mapsto H(z) = z \log z$ we have by

Jensen's inequality

$$\mathcal{H}(\rho_t) = \int_{\mathcal{X}} H(\rho_t(x)) \mathrm{d}x \leq \int_{\mathcal{X}} \left(\int_{\mathcal{X}} H(\rho(y)) K_t(x,y) \mathrm{d}y \right) \mathrm{d}x$$

$$= \int_{\mathcal{X}} H(\rho(y)) \left(\int_{\mathcal{X}} K_t(x,y) \mathrm{d}x \right) \mathrm{d}y = \int_{\mathcal{X}} H(\rho(y)) \mathrm{d}y = \mathcal{H}(\rho).$$

Moreover by continuity $\rho_t \to \rho$ and lower semi-continuity of $\mathcal{H}$ we also have $\mathcal{H}(\rho) \leq \liminf_{t \to 0} \mathcal{H}(\rho_t)$, hence in particular

$$\mathcal{H}(\rho_t) \to \mathcal{H}(\rho) < +\infty \qquad \text{as } t \downarrow 0.$$

Using the identity $\Delta \rho = \operatorname{div}(\rho \nabla \log \rho)$ (valid at least for smooth positive ρ), a very classical computation gives, for $t > 0$,

$$\frac{d}{dt}\mathcal{H}(\rho_t) = \frac{d}{dt} \int_{\mathcal{X}} \rho_t \log \rho_t = \int_{\mathcal{X}} (\log \rho_t + 1)\partial_t \rho_t = \int_{\mathcal{X}} (\log \rho_t + 1)\Delta \rho_t$$

$$= \int_{\mathcal{X}} (\log \rho_t + 1) \operatorname{div}(\rho_t \nabla \log \rho_t) = -\int_{\mathcal{X}} |\nabla \log \rho_t|^2 \rho_t = -\mathcal{I}(\rho_t).$$

Integrating from $t_0 > 0$ to $t > t_0$ and taking $t_0 \to 0$ with $\mathcal{H}(\rho_{t_0}) \to \mathcal{H}(\rho)$, we see that $t \mapsto \mathcal{I}(\rho_t)$ is integrable up to $t_0 = 0$ with

$$\frac{\mathcal{H}(\rho) - \mathcal{H}(\rho_t)}{t} = \lim_{t_0 \downarrow 0} \frac{\mathcal{H}(\rho_{t_0}) - \mathcal{H}(\rho_t)}{t} = \lim_{t_0 \downarrow 0} \frac{1}{t} \int_{t_0}^t \mathcal{I}(\rho_s) \mathrm{d}s = \frac{1}{t} \int_0^t \mathcal{I}(\rho_s) \mathrm{d}s. \qquad \text{(A.3)}$$

In order to estimate the squared Wasserstein distance, we observe that $\partial_s \rho_s = \Delta \rho_s = \operatorname{div}(\rho_s \nabla \log \rho_s)$, hence $(\rho_s, v_s) = (\rho_s, -\nabla \log \rho_s)$ solves the continuity equation

$$\partial_s \rho_s + \operatorname{div}(\rho_s v_s) = 0 \qquad \text{with} \qquad \int_0^t \int_{\mathcal{X}} |v_s|^2 \rho_s \mathrm{d}s = \int_0^t \mathcal{I}(\rho_s) \mathrm{d}s < \infty.$$

By the Benamou-Brenier formula [4, 26] we get

$$\frac{\mathcal{W}_{\mathcal{X}}^2(\rho, \rho_t)}{t^2} \leq \frac{1}{t} \int_0^t |v_s|^2 \rho_s = \frac{1}{t} \int_0^t \mathcal{I}(\rho_s) \mathrm{d}s \qquad \text{(A.4)}$$

Gathering (A.3)(A.4) we get altogether

$$\forall t > 0 : \qquad \frac{1}{t} \left[\mathcal{H}(\rho) - \mathcal{H}(\rho_t) - \frac{1}{2t} \mathcal{W}_{\mathcal{X}}^2(\rho, \rho_t) \right] \geq \frac{1}{2t} \int_0^t \mathcal{I}(\rho_s) \mathrm{d}s = \frac{1}{2} \int_0^1 \mathcal{I}(\rho_{tu}) \mathrm{d}u.$$

Fatou's lemma finally leads to

$$\liminf_{t \downarrow 0} \frac{1}{t} \left[\mathcal{H}(\rho) - \mathcal{H}(\rho_t) - \frac{1}{2t} \mathcal{W}_{\mathcal{X}}^2(\rho, \rho_t) \right] \geq \liminf_{t \downarrow 0} \frac{1}{2} \int_0^1 \mathcal{I}(\rho_{tu}) \mathrm{d}u$$

$$\geq \frac{1}{2} \int_0^1 \liminf_{t \downarrow 0} \mathcal{I}(\rho_{tu}) \mathrm{d}u \geq \frac{1}{2} \int_0^1 \mathcal{I}(\rho) \mathrm{d}u = \frac{1}{2}\mathcal{I}(\rho),$$

where the last inequality follows by lower semi-continuity of $\mathcal{I}$ combined with continuity of the heat flow $\rho_{tu} \to \rho_0 = \rho$ as $t \downarrow 0$ for a.e. $u \in (0,1)$. Note that this also covers the case $\mathcal{I}(\rho) = +\infty$. $\qquad \square$

Proposition A.2. *Let $\overline{\Omega} \subset \mathbb{R}^d$ be a smooth domain. Assume that (ρ, m) solves the continuity equation $\partial_t \rho + \operatorname{div} m = 0$ in $(0, T) \times \overline{\Omega}$ with no-flux boundary condition on $\partial\Omega$, in the sense that*

$$\int_{\overline{\Omega}} \phi \rho_{t_1} - \int_{\overline{\Omega}} \phi \rho_{t_0} = \int_{t_0}^{t_1} \int_{\overline{\Omega}} \nabla \phi \cdot m, \qquad \forall 0 \leq t_0 < t_1 \leq T \text{ and } \phi \in C^1(\overline{\Omega}) \qquad \text{(A.5)}$$

with $t \mapsto \rho_t$ being narrowly continuous. Then $m_\Gamma := m \llcorner \Gamma$ is tangential.
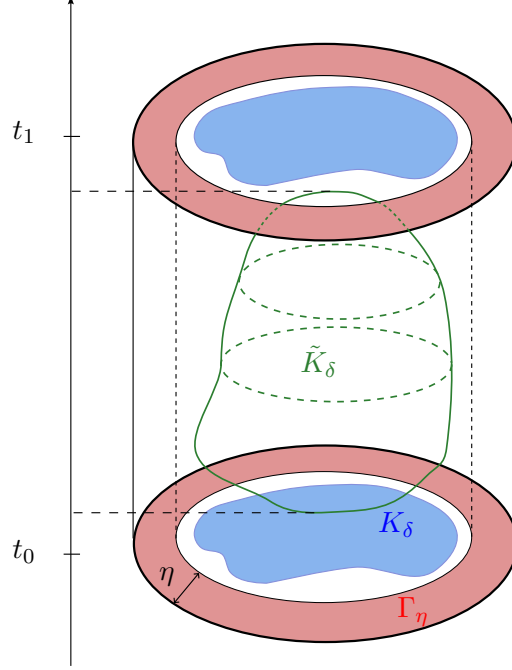
Figure 3: cylindrical exhaustion $(t_0, t_1) \times \Gamma_\eta$ for $\omega_{t_0}, \omega_{t_1}, |m_\Omega|$

This simply means that the no-flux boundary condition $m \cdot \nu = 0$, although encoded in a weak meaning in (A.5), can still be recovered in a strong sense for the boundary momentum $m_\Gamma \cdot \nu = 0$.

Proof. For notational convenience we write $m_\Omega := m \llcorner \Omega, m_\Gamma := m \llcorner \Gamma$ with $m = m_\Omega + m_\Gamma$, and also decompose m_Γ into tangential and normal components $m_\Gamma = \mathfrak{t}_\Gamma + \mathfrak{n}_\Gamma$. We need to show that $\mathfrak{n}_\Gamma = 0$.

Fix $0 \le t_0 < t_1 \le T$ and take a small $\delta > 0$. Being Ω a Polish space, the measure $\omega_{t_0} \in \mathcal{M}^+(\Omega)$ is inner regular and therefore here exists a compact set $K_\delta \subset\subset \Omega$ with $\omega_{t_0}(\Omega \setminus K_\delta) \le \delta$. Similarly, we can assume that $\omega_{t_1}(\Omega \setminus K_\delta) \le \delta$ at time $t = t_1$, and K_δ is at positive distance from $\partial\Omega$. Let now $Q = (t_0, t_1) \times \Omega$. As before, the measure $|m_\Omega|$ is inner regular and there exists a compact set $\tilde{K}_\delta \subset\subset Q$ such that $|m_\Omega|(Q \setminus \tilde{K}_\delta) \le \delta$. By compactness $\tilde{K}_\delta$ is at positive distance from $\partial Q = ([t_0, t_1] \times \partial\Omega) \cup (\{t_0\} \times \overline{\Omega}) \cup (\{t_1\} \times \overline{\Omega})$. In particular there exists $\eta > 0$ such that the smooth interior η-neighborhood $\Gamma_\eta = \{x \in \Omega : \mathsf{d}(x, \partial\Omega) < \eta\}$ of Γ satisfies $\Gamma_\eta \subset \Omega \setminus K_\delta$ and $(t_0, t_1) \times \Gamma_\eta \subset Q \setminus \tilde{K}_\delta$ – see Figure 3. Whence

$$\omega_{t_0}(\Gamma_\eta) \le \delta, \quad \omega_{t_1}(\Gamma_\eta) \le \delta, \quad |m_\Omega|((t_0, t_1) \times \Gamma_\eta) \le \delta.$$

Pick an arbitrary test-function $\psi \in C(\Gamma)$. Since Γ_η is smooth, for any small η we can find a function $\phi^\eta(x) \in C^1(\overline{\Omega})$ such that

(i) $\phi^\eta|_{\partial\Omega} = 0$

(ii) $\partial_\nu \phi^\eta|_{\partial\Omega} = \psi$

(iii) ϕ^η is supported in Γ_η

(iv) we have the bounds $\|\phi^\eta\|_\infty + \|\nabla\phi^\eta\| \le C = C_\psi$ uniformly in $\eta \to 0$.

Roughly speaking, ϕ^η is a sort of (possibly signed) bump function on the annular neighborhood Γ_η of Γ, vanishing on both the inner and outer boundaries, and having prescribed normal growth $\partial_\nu \phi^\eta = \psi$ across the outer boundary $\Gamma = \partial\Omega$. We want to test now $\phi = \phi^\eta$ in the weak

formulation of the continuity equation. Because ϕ^η vanishes on $\partial\Omega$ and outside of $\Gamma_\eta \subset \Omega$ we see that the left-hand side of (A.5)

$$\left| \int_{\overline{\Omega}} \phi^\eta \rho_{t_1} - \int_{\overline{\Omega}} \phi^\eta \rho_{t_0} \right| = \left| \int_{\Gamma_\eta} \phi^\eta \omega_{t_1} - \int_{\Gamma_\eta} \phi^\eta \omega_{t_0} \right| \leq \|\phi^\eta\|_\infty \left(\omega_{t_0}(\Gamma_\eta) + \omega_{t_1}(\Gamma_\eta) \right) \leq 2C_\psi \delta \to 0$$

as $\delta \to 0$. In the right-hand side, observe that by construction $\nabla \phi^\eta$ is normal on Γ with slope $\partial_\nu \phi^\eta = \psi$. Hence the tangential contribution is identically zero, and writing $\mathfrak{n}_\Gamma = \nu(x) \cdot n_\Gamma$ for some scalar measure $n_\Gamma \in \mathcal{M}((0,T) \times \Gamma)$ we get

$$\int_{t_0}^{t_1} \int_{\overline{\Omega}} \nabla \phi^\eta \cdot m = \int_{t_0}^{t_1} \int_{\Omega} \nabla \phi^\eta \cdot m_\Omega + \int_{t_0}^{t_1} \int_{\Gamma} \nabla \phi^\eta \cdot \mathbf{t}_\Gamma + \int_{t_0}^{t_1} \int_{\Gamma} \nabla \phi^\eta \cdot \mathfrak{n}_\Gamma$$

$$= \int_{t_0}^{t_1} \int_{\Omega} \nabla \phi^\eta \cdot m_\Omega + 0 + \int_{t_0}^{t_1} \int_{\Gamma} \partial_\nu \phi^\eta \, n_\Gamma = \int_{t_0}^{t_1} \int_{\Omega} \nabla \phi^\eta \cdot m_\Omega + \int_{t_0}^{t_1} \int_{\Gamma} \psi \, n_\Gamma.$$

As before the first term

$$\left| \int_{t_1}^{t_1} \int_{\Omega} \nabla \phi^\eta \cdot m_\Omega \right| = \left| \int_{t_1}^{t_1} \int_{\Gamma_\eta} \nabla \phi^\eta \cdot m_\Omega \right| \leq \|\nabla \phi^\eta\|_\infty |m_\Omega|((t_0,t_1) \times \Gamma_\eta) \leq C_\psi \delta \to 0.$$

Taking thus $\delta \to 0$ in $\int_{\overline{\Omega}} \phi^\eta \rho_{t_1} - \int_{\overline{\Omega}} \phi^\eta \rho_{t_0} = \int_{t_0}^{t_1} \int_{\overline{\Omega}} \nabla \phi^\eta \cdot m_t \mathrm{d}t$ we obtain

$$0 = \int_{t_0}^{t_1} \int_{\Gamma} \psi \, n_\Gamma$$

for all $\psi \in C(\Gamma)$ and $t_0 < t_1$. Whence $\mathfrak{n}_\Gamma = \nu(x) \cdot n_\Gamma \equiv 0$ and the proof is complete. $\qquad\square$

Lemma A.3. *In $\Omega = \mathbb{R}^d_+$ let us write $x = (x_1, x') \in \mathbb{R}^+ \times \mathbb{R}^{d-1}$. Fix $p > 1$ and take $\varepsilon, \delta \to 0$ with $\varepsilon^{p-1} = \mathcal{O}(\delta^{d-1})$. For $x_0 = (0, x_0') \in \partial\Omega$ let $\Omega_{\varepsilon,\delta} = \{x = (x_1, x') \in (0,\varepsilon) \times B_\delta^{d-1}(x_0')\} \subset \Omega$. Then for any $u \in W^{1,p}_{loc}(\overline{\Omega})$ and a.e. $x_0' \in \mathbb{R}^{d-1}$ there holds*

$$\fint_{\Omega_{\varepsilon,\delta}} |u(x) - u(x_0)|^p = \frac{1}{|\Omega_{\varepsilon,\delta}|} \int_{\Omega_{\varepsilon,\delta}} |u(x) - u(x_0)|^p \to 0 \quad \text{as} \quad \varepsilon, \delta \to 0.$$

Here by $u(x_0)$ we mean $[\mathrm{tr}\, u](x_0')$, the boundary trace of $u \in W^{1,p}$ evaluated at a point x_0' such that $x_0 = (0, x_0')$, which indeed makes sense for a.e. $x_0' \in \mathbb{R}^{d-1}$. This is nothing but a Lebesgue differentiation for boundary points. Compared to standard versions, such as [18, thm. 5.7], it is worth stressing that the shrinking sets $\Omega_{\varepsilon,\delta}$ in our statement may be very thin in the x_1 direction and can have unbounded eccentricity (a typical requirement for standard Lebesgue differentiation). We will typically use this with ε much smaller than δ, i-e averaging on very thin boundary layers.

Proof. For simplicity let us denote $v(x') = \mathrm{tr}\, u(x') = u(0, x')$ and $u_0 := u(x_0) = v(x_0') =: v_0$. Assume first that $u \in C^1(\overline{\Omega})$. Then for $x = (x_1, x') \in \Omega_{\varepsilon,\delta}$ we have by Jensen's inequality

$$|u(x_1, x') - v(x')|^p = \left| \int_0^{x_1} \partial_{x_1} u(z, x') \mathrm{d}z \right|^p \leq x_1^{p-1} \int_0^{x_1} |\partial_{x_1} u(z, x')|^p \mathrm{d}z \leq x_1^{p-1} \int_0^\varepsilon |\nabla u(z, x')|^p \mathrm{d}z.$$

Integrating in $x_1 \in (0, \varepsilon)$ and $x' \in B_\delta = B_\delta^{d-1}(x_0')$ we get

$$\int_{\Omega_{\varepsilon,\delta}} |u(x_1, x') - u(0, x')|^p \mathrm{d}x \leq \frac{\varepsilon^p}{p} \int_{\Omega_{\varepsilon,\delta}} |\nabla u|^p \mathrm{d}x.$$

By triangular inequality with $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ we obtain

$$\frac{1}{2^p} \int_{\Omega_{\varepsilon,\delta}} |u(x) - u_0|^p \leq \int_{\Omega_{\varepsilon,\delta}} |u(x_1, x') - v(x')|^p + \int_{\Omega_{\varepsilon,\delta}} |v(x') - v_0|^p$$

$$\leq \frac{\varepsilon^p}{p} \int_{\Omega_{\varepsilon,\delta}} |\nabla u|^p \mathrm{d}x + \varepsilon \int_{B_\delta} |v(x') - v_0|^p.$$

Dividing by $|\Omega_{\varepsilon,\delta}| = \varepsilon |B_\delta^{d-1}|$ we get

$$\frac{1}{2^p} \fint_{\Omega_{\varepsilon,\delta}} |u(x) - u_0|^p \leq \frac{\varepsilon^p}{p |B_1^{d-1}| \varepsilon \delta^{d-1}} \int_{\Omega_{\varepsilon,\delta}} |\nabla u|^p + \fint_{B_\delta} |v(x') - v_0|^p$$

and therefore, owing to our standing assumption that $\varepsilon^{p-1} = \mathcal{O}(\delta^{d-1})$,

$$\fint_{\Omega_{\varepsilon,\delta}} |u(x) - u_0|^p \leq C \left(\int_{\Omega_{\varepsilon,\delta}} |\nabla u|^p + \fint_{B_\delta} |v(x') - v_0|^p \right)$$

for $C > 0$ only depending on $p > 1$, the dimension d, and the upper bound on $\varepsilon^{p-1}/\delta^{d-1} = \mathcal{O}(1)$, but not on u. By density of $C^1(\overline{\Omega})$ in $W^{1,p}$ and continuity of the trace operator $\mathrm{tr} : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ the same holds for any $u \in W^{1,p}(\Omega)$. Now since $|\nabla u|^p \in L^1$ the first term converges to zero as soon as $|\Omega_{\varepsilon,\delta}| \to 0$, while the second term converges to zero simply by the Lebesgue differentiation theorem applied to $v = \mathrm{tr}\, u$ (here it is important that we chose $\Omega_{\varepsilon,\delta}$ with bounded eccentricity in the $\mathbb{R}^{d-1}$ tangential direction). $\qquad \square$

References

[1] Martial Agueh. Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory. *Advances in Differential Equations*, 10(3):309 – 360, 2005.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with ricci bounds from below. *Inventiones mathematicae*, 195(2):289–391, 2014.

[4] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[5] Adrien Blanchet, Vincent Calvez, and José A Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical patlak–keller–segel model. *SIAM Journal on Numerical Analysis*, 46(2):691–721, 2008.

[6] Marie Bormann, Max von Renesse, and Feng-Yu Wang. Functional inequalities for brownian motion on riemannian manifolds with sticky-reflecting boundary diffusion, 2023.

[7] Nawaf Bou-Rabee and Miranda C Holmes-Cerfon. Sticky brownian motion and its numerical solution. *SIAM review*, 62(1):164–195, 2020.

[8] Yann Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *CR Acad. Sci. Paris Sér. I Math.*, 305:805–808, 1987.

[9] José A Carrillo, Robert J McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matematica Iberoamericana*, 19(3):971–1018, 2003.

[10] Jean-Baptiste Casteras, Léonard Monsaingeon, and Luca Nenna. Large deviations and optimal transport for sticky-reflecting diffusions. *in preparation*, 2024.

[11] Ennio De Giorgi, Antonio Marino, and Mario Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.(8)*, 68(3):180–187, 1980.

[12] Simone Di Marino and Lénaïc Chizat. A tumor growth model of hele-shaw type as a gradient flow. *ESAIM: Control, Optimisation and Calculus of Variations*, 26:103, 2020.

[13] Hans-Jürgen Engelbert and Goran Peskir. Stochastic differential equations for sticky brownian motion. *Stochastics An International Journal of Probability and Stochastic Processes*, 86(6):993–1021, 2014.

[14] Torben Fattler, Martin Grothaus, and Robert Voßhall. Construction and analysis of a sticky reflected distorted brownian motion. 2016.

[15] William Feller. The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics*, pages 468–519, 1952.

[16] Alessio Figalli and Nicola Gigli. A new transportation distance between non-negative measures, with applications to gradients flows with dirichlet boundary conditions. *J. Math. Pures Appl.(9)*, 94(2):107–130, 2010.

[17] Gallouët, Thomas, Laborde, Maxime, and Monsaingeon, Léonard. An unbalanced optimal transport splitting scheme for general advection-reaction-diffusion problems. *ESAIM: COCV*, 25:8, 2019.

[18] LC Evans-RF Gariepy and LC Evans. Measure theory and fine properties of functions, revised edition. *Studies in Advanced Mathematics, CRC Press, Boca Raton, FL*, 2015.

[19] Annegret Glitzky and Alexander Mielke. A gradient structure for systems coupling reaction–diffusion effects in bulk and interfaces. *Zeitschrift für angewandte Mathematik und Physik*, 64(1):29–52, 2013.

[20] Martin Grothaus and Robert Voßhall. Stochastic differential equations with sticky reflection and boundary diffusion. 2017.

[21] Mikaela Iacobelli, Francesco S. Patacchini, and Filippo Santambrogio. Weighted ultrafast diffusion equations: From well-posedness to long-time behaviour. *Archive for Rational Mechanics and Analysis*, 232(3):1165–1206, 2019.

[22] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[23] Vitalii Konarovskyi, Victor Marx, and Max von Renesse. Spectral gap estimates for brownian motion on domains with sticky-reflecting boundary diffusion. *arXiv preprint arXiv:2106.00080*, 2021.

[24] Stanislav Kondratyev, Léonard Monsaingeon, and Dmitry Vorotnikov. A new optimal transport distance on the space of finite Radon measures. *Advances in Differential Equations*, 21(11/12):1117–1164, 2016.

[25] Giovanni Leoni. *A first course in Sobolev spaces*. American Mathematical Soc., 2017.

[26] Stefano Lisini. Characterization of absolutely continuous curves in wasserstein spaces. *Calculus of variations and partial differential equations*, 28(1):85–120, 2007.

[27] Daniel Matthes, Robert J McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations*, 34(11):1352–1397, 2009.

[28] Robert J McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

[29] Robert J McCann. Polar factorization of maps on riemannian manifolds. *Geometric & Functional Analysis GAFA*, 11(3):589–608, 2001.

[30] Léonard Monsaingeon. A new transportation distance with bulk/interface interactions and flux penalization. *Calculus of Variations and Partial Differential Equations*, 60(3):101, 2021.

[31] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. 2001.

[32] Mark A Peletier, Riccarda Rossi, Giuseppe Savaré, and Oliver Tse. Jump processes as generalized gradient flows. *Calculus of Variations and Partial Differential Equations*, 61:1–85, 2022.

[33] Goran Peskir. On boundary behaviour of one-dimensional diffusions: from Brown to Feller and beyond. *William Feller, Selected Papers II*, pages 77–93, 2015.

[34] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 55(58-63):94, 2015.

[35] Filippo Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bulletin of Mathematical Sciences*, 7:87–154, 2017.

[36] Juan Luis Vázquez and Enzo Vitillaro. Heat equation with dynamical boundary conditions of reactive–diffusive type. *Journal of Differential Equations*, 250(4):2143–2161, 2011.

[37] Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

[38] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

JEAN-BAPTISTE CASTERAS (`jeanbaptiste.casteras@gmail.com`).
CMAFcIO, FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DE LISBOA, EDIFICIO C6, PISO 1, CAMPO GRANDE 1749-016 LISBOA, PORTUGAL

LÉONARD MONSAINGEON (`lmonsaingeon@ciencias.ulisboa.pt`).
INSTITUT ÉLIE CARTAN DE LORRAINE, UNIVERSITÉ DE LORRAINE, SITE DE NANCY B.P. 70239, F-54506 VANDOEUVRE-LÈS-NANCY CEDEX, FRANCE
GRUPO DE FÍSICA MATEMÁTICA, DEPARTAMENTO DE MATEMÁTICA, INSTITUTO SUPERIOR TÉCNICO, AV. ROVISCO PAIS 1049-001 LISBOA, PORTUGAL

FILIPPO SANTAMBROGIO (`santambrogio@math.univ-lyon1.fr`).
UNIVERSITE CLAUDE BERNARD LYON 1, ICJ UMR5208, CNRS, ECOLE CENTRALE DE LYON, INSA LYON, UNIVERSITÉ JEAN MONNET, 69622 VILLEURBANNE, FRANCE.