WASSERSTEIN STABILITY OF POROUS MEDIUM-TYPE EQUATIONS ON MANIFOLDS WITH RICCI CURVATURE BOUNDED BELOW

NICOLÒ DE PONTI, MATTEO MURATORI, CARLO ORRIERI

ABSTRACT. Given a complete, connected Riemannian manifold $\mathbb{M}^n$ with Ricci curvature bounded from below, we discuss the stability of the solutions of a porous medium-type equation with respect to the 2-Wasserstein distance. We produce (sharp) stability estimates under negative curvature bounds, which to some extent generalize well-known results by Sturm [35] and Otto-Westdickenberg [32]. The strategy of the proof mainly relies on a quantitative L^1–L^∞ smoothing property of the equation considered, combined with the Hamiltonian approach developed by Ambrosio, Mondino and Savaré in a metric-measure setting [4].

CONTENTS

1. INTRODUCTION

In this paper we investigate the Cauchy problem for the following *porous medium-type* equation:

$$\begin{cases} \partial_t \rho = \Delta P(\rho) & \text{in } \mathbb{M}^n \times \mathbb{R}^+, \\ \rho(\cdot, 0) = \mu_0 \geq 0 & \text{in } \mathbb{M}^n \times \{0\}, \end{cases} \tag{1.1}$$

where μ_0 is a suitable finite, nonnegative Borel measure and P is a nonlinearity whose model case corresponds to $P(\rho) = \rho^m$ with $m > 1$, namely the *porous medium equation* (PME for short). Here $\mathbb{M}^n$ is a smooth, complete, connected, n-dimensional ($n \geq 2$) Riemannian manifold endowed with the standard Riemannian distance d and the Riemannian volume measure $\mathcal{V}$. We denote by Δ the Laplace-Beltrami operator on $\mathbb{M}^n$, which hereafter for simplicity will mostly be referred to as the "Laplacian". The initial datum μ_0 is assumed to belong to $\mathscr{M}_2^M(\mathbb{M}^n)$, namely the space of finite,

1

nonnegative Borel measures on $\mathbb{M}^n$ having mass M and finite second moment, that is

$$\mu_0(\mathbb{M}^n) = M \qquad \text{and} \qquad \int_{\mathbb{M}^n} \mathsf{d}(x, o)^2 \, \mathrm{d}\mu_0(x) < \infty$$

for some (hence all) $o \in \mathbb{M}^n$. As is well known, one can make $\mathscr{M}_2^M(\mathbb{M}^n)$ a complete metric space by endowing it with the 2-Wasserstein distance, which we will denote by $\mathcal{W}_2$ (see Subsection 3.2 for more details).

Our main focus is on a stability property of the evolution (1.1) with respect to $\mathcal{W}_2$, when $\mathbb{M}^n$ is possibly noncompact (with infinite volume) and its Ricci curvature is merely bounded from below. This is strongly motivated by the results obtained by Sturm [35] and Otto-Westdickenberg [32] under the nonnegativity assumption of the Ricci curvature, which we recall below. We point out that by "stability" we mean the possibility to control the $\mathcal{W}_2$-distance between two solutions of (1.1), along the flow, in terms of the $\mathcal{W}_2$-distance of the corresponding initial data. We will refer to this property as "contraction" when the $\mathcal{W}_2$-distance of the initial data cannot be increased by the flow.

To attack the problem we have at our disposal at least two different points of view. On the one hand, one can profit from the recent developments in the theory of nonlinear diffusion equations in non-Euclidean setting, where the connection with the geometry of the underlying structure is taken into account. On the other hand, the theory of optimal transportation can be employed to lift the problem to the space of measures endowed with the Wasserstein distance. The results obtained herein actually take advantage of the combination of techniques borrowed from both the two approaches.

For what concerns the analysis of nonlinear diffusion equations on Riemannian manifolds, we mention the following recent contributions. In [10] the authors consider well-posedness and finite-time extinction phenomena for the fast-diffusion equation (i.e. (1.1) with $P(\rho) = \rho^m$ for $m \in (0, 1)$) on Cartan-Hadamard manifolds, namely simply connected, complete Riemannian manifolds with nonpositive sectional curvature, for sufficiently integrable initial data. In the same geometric setting, in [21] the porous medium equation is investigated when initial data are finite Borel measures, by means of potential techniques. Still in the Cartan-Hadamard setting and for porous medium equation, in [20] the authors study well-posedness and blow-up phenomena for initial data possibly growing at infinity. The asymptotic behaviour for large times is addressed in [22, 23], complementing some results previously obtained in [38] in the hyperbolic space $\mathbb{H}^n$. Surprisingly, not much is known on the asymptotics of the *heat equation* in $\mathbb{H}^n$: we refer to [37] for an account of the state of the art along with some further progress.

With regards to the theory of optimal transport, after the seminal work of Otto et al. [25, 29] a lot of interest has been drawn in the description of certain PDEs as gradient flows in the space of probability measures endowed with the quadratic Wasserstein distance. In fact, when associated with a convex structure, such a formulation turns out to be extremely useful to obtain existence and stability results for a large class of PDEs. To that purpose, a very general theory of gradient flows of geodesically-convex functionals in metric spaces has rigorously been developed by Ambrosio, Gigli and Savaré: we refer to the monograph [1] for a comprehensive treatment of this topic.

Let us first briefly comment on the analysis of the heat equation (at first in $\mathbb{R}^n$), for which the picture is by now quite clear. By setting

$$\mathsf{E}(\mu) := \begin{cases} \int_{\mathbb{R}^n} \rho \log \rho \, \mathrm{d}x & \text{if } \mathrm{d}\mu = \rho(x)\mathrm{d}x \,, \\ +\infty & \text{elsewhere} \,, \end{cases}$$

that is the so-called *relative entropy*, and denoting by $\mathscr{P}_2(\mathbb{R}^n)$ the space of probability measures with finite second moment, the following holds: for every initial datum $\mu_0 \in \mathscr{P}_2(\mathbb{R}^n)$ there exists a unique gradient flow of E in $(\mathscr{P}_2(\mathbb{R}^n), \mathcal{W}_2)$ in the sense of Evolution Variational Inequalities (EVI), whose trajectories coincide with the corresponding solution of the heat equation. The Wasserstein

contraction property of the solutions is then a consequence of the *displacement convexity* of E in $(\mathscr{P}_2(\mathbb{R}^n), \mathcal{W}_2)$. This result was further extended to the Riemannian setting in [33], see also [41, 14], upon taking into account the Ricci curvature of the manifold $\mathbb{M}^n$: it is shown that the bound $\mathrm{Ric} \geq \lambda$ ($\lambda \in \mathbb{R}$) is equivalent to both the λ-convexity of the relative entropy and the following stability property of the generated gradient flow:

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq e^{-\lambda t} \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t \geq 0,$$

where the densities ρ and $\hat{\rho}$ represent the solutions of the heat equation on $\mathbb{M}^n$ starting from μ_0 and $\hat{\mu}_0$, respectively. An equivalence of this form is still missing in the context of nonlinear diffusion, where only partial results can be found in the literature.

As concerns the classical porous medium equation, a gradient flow interpretation was firstly treated in [29]. Then, numerous results have subsequently been obtained in the Euclidean setting even for more general PDEs. For instance, in [11] the authors quantify the Wasserstein contraction for diffusion equations that may also exhibit a nonlocal structure. In the one-dimensional case, contraction estimates for granular-media models are obtained in [26], by exploiting the explicit formulation of the Wasserstein distance. Regularizing effects and decay estimates for porous medium evolutions (with a nonlocal pressure) can be obtained by means of the minimizing movement approximation scheme in $(\mathscr{P}_2(\mathbb{R}^n), \mathcal{W}_2)$, as is shown in [27]. Finally, we refer to [9] for a simple proof of the equivalence between the contraction of the flow and the convexity condition, in which the gradient-flow structure of the problem is in fact not exploited. A related argument (coming from the probabilistic coupling method) can also be found in the recent manuscript [17].

As already mentioned above, for nonlinear diffusions the passage from the Euclidean to the Riemannian setting is not straightforward. The first contribution in this direction was given by Sturm in [35], where the equivalence between the geodesic convexity of the free energy and the curvature-dimension conditions is shown. In this setting, stability estimates for the PME on *nonnegatively* curved manifolds are still a consequence of the geodesic convexity of the free energy, thus complementing, when $\mathrm{Ric} \geq 0$, the results of [33] in the linear case. More precisely, the gradient-flow structure of the PME on $\mathscr{P}_2(\mathbb{M}^n)$ can be derived by introducing the free energy

$$\tilde{\mathsf{E}}(\mu) := \begin{cases} \int_{\mathbb{M}^n} U(\rho) \, \mathrm{d}\mathcal{V} & \text{if } \mathrm{d}\mu = \rho \, \mathrm{d}\mathcal{V}, \\ +\infty & \text{elsewhere}, \end{cases} \tag{1.2}$$

where U is linked to the nonlinearity of the equation through the relation $P(\rho) = \rho U'(\rho) - U(\rho)$. When $\mathbb{M}^n$ satisfies $\mathrm{Ric}_x \geq 0$ for every $x \in \mathbb{M}^n$, it is shown that under the additional assumption $\rho U'(\rho) \geq (1 - 1/n) U(\rho)$, the following contraction property holds along the flow:

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq 0 \qquad \forall t \geq 0.$$

Furthermore, the conditions on U and Ricci turn out to be also necessary for the contraction to hold, and they are equivalent to the displacement convexity of the functional $\tilde{\mathsf{E}}$:

$$\tilde{\mathsf{E}}(\mu^s) \leq (1 - s)\tilde{\mathsf{E}}(\mu^0) + s\tilde{\mathsf{E}}(\mu^1)$$

for every 2-Wasserstein geodesic $\{\mu^s\}_{0 \leq s \leq 1} \subset (\mathscr{P}_2(\mathbb{M}^n), \mathcal{W}_2)$.

Let us recall that the above argument was subsequently revisited in the compact setting by Otto and Westdickenberg in [32] through the so-called *Eulerian* calculus. Recent developments have also been obtained in [31] in the context of weighted Riemannian and Finsler manifolds.

Our main goal is to obtain stability estimates for the porous medium-type evolution (1.1) without imposing the nonnegativity of the Ricci curvature. To that purpose, we need to introduce some key hypotheses both on the manifold and on the form of the nonlinearity we consider.

First of all, we assume that $\mathbb{M}^n$ ($n \geq 3$) supports the following *Sobolev-type* inequality:

$$\|f\|_{L^{2^\star}(\mathbb{M}^n)} \leq C_S \left(\|\nabla f\|_{L^2(\mathbb{M}^n)} + \|f\|_{L^2(\mathbb{M}^n)} \right) \quad \forall f \in W^{1,2}(\mathbb{M}^n), \qquad 2^\star := \frac{2n}{n-2}, \qquad (1.3)$$

and has Ricci curvature bounded from below, that is

$$\mathrm{Ric}_x \geq -K \qquad \forall x \in \mathbb{M}^n \qquad (1.4)$$

for some constant $K \geq 0$, in the sense of quadratic forms. Note that (1.3) is guaranteed on any complete, n-dimensional ($n \geq 3$) Riemannian manifold satisfying (1.4) along with the *noncollapse* condition, see Section 2. The 2-dimensional case can also be dealt with by means of minor modifications: we refer to Remark 2.8. For what concerns the nonlinearity, we assume P to be a $C^1([0, +\infty))$, strictly increasing function satisfying $P(0) = 0$ and the two-sided bound

$$c_0 \, m \, \rho^{m-1} \leq P'(\rho) \leq c_1 \, m \, \rho^{m-1} \qquad \forall \rho \geq 0, \qquad (1.5)$$

for some $c_1 \geq c_0 > 0$ and $m > 1$. In fact the requirement $m > 1$ corresponds to the so-called *slow diffusion* regime. Furthermore, it will also be crucial to ask that P complies with the *McCann* condition

$$\rho \, P'(\rho) - \left(1 - \tfrac{1}{n}\right) P(\rho) \geq 0 \qquad \forall \rho \geq 0. \qquad (1.6)$$

Let us observe that the *pure* porous medium nonlinearity, namely $P(\rho) = \rho^m$, obviously complies with (1.5) and (1.6).

In our main result, that is Theorem 2.4, we show that under the above conditions problem (1.1) admits a unique solution in an appropriate weak sense (see again Section 2 for more details). Moreover, for any pair of initial data $\mu_0, \hat{\mu}_0 \in \mathscr{M}_2^M(\mathbb{M}^n)$, the corresponding solutions $\mu(t) = \rho(t)\mathcal{V}$ and $\hat{\mu}(t) = \hat{\rho}(t)\mathcal{V}$ have a (bounded) density for every $t > 0$ and satisfy the following stability estimate with respect to the 2-Wasserstein distance:

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq \exp\left\{ K \, c_1 \, \mathfrak{C}_m \left[\left(t M^{m-1}\right)^{\frac{2}{2+n(m-1)}} \vee \left(t M^{m-1}\right) \right] \right\} \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t > 0, \qquad (1.7)$$

where a semi-explicit form of the constant $\mathfrak{C}_m > 0$ is also given. Estimate (1.7) seems to be new in the context of diffusion equations on manifolds, due to the presence of a nonlinear time power in the exponent. Moreover, in Theorem 2.5 we exhibit a nontrivial example that shows that our estimate is indeed optimal (for small times). Precisely, in the n-dimensional hyperbolic space $\mathbb{H}_K^n$ of constant sectional curvature $-K$ (thus of Ricci curvature $-(n-1)K$), given two close enough points $x, y \in \mathbb{H}_K^n$ and the associated Dirac measures $\mu_0 = M\delta_x$, $\hat{\mu}_0 = M\delta_y$, there holds

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \geq \left[1 + K \, \kappa \left(t M^{m-1}\right)^{\frac{2}{2+n(m-1)}} \right] \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t \in (0, \bar{t}), \qquad (1.8)$$

for a suitable constant $\kappa = \kappa(n, m) > 0$ and a sufficiently small time $\bar{t} > 0$. As a consequence, we can deduce that the PME *is not a gradient flow* with respect to $\mathcal{W}_2$ on $\mathbb{H}_K^n$, or more generally on negatively-curved manifolds, in the sense of *Evolution Variational Inequalities*. We refer to Remark 2.6 for further details.

1.1. **Strategy.** The strategy we adopt has its roots in the so called *Eulerian* approach employed in [32, 13] and subsequently in [4]. Instead of relying on existence and smoothness of the optimal transport map, the main insight of the Eulerian approach is to directly work in the subspace of smooth densities and to take advantage of the Benamou-Brenier formulation of the Wasserstein distance. The basic idea is to link the contraction property of the Wasserstein distance to the monotonicity of the associated Lagrangian. Moreover, as is discussed in greater detail in [4], the contraction of the distance under the action of the flow is also equivalent to the monotonicity of the associated Hamiltonian functional (in the sense of Fenchel duality). Such equivalence turns out to be more convenient in the context of porous medium flows; we give here a flavor of the strategy,

referring to Section 5 for a more complete discussion. Let us start by writing the 2-Wasserstein distance as an action functional of the following form:

$$\frac{1}{2}\mathcal{W}_2(\rho_0, \hat{\rho}_0) = \inf \left\{ \int_0^1 \mathcal{L}(\rho^s, \tfrac{\mathrm{d}}{\mathrm{d}s}\rho^s)\,\mathrm{d}s :\ s \mapsto \rho^s \ \text{with}\ \rho^0 = \rho_0\,,\ \rho^1 = \hat{\rho}_0 \right\},$$

where

$$\mathcal{L}(\rho, w) = \frac{1}{2}\int_{\mathbb{M}^n} |\nabla\phi|^2 \rho\,\mathrm{d}\mathcal{V}\,, \qquad -\mathrm{div}(\rho\nabla\phi) = w \quad \text{in } \mathbb{M}^n\,.$$

Rather than looking directly at the Lagrangian $\mathcal{L}$, we consider the Hamiltonian functional

$$\mathcal{E}_\rho[\phi] := \frac{1}{2}\int_{\mathbb{M}^n} |\nabla\phi|^2 \rho\,\mathrm{d}\mathcal{V}\,.$$

If $\rho \equiv \rho(t)$ is a solution of (1.1) and $\phi \equiv \phi(t)$ is the solution of the corresponding *linearized* backward flow given by $\frac{\mathrm{d}}{\mathrm{d}t}\phi = -P'(\rho)\Delta\phi$, it is not difficult to check that, at least formally, there holds (see [4, Example 2.4])

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho(t)}[\phi(t)] = \int_{\mathbb{M}^n} P(\rho(t))\,\Gamma_2(\phi(t))\,\mathrm{d}\mathcal{V} + \int_{\mathbb{M}^n} \left[\rho(t)P'(\rho(t)) - P(\rho(t))\right](\Delta\phi(t))^2\,\mathrm{d}\mathcal{V}\,,$$

where Γ_2 is the iterated *carré du champ* operator, whose definition is provided in Subsection 3.1. By exploiting (1.6) and the Bakry-Émery formulation $\mathrm{BE}(0,n)$ of the curvature bound $\mathrm{Ric} \geq 0$ (we refer again to Subsection 3.1), one can deduce the monotonicity of the Hamiltonian along the flow, namely $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho(t)}[\phi(t)] \geq 0$, which is a key step in order to prove the 2-Wasserstein contraction property (see [4, Proposition 2.1] in a simplified framework).

However, in the present setting we are dealing with the more general case in which the Ricci curvature is merely bounded from below. As a consequence, by employing the Bakry-Émery formulation $\mathrm{BE}(-K,n)$, a priori we only have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho(t)}[\phi(t)] \geq -K \int_{\mathbb{M}^n} |\nabla\phi(t)|^2\,P(\rho(t))\,\mathrm{d}\mathcal{V}\,.$$

In order to compare $\rho(t)$ with $P(\rho(t))$, and therefore to close the above differential inequality, the crucial idea is now to exploit a *quantitative* $L^1(\mathbb{M}^n)$–$L^\infty(\mathbb{M}^n)$ smoothing estimate for weak energy solutions of (1.1), see Proposition 4.4. To that purpose, it is necessary to first understand problem (1.1) for more regular initial data, namely

$$\begin{cases} \partial_t \rho = \Delta P(\rho) & \text{in } \mathbb{M}^n \times \mathbb{R}^+\,, \\ \rho(\cdot, 0) = \rho_0 \geq 0 & \text{on } \mathbb{M}^n \times \{0\}\,, \end{cases} \tag{1.9}$$

where $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$; in fact, it will also be essential to deal with a nondegenerate regularization of the equation, which will be addressed in detail in Sections 4 and 5. We point out that smoothing effects are a very important and well-established tool in the theory of a large class of nonlinear diffusion equations: we refer the reader e.g. to the monograph [39]. This way we are able to integrate the differential inequality to get the estimate

$$\mathcal{E}_{\rho(t)}[\phi(t)] \geq \exp\{-K\,C(t,m,n)\}\,\mathcal{E}_{\rho_0}[\phi(0)]\,, \tag{1.10}$$

where an explicit computation of $C(t,m,n) > 0$ is available (see Lemma 5.3). The final step consists of exploiting the dual formulation of the Wasserstein distance for suitable regular curves, and we refer to Subsection 5.1 for a precise description of the strategy that allows one to pass from (1.10) to the stability estimate (1.7).

As for the optimality, we choose $\mathbb{M}^n$ as the hyperbolic space $\mathbb{H}^n_K$ of constant sectional curvature $-K$. The key ingredient to derive (1.8) is a delicate estimate on the Wasserstein distance between suitable radially-symmetric densities centered about two different (sufficiently close) points. To that purpose, we take advantage of a result originally proved by Ollivier [30] in the simpler case of

uniform densities, combined with the behaviour for small times of Barenblatt solutions of the PME in $\mathbb{H}_K^n$, obtained in [38]. All the rigorous computations are carried out in Subsection 5.4.

Let us point out that the extension of the present results to a metric-measure setting appears not to be straightforward, mainly due to the PDE techniques we employ in Section 4. Indeed, the proof of the L^1–L^∞ smoothing estimate, which is a crucial ingredient of our strategy, is not directly applicable. The point is that we take advantage of a uniformly *parabolic* regularization of problem (1.1) in smooth domains, whose analogue in the metric-measure framework is in principle not available (see Remark 4.9). Another key tool we use, in order to show that solutions starting from data in $\mathscr{M}_2^M(\mathbb{M}^n)$ stay in $\mathscr{M}_2^M(\mathbb{M}^n)$, is the so-called *compact-support* property, that we establish again by pure PDE methods (see Proposition 4.5). The counterpart of this result in metric-measure spaces should be investigated by a different approach.

1.2. **Notations.** Throughout, we will deal with a complete and connected Riemannian manifold $(\mathbb{M}^n, \mathfrak{g})$. In the sequel, for simplicity, we will omit the explicit dependence of the geometric quantities on the metric $\mathfrak{g}$. We denote by d the associated Riemannian distance and by $\mathcal{V}$ the Riemannian volume measure. The former, with some abuse of notation, will also be used to denote distance between sets. The symbol $T_x\mathbb{M}^n$ will stand for the tangent space at $x \in \mathbb{M}^n$, endowed with a scalar product $\langle \cdot, \cdot \rangle$ that induces the norm $|\cdot|$.

We define $\mathscr{M}(\mathbb{M}^n)$ as the space of finite, nonnegative Borel measures over $(\mathbb{M}^n, \mathsf{d})$ and $\mathscr{M}^M(\mathbb{M}^n)$ as the space of measures $\mu \in \mathscr{M}(\mathbb{M}^n)$ such that $\mu(\mathbb{M}^n) = M > 0$. If μ also has a finite second moment we write $\mu \in \mathscr{M}_2^M(\mathbb{M}^n)$, and we denote by $\mathcal{W}_2(\mu, \nu)$ the 2-Wasserstein distance between any two elements $\mu, \nu \in \mathscr{M}_2^M(\mathbb{M}^n)$. If the measures have densities w.r.t. $\mathcal{V}$, say ρ_μ and ρ_ν, we will often write $\mathcal{W}_2(\rho_\mu, \rho_\nu)$ in place of $\mathcal{W}_2(\mu, \nu)$.

For simplicity's sake, in the following we use the notations $\mathbb{H}$, $\mathbb{V}$ and $\mathbb{D}$ for the Hilbert spaces

$$\mathbb{H} := L^2(\mathbb{M}^n), \qquad \mathbb{V} := W^{1,2}(\mathbb{M}^n), \qquad \mathbb{D} := \{f \in \mathbb{V} : \Delta f \in \mathbb{H}\}, \tag{1.11}$$

with associated norms $\|f\|_{\mathbb{V}}^2 := \|f\|_{\mathbb{H}}^2 + \|\nabla f\|_{\mathbb{H}}^2$ and $\|f\|_{\mathbb{D}}^2 := \|f\|_{\mathbb{V}}^2 + \|\Delta f\|_{\mathbb{H}}^2$. It is useful to recall that, by an elementary cut-off argument (in case $\mathbb{M}^n$ is noncompact), it can be shown that $\mathbb{V}$ coincides with $W_0^{1,2}(\mathbb{M}^n)$, where the latter symbol denotes the closure of $C_c^\infty(\mathbb{M}^n)$ with respect to $\|\cdot\|_{\mathbb{V}}$.

Given $T > 0$ and two Hilbert spaces X and Y continuously embedded in a Banach space U, we introduce the space of time-dependent functions

$$W^{1,2}((0,T); X, Y) := \left\{u \in W^{1,2}((0,T); U) : u \in L^2((0,T); X), \ \tfrac{\mathrm{d}u}{\mathrm{d}t} \in L^2((0,T); Y)\right\},$$

with associated norm

$$\|u\|_{W^{1,2}((0,T);X,Y)}^2 := \|u\|_{L^2((0,T);X)}^2 + \left\|\tfrac{\mathrm{d}u}{\mathrm{d}t}\right\|_{L^2((0,T);Y)}^2.$$

Let $T > 0$. For any function $F \in C^1(\mathbb{R})$ with $F(0) = 0$, such that $0 < \lambda \le F'(r) \le \lambda^{-1}$ for every $r \in \mathbb{R}$, for some $\lambda > 0$, in agreement with [4, Section 3.3] we introduce the set

$$\mathcal{ND}_F(0, T) := \left\{u \in W^{1,2}((0,T); \mathbb{H}) \cap C^1([0,T]; \mathbb{V}') : F(u) \in L^2((0,T); \mathbb{D})\right\}.$$

As a general rule, we will use superscripts to denote the parameter of curves that are related to geodesics in the Wasserstein space over $(\mathbb{M}^n, \mathsf{d})$ and subscripts to denote the index or parameter of an approximation. Since subscripts are also typically used to refer to initial data of a Cauchy problem as in (1.1) or (1.9), we will try to avoid ambiguity as much as possible.

Finally, when referring to a function $\rho : D \subseteq \mathbb{M}^n \times \mathbb{R}^+ \to \mathbb{R}$ (or to a measure) evaluated at some time t *as a whole*, we will adopt the notation $\rho(t)$ (or $\mu(t)$). As for its time derivative, we will write $\frac{\partial \rho}{\partial t}$ whenever it can be understood as a classical partial derivative; we will write $\frac{\mathrm{d}\rho}{\mathrm{d}t}$ instead if it must be interpreted as the time derivative of ρ as a curve in a suitable Banach space. The notation $\dot{\rho}$ will mostly be used for *metric derivatives*.

List of main notations

$\mathbb{M}^n$	complete, connected, n-dimensional Riemannian manifold
$\mathbb{H}_K^n$	n-dimensional hyperbolic space with sectional curvature $-K$
$\mathcal{V}$	Riemannian volume measure on $\mathbb{M}^n$
Ric_x	Ricci curvature at $x \in \mathbb{M}^n$
$T_x\mathbb{M}^n$	tangent space at $x \in \mathbb{M}^n$
$\exp_x v$	exponential map at $x \in \mathbb{M}^n$ along $v \in T_x\mathbb{M}^n$
$\mathscr{P}(\mathbb{M}^n)$	Borel probability measures on $\mathbb{M}^n$
$\mathscr{P}_2(\mathbb{M}^n)$	Borel probability measures with finite quadratic moment
$\mathscr{M}_2^M(\mathbb{M}^n)$	nonnegative Borel measures with mass M and finite quadratic moment
$\mathcal{W}_2(\mu_0, \mu_1)$	Kantorovich-Rubinstein-Wasserstein distance
$C([0,T]; X)$	continuous curves from $[0,T]$ with values in the metric space X
$\text{Lip}([0,1]; X)$	Lipschitz curves from $[0,1]$ with values in the metric space X
$W^{k,p}(\mathbb{M}^n)$	standard Sobolev spaces in $\mathbb{M}^n$
$\mathbb{H} := L^2(\mathbb{M}^n)$	space of square integrable functions
$\mathbb{V} := W^{1,2}(\mathbb{M}^n)$	standard Sobolev space of order 1
$\mathbb{D} := \{f \in \mathbb{V} : \Delta f \in \mathbb{H}\}$	space of Sobolev functions with square integrable Laplacian
$L_c^\infty(\mathbb{M}^n)$	bounded real functions with compact support in $\mathbb{M}^n$
$C_b(\mathbb{M}^n)$	bounded and continuous real functions in $\mathbb{M}^n$
$\text{Lip}_c(\mathbb{M}^n)$	Lipschitz real functions with compact support in $\mathbb{M}^n$
$W_c^{1,2}((0,T); L^2(\mathbb{M}^n))$	Sobolev space of $L^2(\mathbb{M}^n)$-valued functions with compact support in time
$\Gamma, \Gamma_2, \mathbf{\Gamma}_2$	classical, iterated and nonlocal *carré du champ* operator, respectively, see Section 3.1
$\mathcal{E}_\rho[f]$	weighted Dirichlet energy (Hamiltonian functional), see (3.7)
$\mathcal{E}_\rho^*[\ell]$	dual of the Hamiltonian functional, see (3.8)
$Q_s\varphi$	Hopf-Lax semigroup starting from φ, see (3.15)

2. STATEMENT OF THE MAIN RESULTS

We consider the following nonlinear diffusion equation:

$$\begin{cases} \partial_t \rho = \Delta P(\rho) & \text{in } \mathbb{M}^n \times \mathbb{R}^+ , \\ \rho(\cdot, 0) = \mu_0 \geq 0 & \text{in } \mathbb{M}^n \times \{0\} , \end{cases} \tag{2.1}$$

where $\mu_0 \in \mathscr{M}_2^M(\mathbb{M}^n)$ and $\rho \mapsto P(\rho)$ is a suitable $C^1([0, +\infty))$ function of *porous medium type*. We require that $\mathbb{M}^n$ and P satisfy a precise set of hypotheses.

Hypotheses 2.1 (Manifold). *We assume throughout that $\mathbb{M}^n$ ($n \geq 3$) is a smooth, complete and connected Riemannian manifold. Moreover, it will comply with either one or more of the following conditions:*

- *The Ricci curvature is uniformly bounded from below, i.e. there exists $K \geq 0$ such that*

$$\text{Ric}_x(v,v) \geq -K|v|^2 \qquad \forall x \in \mathbb{M}^n \text{ and } v \in T_x\mathbb{M}^n ; \tag{H1}$$

- *For some $C_S > 0$ there holds the Sobolev-type inequality*

$$\|f\|_{L^{2^\star}(\mathbb{M}^n)} \leq C_S \left(\|\nabla f\|_{L^2(\mathbb{M}^n)} + \|f\|_{L^2(\mathbb{M}^n)} \right) \quad \forall f \in W^{1,2}(\mathbb{M}^n) , \qquad \text{with } 2^\star := \frac{2n}{n-2} . \tag{H2}$$

A result originally due to Varopoulos [36] asserts that (H2) does hold on any complete, n-dimensional ($n \geq 3$) Riemannian manifold satisfying (H1) along with the *noncollapse* condition

$$\inf_{x \in \mathbb{M}^n} \mathcal{V}(B_1(x)) > 0 , \tag{2.2}$$

where $B_1(x) := \{y \in \mathbb{M}^n : \mathsf{d}(x,y) < 1\}$. We refer in particular to [24, Theorem 3.2] (in fact B_1 could be replaced by B_r for any $r > 0$). Condition (2.2) is also necessary for (H2) to hold, see [24,

Lemma 2.2]. Note that (H1) and (2.2) are for free on any *compact* Riemannian manifold, a simple subcase of the frameworks we will work within. On the other hand, if $\mathbb{M}^n$ is noncompact and has finite volume, or more in general has an end with finite volume, then (2.2) (and therefore (H2)) necessarily fails.

As concerns the nonlinearity P appearing in (2.1), we introduce the following set of hypotheses. We write them separately in order to be able to single out the specific assumption(s) needed for each result we will prove.

Hypotheses 2.2 (Nonlinearity). *We assume throughout that $P \in C^1([0, +\infty))$. Moreover, it will comply with either one or more of the following conditions:*

$$P(0) = 0 \text{ and the map } \rho \mapsto P(\rho) \text{ is strictly increasing}; \tag{H3}$$

there exist $c_1 \geq c_0 > 0$ and $m > 1$ such that

$$c_0 \, m \, \rho^{m-1} \leq P'(\rho) \leq c_1 \, m \, \rho^{m-1} \qquad \forall \rho \geq 0 \,, \tag{H4}$$

$$\rho \, P'(\rho) - \left(1 - \tfrac{1}{n}\right) P(\rho) \geq 0 \qquad \forall \rho \geq 0 \,. \tag{H5}$$

It is straightforward to check that (H5) is implied by (H4) provided $c_1 \leq c_0 \, m \, \frac{n}{n-1}$.

Let us firstly notice that the choice $P(\rho) = \rho^m$ for some $m > 1$ (corresponding to the PME) obviously implies (H3), (H4) and (H5). We point out that condition (H4) is essential to establish the smoothing effect (see (2.5)) and the compact-support property (see Proposition 4.5), while (H5) is a key tool to develop the Hamiltonian approach in its abstract formulation (we refer to Lemma 5.2).

We start by providing a good notion of weak solution of (2.1) for initial data in $\mathscr{M}_2^M(\mathbb{M}^n)$ and for a general nonlinearity P, which is inspired by the (wide) existing literature, see Section 4.

Definition 2.3 (Weak Wasserstein solutions). *Let P comply with assumption* (H3). *Given $\mu_0 \in \mathscr{M}_2^M(\mathbb{M}^n)$, we say that a nonnegative measurable function ρ is a Wasserstein solution of* (2.1) *if, for every $T > \tau > 0$, there hold*

$$\rho, P(\rho) \in L^2(\mathbb{M}^n \times (\tau, T)) \,, \quad \nabla P(\rho) \in L^2(\mathbb{M}^n \times (\tau, T)) \,, \tag{2.3}$$

$$\int_0^T \int_{\mathbb{M}^n} \rho \, \partial_t \eta \, \mathrm{d}\mathcal{V}\mathrm{d}t = \int_0^T \int_{\mathbb{M}^n} \langle \nabla P(\rho) \,, \nabla \eta \rangle \, \mathrm{d}\mathcal{V}\mathrm{d}t \tag{2.4}$$

for every $\eta \in W_c^{1,2}((0,T); L^2(\mathbb{M}^n))$ with $\nabla \eta \in L^2((0,T); L^2(\mathbb{M}^n))$, and

$$\mu \in C\big([0,T); \big(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2\big)\big) \text{ with } \mu(0) = \mu_0 \,,$$

where $\mu(t) = \rho(t)\mathcal{V}$ for $t > 0$.

We are now in position to state our main results, which will be proved in Section 5.

Theorem 2.4 (Wasserstein stability). *Let $\mathbb{M}^n$ ($n \geq 3$) comply with assumptions* (H1) *and* (H2). *Let moreover P comply with assumptions* (H3), (H4) *and* (H5). *Let $\mu_0 \in \mathscr{M}_2^M(\mathbb{M}^n)$. Then there exists a unique Wasserstein solution ρ of* (2.1), *which satisfies the smoothing estimate*

$$\|\rho(t)\|_{L^\infty(\mathbb{M}^n)} \leq C \left(t^{-\frac{n}{2+n(m-1)}} M^{\frac{2}{2+n(m-1)}} + M\right) \qquad \forall t > 0 \,, \tag{2.5}$$

where $C \geq 1$ is a constant depending only on C_S, n, c_0 and independent of m ranging in a bounded subset of $(1, +\infty)$. Furthermore, if $\hat{\rho}$ is the Wasserstein solution of (2.1) *corresponding to another initial datum $\hat{\mu}_0 \in \mathscr{M}_2^M(\mathbb{M}^n)$, the stability estimate*

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq \exp\left\{K \, c_1 \, \mathfrak{C}_m \left[\left(tM^{m-1}\right)^{\frac{2}{2+n(m-1)}} \vee \left(tM^{m-1}\right)\right]\right\} \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t > 0 \tag{2.6}$$

holds, where $\mathfrak{C}_m := C^{m-1} \, 2^{m-2} \, [2 + n(m-1)]$.

In fact (2.6) is sharp, as $t \downarrow 0$, in the hyperbolic space $\mathbb{H}^n_K$ of sectional curvature $-K$, i.e. of Ricci curvature $-(n-1)K$.

Theorem 2.5 (Optimality). *Estimate* (2.6) *is optimal in* $\mathbb{M}^n = \mathbb{H}^n_K$, *for* $P(\rho) = \rho^m$, *with the choices* $\mu_0 = M\delta_x$ *and* $\hat{\mu}_0 = M\delta_y$, *provided the points* $x, y \in \mathbb{H}^n_K$ *are close enough. More precisely, upon setting* $\delta := \mathsf{d}(x, y) > 0$, *there exist constants* $\kappa = \kappa(n, m) > 0$, $\overline{\delta} = \overline{\delta}(n, K, m) > 0$ *and* $\overline{t} = \overline{t}(\delta, n, K, m, M) > 0$ *such that if* $\delta \in (0, \overline{\delta})$ *then*

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \geq \left[1 + K\, \kappa\, \left(tM^{m-1} \right)^{\frac{2}{2+n(m-1)}} \right] \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t \in (0, \overline{t}). \tag{2.7}$$

The proof of Theorem 2.5 will be provided in Subsection 5.4. Some comments regarding both Theorem 2.4 and Theorem 2.5 are now in order.

Remark 2.6 (The PME, the heat equation and gradient flows). As mentioned above, the explicit choice $P(\rho) = \rho^m$ corresponds to the well-known *porous medium equation* (PME). In this case estimate (2.6) holds with $c_1 = 1$. In particular, if we let $m \downarrow 1$, thanks to the fact that $\mathfrak{C}_m \to 1$ we recover exactly the following stability estimate for the heat equation:

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq e^{Kt}\, \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t > 0. \tag{2.8}$$

We recall that the Ricci bound (H1) is *equivalent* to the $(-K)$-gradient flow formulation of the heat equation with respect to the relative entropy in $(\mathscr{P}_2(\mathbb{M}^n), \mathcal{W}_2)$, from which (2.8) follows: we refer to [33, Theorem 1.1 and Corollary 1.4] for more details. We stress that, as a byproduct of Theorem 2.5, we can deduce that in general on negatively-curved manifolds the porous medium equation *cannot* be seen as the gradient flow of some λ-convex functional with respect to the 2-Wasserstein distance, at least in the sense of Evolution Variational Inequalities (see [1]). Indeed, if it was, then the estimate

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq e^{\lambda t}\, \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t > 0$$

would hold for some $\lambda \in \mathbb{R}$, thus contradicting (2.7). On the other hand, it is known that the PME *can* indeed be seen as the gradient flow of the free energy (1.2) in the case where the Ricci curvature is nonnegative (we refer to [35] and [29, 32]), so that (2.6) holds with $K = 0$.

Remark 2.7 (The Cartan-Hadamard case). If, in place of (H2), the manifold $\mathbb{M}^n$ supports a *Euclidean* Sobolev inequality, namely

$$\|f\|_{L^{2^\star}(\mathbb{M}^n)} \leq C_S\, \|\nabla f\|_{L^2(\mathbb{M}^n)} \qquad \forall f \in C^1_c(\mathbb{M}^n), \tag{2.9}$$

then it is not difficult to deduce that (2.6) turns into the following estimate:

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq \exp\left\{ K\, c_1\, \mathfrak{C}_m\, \left(tM^{m-1} \right)^{\frac{2}{2+n(m-1)}} \right\} \mathcal{W}_2(\mu_0, \hat{\mu}_0) \qquad \forall t > 0. \tag{2.10}$$

This is a simple consequence of our method of proof, since in that case the smoothing effect (2.5) holds with no additional M term on the right-hand side, which causes the linear term to appear at the exponent of (2.6). Note that when $K > 0$ estimate (2.10) improves (2.6) for long times (the just mentioned leading linear term disappears) and remains qualitatively the same for small times. We recall that (2.9) does hold, for instance, on any *Cartan-Hadamard* manifold, that is a complete, simply connected Riemannian manifold with everywhere nonpositive sectional curvature (see [21] and references therein).

Remark 2.8 (The 2-dimensional case). The results of Theorem 2.4 can also be extended to the dimension $n = 2$. In that case, the Sobolev inequality should be replaced by the Gagliardo-Nirenberg inequality

$$\|f\|_{L^r(\mathbb{M}^2)} \leq C_{GN} \left(\|\nabla f\|_{L^2(\mathbb{M}^2)} + \|f\|_{L^2(\mathbb{M}^2)} \right)^{\frac{r-s}{r}} \|f\|_{L^s(\mathbb{M}^2)}^{\frac{s}{r}} \qquad \forall f \in W^{1,2}(\mathbb{M}^2) \cap L^s(\mathbb{M}^2), \tag{2.11}$$

for some $r > s > 0$ and $C_{GN} > 0$. We recall that, by [5, Theorem 3.3], the validity of (2.11) for *some* $r > s > 0$ yields the validity of the same inequality for *every* $r > s > 0$. In particular, this allows us to reproduce the proof of Proposition 4.4 also for $n = 2$, starting from (2.11) in place of (4.35). The rest of the results we need in order to establish Theorem 2.4 also hold for $n = 2$. Note that, again, inequality (2.11) is satisfied (e.g. with $r > 2$ and $s = r - 2$) on any 2-dimensional Riemannian manifold complying with (H1) and (2.2): this is a simple consequence of [24, Lemma 2.1 and Theorem 3.2]. As concerns the optimality result contained in Theorem 2.5, we just observe that its proof follows with no modifications in the case $n = 2$ as well (see Subsection 5.4).

3. Geometric and functional preliminaries

In this section we recall some basic results concerning the Γ-calculus, curvature conditions, the Wasserstein distance(s) and the Hopf-Lax semigroup. We also resume a crucial density result for Wasserstein geodesics, which will be needed in the sequel.

3.1. **The Bakry-Émery curvature condition.** Let $(\mathbb{M}^n, \mathcal{B}, \mathcal{V})$ be the measure space given by the Riemannian manifold $\mathbb{M}^n$, the σ-algebra of Borel sets $\mathcal{B}$ and the volume measure $\mathcal{V}$ associated with the metric. Given a diffusion operator L on $(\mathbb{M}^n, \mathcal{B}, \mathcal{V})$ and a suitable algebra of functions $\mathcal{A}$, it is by now standard to define the *carré du champ* operator

$$\Gamma(f, g) := \frac{1}{2} \left(L(fg) - fLg - gLf \right), \qquad f, g \in \mathcal{A},$$

along with the *iterated carré du champ* operator

$$\Gamma_2(f, g) := \frac{1}{2} \left(L(\Gamma(f, g)) - \Gamma(f, Lg) - \Gamma(g, Lf) \right), \qquad f, g \in \mathcal{A}. \tag{3.1}$$

The introduction of these tools is motivated by the fact that they carry the geometric information on the measure space $(\mathbb{M}^n, \mathcal{B}, \mathcal{V})$, being at the same time very suitable for computations. For more details, we refer the reader to the monograph [7] and to the original paper by Bakry and Émery [6].

In the present setting we fix once for all L as the unique self-adjoint extension in $L^2(\mathbb{M}^n)$ of the Laplace-Beltrami operator $L = \Delta$. In this case, it is apparent that

$$\Gamma(f, g) = \langle \nabla f, \nabla g \rangle, \quad \Gamma(f) := \Gamma(f, f) = |\nabla f|^2, \quad \Gamma_2(f) := \Gamma_2(f, f) = \frac{1}{2} \Delta |\nabla f|^2 - \langle \nabla f, \nabla \Delta f \rangle. \tag{3.2}$$

We recall that, thanks to [34, Theorem 2.4], the operator $(-\Delta)$ defined in $C_c^\infty(\mathbb{M}^n)$ is essentially self-adjoint on *any* complete Riemannian manifold, i.e. $\mathbb{D}$ coincides with the closure of $C_c^\infty(\mathbb{M}^n)$ with respect to the norm $\| \cdot \|_{\mathbb{D}}$.

When the Ricci curvature of $\mathbb{M}^n$ is uniformly bounded from below by a constant $\lambda \in \mathbb{R}$, by applying the Bochner-Lichnerowicz formula it follows that (for all sufficiently regular function f)

$$\Gamma_2(f) \geq \lambda \Gamma(f) + \frac{1}{n}(\Delta f)^2, \tag{3.3}$$

which goes under the name of *Bakry-Émery curvature-dimension condition* $\mathrm{BE}(\lambda, n)$. It is possible to show that in fact the converse implication is also true: if a Riemannian manifold $\mathbb{M}^n$ satisfies the condition $\mathrm{BE}(\lambda, N)$, then $n \leq N$ and $\mathrm{Ric} \geq \lambda$, see [7, Subsection 1.16 and Sections C.5, C.6] for further details.

Let us now introduce the (local) Dirichlet form $\mathcal{E} : \mathbb{H} \to [0, +\infty]$ by setting

$$\mathcal{E}(f) := \int_{\mathbb{M}^n} \Gamma(f) \, \mathrm{d}\mathcal{V} = \int_{\mathbb{M}^n} |\nabla f|^2 \, \mathrm{d}\mathcal{V}, \tag{3.4}$$

with proper domain $\mathbb{V}$. In addition, according to [3], it is convenient to define a suitable "integral" version of the Γ_2 operator, in the following form:

$$
\begin{aligned}
\mathbf{\Gamma}_2[f;\rho] &:= \int_{\mathbb{M}^n} \left[\frac{1}{2} \Gamma(f)\,\Delta\rho - \Gamma(f,\Delta f)\,\rho \right] \mathrm{d}\mathcal{V} \\
&= \int_{\mathbb{M}^n} \left[\frac{1}{2} \Gamma(f)\,\Delta\rho + \Gamma(f,\rho)\,\Delta f + (\Delta f)^2\,\rho \right] \mathrm{d}\mathcal{V} \qquad \forall (f,\rho) \in \mathbb{D}_\infty \,,
\end{aligned}
\tag{3.5}
$$

where $\mathbb{D}_\infty$ stands for the algebra of functions defined as $\mathbb{D}_\infty := \mathbb{D} \cap L^\infty(\mathbb{M}^n)$. Note that, formally, (3.5) is obtained upon choosing $g = f$ in (3.1), multiplying by ρ and integrating by parts. The introduction of the multilinear form $\mathbf{\Gamma}_2$ provides a weak version of the Bakry-Émery condition: for every $(f,\rho) \in \mathbb{D}_\infty$ with $\rho \geq 0$ there holds

$$
\mathbf{\Gamma}_2[f;\rho] \geq \lambda \int_{\mathbb{M}^n} \Gamma(f)\,\rho\,\mathrm{d}\mathcal{V} + \frac{1}{n} \int_{\mathbb{M}^n} (\Delta f)^2\,\rho\,\mathrm{d}\mathcal{V}\,.
\tag{3.6}
$$

On a Riemannian manifold, the two formulations (3.3) and (3.6) turn out to be equivalent, and we will refer to both of them as $\mathrm{BE}(\lambda, n)$. For a proof of such equivalence see e.g. [3, Subsection 2.2].

In order to deal with "variational" solutions of (2.1), we will also consider a weighted version of the Dirichlet energy (3.4). More precisely, given $\rho \in L^\infty(\mathbb{M}^n)$ with $\rho \geq 0$, we set $\mathcal{E}_\rho : \mathbb{V} \to [0, +\infty)$ as

$$
\mathcal{E}_\rho[f] := \int_{\mathbb{M}^n} \Gamma(f)\,\rho\,\mathrm{d}\mathcal{V}\,.
\tag{3.7}
$$

The associated dual weighted Dirichlet energy $\mathcal{E}_\rho^* : \mathbb{V}' \to [0, +\infty]$ is defined as

$$
\frac{1}{2} \mathcal{E}_\rho^*[\ell] := \sup_{f \in \mathbb{V}} {}_{\mathbb{V}'}\langle \ell, f \rangle_{\mathbb{V}} - \frac{1}{2} \mathcal{E}_\rho[f]\,,
\tag{3.8}
$$

where we denoted by $\mathbb{V}'$ the dual space of $\mathbb{V}$.

3.2. **The Wasserstein space.** Let (X, d) be a complete metric space. We say that a curve $\gamma : [0, 1] \to (X, \mathsf{d})$ belongs to $\mathrm{AC}^2([0, 1]; (X, \mathsf{d}))$ if there exists a function $w \in L^2((0, 1))$ such that

$$
\mathsf{d}(\gamma(s), \gamma(t)) \leq \int_s^t w(r)\,\mathrm{d}r \qquad \text{for every } 0 \leq s \leq t \leq 1\,.
\tag{3.9}
$$

When $\gamma \in \mathrm{AC}^2([0, 1]; (X, \mathsf{d}))$ its *metric velocity*, defined as

$$
|\dot{\gamma}|(r) := \lim_{h \to 0} \frac{\mathsf{d}(\gamma(r + h), \gamma(r))}{|h|}\,,
$$

exists for a.e. $r \in (0, 1)$. Moreover, $|\dot{\gamma}|$ belongs to $L^2((0, 1))$ and provides the *minimal* function w, up to negligible sets, such that (3.9) holds (see [1, Theorem 1.1.2]).

A (constant-speed) *geodesic* is a curve γ satisfying

$$
\mathsf{d}(\gamma(0), \gamma(1)) = \int_0^1 |\dot{\gamma}|(r)\,\mathrm{d}r\,,
$$

or equivalently

$$
\mathsf{d}(\gamma(s), \gamma(t)) = \mathsf{d}(\gamma(0), \gamma(1))(t - s) \qquad \text{for every } 0 \leq s \leq t \leq 1\,;
$$

in particular, a geodesic is a Lipschitz curve.

We say that a measure $\mu \in \mathscr{M}^M(\mathbb{M}^n)$ has finite p-moment, $p \geq 1$, and we write $\mu \in \mathscr{M}_p^M(\mathbb{M}^n)$, if there exists a point $o \in \mathbb{M}^n$ such that

$$
\int_{\mathbb{M}^n} \mathsf{d}(x, o)^p\,\mathrm{d}\mu(x) < \infty\,.
$$

We define the *p-Wasserstein* cost between two measures $\mu^0, \mu^1 \in \mathscr{M}(\mathbb{M}^n)$ as

$$\mathcal{W}_p^p(\mu^0, \mu^1) := \inf_\pi \int_{\mathbb{M}^n \times \mathbb{M}^n} \mathsf{d}(x,y)^p \, \mathrm{d}\pi(x,y) \,,$$

where the infimum is taken among all the transport plans π between μ^0 and μ^1. The latter are measures $\pi \in \mathscr{M}(\mathbb{M}^n \times \mathbb{M}^n)$ such that $\pi(A \times \mathbb{M}^n) = \mu^0(A)$ and $\pi(\mathbb{M}^n \times B) = \mu^1(B)$ for every Borel sets $A, B \subset \mathbb{M}^n$. We observe that $\mathcal{W}_p(\mu^0, \mu^1) = \infty$ whenever $\mu^0(\mathbb{M}^n) \neq \mu^1(\mathbb{M}^n)$. Another elementary fact is that

$$\mu^0 \in \mathscr{M}_p^M(\mathbb{M}^n) \quad \text{and} \quad \mathcal{W}_p(\mu^0, \mu^1) < \infty \quad \text{implies} \quad \mu^1 \in \mathscr{M}_p^M(\mathbb{M}^n) \,. \tag{3.10}$$

We are mainly interested in the cases $p = 2$ and $p = 1$. As regards the 1-Wasserstein distance, we will only use these two well-known properties (see [41, Chapter 6]):

$$\mathcal{W}_1(\mu^0, \mu^1) \leq \mathcal{W}_2(\mu^0, \mu^1) \qquad \text{for every } \mu_1, \mu_2 \in \mathscr{P}_2(\mathbb{M}^n) \,, \tag{3.11}$$

$$\mathcal{W}_1(\mu^0, \mu^1) = \sup \left\{ \int_{\mathbb{M}^n} f \, \mathrm{d}\mu^1 - \int_{\mathbb{M}^n} f \, \mathrm{d}\mu^0 : \quad f : \mathbb{M}^n \to \mathbb{R}, \ f \text{ is 1-Lipschitz} \right\}. \tag{3.12}$$

When $p = 2$, it can be shown that for every $M \in (0, +\infty)$ the space $(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2)$ is a metric space, called the *2-Wasserstein* (or simply Wasserstein) *space* of mass M over $\mathbb{M}^n$, which inherits many geometric properties of the ambient space $\mathbb{M}^n$. In particular, it is complete, separable and geodesic (for a proof of these facts we refer again to [41, Chapter 6]).

Here we will mostly work with the *dual characterization* of the Wasserstein distance due to Kantorovich, which asserts that (see e.g. [41, Theorem 5.10(i)]) for any $\mu^0, \mu^1 \in \mathscr{M}^M(\mathbb{M}^n)$ there holds

$$\frac{1}{2} \mathcal{W}_2^2(\mu^0, \mu^1) = \sup_{\substack{\varphi, \psi \in C_b(\mathbb{M}^n): \\ \psi(x) \leq \varphi(y) + \frac{1}{2}\mathsf{d}(x,y)^2 \ \forall x,y \in \mathbb{M}^n}} \left\{ \int_{\mathbb{M}^n} \psi \, \mathrm{d}\mu^1 - \int_{\mathbb{M}^n} \varphi \, \mathrm{d}\mu^0 \right\}. \tag{3.13}$$

From (3.13) it is clear that, for any fixed φ, the best possible choice of ψ is provided by

$$\psi(x) = Q_1\varphi(x) := \inf_{y \in \mathbb{M}^n} \varphi(y) + \frac{1}{2}\mathsf{d}(x,y)^2 \qquad \forall x \in \mathbb{M}^n \,. \tag{3.14}$$

Thanks to (3.14), by means of a cut-off argument it is not difficult to show that the supremum in (3.13) can actually be taken over the space $C_c(\mathbb{M}^n)$ of continuous and compactly-supported functions. A local regularization procedure then ensures that one can replace $C_c(\mathbb{M}^n)$ with the space of *Lipschitz* and *compactly-supported* functions $\mathrm{Lip}_c(\mathbb{M}^n)$.

In our framework it is convenient to see $Q_1\varphi$ as an endpoint of the *Hopf-Lax* evolution semigroup starting from φ. We recall that the latter is given by the family of maps $Q_s : \mathrm{Lip}_c(\mathbb{M}^n) \to \mathrm{Lip}_c(\mathbb{M}^n)$, $s \geq 0$, defined as

$$Q_s\varphi(x) := \inf_{y \in \mathbb{M}^n} \varphi(y) + \frac{\mathsf{d}(x,y)^2}{2s} \quad \forall s > 0 \,, \qquad Q_0\varphi(x) := \varphi(x) \qquad \forall x \in \mathbb{M}^n \,. \tag{3.15}$$

It is readily seen that $Q_s\varphi$ satisfies

$$\inf_{\mathbb{M}^n} \varphi \leq Q_s\varphi(x) \leq \varphi(x) \qquad \forall s \geq 0 \,, \ \forall x \in \mathbb{M}^n \,.$$

More importantly, since $\mathbb{M}^n$ is a geodesic space, it can be shown (see [2, Theorem 3.6]) that $(s, x) \mapsto Q_s\varphi(x)$ is the Lipschitz solution of the Hopf-Lax (or Hamilton-Jacobi) problem

$$\begin{cases} \frac{\partial}{\partial s} Q_s\varphi(x) = -\frac{1}{2} |\nabla Q_s\varphi|^2(x) & \text{for a.e. } (x,s) \in \mathbb{M}^n \times \mathbb{R}^+ \,, \\ Q_0\varphi = \varphi \,. \end{cases} \tag{3.16}$$

We can subsume the above discussion in the following.

Proposition 3.1. *Let $\mu^0, \mu^1 \in \mathscr{M}^M(\mathbb{M}^n)$. Then*

$$\frac{1}{2} \mathcal{W}_2^2(\mu^0, \mu^1) = \sup_{\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)} \left\{ \int_{\mathbb{M}^n} Q_1 \varphi \, \mathrm{d}\mu^1 - \int_{\mathbb{M}^n} \varphi \, \mathrm{d}\mu^0 \right\}.$$

We now recall a useful characterization of the convergence in the 2-Wasserstein space, whose proof can be found in [1, Proposition 7.1.5].

Proposition 3.2. *Let $\mu \in \mathscr{M}_2^M(\mathbb{M}^n)$ and $\{\mu_j\}_{j \in \mathbb{N}} \subset \mathscr{M}_2^M(\mathbb{M}^n)$. Then*

$$\lim_{j \to \infty} \mathcal{W}_2(\mu_j, \mu) = 0$$

if and only if $\mu_j \rightharpoonup \mu$ narrowly (i.e. tested against any function of $C_b(\mathbb{M}^n)$) and $\{\mu_j\}_{j \in \mathbb{N}}$ has equi-integrable second moments, namely there exists a point $o \in \mathbb{M}^n$ such that

$$\lim_{k \to \infty} \limsup_{j \to \infty} \int_{\mathbb{M}^n \setminus B_k(o)} \mathsf{d}(x, o)^2 \, \mathrm{d}\mu_j(x) = 0.$$

In the proof of Theorem 2.4 it will be crucial to connect any two given measures $\mu^0, \mu^1 \in \mathscr{M}_2^M(\mathbb{M}^n)$ through a curve in the 2-Wasserstein space that satisfies some additional regularity properties, according to the following definition.

Definition 3.3. *Let $\mu \equiv \{\mu^s\}_{s \in [0,1]}$ be a curve with values in $\mathscr{M}_2^M(\mathbb{M}^n)$. We say that μ is a* regular *curve if $\mu^s = \rho^s \mathcal{V}$ and the following hold:*

   (i) *There exists a constant $R > 0$ such that $\|\rho^s\|_{L^\infty(\mathbb{M}^n)} \leq R$ for every $s \in [0,1]$;*

   (ii) *$\mu \in \mathrm{Lip}\big([0,1]; \big(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2\big)\big)$;*

   (iii) *$\sqrt{\rho^s} \in \mathbb{V}$ and there exists a constant E such that*

$$\int_{\mathbb{M}^n} \Gamma\big(\sqrt{\rho^s}\big) \, \mathrm{d}\mathcal{V} \leq E \qquad \forall s \in [0,1].$$

Remark 3.4. *If $\mu = \rho \mathcal{V}$ is a regular curve, in particular $\rho^s \in \mathbb{V}$ for every $s \in [0,1]$. Moreover, thanks to [4, Lemma 8.1], condition (ii) ensures that $\rho \in \mathrm{Lip}([0,1]; \mathbb{V}')$.*

The following density result, whose proof is contained in [4, Lemma 12.2], allows one to approximate Wasserstein geodesics by means of regular curves.

Lemma 3.5. *Let $\mathbb{M}^n$ satisfy (H1) and $\mu^0, \mu^1 \in \mathscr{M}_2^M(\mathbb{M}^n)$. Then there exist a geodesic $\{\mu^s\}_{s \in [0,1]}$ connecting μ^0 and μ^1 and a sequence of regular curves $\{\mu_j^s\}_{j \in \mathbb{N}, s \in [0,1]} \subset \mathscr{M}_2^M(\mathbb{M}^n)$ such that*

$$\lim_{j \to \infty} \mathcal{W}_2\big(\mu_j^s, \mu^s\big) = 0 \qquad \forall s \in [0,1] \tag{3.17}$$

and

$$\limsup_{j \to \infty} \int_0^1 |\dot{\mu}_j^s|^2 \, \mathrm{d}s \leq \mathcal{W}_2^2(\mu^0, \mu^1). \tag{3.18}$$

Furthermore, if $\mu^0 = \rho^0 \mathcal{V}$ and $\mu^1 = \rho^1 \mathcal{V}$ with $\rho^0, \rho^1 \in L_c^\infty(\mathbb{M}^n)$, then $\mu^s = \rho^s \mathcal{V}$ with ρ^s uniformly (w.r.t. s) bounded and compactly supported, and in addition to (3.17)–(3.18) also the following hold:

$$\lim_{j \to \infty} \big\|\rho_j^s - \rho^s\big\|_{L^p(\mathbb{M}^n)} = 0 \qquad \forall p \in [1, \infty), \quad \forall s \in [0,1], \tag{3.19}$$

$$\limsup_{j \to \infty} \sup_{s \in [0,1]} \big\|\rho_j^s\big\|_{L^\infty(\mathbb{M}^n)} < \infty. \tag{3.20}$$

To conclude, given a regular curve $\mu^s = \rho^s \mathrm{d}\mathcal{V}$ in the sense of Definition 3.3 (not necessarily a geodesic), by combining [4, Theorem 6.6, formula (6.11)] and [4, Theorem 8.2, formula (8.7)] we can deduce that the following key identity holds:

$$\int_0^1 |\dot{\mu}^s|^2 \, \mathrm{d}s = \int_0^1 \mathcal{E}_{\rho^s}^* \big[\tfrac{\mathrm{d}}{\mathrm{d}s} \rho^s\big] \, \mathrm{d}s, \tag{3.21}$$

where $\mathcal{E}_\rho^*$ is the dual weighted Dirichlet energy introduced in (3.8). Note that the r.h.s. of (3.21) does make sense, in view of Remark 3.4.

4. Fundamental properties of porous medium-type equations on manifolds

This section is devoted to the study of (2.1) for more regular initial data, that is the problem

$$\begin{cases} \partial_t \rho = \Delta P(\rho) & \text{in } \mathbb{M}^n \times \mathbb{R}^+ \,, \\ \rho(\cdot, 0) = \rho_0 \geq 0 & \text{on } \mathbb{M}^n \times \{0\} \,, \end{cases} \tag{4.1}$$

with $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$. To begin with, we will introduce the notion of *weak energy* solution and then discuss some crucial related properties. In particular, we will focus on the smoothing effect and on a bound on the support of such solutions (when the initial data are compactly supported). Inspired by [4], for a restricted class of nonlinearities we will also give an alternative (variational) notion of solution and consequently prove the equivalence with the weak-energy one. Finally, with regards to the Hamiltonian strategy mentioned in the Introduction, we will discuss well-posedness results for the *forward linearized* equation associated with (4.1) and for the related *backward adjoint* equation.

For convenience, in the following we make the additional (implicit) assumption that $\mathbb{M}^n$ is *non-compact* and with *infinite volume*, as well as in Subsection 5.2. Note that for our purposes there is no point in considering noncompact manifolds with *finite* volume, since most of our results require the validity of the Sobolev inequality (H2), which does not hold on such manifolds.

The simple modifications required to deal with *compact* manifolds will be shortly addressed in Subsection 5.3.

4.1. **Weak energy solutions.** The concept of *weak energy* solution of (4.1) has been proved to be well suited for porous medium-type equations: see e.g. [40, Subsections 5.3.2 and 11.2.1], [16, Section 3], [19, Subsections 3.1 and 3.2] or [21, Section 2]. Here we mostly take inspiration from [16, Section 3]: there the framework is purely Euclidean, but the basic definitions and properties are straightforwardly adaptable to the Riemannian setting.

Even if in Subsection 1.2 we introduced the more synthetic notations (1.11), here we keep the standard notations typically used in the PDE framework.

Definition 4.1 (Weak energy solutions). *Let P comply with assumption* (H3). *Given a nonnegative* $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$, *we say that a nonnegative measurable function ρ is a* weak energy solution *of* (4.1) *if, for every $T > 0$, there hold*

$$\rho, P(\rho) \in L^2(\mathbb{M}^n \times (0, T)) \,, \qquad \nabla P(\rho) \in L^2(\mathbb{M}^n \times (0, T))$$

and

$$\int_0^T \int_{\mathbb{M}^n} \rho \, \partial_t \eta \, \mathrm{d}\mathcal{V} \mathrm{d}t = - \int_{\mathbb{M}^n} \rho_0(x) \, \eta(x, 0) \, \mathrm{d}\mathcal{V}(x) + \int_0^T \int_{\mathbb{M}^n} \langle \nabla P(\rho) \,, \nabla \eta \rangle \, \mathrm{d}\mathcal{V} \mathrm{d}t \tag{4.2}$$

for every $\eta \in W^{1,2}((0, T); L^2(\mathbb{M}^n))$ with $\nabla \eta \in L^2((0, T); L^2(\mathbb{M}^n))$ such that $\eta(T) = 0$.

Existence and uniqueness of weak energy solutions, at least for the class of data $L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$, is by now a well-established issue (see e.g. the references quoted above). Nevertheless, since it will be very useful to our later purposes, we recall here the approximation procedure that allows one to construct such solutions: the essential idea is to approximate the possibly degenerate nonlinearity $P \in C^1([0, +\infty))$ by means of suitable *nondegenerate* nonlinearities. More precisely, for every $\varepsilon > 0$ we define a function P_ε by

$$(P_\varepsilon)'(\rho) := \begin{cases} P'(\rho) + \varepsilon & \text{if } \rho \in \left[0, \frac{1}{\varepsilon}\right] \,, \\ \left[P'\left(\frac{1}{\varepsilon}\right) \wedge P'(\rho)\right] + \varepsilon & \text{if } \rho > \frac{1}{\varepsilon} \,, \end{cases} \qquad P_\varepsilon(0) = 0 \,. \tag{4.3}$$

In the following simple lemma, we collect some crucial properties enjoyed by P_ε.

Lemma 4.2. *Let P comply with* (H3). *We have that $P_\varepsilon \in C^1([0, +\infty))$ with the estimates*

$$P_\varepsilon(\rho) \le P(\rho) + \varepsilon\rho \qquad \forall \rho \ge 0\,, \tag{4.4}$$

$$(P_\varepsilon)'(\rho) \ge P'(\rho) \qquad \forall \rho \in \left[0, \tfrac{1}{\varepsilon}\right]\,, \tag{4.5}$$

and

$$\varepsilon \le (P_\varepsilon)'(\rho) \le \max_{\rho \in [0, 1/\varepsilon]} P'(\rho) + \varepsilon \qquad \forall \rho \ge 0\,. \tag{4.6}$$

In particular, P_ε is also strictly increasing. Moreover, if (H5) *holds then*

$$\rho\, (P_\varepsilon)'(\rho) - \left(1 - \tfrac{1}{n}\right) P_\varepsilon(\rho) \ge 0 \qquad \forall \rho \ge 0\,. \tag{4.7}$$

Proof. The fact that P_ε is $C^1([0, +\infty))$ easily follows from (4.3) and the continuity of the minimum operator, since P' is continuous. By integration, we obtain

$$P_\varepsilon(\rho) = \begin{cases} P(\rho) + \varepsilon\rho & \text{if } \rho \in \left[0, \tfrac{1}{\varepsilon}\right]\,, \\ P\left(\tfrac{1}{\varepsilon}\right) + \varepsilon\rho + \displaystyle\int_{\frac{1}{\varepsilon}}^{\rho} P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(s)\,\mathrm{d}s & \text{if } \rho > \tfrac{1}{\varepsilon}\,. \end{cases} \tag{4.8}$$

Inequality (4.4) is obvious (being an identity) in the interval $\left[0, \tfrac{1}{\varepsilon}\right]$, while for $\rho > \tfrac{1}{\varepsilon}$ it is a consequence of the trivial inequality $P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(s) \le P'(s)$. The bounds (4.5) and (4.6) are a direct consequence of the definition of $(P_\varepsilon)'$ together with the properties $P' \ge 0$ and

$$\left[P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(\rho)\right] + \varepsilon \le P'\left(\tfrac{1}{\varepsilon}\right) + \varepsilon \le \max_{\rho \in [0, 1/\varepsilon]} P'(\rho) + \varepsilon\,.$$

It remains to prove (4.7) under (H5). Its validity is clear in the interval $\left[0, \tfrac{1}{\varepsilon}\right]$, so let us assume without loss of generality that $\rho > \tfrac{1}{\varepsilon}$. Using the explicit expression (4.8), it holds:

$$\begin{aligned} &\rho\, (P_\varepsilon)'(\rho) - \left(1 - \tfrac{1}{n}\right) P_\varepsilon(\rho) \\ &= \rho \left[P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(\rho)\right] + \varepsilon\rho - \left(1 - \tfrac{1}{n}\right)\left[P\left(\tfrac{1}{\varepsilon}\right) + \varepsilon\rho + \int_{\frac{1}{\varepsilon}}^{\rho} P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(s)\,\mathrm{d}s\right] \\ &\ge \rho \left[P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(\rho)\right] - \left(1 - \tfrac{1}{n}\right)\left[P\left(\tfrac{1}{\varepsilon}\right) + \int_{\frac{1}{\varepsilon}}^{\rho} P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(s)\,\mathrm{d}s\right]\,. \end{aligned} \tag{4.9}$$

Now, if $P'\left(\tfrac{1}{\varepsilon}\right) \ge P'(\rho)$ the result follows by noticing that $P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(s) \le P'(s)$, integrating in the right-most side of (4.9) and taking advantage of assumption (H5). If instead $P'\left(\tfrac{1}{\varepsilon}\right) < P'(\rho)$, we exploit the property $P'\left(\tfrac{1}{\varepsilon}\right) \wedge P'(s) \le P'\left(\tfrac{1}{\varepsilon}\right)$ and observe that in this case the right-most side of (4.9) is bounded from below by

$$\begin{aligned} &\rho P'\left(\tfrac{1}{\varepsilon}\right) - \left(1 - \tfrac{1}{n}\right) P\left(\tfrac{1}{\varepsilon}\right) - \left(1 - \tfrac{1}{n}\right)\left(\rho - \tfrac{1}{\varepsilon}\right) P'\left(\tfrac{1}{\varepsilon}\right) \\ &= \tfrac{1}{\varepsilon} P'\left(\tfrac{1}{\varepsilon}\right) - \left(1 - \tfrac{1}{n}\right) P\left(\tfrac{1}{\varepsilon}\right) + \tfrac{1}{n}\left(\rho - \tfrac{1}{\varepsilon}\right) P'\left(\tfrac{1}{\varepsilon}\right) \ge 0\,, \end{aligned}$$

where we have used again assumption (H5) at $\rho = \tfrac{1}{\varepsilon}$ along with the fact that $P' \ge 0$. $\square$

Note that if P complies with the left-hand bound in (H4) so does P_ε' in the interval $[0, 1/\varepsilon]$, thanks to (4.5). This bound is crucial to establish the *smoothing effect*, which is a key ingredient to our strategy (see Proposition 4.4 below). Accordingly, we thus address the following approximate version of (4.1):

$$\begin{cases} \partial_t \rho_\varepsilon = \Delta P_\varepsilon(\rho_\varepsilon) & \text{in } \mathbb{M}^n \times \mathbb{R}^+\,, \\ \rho_\varepsilon(\cdot, 0) = \rho_0 & \text{on } \mathbb{M}^n \times \{0\}\,. \end{cases} \tag{4.10}$$

We stress that the notation ρ_ε for the solution of (4.10) only makes sense for $\varepsilon > 0$, and it should not be confused with the initial datum ρ_0. It may help to keep in mind that ρ_ε is supposed to be a "regularization" of ρ, the latter being the solution of (4.1).

Problem (4.10) can be interpreted both from the viewpoint of *linear* and *nonlinear* theory, in the sense that P_ε is a nonlinear function but it is "uniformly elliptic", hence one expects that the solutions of (4.10) enjoy, to some extent, properties similar to those satisfied by the solutions of the *heat equation* (we refer to Propositions 4.7 and 4.8 below). We will mainly take advantage of the linear interpretation in Section 5, in agreement with the approach of [4]. The nonlinear interpretation is exploited in the present section.

Proposition 4.3 (Existence, uniqueness, properties of weak energy solutions). *Let P comply with* (H3). *Given a nonnegative $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$, there exists a* unique *weak energy solution ρ of* (4.1), *which enjoys the following additional properties:*

- L^1-*continuity:* $\{\rho(t)\}_{t \geq 0}$ *is a continuous curve with values in $L^1(\mathbb{M}^n)$;*
- *Energy inequality:* ρ *satisfies*

$$\int_0^T \int_{\mathbb{M}^n} |\nabla P(\rho)|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t + \int_{\mathbb{M}^n} \Psi(\rho(x,T)) \, \mathrm{d}\mathcal{V}(x) \leq \int_{\mathbb{M}^n} \Psi(\rho_0) \, \mathrm{d}\mathcal{V} \qquad \forall T > 0, \tag{4.11}$$

  where $\Psi(\rho) := \int_0^\rho P(r) \, \mathrm{d}r$;
- *Nonexpansivity of the L^p norms: for every $p \in [1, \infty]$ there holds*

$$\|\rho(t)\|_{L^p(\mathbb{M}^n)} \leq \|\rho_0\|_{L^p(\mathbb{M}^n)} \qquad \forall t > 0; \tag{4.12}$$

- *Mass conservation: if in addition $\mathbb{M}^n$ satisfies* (H1) *then*

$$\int_{\mathbb{M}^n} \rho(x,t) \, \mathrm{d}\mathcal{V}(x) = \int_{\mathbb{M}^n} \rho_0 \, \mathrm{d}\mathcal{V} \qquad \forall t > 0; \tag{4.13}$$

- *Approximation: if $\varepsilon > 0$ and ρ_ε is the weak energy solution of* (4.10), *where $P_\varepsilon(\rho)$ is defined in* (4.3), *then*

$$\lim_{\varepsilon \downarrow 0} \|\rho_\varepsilon(t) - \rho(t)\|_{L^1_{\mathrm{loc}}(\mathbb{M}^n)} = 0 \qquad \forall t > 0; \tag{4.14}$$

  if in addition (H1) *is satisfied, then*

$$\lim_{\varepsilon \downarrow 0} \|\rho_\varepsilon(t) - \rho(t)\|_{L^1(\mathbb{M}^n)} = 0 \qquad \forall t > 0; \tag{4.15}$$

- L^1-*contraction: if $\hat\rho$ is the weak energy solution corresponding to another nonnegative initial datum $\hat\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$, then*

$$\|\rho(t) - \hat\rho(t)\|_{L^1(\mathbb{M}^n)} \leq \|\rho_0 - \hat\rho_0\|_{L^1(\mathbb{M}^n)} \qquad \forall t > 0. \tag{4.16}$$

Proof. We start by recalling that uniqueness of weak energy solutions follows from a standard trick due to Oleĭnik: given $T > 0$, one plugs the (admissible) test function

$$\eta(x,t) = \int_t^T [P(\rho(x,s)) - P(\hat\rho(x,s))] \, \mathrm{d}s, \qquad (x,t) \in \mathbb{M}^n \times [0,T],$$

into the weak formulation satisfied by the difference between ρ and $\hat\rho$ (the latter being two possibly different solutions corresponding to the same initial datum), thus obtaining

$$\int_0^T \int_{\mathbb{M}^n} (\rho - \hat\rho) \, (P(\rho) - P(\hat\rho)) \, \mathrm{d}\mathcal{V}\mathrm{d}t$$
$$= \int_0^T \int_{\mathbb{M}^n} \left\langle \nabla \left[P(\rho(x,t)) - P(\hat\rho(x,t))\right], \int_t^T \nabla \left[P(\rho(x,s)) - P(\hat\rho(x,s))\right] \mathrm{d}s \right\rangle \mathrm{d}\mathcal{V}(x)\mathrm{d}t. \tag{4.17}$$

A simple time integration in (4.17) yields

$$\int_0^T \int_{\mathbb{M}^n} (\rho - \hat\rho) \, (P(\rho) - P(\hat\rho)) \, \mathrm{d}\mathcal{V}\mathrm{d}t + \frac{1}{2} \int_{\mathbb{M}^n} \left| \int_0^T \nabla \left[P(\hat\rho(x,s)) - P(\rho(x,s))\right] \mathrm{d}s \right|^2 \mathrm{d}\mathcal{V}(x) = 0,$$

which ensures that $\rho = \hat\rho$ given the strict monotonicity of $\rho \mapsto P(\rho)$ and the arbitrariness of T. Note that here we have only used the validity of (4.2) for functions $\eta \in W^{1,2}((0,T);W^{1,2}(\mathbb{M}^n))$. Furthermore, the fact that $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$ is unimportant. These observations will be useful in the proof of Proposition 4.7 below.

As concerns the construction of a weak energy solution, we will not provide a complete proof since the procedure is quite standard: see e.g. [40, Theorem 5.7 and Lemma 5.8] or [19, Theorems 3.4 and 3.7] in Euclidean or weighted-Euclidean contexts. The basic idea consists first of solving problem (4.1) in a sequence D_k of bounded regular domains that form an exhaustion for $\mathbb{M}^n$ (see the proof of Lemma 4.5 below for more details on such a sequence), with homogeneous Dirichlet boundary conditions on ∂D_k. In order to do this, it is convenient to make a further approximation by replacing P with P_ε: let us denote by $\rho_{\varepsilon,k}$ the corresponding solutions, which are therefore regular enough (up to approximating also the initial datum ρ_0 and approximating further P_ε in case P' is merely continuous – we skip this passages). A first key estimate is provided by the energy inequality itself, which is obtained upon multiplying the differential equation by $P_\varepsilon(\rho_{\varepsilon,k})$ and integrating by parts:

$$\int_0^T \int_{D_k} |\nabla P_\varepsilon(\rho_{\varepsilon,k})|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t + \int_{D_k} \Psi_\varepsilon(\rho_{\varepsilon,k}(x,T)) \, \mathrm{d}\mathcal{V}(x) = \int_{D_k} \Psi_\varepsilon(\rho_0) \, \mathrm{d}\mathcal{V} \qquad \forall T > 0, \qquad (4.18)$$

where $\Psi_\varepsilon(\rho) := \int_0^\rho P_\varepsilon(r) \, \mathrm{d}r$. Note that for the moment the energy inequality is in fact an identity. Another crucial estimate involves time derivatives and is obtained by multiplying the differential equation by $\zeta \, P_\varepsilon'(\rho_{\varepsilon,k}) \, \partial_t \rho_{\varepsilon,k}$ and again integrating by parts, where $\zeta \in C_c^\infty((0,+\infty))$ is any cut-off function that depends only on time and satisfies $0 \le \zeta \le 1$; this yields

$$\int_0^T \int_{D_k} \zeta \, |\partial_t \Upsilon_\varepsilon(\rho_{\varepsilon,k})|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t = \frac{1}{2} \int_0^T \int_{D_k} \zeta' \, |\nabla P_\varepsilon(\rho_{\varepsilon,k})|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t \qquad \forall T > 0, \qquad (4.19)$$

where $\Upsilon_\varepsilon(\rho) := \int_0^\rho \sqrt{P_\varepsilon'(r)} \, \mathrm{d}r$. Finally, by using $\rho_{\varepsilon,k}$ itself as a test function we obtain

$$\int_0^T \int_{D_k} |\nabla \Upsilon_\varepsilon(\rho_{\varepsilon,k})|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t + \frac{1}{2} \int_{D_k} \rho_{\varepsilon,k}(x,T)^2 \, \mathrm{d}\mathcal{V}(x) = \frac{1}{2} \int_{D_k} \rho_0^2 \, \mathrm{d}\mathcal{V} \qquad \forall T > 0; \qquad (4.20)$$

a similar computation ensures that in fact all $L^p(D_k)$ norms do not increase:

$$\|\rho_{\varepsilon,k}(t)\|_{L^p(D_k)} \le \|\rho_0\|_{L^p(D_k)} \qquad \forall t > 0, \quad \forall p \in [1,\infty]. \qquad (4.21)$$

If $\hat\rho_{\varepsilon,k}$ is another (approximate) solution corresponding to a different nonnegative $\hat\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$, the L^1-contraction property simply follows upon multiplying the differential equation satisfied by $(\rho_{\varepsilon,k} - \hat\rho_{\varepsilon,k})$ formally by the test function $\mathrm{sign}(\rho_{\varepsilon,k} - \hat\rho_{\varepsilon,k})$ and integrating: this leads to

$$\|\rho_{\varepsilon,k}(t) - \hat\rho_{\varepsilon,k}(t)\|_{L^1(D_k)} \le \|\rho_0 - \hat\rho_0\|_{L^1(D_k)} \qquad \forall t > 0.$$

Actually, to be more rigorous, the sign function should further be approximated by regular nondecreasing functions, see [40, Proposition 3.5]. We are now ready to pass to the limit into the weak formulation satisfied by each $\rho_{\varepsilon,k}$, which reads

$$\int_0^T \int_{D_k} \rho_{\varepsilon,k} \, \partial_t \eta \, \mathrm{d}\mathcal{V}\mathrm{d}t = - \int_{D_k} \rho_0(x) \, \eta(x,0) \, \mathrm{d}\mathcal{V}(x) + \int_0^T \int_{D_k} \langle \nabla P_\varepsilon(\rho_{\varepsilon,k}), \nabla \eta \rangle \, \mathrm{d}\mathcal{V}\mathrm{d}t \qquad (4.22)$$

for every $T > 0$ and every $\eta \in W^{1,2}((0,T);L^2(D_k)) \cap L^2((0,T);W_0^{1,2}(D_k))$ such that $\eta(T) = 0$. Indeed, the energy estimate (4.18) ensures that $\{\nabla P_\varepsilon(\rho_{\varepsilon,k})\}_{\varepsilon>0}$ weakly converges (up to subsequences) as $\varepsilon \downarrow 0$ to some vector field $\vec{w}$ in $L^2(D_k \times (0,T))$, whereas (4.21) yields weak convergence of $\{\rho_{\varepsilon,k}\}_{\varepsilon>0}$ for instance in $L^2(D_k \times (0,T))$ to some limit function ρ_k, still up to subsequences. On the other hand, estimates (4.18)–(4.20) guarantee that $\{\Upsilon_\varepsilon(\rho_{\varepsilon,k})\}_{\varepsilon>0}$ is locally bounded in $H^1(D_k \times (0,T))$; in particular it admits a subsequence that converges pointwise almost everywhere. Since Υ_ε, $\Upsilon_\varepsilon^{-1}$ and P_ε are continuous, monotone increasing functions converging pointwise (and therefore locally uniformly) as $\varepsilon \downarrow 0$ to their continuous limits $\Upsilon(\rho) := \int_0^\rho \sqrt{P'(r)} \, \mathrm{d}r$, Υ^{-1} and P, respectively, we

can assert that also $\{\rho_{\varepsilon,k}\}_{\varepsilon>0}$ and $\{P_\varepsilon(\rho_{\varepsilon,k})\}_{\varepsilon>0}$ converge pointwise, up to subsequences. This is the key to guarantee the identification $\vec{w} = \nabla P(\rho_k)$, so that by letting $\varepsilon \downarrow 0$ in (4.22) we end up with

$$\int_0^T \int_{D_k} \rho_k \, \partial_t \eta \, \mathrm{d}\mathcal{V}\mathrm{d}t = -\int_{D_k} \rho_0(x)\,\eta(x,0)\,\mathrm{d}\mathcal{V}(x) + \int_0^T \int_{D_k} \langle \nabla P(\rho_k)\,, \nabla \eta \rangle \, \mathrm{d}\mathcal{V}\mathrm{d}t\,,$$

which is valid for every $T > 0$ and the same type of test functions η as in (4.22). Note that all the above estimates pass to the limit as $\varepsilon \downarrow 0$ e.g. by lower semicontinuity, yielding

$$\int_0^T \int_{D_k} |\nabla P(\rho_k)|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t + \int_{D_k} \Psi(\rho_k(x,T))\,\mathrm{d}\mathcal{V}(x) \le \int_{D_k} \Psi(\rho_0)\,\mathrm{d}\mathcal{V} \qquad \forall T > 0\,, \qquad (4.23)$$

$$\int_0^T \int_{D_k} \zeta\,|\partial_t \Upsilon(\rho_k)|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t \le \frac{\max_{\mathbb{R}^+} |\zeta'|}{2} \int_{D_k} \Psi(\rho_0)\,\mathrm{d}\mathcal{V} \qquad \forall T > 0\,, \qquad (4.24)$$

$$\int_0^T \int_{D_k} |\nabla \Upsilon(\rho_k)|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t + \frac{1}{2}\int_{D_k} \rho_k(x,T)^2 \, \mathrm{d}\mathcal{V}(x) \le \frac{1}{2}\int_{D_k} \rho_0^2 \, \mathrm{d}\mathcal{V} \qquad \forall T > 0\,, \qquad (4.25)$$

$$\|\rho_k(t)\|_{L^p(D_k)} \le \|\rho_0\|_{L^p(D_k)} \qquad \forall t > 0\,, \quad \forall p \in [1,\infty]\,, \qquad (4.26)$$

$$\|\rho_k(t) - \hat{\rho}_k(t)\|_{L^1(D_k)} \le \|\rho_0 - \hat{\rho}_0\|_{L^1(D_k)} \qquad \forall t > 0\,. \qquad (4.27)$$

At this point we are allowed to let $k \to \infty$, so that D_k will eventually become the whole manifold $\mathbb{M}^n$. By exploiting estimates (4.23)–(4.27) and reasoning similarly to the previous step, we can easily deduce that $\{\rho_k\}_{k\in\mathbb{N}}$ (extended to zero in $\mathbb{M}^n \setminus D_k$) suitably converges as $k \to \infty$ to the energy solution ρ of (4.1), which therefore satisfies (4.11), (4.12) and (4.16) (upon repeating the same procedure starting from $\hat{\rho}_0$), along with

$$\int_0^T \int_{\mathbb{M}^n} \zeta\,|\partial_t \Upsilon(\rho)|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t \le \frac{\max_{\mathbb{R}^+} |\zeta'|}{2} \int_{\mathbb{M}^n} \Psi(\rho_0)\,\mathrm{d}\mathcal{V} \qquad \forall T > 0\,, \qquad (4.28)$$

$$\int_0^T \int_{\mathbb{M}^n} |\nabla \Upsilon(\rho)|^2 \, \mathrm{d}\mathcal{V}\mathrm{d}t + \frac{1}{2}\int_{\mathbb{M}^n} \rho(x,T)^2 \, \mathrm{d}\mathcal{V}(x) \le \frac{1}{2}\int_{\mathbb{M}^n} \rho_0^2 \, \mathrm{d}\mathcal{V} \qquad \forall T > 0\,. \qquad (4.29)$$

Note that since $P \in C^1([0,+\infty))$ and $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$ the r.h.s. of (4.28) is surely finite. We are thus left with proving L^1-continuity, mass conservation and (4.14)–(4.15).

In order to establish the mass-conservation property, we take advantage of a recent result contained in [8], which ensures that under (H1) for every $R \ge 1$ there exist positive constants C, γ independent of R and a nonnegative function $\phi_R \in C_c^\infty(\mathbb{M}^n)$ such that $\phi_R = 1$ in $B_R(o)$, $\mathrm{supp}\,\phi_R \subset B_{\gamma R}(o)$ (let $o \in \mathbb{M}^n$ be a fixed pole), $\phi_R \le 1$ and $|\Delta \phi_R| \le C/R$. See in particular [8, Corollary 2.3]. So let us plug into (4.2) the test function $\eta(x,t) = \phi_R(x)\xi(t)$, where $\xi \in C_c^\infty([0,T))$ with $\xi(0) = 1$; we obtain

$$\int_0^T \int_{\mathbb{M}^n} \rho\,\phi_R\,\xi' \, \mathrm{d}\mathcal{V}\mathrm{d}t = -\int_{\mathbb{M}^n} \rho_0\,\phi_R \, \mathrm{d}\mathcal{V} + \int_0^T \int_{\mathbb{M}^n} \xi \, \langle \nabla P(\rho)\,, \nabla \phi_R \rangle \, \mathrm{d}\mathcal{V}\mathrm{d}t\,. \qquad (4.30)$$

If we suitably let $\xi \to \chi_{[0,T]}$ and we integrate by parts the second term in the r.h.s. of (4.30), we end up with

$$\int_{\mathbb{M}^n} \rho(x,T)\,\phi_R(x)\,\mathrm{d}\mathcal{V}(x)\mathrm{d}t = \int_{\mathbb{M}^n} \rho_0\,\phi_R \, \mathrm{d}\mathcal{V} + \int_0^T \int_{\mathbb{M}^n} P(\rho)\,\Delta \phi_R \, \mathrm{d}\mathcal{V}\mathrm{d}t\,.$$

By letting $R \to \infty$, exploiting the integrability properties of ρ (note that $P(\rho) \in L^1(\mathbb{M}^n \times (0,T))$) along with the above estimate on $\Delta \phi_R$ and the arbitrariness of T, we deduce (4.13).

As concerns L^1-continuity, as a first step we point out that it could be proved by means of an alternative construction of weak energy solutions that takes advantage of time-discretization and the Crandall-Liggett Theorem: see e.g. [16, Remark 3.7]. More comments on such a construction will be made in Remark 4.9 at the end of this section. However, in the present framework it can be

obtained in a more direct fashion, at least under (H1). Indeed, if we let $\zeta \to \chi_{[0,T]}$ in (4.19), upon a passage to the limit as $\varepsilon \downarrow 0$ and $k \to \infty$ we infer that

$$\int_0^T \int_{\mathbb{M}^n} |\partial_t \Upsilon(\rho)|^2 \, \mathrm{d}\mathcal{V} \mathrm{d}t + \frac{1}{2} \int_{\mathbb{M}^n} |\nabla P(\rho(x,T))|^2 \, \mathrm{d}\mathcal{V}(x) \le \frac{1}{2} \int_{\mathbb{M}^n} |\nabla P(\rho_0)|^2 \, \mathrm{d}\mathcal{V} \qquad \forall T > 0 \,.$$

This in particular ensures that, at least for initial data $\rho_0 \in C_c^1(\mathbb{M}^n)$, the curve $t \mapsto \Upsilon(\rho(t))$ is in $W^{1,2}\big((0,T); L^2(\mathbb{M}^n)\big)$, which further guarantees that $\rho(t) \to \rho_0$ as $t \downarrow 0$ in $L_{\mathrm{loc}}^1(\mathbb{M}^n)$ (recall the uniform boundedness of ρ); on the other hand, the just proved mass-conservation property implies $\|\rho(t)\|_{L^1(\mathbb{M}^n)} = \|\rho_0\|_{L^1(\mathbb{M}^n)}$ for all $t > 0$, so that the convergence does occur in $L^1(\mathbb{M}^n)$. By virtue of the contraction estimate (4.16), the L^1-continuity of $t \mapsto \rho(t)$ at $t = 0$ yields the L^1-continuity at any other time, so that in fact $\rho \in C([0, +\infty); L^1(\mathbb{M}^n))$. This holds provided $\rho_0 \in C_c^1(\mathbb{M}^n)$: for a general initial datum $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$, if we take a sequence $\{\rho_{j,0}\}_{j \in \mathbb{N}} \subset C_c^1(\mathbb{M}^n)$ such that $\rho_{j,0} \to \rho_0$ in $L^1(\mathbb{M}^n)$, with $\rho_{j,0} \ge 0$, still the contraction estimate (4.16) ensures that the corresponding sequence of energy solutions $\{\rho_j\}_{j \in \mathbb{N}}$ converges to ρ in $L^\infty(\mathbb{R}^+; L^1(\mathbb{M}^n))$, hence also $t \mapsto \rho(t)$ belongs to $C([0, +\infty); L^1(\mathbb{M}^n))$ (so that a posteriori we have the right to write all the above estimates for *every* rather than *almost every* t or T).

Let us finally establish the approximation properties (4.14)–(4.15). Given $\varepsilon > 0$, if ρ_ε is the weak energy solution of (4.10) then it satisfies (4.11) (with $P \equiv P_\varepsilon$ and $\Psi \equiv \Psi_\varepsilon$), (4.12), (4.13) and (4.28)–(4.29) (with $\Upsilon \equiv \Upsilon_\varepsilon$ and $\Psi \equiv \Psi_\varepsilon$): by proceeding as in the first part of the proof, one can easily infer that $\{\rho_\varepsilon\}_{\varepsilon>0}$ converges pointwise almost everywhere in $\mathbb{M}^n \times \mathbb{R}^+$ as $\varepsilon \downarrow 0$ to ρ, up to subsequences. This implies convergence in $L_{\mathrm{loc}}^1(\mathbb{M}^n)$ for a.e. $t \in \mathbb{R}^+$, given the uniform boundedness of $\{\rho_\varepsilon\}_{\varepsilon>0}$. In order to show that such convergence occurs at *every* t, note that by (4.28) the family $\{\Upsilon_\varepsilon(\rho_\varepsilon)\}_{\varepsilon>0}$ is equicontinuous with values in $L^2(\mathbb{M}^n)$, at least for times bounded away from zero:

$$\|\Upsilon_\varepsilon(\rho_\varepsilon(t)) - \Upsilon_\varepsilon(\rho_\varepsilon(s))\|_{L^2(\mathbb{M}^n)} \le \sqrt{t - s} \, \|\partial_t \Upsilon_\varepsilon(\rho_\varepsilon)\|_{L^2(\mathbb{M}^n \times (s,t))} \qquad \forall t > s > 0 \,;$$

by the Ascoli-Arzelà theorem we then deduce that $\{\Upsilon_\varepsilon(\rho_\varepsilon(t))\}_{\varepsilon>0}$ converges locally in $L^2(\mathbb{M}^n)$ to $\Upsilon(\rho(t))$ for every $t > 0$, whence the convergence of $\{\rho_\varepsilon(t)\}_{\varepsilon>0}$ in $L_{\mathrm{loc}}^1(\mathbb{M}^n)$, thanks to the just recalled uniform boundedness of $\{\rho_\varepsilon\}_{\varepsilon>0}$. Finally, the global convergence under (H1) is again a consequence of mass conservation. $\qquad\square$

As mentioned above, a fundamental ingredient to the strategy of proof of Theorem 2.4 (see Section 5) is the smoothing effect, namely a quantitative $L^1(\mathbb{M}^n)$–$L^\infty(\mathbb{M}^n)$ regularization property of the nonlinear evolution that depends only on the L^1 norm of the initial datum. To this end we need to ask some crucial extra assumptions: the validity of the Sobolev-type inequality (H2) and a bound from below on the degeneracy of P given by the left-hand side of (H4). The proof is largely inspired from [16, Section 4], where a Moser-type iteration is exploited (see also references quoted therein); nevertheless, here we are also interested in keeping track of the dependence of the multiplying constants on m as $m \downarrow 1$.

Proposition 4.4 (Smoothing effect). *Let $\mathbb{M}^n$ ($n \ge 3$) comply with* (H2). *Let P comply with* (H3) *and the left-hand inequality in* (H4). *Let $\varepsilon > 0$ and $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$ be nonnegative. Then the weak energy solution ρ_ε of* (4.10), *where P_ε is defined by* (4.3), *satisfies the* smoothing estimate

$$\|\rho_\varepsilon(t)\|_{L^\infty(\mathbb{M}^n)} \le C \left(t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_{L^1(\mathbb{M}^n)}^{\frac{2}{2+n(m-1)}} + \|\rho_0\|_{L^1(\mathbb{M}^n)} \right) \qquad \forall t > 0 \tag{4.31}$$

provided

$$\|\rho_0\|_{L^\infty(\mathbb{M}^n)} \le \frac{1}{\varepsilon} \,, \tag{4.32}$$

where $C \ge 1$ is a constant depending only on c_0, C_S, n and independent of m ranging in a bounded subset of $(1, +\infty)$. As a consequence, if ρ is the weak energy solution of (4.1) *starting from the*

same initial datum, there holds

$$\|\rho(t)\|_{L^\infty(\mathbb{M}^n)} \leq C \left(t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_{L^1(\mathbb{M}^n)}^{\frac{2}{2+n(m-1)}} + \|\rho_0\|_{L^1(\mathbb{M}^n)} \right) \qquad \forall t > 0. \tag{4.33}$$

Proof. Given $t > 0$, we consider the sequence of time steps $t_j := (1 - 2^{-j})t$, for all $j \in \mathbb{N}$, so that $t_0 = 0$ and $t_\infty = t$. Associated with $\{t_j\}_{j \in \mathbb{N}}$, we take an increasing sequence of exponents $\{p_j\}_{j \in \mathbb{N}}$ to be defined later, such that $p_0 \geq 2$ and $p_\infty = \infty$. Throughout, we will work with the approximate solutions $\{\rho_{\varepsilon,k}\}_{\varepsilon>0,k\in\mathbb{N}}$ defined in the proof of Proposition 4.3, so that the computations we will perform below are justified. The key starting point consists of multiplying the differential equation in (4.10) by the $(p_j - 1)$-th power of $\rho_{\varepsilon,k}$, integrating by parts in $D_k \times [t_j, t_{j+1}]$, using (1.5) (only the bound from below) and (4.5) along with (4.21) and (4.32), so as to obtain

$$\frac{4\,c_0\,m\,p_j\,(p_j - 1)}{(m + p_j - 1)^2} \int_{t_j}^{t_{j+1}} \int_{D_k} \left| \nabla \left(\rho_{\varepsilon,k}^{(m+p_j-1)/2} \right) \right|^2 \mathrm{d}\mathcal{V}\mathrm{d}t \leq$$

$$p_j\,(p_j - 1) \int_{t_j}^{t_{j+1}} \int_{D_k} \rho_{\varepsilon,k}^{p_j-2}\, P_\varepsilon'(\rho_{\varepsilon,k})\, |\nabla \rho_{\varepsilon,k}|^2 \,\mathrm{d}\mathcal{V}\mathrm{d}t = \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{p_j} - \|\rho_{\varepsilon,k}(t_{j+1})\|_{p_j}^{p_j} \leq \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{p_j}. \tag{4.34}$$

For readability's sake, we set $\| \cdot \|_{L^p(D_k)} = \| \cdot \|_p$. Before proceeding further, it is convenient to recall (see [5, Theorem 3.1]) that the Sobolev-type inequality (H2) can equivalently be rewritten in a "Gagliardo-Nirenberg" form as

$$\|f\|_{L^r(\mathbb{M}^n)} \leq \widetilde{C}_S \left(\|\nabla f\|_{L^2(\mathbb{M}^n)} + \|f\|_{L^2(\mathbb{M}^n)} \right)^{\vartheta(s,r,n)} \|f\|_{L^s(\mathbb{M}^n)}^{1-\vartheta(s,r,n)} \qquad \forall f \in W^{1,2}(\mathbb{M}^n) \cap L^s(\mathbb{M}^n)$$

$$\text{for every } 0 < s < r \leq 2^\star, \qquad \text{where } \vartheta = \vartheta(s,r,N) := \frac{2n\,(r - s)}{r\,[2n - s(n - 2)]} \in (0, 1) \tag{4.35}$$

and $\widetilde{C}_S$ is another positive constant that can be taken independent of r, s. Taking advantage of Young's inequality, it is not difficult to show that (4.35) implies

$$\|f\|_{L^r(\mathbb{M}^n)} \leq \widetilde{C}_S \left(\|\nabla f\|_{L^2(\mathbb{M}^n)} + \|f\|_{L^s(\mathbb{M}^n)} \right)^{\vartheta(s,r,n)} \|f\|_{L^s(\mathbb{M}^n)}^{1-\vartheta(s,r,n)} \qquad \forall f \in W^{1,2}(\mathbb{M}^n) \cap L^s(\mathbb{M}^n)$$

$$\text{for every } 0 < s < r \leq 2^\star \text{ with } s \leq 2, \tag{4.36}$$

for a possibly different positive constant $\widetilde{C}_S$ as above that we do not relabel. We are now in position to handle the l.h.s. of (4.34) by applying (4.36) to the function

$$f = \rho_{\varepsilon,k}^{(m+p_j-1)/2}(t),$$

which yields (we can suppose that the solution is not identically zero)

$$\frac{2\,c_0\,m\,p_j\,(p_j - 1)}{\widetilde{C}_S^{\frac{2}{\vartheta}}\,(m + p_j - 1)^2} \int_{t_j}^{t_{j+1}} \frac{\|\rho_{\varepsilon,k}(t)\|_{r(m+p_j-1)/2}^{(m+p_j-1)/\vartheta}}{\|\rho_{\varepsilon,k}(t)\|_{s(m+p_j-1)/2}^{(1-\vartheta)(m+p_j-1)/\vartheta}} \,\mathrm{d}t$$

$$\leq \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{p_j} + \frac{4\,c_0\,m\,p_j\,(p_j - 1)}{(m + p_j - 1)^2} \int_{t_j}^{t_{j+1}} \|\rho_{\varepsilon,k}(t)\|_{s(m+p_j-1)/2}^{m+p_j-1} \,\mathrm{d}t. \tag{4.37}$$

Upon making the (feasible) choices

$$s = \frac{2p_j}{m + p_j - 1}, \qquad r = 2 + \frac{2s}{n} = 2\,\frac{(n + 2)p_j + n(m - 1)}{n(m + p_j - 1)},$$

recalling the recursive definition of $\{t_j\}_{j\in\mathbb{N}}$ and using (4.21), from (4.37) we can infer that

$$\frac{c_0\, m\, p_j\, (p_j-1)\, t}{\widetilde{C}_S^{\frac{2}{\vartheta}}\, 2^j(m+p_j-1)^2}\, \frac{\|\rho_{\varepsilon,k}(t_{j+1})\|_{p_{j+1}}^{p_{j+1}}}{\|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{2p_j/n}} \le \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{p_j} + \frac{2\, c_0\, m\, p_j\, (p_j-1)\, t}{2^j(m+p_j-1)^2}\, \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{m+p_j-1}\,, \quad (4.38)$$

where p_j is also defined recursively by

$$p_{j+1} = \frac{n+2}{n}\, p_j + m-1 \quad\Longrightarrow\quad p_j = \left[p_0 + \frac{n(m-1)}{2}\right]\left(\frac{n+2}{n}\right)^j - \frac{n(m-1)}{2} \quad \forall j\in\mathbb{N}. \tag{4.39}$$

From here on, we will denote by H a generic positive constant that depends only on $c_0, \widetilde{C}_S, n, p_0$ and is independent of m ranging in a bounded subset of $(1,+\infty)$, which may vary from line to line. Hence estimate (4.38) can be rewritten as

$$\|\rho_{\varepsilon,k}(t_{j+1})\|_{p_{j+1}}^{p_{j+1}} \le H\left(\frac{2^j}{t}\, \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{\frac{n+2}{n}\, p_j} + \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{\frac{n+2}{n}\, p_j + m-1}\right). \tag{4.40}$$

By combining (4.21), the monotonicity of $\{p_j\}_{j\in\mathbb{N}}$, interpolation and Young's inequalities, we easily obtain:

$$\|\rho_{\varepsilon,k}(t_j)\|_{p_j} \le \|\rho_0\|_\infty + \|\rho_0\|_{p_0}\,,$$

whence from (4.40) there follows

$$\|\rho_{\varepsilon,k}(t_{j+1})\|_{p_{j+1}} \le H^{\frac{j+1}{p_{j+1}}}\left[t^{-1} + \left(\|\rho_0\|_\infty + \|\rho_0\|_{p_0}\right)^{m-1}\right]^{\frac{1}{p_{j+1}}} \|\rho_{\varepsilon,k}(t_j)\|_{p_j}^{\frac{n+2}{n}\frac{p_j}{p_{j+1}}}\,. \tag{4.41}$$

Iterating (4.41) and exploiting again (4.21) (in the l.h.s. of (4.41)) yields

$$\|\rho_{\varepsilon,k}(t)\|_{p_{j+1}} \le H^{\frac{\sum_{h=1}^{j+1} h\left(\frac{n+2}{n}\right)^{j+1-h}}{p_{j+1}}}\left[t^{-1} + \left(\|\rho_0\|_\infty + \|\rho_0\|_{p_0}\right)^{m-1}\right]^{\frac{\sum_{h=0}^{j}\left(\frac{n+2}{n}\right)^h}{p_{j+1}}} \|\rho_0\|_{p_0}^{\left(\frac{n+2}{n}\right)^{j+1}\frac{p_0}{p_{j+1}}}\,;$$

by letting $j\to\infty$, recalling (4.39), we thus end up with

$$\|\rho_{\varepsilon,k}(t)\|_\infty \le H\left[t^{-1} + \left(\|\rho_0\|_\infty + \|\rho_0\|_{p_0}\right)^{m-1}\right]^{\frac{n}{2p_0+n(m-1)}} \|\rho_0\|_{p_0}^{\frac{2p_0}{2p_0+n(m-1)}}\,,$$

whence

$$\|\rho_{\varepsilon,k}(t)\|_\infty \le H\left[t^{-\frac{n}{2p_0+n(m-1)}} + \left(\|\rho_0\|_\infty + \|\rho_0\|_{p_0}\right)^{\frac{n(m-1)}{2p_0+n(m-1)}}\right] \|\rho_0\|_{p_0}^{\frac{2p_0}{2p_0+n(m-1)}}\,. \tag{4.42}$$

At this point we need to take advantage of the following version of Young's inequality:

$$A^\theta\, B^{1-\theta} \le \epsilon\,\theta\, A + \epsilon^{-\frac{\theta}{1-\theta}}\, (1-\theta)\, B \qquad \forall A, B, \epsilon > 0\,, \quad \forall\theta\in(0,1)\,.$$

Upon choosing

$$A = \|\rho_0\|_\infty + \|\rho_0\|_{p_0}\,, \quad B = \|\rho_0\|_{p_0}\,, \quad \theta = \frac{n(m-1)}{2p_0+n(m-1)}\,, \quad \epsilon = \left(H\,\theta\,2^{1+\frac{\theta}{m-1}}\right)^{-1}\,,$$

from (4.42) we infer that

$$\|\rho_{\varepsilon,k}(t)\|_\infty \le \frac{\|\rho_0\|_\infty}{2^{1+\frac{\theta}{m-1}}} + H\, t^{-\frac{\theta}{m-1}}\, \|\rho_0\|_{p_0}^{1-\theta} + \left[2^{-1-\frac{\theta}{m-1}} + H\left(H\,\theta\,2^{1+\frac{\theta}{m-1}}\right)^{\frac{\theta}{1-\theta}}\right]\|\rho_0\|_{p_0}\,; \tag{4.43}$$

since θ stays bounded away from 1 and $\theta/(m-1)$ stays bounded as m ranges in a bounded subset of $(1,+\infty)$, we can equivalently rewrite (4.43) as

$$\|\rho_{\varepsilon,k}(t)\|_\infty \le \frac{\|\rho_0\|_\infty}{2^{1+\frac{\theta}{m-1}}} + H\, t^{-\frac{\theta}{m-1}}\, \|\rho_0\|_{p_0}^{1-\theta} + H\, \|\rho_0\|_{p_0} \qquad \forall t > 0\,. \tag{4.44}$$

In order to remove the dependence of the r.h.s. of (4.44) on $\|\rho_0\|_\infty$, we can use a time-shift argument, namely for each $j \in \mathbb{N}$ we consider (4.44) evaluated at $t \equiv t/2^j$ with time origin shifted from 0 to $t/2^{j+1}$ (we implicitly rely on the uniqueness of energy solutions). This, along with (4.21), ensures that

$$\left\|\rho_{\varepsilon,k}\big(t/2^j\big)\right\|_\infty \leq \frac{\left\|\rho_{\varepsilon,k}\big(t/2^{j+1}\big)\right\|_\infty}{2^{1+\frac{\theta}{m-1}}} + 2^{\frac{\theta(j+1)}{m-1}} H \, t^{-\frac{\theta}{m-1}} \|\rho_0\|_{p_0}^{1-\theta} + H \, \|\rho_0\|_{p_0} \qquad \forall j \in \mathbb{N}. \quad (4.45)$$

By iterating (4.45) from $j = 0$ to $j = J \in \mathbb{N}$, we obtain:

$$\|\rho_{\varepsilon,k}(t)\|_\infty \leq \frac{\|\rho_0\|_\infty}{2^{\left(1+\frac{\theta}{m-1}\right)(J+1)}} + 2^{\frac{\theta}{m-1}} H \, t^{-\frac{\theta}{m-1}} \|\rho_0\|_{p_0}^{1-\theta} \sum_{j=0}^{J} 2^{-j} + H \, \|\rho_0\|_{p_0} \sum_{j=0}^{J} 2^{-\left(1+\frac{\theta}{m-1}\right)j},$$

so that taking limits as $J \to \infty$ yields

$$\|\rho_{\varepsilon,k}(t)\|_\infty \leq H \left(t^{-\frac{\theta}{m-1}} \|\rho_0\|_{p_0}^{1-\theta} + \|\rho_0\|_{p_0} \right) \qquad \forall t > 0. \quad (4.46)$$

We finally need to extend the just proved estimate to the case $p_0 = 1$, the one we are primarily interested in. Given any $p_0 \geq 2$ as above (fixed), let us plug the interpolation inequality

$$\|\rho_0\|_{p_0} \leq \|\rho_0\|_\infty^{1-\frac{1}{p_0}} \|\rho_0\|_1^{\frac{1}{p_0}}$$

into (4.46):

$$\|\rho_{\varepsilon,k}(t)\|_\infty \leq C \, \|\rho_0\|_\infty^{\frac{2(p_0-1)}{2p_0+n(m-1)}} \left(t^{-\frac{n}{2p_0+n(m-1)}} \|\rho_0\|_1^{\frac{2}{2p_0+n(m-1)}} + \|\rho_0\|_\infty^{\frac{n(m-1)(p_0-1)}{p_0[2p_0+n(m-1)]}} \|\rho_0\|_1^{\frac{1}{p_0}} \right) \quad \forall t > 0, \quad (4.47)$$

where C stands for a generic positive constant as in the statement. By exploiting again a time-shift argument, it is readily seen that (4.47) entails, for all $j \in \mathbb{N}$,

$$\left\|\rho_{\varepsilon,k}\big(t/2^j\big)\right\|_\infty \leq 2^{\frac{n(j+1)}{2p_0+n(m-1)}} C \left\|\rho_{\varepsilon,k}\big(t/2^{j+1}\big)\right\|_\infty^{\frac{2(p_0-1)}{2p_0+n(m-1)}}$$
$$\times \left(t^{-\frac{n}{2p_0+n(m-1)}} \|\rho_0\|_1^{\frac{2}{2p_0+n(m-1)}} + \|\rho_0\|_\infty^{\frac{n(m-1)(p_0-1)}{p_0[2p_0+n(m-1)]}} \|\rho_0\|_1^{\frac{1}{p_0}} \right). \quad (4.48)$$

Since

$$\frac{2(p_0-1)}{2p_0+n(m-1)} \leq 1 - \frac{1}{p_0},$$

a straightforward iteration of (4.48) ensures that

$$\|\rho_{\varepsilon,k}(t)\|_\infty \leq C \left(t^{-\frac{n}{2p_0+n(m-1)}} \|\rho_0\|_1^{\frac{2}{2p_0+n(m-1)}} + \|\rho_0\|_\infty^{\frac{n(m-1)(p_0-1)}{p_0[2p_0+n(m-1)]}} \|\rho_0\|_1^{\frac{1}{p_0}} \right)^{\frac{2p_0+n(m-1)}{2+n(m-1)}},$$
$$\leq C \left(t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_1^{\frac{2}{2+n(m-1)}} + \|\rho_0\|_\infty^{\frac{n(m-1)(p_0-1)}{p_0[2+n(m-1)]}} \|\rho_0\|_1^{\frac{2p_0+n(m-1)}{p_0[2+n(m-1)]}} \right). \quad (4.49)$$

By applying a Young-type inequality similar to the one that led us to (4.43), from (4.49) we easily deduce that

$$\|\rho_{\varepsilon,k}(t)\|_\infty \leq \frac{\|\rho_0\|_\infty}{2^{1+\frac{n}{2+n(m-1)}}} + C \, t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_1^{\frac{2}{2+n(m-1)}}$$
$$+ C \left(C \frac{n(m-1)(p_0-1)}{p_0[2+n(m-1)]} 2^{1+\frac{n}{2+n(m-1)}} \right)^{\frac{n(m-1)(p_0-1)}{2p_0+n(m-1)}} \|\rho_0\|_1 \quad (4.50)$$
$$\leq \frac{\|\rho_0\|_\infty}{2^{1+\frac{n}{2+n(m-1)}}} + C \, t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_1^{\frac{2}{2+n(m-1)}} + C \, \|\rho_0\|_1 \qquad \forall t > 0.$$

Estimate (4.50) is completely analogous to (4.44), so that by reasoning in the same fashion we end up with

$$\|\rho_{\varepsilon,k}(t)\|_{L^\infty(D_k)} \le C \left(t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_{L^1(D_k)}^{\frac{2}{2+n(m-1)}} + \|\rho_0\|_{L^1(D_k)} \right) \qquad \forall t > 0. \tag{4.51}$$

Recalling the convergence results encompassed by Proposition 4.3, the smoothing effect (4.31) follows by letting $k \to \infty$ in (4.51), whereas (4.33) follows by letting $\varepsilon \downarrow 0$ in (4.31). $\qquad \square$

The next proposition establishes that solutions starting from bounded and compactly-supported data stay with compact support, at least for short times. It is a consequence of the power degeneracy of P induced by assumption (H4) (here we need both sides), hence it is a purely *nonlinear* effect. We stress that this property will be crucial in order to show two essential facts: solutions starting from data in $\mathscr{M}_2^M(\mathbb{M}^n)$ belong to $\mathscr{M}_2^M(\mathbb{M}^n)$ for all times and they form a *continuous* curve with values in $(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2)$.

Proposition 4.5 (Compactness of the support). *Let P comply with* (H3) *and* (H4). *Let $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$ be nonnegative with* compact support. *Then there exist $t_1 > 0$ and a compact set $B \subset \mathbb{M}^n$, depending on ρ_0, m, c_0, c_1 and $\mathbb{M}^n$, such that the weak energy solution ρ to* (4.1) *satisfies*

$$\operatorname{supp} \rho(t) \subset B \qquad \forall t \in [0, t_1]. \tag{4.52}$$

Proof. Since $\mathbb{M}^n$ is a smooth, complete, connected and noncompact Riemannian manifold, it is well known that it admits a regular exhaustion, namely a sequence of open sets $D_k \subset \mathbb{M}^n$ such that $\overline{D}_k$ is a smooth, compact manifold with boundary (for all $k \in \mathbb{N}$) and there hold

$$\overline{D}_k \Subset D_{k+1} \qquad \text{and} \qquad \bigcup_{k=1}^\infty D_k = \mathbb{M}^n.$$

In particular, ∂D_k is a smooth $(n-1)$-dimensional, compact, orientable submanifold of $\mathbb{M}^n$, with a natural orientation given by the outward-pointing normal field w.r.t. D_k. For such a construction we refer e.g. to [28, Proposition 2.28, Theorem 6.10, Propositions 15.24 and 15.33]. Given $\epsilon > 0$, let us define the set of all points inside D_k whose distance from ∂D_k is smaller than ϵ, that is

$$D_k^\epsilon := \{ x \in D_k : \ \mathsf{d}(x, \partial D_k) < \epsilon \}.$$

Since ∂D_k enjoys the above recalled regularity properties, if ϵ is sufficiently small then each $x \in D_k^\epsilon$ admits a unique projection $\pi(x)$ onto ∂D_k. Hence every such point is uniquely identified by the pair $\Pi(x) := (\pi(x), \delta(x))$, where $\delta(x)$ is the geodesic distance from x to $\pi(x)$ (or equivalently to ∂D_k). Moreover, the map Π is a diffeomorphism between D_k^ϵ and $\partial D_k \times (0, \epsilon)$, so that one can use $\delta = \delta(x)$ and $\pi = \pi(x)$ as coordinates that span the whole D_k^ϵ (see e.g. [15]). It is not difficult to check that δ being a geodesic coordinate, the Laplacian of a regular function ϕ (defined on D_k^ϵ) that depends only on δ reads

$$\Delta\phi(\pi, \delta) = \phi''(\delta) + \mathsf{m}(\pi, \delta)\, \phi'(\delta) \qquad \forall (\pi, \delta) \in \partial D_k \times (0, \epsilon), \tag{4.53}$$

where $\mathsf{m}(\pi, \delta)$ is also regular (in fact it is the Laplacian of the distance function itself).

Taking advantage of such framework, first of all we pick k so large that $\operatorname{supp} \rho_0 \subset D_{k-1}$ and $\epsilon > 0$ so small that, alongside with the unique-projection property, there holds $D_{k-1} \cap D_k^\epsilon = \emptyset$. Then we define

$$\Sigma_\epsilon := \Pi^{-1}(\partial D_k \times \{\epsilon\}),$$

namely the set of points inside D_k whose distance to ∂D_k is equal to ϵ, which describes a smooth submanifold having analogous properties to ∂D_k (note that, since one has the right to choose ϵ

arbitrarily small, Π can smoothly be extended up to $\partial D_k \times \{\epsilon\}$). We also define Ω_ϵ to be the regular domain enclosed by Σ_ϵ. Now let us consider the Cauchy-Dirichlet problem

$$\begin{cases} \partial_t u = \Delta P(u) & \text{in } \mathbb{M}^n \setminus \Omega_\epsilon \times (0, t_1), \\ u = \|\rho_0\|_\infty & \text{on } \Sigma_\epsilon \times (0, t_1), \\ u = 0 & \text{on } \mathbb{M}^n \setminus \Omega_\epsilon \times \{0\}, \end{cases} \tag{4.54}$$

where $t_1 > 0$ is a small enough time to be chosen later. Since $\rho \leq \|\rho_0\|_\infty$ in $\mathbb{M}^n \times \mathbb{R}^+$ and $\operatorname{supp} \rho_0 \subset \Omega_\epsilon$, it is apparent that ρ is a *subsolution* of (4.54). Our aim is to construct a *supersolution* which depends spatially only on δ and has compact support for all $t \in [0, t_1]$. The candidate profile is modeled after Euclidean planar *traveling waves* for the porous medium equation, see [40, Section 4.3]. That is, we consider the following function:

$$\overline{u}(\delta, t) := P^{-1}\left(\left[C_1 \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+ \right]^{\frac{m}{m-1}} \right) \qquad \forall (\delta, t) \in (0, \epsilon) \times [0, t_1], \tag{4.55}$$

where C_1 and C_2 are positive constants to be selected. In view of the assumptions on P, it is not difficult to deduce the following inequalities:

$$\left(\frac{v}{c_1} \right)^{\frac{1}{m}} \leq P^{-1}(v) \leq \left(\frac{v}{c_0} \right)^{\frac{1}{m}} \qquad \forall v \geq 0, \tag{4.56}$$

$$\left[P^{-1} \right]'(v) \geq \frac{c_0^{1 - \frac{1}{m}}}{c_1 m} v^{\frac{1}{m} - 1} \qquad \forall v > 0. \tag{4.57}$$

Clearly $\overline{u}(\delta, 0) \geq 0$ and, thanks to (4.56),

$$\overline{u}(\epsilon, t) \geq c_1^{-\frac{1}{m}} \left[C_1 \frac{\epsilon}{2} \right]^{\frac{1}{m-1}} \qquad \forall t \geq 0;$$

hence a first requirement to make sure that $\overline{u}$ complies with the boundary condition in (4.54) is

$$C_1 \geq \frac{2}{\epsilon} c_1^{\frac{m-1}{m}} \|\rho_0\|_\infty^{m-1}. \tag{4.58}$$

Let us now compute the derivatives of $\overline{u}$ and $P(\overline{u})$ needed to construct a supersolution:

$$\partial_t \overline{u}(\delta, t) = C_2 C_1^{\frac{m}{m-1}} \frac{m}{m-1} \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+^{\frac{1}{m-1}} \left[P^{-1} \right]'\left(\left[C_1 \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+ \right]^{\frac{m}{m-1}} \right)$$

$$\overset{(4.57)}{\geq} C_2 C_1^{\frac{1}{m-1}} \frac{c_0^{\frac{m-1}{m}}}{(m-1)c_1} \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+^{\frac{2-m}{m-1}}, \tag{4.59}$$

$$\partial_\delta (P(\overline{u}))(\delta, t) = C_1^{\frac{m}{m-1}} \frac{m}{m-1} \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+^{\frac{1}{m-1}}, \tag{4.60}$$

$$\partial_{\delta\delta} (P(\overline{u}))(\delta, t) = C_1^{\frac{m}{m-1}} \frac{m}{(m-1)^2} \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+^{\frac{2-m}{m-1}}. \tag{4.61}$$

We pick t_1 in such a way that the distance of the support of $\overline{u}$ from ∂D_k is not smaller than $\epsilon/4$ for all $t \in [0, t_1]$, namely

$$t_1 = \frac{\epsilon}{4C_2}. \tag{4.62}$$

Let σ denote the maximum of $\mathsf{m}(\pi, \delta)$ in the region $E_\epsilon := \partial D_k \times [\epsilon/4, \epsilon]$. Because $\overline{u}$ is nondecreasing in δ and (4.62) ensures that the support of $\overline{u}$ lies in E_ϵ, in order to guarantee that the latter is a (weak) supersolution of the differential equation in (4.54) it suffices to ask that (recalling (4.53))

$$\partial_t \overline{u}(\delta, t) \geq \partial_{\delta\delta} P(\overline{u})(\delta, t) + \sigma \partial_\delta P(\overline{u})(\delta, t) \qquad \forall (\delta, t) \in [\epsilon/4, \epsilon] \times [0, t_1]. \tag{4.63}$$

Thanks to (4.59)–(4.61), after some simplifications we find that (4.63) holds if

$$C_2 \frac{c_0^{\frac{m-1}{m}}}{(m-1)c_1} \geq C_1 \frac{m}{(m-1)^2} \left[1 + (m-1)\sigma \left(C_2 t + \delta - \frac{\epsilon}{2} \right)_+ \right] \qquad \forall (\delta, t) \in [\epsilon/4, \epsilon] \times [0, t_1] \,,$$

the latter inequality being in turn implied by

$$C_2 \geq C_1 \frac{c_1 \, m}{(m-1)c_0^{\frac{m-1}{m}}} \left[1 + \frac{3(m-1)\sigma\epsilon}{4} \right]. \tag{4.64}$$

Hence by choosing C_1 as in (4.58), C_2 as in (4.64) and finally t_1 as in (4.62), we infer that (4.55) is indeed a supersolution of (4.54) (obviously extended in $\mathbb{M}^n \setminus D_k$). By comparison we can therefore assert that $\rho \leq \overline{u}$ in $\mathbb{M}^n \setminus \Omega_\epsilon \times [0, t_1]$, which yields (4.52) with $B = \overline{D}_k$.

As concerns the comparison principle we have just applied, let us point out that in order to justify it rigorously one should know a priori that ρ is also a strong solution, namely that it has an $L^1(\mathbb{M}^n)$ time derivative: see [40, Section 8.2], we refer in particular to the analogue of [40, Lemma 8.11] in our framework. On the other hand $\overline{u}$ is a strong supersolution by construction. To circumvent this issue, it is enough (for instance) to exploit the fact that ρ can always be seen as the limit of solutions ρ_j to homogeneous Dirichlet problems set up on each D_j (recall the proof of Proposition 4.3). Since every ρ_j is a strong solution in D_j (see e.g. [40, Corollary 8.3] in the Euclidean setting) and $\overline{u}$ clearly satisfies homogeneous Dirichlet boundary conditions on ∂D_j for j large enough, one obtains $\rho_j \leq \overline{u}$ in $D_j \setminus \Omega_\epsilon \times [0, t_1]$ for every $j \in \mathbb{N}$ by proceeding as above, and then lets $j \to \infty$. $\quad\square$

4.2. Variational solutions, linearized and adjoint equation.
For the purposes of proving Theorem 2.4, we first introduce a suitable (variational) notion of solution of the approximate problem (4.10) and we show its equivalence with the notion of weak energy solution discussed in the previous subsection. Hereafter we identify $\mathbb{H}$ with its dual $\mathbb{H}'$ and consider the following Hilbert triple:

$$\mathbb{V} \hookrightarrow \mathbb{H} \equiv \mathbb{H}' \hookrightarrow \mathbb{V}'.$$

Problem (4.10) reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = \Delta(P_\varepsilon(\rho)), \qquad \rho(0) = \rho_0 \,, \tag{4.65}$$

where ρ is seen as a curve with values in $\mathbb{H}$ and, accordingly, Δ is the realization of the (self-adjoint) Laplace-Beltrami operator in $\mathbb{H}$. In agreement with the notations of Subsection 1.2, for every $\varepsilon > 0$ and $T > 0$ we recall the definition of the set $\mathcal{ND}(0, T)$ associated with P_ε:

$$\mathcal{ND}_{P_\varepsilon}(0, T) := \left\{ u \in W^{1,2}((0, T); \mathbb{H}) \cap C^1([0, T]; \mathbb{V}') : \; u \geq 0, \; P_\varepsilon(u) \in L^2((0, T); \mathbb{D}) \right\}.$$

Note that the nonlinearity P_ε falls within the class of functions considered in [4, Subsection 3.3], in the more general framework of Dirichlet forms.

Definition 4.6 (Strong variational solutions). *Let P comply with* (H3) *and P_ε ($\varepsilon > 0$) be defined by* (4.3). *Let $\rho_0 \in \mathbb{H}$, with $\rho_0 \geq 0$, and $T > 0$. We say that a curve $\rho \in W^{1,2}((0, T); \mathbb{V}, \mathbb{V}')$, with $\rho \geq 0$, is a* strong variational solution *of* (4.65) *in the time interval* $(0, T)$ *if there holds*

$$- \,_{\mathbb{V}'}\!\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\rho(t), \eta \right\rangle_{\mathbb{V}} = \int_{\mathbb{M}^n} \left\langle \nabla P_\varepsilon(\rho(t)), \nabla \eta \right\rangle \mathrm{d}\mathcal{V} \qquad \text{for a.e. } t \in (0, T), \quad \forall \eta \in \mathbb{V}, \tag{4.66}$$

and $\lim_{t\downarrow 0} \rho(t) = \rho_0$ *in* $\mathbb{H}$.

We point out that Definition 4.6 does make sense since $\rho \in W^{1,2}((0, T); \mathbb{V}, \mathbb{V}')$ implies $\rho \in C([0, T]; \mathbb{H})$, see [4, formula (3.28)] (this is indeed a rather general fact).

The following well-posedness result is established by [4, Theorem 3.4].

Proposition 4.7 (Existence of strong variational solutions)**.** *Let P comply with* (H3) *and P_ε ($\varepsilon > 0$) be defined by* (4.3)*. Let $\rho_0 \in \mathbb{H}$, with $\rho_0 \geq 0$, and $T > 0$. Then there exists a* unique *strong variational solution of* (4.65)*, in the sense of Definition 4.6. If in addition $\rho_0 \in \mathbb{V}$ then $\rho \in \mathcal{ND}_{P_\varepsilon}(0, T)$.*

Weak energy solutions and strong variational solutions in fact coincide.

Proposition 4.8 (Equivalent notions of solution)**.** *Let P comply with* (H3) *and P_ε ($\varepsilon > 0$) be defined by* (4.3)*. Let $T > 0$. Then for any nonnegative $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n)$ the weak energy solution of* (4.10) *(provided by Proposition 4.3) and the strong variational solution of* (4.65) *(provided by Proposition 4.7) are equal, up to $t = T > 0$.*

Proof. Let us denote by $\hat{\rho}$ the solution constructed in Proposition 4.7. Thanks to the integrability properties of $\hat{\rho}$ and the C^1 regularity of the map $\rho \mapsto P_\varepsilon(\rho)$, we know that $\hat{\rho} \in L^2(\mathbb{M}^n \times (0, T))$, which is equivalent to $P_\varepsilon(\hat{\rho}) \in L^2(\mathbb{M}^n \times (0, T))$, and $\nabla \hat{\rho} \in L^2(\mathbb{M}^n \times (0, T))$, which is equivalent to $\nabla P_\varepsilon(\hat{\rho}) \in L^2(\mathbb{M}^n \times (0, T))$. By (4.66), for any curve $\eta \in W^{1,2}((0, T); W^{1,2}(\mathbb{M}^n))$ with $\eta(T) = 0$ there holds

$$- {}_{\mathbb{V}'}\!\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t), \eta(t) \right\rangle_{\mathbb{V}} = \int_{\mathbb{M}^n} \langle \nabla P_\varepsilon(\hat{\rho}(t)), \nabla \eta(t) \rangle \, \mathrm{d}\mathcal{V} \qquad \text{for a.e. } t \in (0, T) \, ; \qquad (4.67)$$

since both $\hat{\rho}$ and η are continuous curves with values in $L^2(\mathbb{M}^n)$, integrating (4.67) between $t = 0$ and $t = T$ yields

$$\int_0^T \int_{\mathbb{M}^n} \hat{\rho} \, \partial_t \eta \, \mathrm{d}\mathcal{V}\mathrm{d}t + \int_{\mathbb{M}^n} \rho_0(x) \, \eta(x, 0) \, \mathrm{d}\mathcal{V}(x) = \int_0^T \int_{\mathbb{M}^n} \langle \nabla P_\varepsilon(\hat{\rho}), \nabla \eta \rangle \, \mathrm{d}\mathcal{V}\mathrm{d}t \, ,$$

which shows that $\hat{\rho}$ is also a weak energy solution of (4.10) starting from ρ_0 and therefore it coincides with the one provided by Proposition 4.3, up to the observations made in the first part of the corresponding proof. □

Remark 4.9 (On possibly different constructions of weak energy solutions)**.** In Subsection 4.1 we used a well-established approach to prove existence of weak energy solutions of (4.1), which consists in the first place of solving *evolution* problems associated with nondegenerate nonlinearities on regular domains. As shown above, this technique is suitable to prove several key estimates, especially the smoothing effect of Proposition 4.4. Nevertheless, we mention that there exists at least another fruitful method, which relies first on solving a *discretized* version (in time) of problem (4.1) by means of the Crandall-Liggett Theorem (see [40, Chapter 10] in the Euclidean context). This is precisely the technique employed in [4, Section 3.3] to construct solutions of (4.1) in the general setting considered therein; the advantage of such an approach is that it also works in nonsmooth frameworks (like metric-measure spaces). However, in that case the proof of the smoothing effect is less trivial and should be investigated further (one can no longer differentiate L^p norms along the flow), for instance by taking advantage of the abstract tools developed in [12], which a priori work upon assuming the validity of the stronger *Euclidean* Sobolev inequality (2.9).

To implement the Hamiltonian approach described in the Introduction, it is necessary to study the linearization of (4.65) along with its formal adjoint. More precisely, in the variational setting $\mathbb{V} \hookrightarrow \mathbb{H} \hookrightarrow \mathbb{V}'$ described above, we can consider the *forward linearized* equation

$$\frac{\mathrm{d}}{\mathrm{d}t} w = \Delta \big[P_\varepsilon'(\rho) \, w \big] \, , \qquad w(0) = w_0 \, , \qquad (4.68)$$

and the *backward adjoint* equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi = -P_\varepsilon'(\rho) \, \Delta \phi \, , \qquad \phi(T) = \phi_T \, . \qquad (4.69)$$

Following [4, Theorem 4.5], we begin with rephrasing in our setting a well-posedness result for (4.68). Hereafter we denote by $\mathbb{D}'$ the dual of $\mathbb{D}$, recalling that $\mathbb{H} \hookrightarrow \mathbb{V}' \hookrightarrow \mathbb{D}'$ with continuous and dense inclusions.

Theorem 4.10 (Forward linearized equation). *Let P comply with* (H3) *and P_ε ($\varepsilon > 0$) be defined by* (4.3). *Let $T > 0$. For every nonnegative $\rho \in L^2((0,T); \mathbb{H})$ and for every $w_0 \in \mathbb{V}'$, there exists a unique weak solution $w \in W^{1,2}((0,T); \mathbb{H}, \mathbb{D}')$ of* (4.68), *in the sense that it satisfies*

$$_{\mathbb{V}'}\langle w(r), \theta(r) \rangle_{\mathbb{V}} - \int_0^r \int_{\mathbb{M}^n} \left[\partial_t \theta(t) + P_\varepsilon'(\rho(t)) \, \Delta \theta(t) \right] w(t) \, \mathrm{d}\mathcal{V} \mathrm{d}t = {}_{\mathbb{V}'}\langle w_0, \theta(0) \rangle_{\mathbb{V}} \qquad \forall r \in [0,T] \ (4.70)$$

for every $\theta \in W^{1,2}((0,T); \mathbb{D}, \mathbb{H})$.

As concerns (4.69) we have the following result, whose proof can be found in [4, Theorem 4.1].

Theorem 4.11 (Backward adjoint equation). *Let P comply with* (H3) *and P_ε ($\varepsilon > 0$) be defined by* (4.3). *Let $T > 0$. For every nonnegative $\rho \in L^2((0,T); \mathbb{H})$ and for every $\phi_T \in \mathbb{V}$, there exists a unique strong solution $\phi \in W^{1,2}((0,T); \mathbb{D}, \mathbb{H})$ of* (4.69). *Moreover, if $\phi_T \in L^\infty(\mathbb{M}^n) \cap \mathbb{V}$ then $\|\phi(t)\|_{L^\infty(\mathbb{M}^n)} \leq \|\phi_T\|_{L^\infty(\mathbb{M}^n)}$ for every $t \in [0,T]$.*

5. PROOF OF THE MAIN RESULTS

This section is entirely devoted to proving Theorems 2.4 and 2.5. After a brief introduction to the strategy of proof of Theorem 2.4 (Subsection 5.1), we will first treat the noncompact case (Subsection 5.2) and then shortly address the compact case (Subsection 5.3). Finally, in Subsection 5.4 we will show that our estimate is optimal for small times, namely Theorem 2.5.

5.1. **Outline of the strategy.** The idea is to prove the stability estimate (2.6) for a suitable approximation of problem (2.1), passing to the limit in the approximation scheme only at the very end. Let us briefly sketch the main steps of the proof.

1. We firstly consider the "elliptic" nonlinearity P_ε as in (4.3) and introduce a regular initial density ρ_0 belonging to $L_c^\infty(\mathbb{M}^n) \cap \mathbb{V}$. We denote by ρ, ϕ and w the solutions of the approximated problems (4.10), (4.68) and (4.69), respectively (for the moment for simplicity we omit the subscript ε).

2. We estimate the derivative $\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_{\rho(t)}[\phi(t)]$ of the Hamiltonian functional defined in (3.7). Here it is essential to exploit the lower bound on the Ricci curvature in the Bakry-Émery form (3.6), which allows us to deduce that (Lemma 5.2)

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_{\rho(t)}[\phi(t)] \geq -K \int_{\mathbb{M}^n} \Gamma(\phi(t)) \, P_\varepsilon(\rho(t)) \, \mathrm{d}\mathcal{V}.$$

We then use the smoothing effect provided by Proposition 4.4 to integrate the above differential inequality; this yields the estimate

$$\mathcal{E}_{\rho(t)}[\phi(t)] \geq \exp(-K \, C(t,m,n)) \, \mathcal{E}_{\rho_0}[\phi(0)],$$

where an explicit computation of $C(t,m,n) > 0$ is given in Lemma 5.3.

3. We take a pair of initial data $\rho_0^0, \rho_0^1 \in L_c^\infty(\mathbb{M}^n) \cap \mathbb{V}$ and connect them by a *regular* curve $\{\rho_0^s\}_{s \in [0,1]}$ (in the sense of Definition 3.3). For any ρ_0^s, hereafter $t \mapsto \rho^s(t)$ will stand for the corresponding solution of (4.10) and ϕ^s for a solution of (4.69) with $\rho \equiv \rho^s$. We then denote by $(s,x) \mapsto Q_s\varphi(x)$ the (Lipschitz) solution of the Hopf-Lax problem (3.16) starting from an arbitrary $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$ and by $w^s(t) \equiv t \mapsto \frac{\mathrm{d}}{\mathrm{d}s} \rho^s(t)$ the solution of the linearized equation (4.68). For every $t > 0$ we compute the Wasserstein distance $\mathcal{W}_2(\rho^0(t), \rho^1(t))$ in

the (Kantorovich) formulation recalled by Proposition 3.1 in terms of the Hamiltonian. The "duality" relation between ϕ^s and w^s (Lemma 5.4) guarantees that

$$\int_{\mathbb{M}^n} Q_1\varphi\, \rho^1(t)\, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi\, \rho^0(t)\, \mathrm{d}\mathcal{V} = \int_0^1 \left(-\frac{1}{2}\mathcal{E}_{\rho^s(t)}[\phi^s(t)] + {}_{\mathbb{V}'}\langle w^s(0), \phi^s(0)\rangle_{\mathbb{V}} \right) \mathrm{d}s\,,$$

where the final datum of ϕ^s is given at time $T \equiv t$ by $\phi^s(t) = Q_s\varphi$.

4. By exploiting the regularity of the curve $s \mapsto \rho_0^s \mathcal{V} =: \mu^s$, we can take advantage of the key identity

$$\int_0^1 |\dot{\mu}^s|^2\, \mathrm{d}s = \int_0^1 \mathcal{E}_{\rho_0^s}^*\left[\tfrac{\mathrm{d}}{\mathrm{d}s}\rho_0^s\right] \mathrm{d}s\,.$$

By combining the latter with the estimate obtained in Step 2 and recalling the definition (3.8) of the (Fenchel) *dual* Hamiltonian $\mathcal{E}_\rho^*$, we can deduce that

$$\int_{\mathbb{M}^n} Q_1\varphi\, \rho^1(t)\, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi\, \rho^0(t)\, \mathrm{d}\mathcal{V} \le \frac{1}{2}\exp\{K\,C(t,m,n)\} \int_0^1 |\dot{\mu}^s|^2\, \mathrm{d}s\,;$$

this is the content of Lemma 5.5.

5. We use Lemma 3.5, which ensures that the right-hand side can be made arbitrarily close to the squared Wasserstein distance between ρ_0^0 and ρ_0^1 (this in fact implies a further approximation of the initial data). As a consequence, we end up with

$$\int_{\mathbb{M}^n} Q_1\varphi\, \rho_\varepsilon^1(t)\, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi\, \rho_\varepsilon^0(t)\, \mathrm{d}\mathcal{V} \le \frac{1}{2}\exp\{K\,C(t,m,n)\}\, \mathcal{W}_2^2\big(\rho_0^0, \rho_0^1\big)\,,$$

where we have reintroduced the dependence on ε in view of the last passage to the limit.

6. By virtue of (4.15), we are allowed to first pass to the limit as $\varepsilon \downarrow 0$ and then take the supremum over all $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$, which yields

$$\mathcal{W}_2\big(\rho^0(t), \rho^1(t)\big) \le \exp\{K\,C(t,m,n)\}\, \mathcal{W}_2\big(\rho_0^0, \rho_0^1\big)\,.$$

7. We exploit Proposition 4.5 in order to show that such solutions do belong to $\mathscr{M}_2^M(\mathbb{M}^n)$ for all times; here we apply inductively the stability estimate itself in the form $\mathcal{W}_2(\rho(t), \rho(t+\tau))$, for small $\tau > 0$, along with (3.10). Then, upon approximating the initial data, we show that the stability estimate extends to the whole class $\mathscr{M}_2^M(\mathbb{M}^n)$.

8. As a final step, we prove that the solutions constructed above are indeed weak Wasserstein solutions, in the sense of Definition 2.3. This basically follows from the smoothing effect (2.5) and the energy inequality (4.11). Uniqueness of Wasserstein solutions is also a direct consequence of the uniqueness result for weak energy solutions, together with their regularity properties.

5.2. The noncompact case.
Throughout this whole subsection we will assume again that $\mathbb{M}^n$ is in addition noncompact and with infinite volume, hence we will carry out the proof of Theorem 2.4 in this case only. We will then discuss in Subsection 5.3 the (simple) modifications required to deal with compact manifolds.

Let ρ be a weak energy solution of (4.10) and let ϕ be a strong variational solution of the associated *backward adjoint* problem, according to Theorem 4.11. Upon recalling (3.7), we define the *Hamiltonian* functional as

$$\mathcal{E}_{\rho(t)}[\phi(t)] := \int_{\mathbb{M}^n} \Gamma(\phi(t))\, \rho(t)\, \mathrm{d}\mathcal{V}\,.$$

Following [4], we firstly connect the time derivative of the Hamiltonian with the *carré du champ* operators defined in (3.2) and (3.5) (see [4, Theorem 11.1 and Lemma 11.2] for a detailed proof).

Lemma 5.1. *Let P comply with* (H3), P_ε ($\varepsilon > 0$) *be defined by* (4.3) *and $T > 0$. Let $\rho \in \mathcal{ND}_{P_\varepsilon}(0, T)$ be a bounded solution of* (4.10), *provided by Proposition 4.7. Let $\phi \in W^{1,2}((0, T); \mathbb{D}, \mathbb{H})$ be a bounded strong solution of* (4.69), *provided by Theorem 4.11. Then the map $t \mapsto \mathcal{E}_{\rho(t)}[\phi(t)]$ is absolutely continuous in $[0, T]$ and satisfies the identity*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho(t)}[\phi(t)] = \mathbf{\Gamma}_2[\phi(t); P_\varepsilon(\rho(t))] + \int_{\mathbb{M}^n} R(\rho(t))\,(\Delta\phi(t))^2\,\mathrm{d}\mathcal{V} \qquad a.e \text{ in } (0, T)\,,$$

where

$$R(\rho) := \rho\,(P_\varepsilon)'(\rho) - P_\varepsilon(\rho) \qquad \forall \rho \geq 0\,.$$

By requiring the additional assumption (H5) on the nonlinearity, we are able to exploit the curvature bound (H1) in the Bakry-Émery form (3.6).

Lemma 5.2. *Let the hypotheses of Lemma 5.1 hold. Assume in addition that $\mathbb{M}^n$ ($n \geq 3$) complies with* (H1) *and P complies with* (H5). *Then*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho(t)}[\phi(t)] \geq -K\int_{\mathbb{M}^n}\Gamma(\phi(t))\,P_\varepsilon(\rho(t))\,\mathrm{d}\mathcal{V} \qquad a.e \text{ in } (0, T)\,. \tag{5.1}$$

Proof. By combining Lemma 5.1 and the Bakry-Émery condition (3.6) with $f \equiv \phi(t)$ and $\rho \equiv P_\varepsilon(\rho(t))$, we deduce that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho(t)}[\phi(t)] \geq -K\int_{\mathbb{M}^n}\Gamma(\phi(t))\,P_\varepsilon(\rho(t))\,\mathrm{d}\mathcal{V}$$
$$+ \int_{\mathbb{M}^n}\left[\rho(t)\,(P_\varepsilon)'(\rho(t)) - \left(1 - \tfrac{1}{n}\right)P_\varepsilon(\rho(t))\right](\Delta\phi(t))^2\,\mathrm{d}\mathcal{V}\,.$$

The conclusion follows upon taking advantage of (4.7). $\qquad\square$

If $K > 0$ in general it is not clear how to bound the r.h.s. of (5.1) in terms of the Hamiltonian itself. Nevertheless, if P complies with (H4) and the Sobolev-type inequality (H2) holds, the smoothing effect provided by Proposition 4.4 allows us to do so.

Lemma 5.3. *Let $\mathbb{M}^n$ ($n \geq 3$) comply with* (H1) *and* (H2). *Let P comply with* (H3), (H4) *and* (H5). *Let $T > 0$ and $\rho_\varepsilon \in \mathcal{ND}_{P_\varepsilon}(0, T)$ be the (weak energy) solution of* (4.10) *corresponding to some nonnegative $\rho_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n) \cap W^{1,2}(\mathbb{M}^n)$ with $\|\rho_0\|_{L^1(\mathbb{M}^n)} =: M$ (recall Proposition 4.8), where P_ε ($\varepsilon > 0$) is defined by* (4.3). *Let $\phi \in W^{1,2}((0, T); \mathbb{D}, \mathbb{H})$ be a bounded solution of* (4.69) *provided by Theorem 4.11. Suppose that ε is so small that*

$$\|\rho_0\|_{L^\infty(\mathbb{M}^n)} \leq \frac{1}{\varepsilon}\,.$$

Then

$$\mathcal{E}_{\rho_\varepsilon(t)}[\phi(t)] \geq \exp\left\{-2K\,c_1\,\mathfrak{C}_m\left[\left(tM^{m-1}\right)^{\frac{2}{2+n(m-1)}}\vee\left(tM^{m-1}\right) + \frac{\varepsilon}{c_1\mathfrak{C}_m}t\right]\right\}\mathcal{E}_{\rho_0}[\phi(0)] \qquad \forall t \geq 0\,, \tag{5.2}$$

where $C > 0$ is the same constant appearing in (4.31) *and*

$$\mathfrak{C}_m := C^{m-1}\,2^{m-2}\,[2 + n(m-1)]\,. \tag{5.3}$$

Proof. By combining inequalities (4.4) and (5.1), we obtain:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\rho_\varepsilon(t)}[\phi(t)] \geq -K\int_{\mathbb{M}^n}\Gamma(\phi(t))\left[P(\rho_\varepsilon(t)) + \varepsilon\rho_\varepsilon(t)\right]\mathrm{d}\mathcal{V}\,. \tag{5.4}$$

Thanks to Proposition 4.4, we know that

$$
\begin{aligned}
\|\rho_\varepsilon(t)\|_{L^\infty(\mathbb{M}^n)}^{m-1} &\le C^{m-1} \left(t^{-\frac{n}{2+n(m-1)}} \|\rho_0\|_{L^1(\mathbb{M}^n)}^{\frac{2}{2+n(m-1)}} + \|\rho_0\|_{L^1(\mathbb{M}^n)} \right)^{m-1} \\
&= C^{m-1} M^{m-1} g_m\!\left(t\, M^{m-1} \right) \qquad \forall t > 0\,,
\end{aligned}
\tag{5.5}
$$

where

$$
g_m(s) := \left(s^{-\frac{n}{2+n(m-1)}} + 1 \right)^{m-1} \qquad \forall s > 0\,.
$$

It is apparent that

$$
g_m(s) \le \begin{cases} 2^{m-1}\, s^{-\frac{n(m-1)}{2+n(m-1)}} & \text{if } s \in (0,1)\,, \\ 2^{m-1} & \text{if } s \ge 1\,. \end{cases}
\tag{5.6}
$$

If we plug (5.5) in (5.4) and recall that $P(\rho)/\rho \le c_1\, \rho^{m-1}$, we find:

$$
\begin{aligned}
\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}_{\rho_\varepsilon(t)}[\phi(t)] &\ge - K \int_{\mathbb{M}^n} \Gamma(\phi(t))\, \rho_\varepsilon(t) \left[c_1\, \rho_\varepsilon(t)^{m-1} + \varepsilon \right] \mathrm{d}\mathcal{V} \\
&\ge - K \left[c_1\, C^{m-1} M^{m-1} g_m\!\left(t\, M^{m-1} \right) + \varepsilon \right] \mathcal{E}_{\rho_\varepsilon(t)}[\phi(t)]\,;
\end{aligned}
\tag{5.7}
$$

by integrating (5.7) we therefore obtain

$$
\mathcal{E}_{\rho_\varepsilon(t)}[\phi(t)] \ge \exp\left\{ -2K \left(c_1\, C^{m-1} \int_0^{t\, M^{m-1}} g_m(s)\, \mathrm{d}s + \varepsilon t \right) \right\} \mathcal{E}_{\rho_0}[\phi(0)] \qquad \forall t \ge 0\,.
\tag{5.8}
$$

In order to suitably simplify (5.8), by exploiting (5.6) we easily infer that

$$
\int_0^\tau g_m(s)\, \mathrm{d}s \le \begin{cases} 2^{m-1} \frac{2+n(m-1)}{2}\, \tau^{\frac{2}{2+n(m-1)}} & \text{if } \tau \in (0,1)\,, \\ 2^{m-1} \left[\tau + \frac{n(m-1)}{2} \right] & \text{if } \tau \ge 1\,, \end{cases}
$$

which implies

$$
\int_0^\tau g_m(s)\, \mathrm{d}s \le 2^{m-2}\, [2 + n(m-1)] \left(\tau^{\frac{2}{2+n(m-1)}} \vee \tau \right) \qquad \forall \tau > 0\,,
$$

whence (5.2). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

In the following, we will connect any two (sufficiently regular) initial data ρ_0^0 and ρ_0^1 with a *regular* curve $\{\rho_0^s\}_{s \in [0,1]}$ (in the sense of Definition 3.3) and consider the corresponding solution $t \mapsto \rho_\varepsilon^s(t)$ of (4.10) with initial datum ρ_0^s, that is

$$
\begin{cases} \partial_t \rho_\varepsilon^s = \Delta P_\varepsilon(\rho_\varepsilon^s) & \text{in } \mathbb{M}^n \times \mathbb{R}^+\,, \\ \rho_\varepsilon^s(0) = \rho_0^s & \text{on } \mathbb{M}^n \times \{0\}\,. \end{cases}
\tag{5.9}
$$

Reasoning as in [4], we will exploit the lower bound on the Hamiltonian ensured by Lemma 5.3 in order to prove the stability estimate (2.6). We start by studying the quantity

$$
s \mapsto \int_{\mathbb{M}^n} Q_s \varphi\, \rho_\varepsilon^s(t)\, \mathrm{d}\mathcal{V}\,,
$$

where $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$ is arbitrary but fixed and $[0,1] \times \mathbb{M}^n \ni (s,x) \mapsto Q_s\varphi(x)$ is the (Lipschitz and compactly-supported) solution of the Hopf-Lax problem (3.16). To this aim, for (almost) every $s \in (0,1)$ we also introduce the solution w^s of the linearized equation (4.68) starting from $\frac{\mathrm{d}}{\mathrm{d}s}\rho_0^s$:

$$
\begin{cases} \partial_t w^s = \Delta[P_\varepsilon'(\rho_\varepsilon^s)\, w^s] & \text{in } \mathbb{M}^n \times \mathbb{R}^+\,, \\ w^s(0) = \frac{\mathrm{d}}{\mathrm{d}s}\rho_0^s & \text{on } \mathbb{M}^n \times \{0\}\,. \end{cases}
\tag{5.10}
$$

Thanks to Theorem 4.10 and Remark 3.4, if $\{\rho_0^s\}_{s\in[0,1]}$ is a regular curve we can guarantee that (5.10) admits a weak solution, at least for almost every $s \in (0,1)$. Moreover, [4, Theorem 4.6] ensures that $w^s(t) = \frac{\mathrm{d}}{\mathrm{d}s}\rho_\varepsilon^s(t)$. (4.70) with initial datum $w_0 = \Delta P(\rho_0)$.

Lemma 5.4. *Let P comply with* (H3) *and P_ε ($\varepsilon > 0$) be defined by* (4.3). *Given a regular curve $\{\rho_0^s\}_{s\in[0,1]}$ and $T > 0$, let $\rho_\varepsilon^s \in \mathcal{ND}_{P_\varepsilon}(0,T)$ be the corresponding (weak energy) solution of* (5.9). *Then, for every $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$ and every $t \in (0,T)$, the map $s \mapsto \int_{\mathbb{M}^n} Q_s\varphi\, \rho_\varepsilon^s(t)\, \mathrm{d}\mathcal{V}$ is Lipschitz continuous in $[0,1]$ and satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}s} \int_{\mathbb{M}^n} Q_s\varphi\, \rho_\varepsilon^s(t)\, \mathrm{d}\mathcal{V} = -\frac{1}{2} \int_{\mathbb{M}^n} \Gamma(Q_s\varphi)\, \rho_\varepsilon^s(t)\, \mathrm{d}\mathcal{V} + {}_{\mathbb{V}'}\langle w^s(t), Q_s\varphi\rangle_{\mathbb{V}} \qquad \text{for a.e. } s \in (0,1)\,, \quad (5.11)$$

where $(s,x) \mapsto Q_s\varphi(x)$ is the (Lipschitz and compactly-supported) solution of the Hopf-Lax problem (3.16) *and $w^s(t) = \frac{\mathrm{d}}{\mathrm{d}s}\rho_\varepsilon^s(t)$ is the weak solution of* (5.10) *provided by Theorem 4.10.*

Moreover, if we denote by $r : (0,t) \mapsto \phi^s(r)$ the solution of the backward adjoint problem (4.69) *corresponding to $\rho \equiv \rho_\varepsilon^s$ with final condition $\phi^s(t) = Q_s\varphi$, provided by Theorem 4.11, the following identities hold:*

$${}_{\mathbb{V}'}\langle w^s(t), Q_s\varphi\rangle_{\mathbb{V}} = {}_{\mathbb{V}'}\langle w^s(t), \phi^s(t)\rangle_{\mathbb{V}} = {}_{\mathbb{V}'}\langle w^s(0), \phi^s(0)\rangle_{\mathbb{V}} = {}_{\mathbb{V}'}\langle \tfrac{\mathrm{d}}{\mathrm{d}s}\rho_0^s, \phi^s(0)\rangle_{\mathbb{V}} \quad \text{for a.e. } s \in (0,1)\,.$$
$$(5.12)$$

Proof. The map $s \mapsto \int_{\mathbb{M}^n} Q_s\varphi\, \rho_\varepsilon^s(t)\, \mathrm{d}\mathcal{V}$ is Lipschitz continuous by virtue of the Lipschitz-continuity of $(s,x) \mapsto Q_s\varphi(x)$ (plus the boundedness of its support) and the Lipschitz-continuity of the curve $s \mapsto \rho_0^s$ with values in $\mathbb{V}'$ (recall Remark 3.4) along with the fact that the semigroup generated by (5.9) turns out to be also a contraction with respect to $\|\cdot\|_{\mathbb{V}'}$. For more details we refer the reader to [4, Proof of Theorem 12.5]. Once we have observed this, identity (5.11) is a direct consequence of (3.16) and the equality $w^s(t) = \frac{\mathrm{d}}{\mathrm{d}s}\rho_\varepsilon^s(t)$ (for a.e. $s \in (0,1)$ independently of t), which can rigorously be proved by proceeding as in [4, Theorem 4.6].

As concerns (5.12), it is enough to observe that it is nothing but formula (4.70) with $\rho \equiv \rho_\varepsilon^s$, $w \equiv w^s$ and $\theta \equiv \phi^s$ (actually with r and t interchanged). $\qquad \square$

Lemma 5.5. *Let $\mathbb{M}^n$ ($n \geq 3$) comply with assumptions* (H1) *and* (H2). *Let moreover P comply with assumptions* (H3), (H4), (H5) *and P_ε ($\varepsilon > 0$) be defined by* (4.3). *Let ρ_ε^0 and ρ_ε^1 be any two (weak energy) solutions of* (4.10) *corresponding to the initial data ρ_0^0 and ρ_0^1, respectively, both nonnegative, belonging to $L_c^\infty(\mathbb{M}^n) \cap W^{1,2}(\mathbb{M}^n)$ and having the same mass $M > 0$. Suppose that $\{\rho_0^s\}_{s\in[0,1]}$ is any regular curve (in the sense of Definition 3.3) connecting ρ_0^0 with ρ_0^1, which satisfies*

$$\|\rho_0^s\|_{L^\infty(\mathbb{M}^n)} \leq \frac{1}{\varepsilon} \qquad \forall s \in [0,1]\,. \tag{5.13}$$

Then for every $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$ there holds

$$\int_{\mathbb{M}^n} Q_1\varphi\, \rho_\varepsilon^1(t)\, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi\, \rho_\varepsilon^0(t)\, \mathrm{d}\mathcal{V}$$
$$\leq \frac{1}{2} \exp\left\{ 2K\, c_1\, \mathfrak{C}_m \left[\left(tM^{m-1}\right)^{\frac{2}{2+n(m-1)}} \vee \left(tM^{m-1}\right) + \frac{\varepsilon}{c_1\mathfrak{C}_m}t \right] \right\} \int_0^1 |\dot\mu^s|^2\, \mathrm{d}s\,, \tag{5.14}$$

where $\mu^s := \rho_0^s \mathcal{V}$ and $\{Q_s\varphi\}_{s\in[0,1]}$ is the (Lipschitz and compactly-supported) solution of the Hopf-Lax problem (3.16) *and the constant $\mathfrak{C}_m$ is defined in* (5.3).

Proof. We follow the line of proof of [4, Theorem 12.5], keeping the same notations as in Lemma 5.4. By combining (5.11) and (5.12), we obtain:

$$\int_{\mathbb{M}^n} Q_1\varphi\, \rho_\varepsilon^1(t)\, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi\, \rho_\varepsilon^0(t)\, \mathrm{d}\mathcal{V} = \int_0^1 \left(-\frac{1}{2} \int_{\mathbb{M}^n} \Gamma(\phi^s(t))\, \rho_\varepsilon^s(t)\, \mathrm{d}\mathcal{V} + {}_{\mathbb{V}'}\langle \tfrac{\mathrm{d}}{\mathrm{d}s}\rho_0^s, \phi^s(0)\rangle_{\mathbb{V}} \right) \mathrm{d}s$$
$$= \int_0^1 \left(-\frac{1}{2} \mathcal{E}_{\rho_\varepsilon^s(t)}[\phi^s(t)] + {}_{\mathbb{V}'}\langle \tfrac{\mathrm{d}}{\mathrm{d}s}\rho_0^s, \phi^s(0)\rangle_{\mathbb{V}} \right) \mathrm{d}s\,.$$

Now we can apply, at every $s \in [0,1]$, estimate (5.2) from Lemma 5.3 with $\rho_\varepsilon(t) \equiv \rho_\varepsilon^s(t)$ and $\phi(t) \equiv \phi^s(t)$, under assumption (5.13). This yields, upon recalling (3.8),

$$
\begin{aligned}
&\int_{\mathbb{M}^n} Q_1 \varphi \, \rho_\varepsilon^1(t) \, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi \, \rho_\varepsilon^0(t) \, \mathrm{d}\mathcal{V} \\
&\leq \int_0^1 \left(-\frac{1}{2} e^{-2K c_1 \mathfrak{C}_m \left[(tM^{m-1})^{\frac{2}{2+n(m-1)}} \vee (tM^{m-1}) + \frac{\varepsilon}{c_1 \mathfrak{C}_m} t \right]} \mathcal{E}_{\rho_0^s}[\phi^s(0)] + {}_{\mathbb{V}'}\langle \tfrac{\mathrm{d}}{\mathrm{d}s} \rho_0^s, \phi^s(0) \rangle_{\mathbb{V}} \right) \mathrm{d}s \\
&= e^{2K c_1 \mathfrak{C}_m \left[(tM^{m-1})^{\frac{2}{2+n(m-1)}} \vee (tM^{m-1}) + \frac{\varepsilon}{c_1 \mathfrak{C}_m} t \right]} \int_0^1 \left(-\frac{1}{2} \mathcal{E}_{\rho_0^s}[\psi^{s,t}] + {}_{\mathbb{V}'}\langle \tfrac{\mathrm{d}}{\mathrm{d}s} \rho_0^s, \psi^{s,t} \rangle_{\mathbb{V}} \right) \mathrm{d}s \\
&\leq e^{2K c_1 \mathfrak{C}_m \left[(tM^{m-1})^{\frac{2}{2+n(m-1)}} \vee (tM^{m-1}) + \frac{\varepsilon}{c_1 \mathfrak{C}_m} t \right]} \int_0^1 \frac{1}{2} \mathcal{E}_{\rho_0^s}^*[\tfrac{\mathrm{d}}{\mathrm{d}s} \rho_0^s] \, \mathrm{d}s \,,
\end{aligned}
\tag{5.15}
$$

where we have set

$$
\psi^{s,t} := e^{-2K c_1 \mathfrak{C}_m \left[(tM^{m-1})^{\frac{2}{2+n(m-1)}} \vee (tM^{m-1}) + \frac{\varepsilon}{c_1 \mathfrak{C}_m} t \right]} \phi^s(0) \,.
$$

Estimate (5.14) thus follows from (5.15) in view of (3.21). □

In order to prove Theorem 2.4, we need first to approximate the geodesic connecting μ_0 and $\hat{\mu}_0$ in $(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2)$ by regular curves, let $\varepsilon \to 0$ in (4.10) and finally pass to the limit in the approximation of the measures μ_0 and $\hat{\mu}_0$ by bounded and compactly supported densities as in Lemma 5.5.

Proof of Theorem 2.4 (noncompact case). To begin with, we suppose that $\mu_0 = \rho_0 \mathcal{V}$ and $\hat{\mu}_0 = \hat{\rho}_0 \mathcal{V}$, where ρ_0 and $\hat{\rho}_0$ are initial data complying with the assumptions of Lemma 5.5: we will remove this hypothesis only at the very end of the proof. By virtue of Lemma 3.5, we know that there exists a sequence of regular curves $\{\rho_j^s\}_{j \in \mathbb{N}, s \in [0,1]}$ satisfying (3.17)–(3.20) (let $\rho^1 = \rho_0$ and $\rho^0 = \hat{\rho}_0$ according to the corresponding notations). Given $\varepsilon > 0$, if we denote by $t \mapsto (\rho_j^s)_\varepsilon(t)$ each weak energy solution of (4.10) starting from $\rho_0 \equiv \rho_j^s$, then by Lemma 5.5 we know that

$$
\begin{aligned}
&\int_{\mathbb{M}^n} Q_1 \varphi \, (\rho_j^1)_\varepsilon(t) \, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi \, (\rho_j^0)_\varepsilon(t) \, \mathrm{d}\mathcal{V} \\
&\leq \frac{1}{2} \exp \left\{ 2K c_1 \mathfrak{C}_m \left[(tM^{m-1})^{\frac{2}{2+n(m-1)}} \vee (tM^{m-1}) + \frac{\varepsilon}{c_1 \mathfrak{C}_m} t \right] \right\} \int_0^1 |\dot{\mu}_j^s|^2 \, \mathrm{d}s
\end{aligned}
\tag{5.16}
$$

for every $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$, provided

$$
\|\rho_j^s\|_{L^\infty(\mathbb{M}^n)} \leq \frac{1}{\varepsilon} \qquad \forall s \in [0,1] \,.
\tag{5.17}
$$

Let us pass to the limit in (5.16) as $j \to \infty$. In the sequel, we denote by ρ_ε and $\hat{\rho}_\varepsilon$ the weak energy solutions of (4.10) starting from ρ_0 and $\hat{\rho}_0$, respectively. Thanks to (3.18), (3.19) (with $p = 1$) and the L^1-contraction property (4.16) of weak energy solutions, which guarantees that $(\rho_j^0)_\varepsilon(t) \to \rho_\varepsilon(t)$ and $(\rho_j^1)_\varepsilon(t) \to \hat{\rho}_\varepsilon(t)$ in $L^1(\mathbb{M}^n)$, we deduce that

$$
\begin{aligned}
&\int_{\mathbb{M}^n} Q_1 \varphi \, \rho_\varepsilon(t) \, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi \, \hat{\rho}_\varepsilon(t) \, \mathrm{d}\mathcal{V} \\
&\leq \frac{1}{2} \exp \left\{ 2K c_1 \mathfrak{C}_m \left[(tM^{m-1})^{\frac{2}{2+n(m-1)}} \vee (tM^{m-1}) + \frac{\varepsilon}{c_1 \mathfrak{C}_m} t \right] \right\} \mathcal{W}_2^2(\rho_0, \hat{\rho}_0)
\end{aligned}
\tag{5.18}
$$

upon requiring

$$
\limsup_{j \to \infty} \sup_{s \in [0,1]} \|\rho_j^s\|_{L^\infty(\mathbb{M}^n)} \leq \frac{1}{2\varepsilon}
$$

in view of (5.17), which holds for ε small enough thanks to (3.20). We are now in position to let $\varepsilon \downarrow 0$. The r.h.s. of (5.18) is clearly stable as $\varepsilon \downarrow 0$. In order to pass to the limit in the l.h.s. we need to exploit Proposition 4.3: in particular, formula (4.15) ensures that $\{\rho_\varepsilon(t)\}_{\varepsilon > 0}$ and $\{\hat{\rho}_\varepsilon(t)\}_{\varepsilon > 0}$ converge in $L^1(\mathbb{M}^n)$ to $\rho(t)$ and $\hat{\rho}(t)$, respectively, so that (5.18) yields

$$\int_{\mathbb{M}^n} Q_1 \varphi \, \rho(t) \, \mathrm{d}\mathcal{V} - \int_{\mathbb{M}^n} \varphi \, \hat{\rho}(t) \, \mathrm{d}\mathcal{V} \le \frac{1}{2} \exp\left\{ 2K \, c_1 \, \mathfrak{C}_m \left[\left(t M^{m-1} \right)^{\frac{2}{2+n(m-1)}} \vee \left(t M^{m-1} \right) \right] \right\} \mathcal{W}_2^2(\rho_0, \hat{\rho}_0) \, . \tag{5.19}$$

If we take the supremum of the l.h.s. of (5.19) over all $\varphi \in \mathrm{Lip}_c(\mathbb{M}^n)$, then by virtue of Proposition 3.1 we obtain

$$\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \le \exp\left\{ K \, c_1 \, \mathfrak{C}_m \left[\left(t M^{m-1} \right)^{\frac{2}{2+n(m-1)}} \vee \left(t M^{m-1} \right) \right] \right\} \mathcal{W}_2(\rho_0, \hat{\rho}_0) \qquad \forall t > 0 \, , \quad (5.20)$$

namely (2.6) restricted to initial data $\rho_0, \hat{\rho}_0 \in L_c^\infty(\mathbb{M}^n) \cap W^{1,2}(\mathbb{M}^n)$. It is apparent that estimate (5.20) remains true in the wider class $\rho_0, \hat{\rho}_0 \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n) \cap \mathscr{M}_2^M$: indeed by local regularization and a standard truncation argument, one can pick sequences of nonnegative initial data of mass M belonging to $L_c^\infty(\mathbb{M}^n) \cap W^{1,2}(\mathbb{M}^n)$ which converge to ρ_0 and $\hat{\rho}_0$, respectively, both in $L^1(\mathbb{M}^n)$ and in $(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2)$ (recall Proposition 3.2). Thanks again to (4.16), i.e. the stability of solutions in $L^1(\mathbb{M}^n)$, this suffices to pass to the limit in (5.19) and hence in (5.20).

We still have to prove that $\rho(t) \in \mathscr{M}_2^M(\mathbb{M}^n)$ for all $t > 0$, since the mass-conservation property (4.13) only ensures that $\rho(t) \in \mathscr{M}^M(\mathbb{M}^n)$. To this aim, we take advantage of Proposition 4.5: from the latter we know that if $\rho_0 \in L_c^\infty(\mathbb{M}^n)$ then the weak energy solution $\rho(t)$ of (4.1) stays (uniformly) bounded with (uniform) compact support in a suitable time interval $[0, t_1]$, so that in particular $\rho(t) \in \mathscr{M}_2^M(\mathbb{M}^n)$ for all $t \in [0, t_1]$. Let $\tau \in (0, t_1]$. Since $\{\rho(t + \tau)\}_{t \ge 0}$ is the weak energy solution of (4.1) starting from $\rho(\tau) \in L^1(\mathbb{M}^n) \cap L^\infty(\mathbb{M}^n) \cap \mathscr{M}_2^M(\mathbb{M}^n)$, estimate (5.20) applied to $\hat{\rho}(t) = \rho(t + \tau)$ guarantees that

$$\mathcal{W}_2(\rho(t), \rho(t + \tau)) \le \exp\left\{ K \, c_1 \, \mathfrak{C}_m \left[\left(t M^{m-1} \right)^{\frac{2}{2+n(m-1)}} \vee \left(t M^{m-1} \right) \right] \right\} \mathcal{W}_2(\rho_0, \rho(\tau)) < \infty \quad \forall t > 0 \, , \tag{5.21}$$

whence $\rho(t) \in \mathscr{M}_2^M(\mathbb{M}^n)$ also for all $t \in (t_1, 2t_1]$ upon recalling (3.10). It is then clear how one can set up an induction procedure to establish that in fact $\rho(t) \in \mathscr{M}_2^M(\mathbb{M}^n)$ for *all* $t > 0$. Furthermore, $\rho \in C([0, +\infty); \mathscr{M}_2^M(\mathbb{M}^n))$. Indeed, the just mentioned property of compactness of the support for short times and the L^1-continuity ensured by Proposition 4.3 easily imply, along with Proposition 3.2, that

$$\lim_{t \downarrow 0} \mathcal{W}_2(\rho(t), \rho_0) = 0 \, . \tag{5.22}$$

Hence by combining (5.21) (understood for all $t, \tau > 0$) and (5.22), we deduce that for every $t_0 > 0$ there holds

$$\lim_{t \to t_0} \mathcal{W}_2(\rho(t), \rho(t_0))$$
$$\le \exp\left\{ K \, c_1 \, \mathfrak{C}_m \left[\left(t_0 M^{m-1} \right)^{\frac{2}{2+n(m-1)}} \vee \left(t_0 M^{m-1} \right) \right] \right\} \lim_{t \to t_0} \mathcal{W}_2(\rho(|t - t_0|), \rho_0) = 0 \, .$$

We have therefore shown the validity of Theorem 2.4 under the additional assumptions $\mu_0 = \rho_0 \mathcal{V}$ and $\hat{\mu}_0 = \hat{\rho}_0 \mathcal{V}$ with $\rho_0, \hat{\rho}_0 \in L_c^\infty(\mathbb{M}^n) \cap W^{1,2}(\mathbb{M}^n)$. In order to be able to deal with general initial data as in the statement, first of all we take a sequence of nonnegative functions $(\rho_{j,0}, \hat{\rho}_{j,0}) \in [L_c^\infty(\mathbb{M}^n) \cap W^{1,2}(\mathbb{M}^n)]^2$ of mass M such that

$$\lim_{j \to \infty} \rho_{j,0} = \mu_0 \qquad \text{and} \qquad \lim_{j \to \infty} \hat{\rho}_{j,0} = \hat{\mu}_0 \qquad \text{in } \left(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2 \right) , \tag{5.23}$$

which exists as a consequence of Definition 3.3, Remark 3.4 and Lemma 3.5 (only applied at the endpoints $s = 0, 1$): the additional property of the compactness of the support can be obtained again

by a straightforward truncation argument. Estimate (5.20) applied to the corresponding sequences of solutions, which we denote by $\{(\rho_j, \hat{\rho}_j)\}_{j\in\mathbb{N}}$, yields

$$
\begin{aligned}
\mathcal{W}_2(\rho_j(t), \rho_i(t)) &\leq \exp\left\{K\, c_1\, \mathfrak{C}_m \left[\left(tM^{m-1}\right)^{\frac{2}{2+n(m-1)}} \vee \left(tM^{m-1}\right)\right]\right\} \mathcal{W}_2(\rho_{j,0}, \rho_{i,0}),\\
\mathcal{W}_2(\hat{\rho}_j(t), \hat{\rho}_i(t)) &\leq \exp\left\{K\, c_1\, \mathfrak{C}_m \left[\left(tM^{m-1}\right)^{\frac{2}{2+n(m-1)}} \vee \left(tM^{m-1}\right)\right]\right\} \mathcal{W}_2(\hat{\rho}_{j,0}, \hat{\rho}_{i,0}),
\end{aligned}
\tag{5.24}
$$

for every $t > 0$ and $i, j \in \mathbb{N}$, whereas the smoothing effect (4.33) ensures that

$$
\|\rho_j(t)\|_{L^\infty(\mathbb{M}^n)} \vee \|\hat{\rho}_j(t)\|_{L^\infty(\mathbb{M}^n)} \leq C\left(t^{-\frac{n}{2+n(m-1)}} M^{\frac{2}{2+n(m-1)}} + M\right) \qquad \forall t > 0,\ \forall j \in \mathbb{N}. \tag{5.25}
$$

From (5.23) and (5.24) we infer that $\{\rho_j\}_{j\in\mathbb{N}}$ and $\{\hat{\rho}_j\}_{j\in\mathbb{N}}$ are Cauchy sequences in the space $C([0,T); (\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2))$ for every $T > 0$, hence they converge to two corresponding curves ρ and $\hat{\rho}$, respectively, both in $C([0,T); (\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2))$ for all $T > 0$. By construction estimates (5.25) and (5.20) (applied to $\rho \equiv \rho_j$ and $\hat{\rho} \equiv \hat{\rho}_j$) are preserved at the limit, ensuring the validity of (2.5)–(2.6). We are thus left with proving that ρ and $\hat{\rho}$ are indeed Wasserstein solutions of (2.1) in the sense of Definition 2.3, i.e. they comply with (2.3) and (2.4). Of course it is enough to show it for ρ only. Since the latter satisfies (2.5) and $\|\rho(t)\|_{L^1(\mathbb{M}^n)} = M$ for all $t > 0$, the first property in (2.3) is trivially fulfilled. In order to establish the second one and (2.4), we take advantage of the energy estimate (4.11) applied to each $\rho \equiv \rho_j$ (with time origin shifted from 0 to $\tau \in (0, T)$) combined with (H4) and (5.25), which yield

$$
\begin{aligned}
\int_\tau^T \int_{\mathbb{M}^n} |\nabla P(\rho_j)|^2 \,\mathrm{d}\mathcal{V}\mathrm{d}t + \int_{\mathbb{M}^n} \Psi(\rho_j(x,T)) \,\mathrm{d}\mathcal{V}(x) &\leq \frac{c_1}{m+1} \int_{\mathbb{M}^n} \rho_j(x,\tau)^{m+1} \,\mathrm{d}\mathcal{V}(x)\\
&\leq \frac{c_1\, C^m M}{m+1} \left(\tau^{-\frac{n}{2+n(m-1)}} M^{\frac{2}{2+n(m-1)}} + M\right)^m.
\end{aligned}
\tag{5.26}
$$

Starting from (5.26), using in a similar way the analogues of (4.28)–(4.29) with $\rho \equiv \rho_j$ and the time origin shifted from 0 to τ, one can reason as in the proof of Proposition 4.3 to deduce that $\{\rho_j\}_{j\in\mathbb{N}}$ converges to ρ and $\{\nabla P(\rho_j)\}_{j\in\mathbb{N}}$ converges to $\nabla P(\rho)$ weakly in $L^2(\mathbb{M}^n \times (\tau, T))$ as $j \to \infty$, whence the validity of (2.4) upon passing to the limit in the weak formulation satisfied by every ρ_j.

Finally, the uniqueness of Wasserstein solutions is a simple consequence of the uniqueness of weak energy solutions (Proposition 4.3) and the continuity in $(\mathscr{M}_2^M(\mathbb{M}^n), \mathcal{W}_2)$ down to $t = 0$. Indeed, if ρ and $\hat{\rho}$ are two Wasserstein solutions starting from the same initial datum, they can be seen as weak energy solutions starting from the initial data $\rho(\tau)$ and $\hat{\rho}(\tau)$, respectively, for every $\tau > 0$. In particular, there holds

$$
\mathcal{W}_2(\rho(t), \hat{\rho}(t)) \leq \exp\left\{K\, c_1\, \mathfrak{C}_m \left[\left(tM^{m-1}\right)^{\frac{2}{2+n(m-1)}} \vee \left(tM^{m-1}\right)\right]\right\} \mathcal{W}_2(\rho(\tau), \hat{\rho}(\tau)) \qquad \forall t > \tau > 0, \tag{5.27}
$$

whence $\mathcal{W}_2(\rho(t), \hat{\rho}(t)) = 0$ upon letting $\tau \downarrow 0$ in (5.27). $\qquad\square$

5.3. **The compact case.** If $\mathbb{M}^n$ is a *compact* manifold, the construction of the Wasserstein solutions of (2.1) performed in Subsection 4.1 is in fact easier with respect to the one performed in the noncompact case. Indeed, in the proofs of Propositions 4.3 and 4.4, there is no need to fill $\mathbb{M}^n$ with a regular exhaustion $\{D_k\}_{k\in\mathbb{N}}$: it is enough to solve the approximate problems (i.e. the ones associated with the nonlinearity P_ε) directly on the compact manifold, where integrations by parts are always justified. Moreover, mass conservation is plain because space-constant functions are admissible test functions in the weak formulation (4.2). The compact-support property established in Proposition 4.5 is clearly for free.

As concerns the variational framework considered in Subsection 4.2, some less trivial modifications have to be implemented. That is, one defines the space

$$\mathbb{V}'_{\mathcal{E}} := \left\{ \ell \in \mathbb{V}' : |_{\mathbb{V}'}\langle \ell, f \rangle_{\mathbb{V}}| \leq C \sqrt{\mathcal{E}(f)} \text{ for every } f \in \mathbb{V}, \text{ for some } C > 0 \right\}$$

endowed with the norm

$$\|\ell\|_{\mathbb{V}'_{\mathcal{E}}} := \sup_{f \in \mathbb{V}:\, \mathcal{E}(f) \neq 0} \frac{|_{\mathbb{V}'}\langle \ell, f \rangle_{\mathbb{V}}|}{\sqrt{\mathcal{E}(f)}},$$

and the space

$$\mathbb{D}'_{\mathcal{E}} := \left\{ \ell \in \mathbb{D}' : |_{\mathbb{D}'}\langle \ell, f \rangle_{\mathbb{D}}| \leq C \|\Delta f\|_{\mathbb{H}} \text{ for every } f \in \mathbb{D}, \text{ for some } C > 0 \right\},$$

endowed with the norm

$$\|\ell\|_{\mathbb{D}'_{\mathcal{E}}} := \sup_{f \in \mathbb{D}:\, \Delta f \neq 0} \frac{|_{\mathbb{D}'}\langle \ell, f \rangle_{\mathbb{D}}|}{\|\Delta f\|_{\mathbb{H}}}.$$

Upon replacing $\mathbb{V}'$ with $\mathbb{V}'_{\mathcal{E}}$ and $\mathbb{D}'$ with $\mathbb{D}'_{\mathcal{E}}$, respectively, the results stated in Subsections 4.2 and 5.2 continue to hold. Here we refer again to the machinery developed in [4].

We point out that, in view of the standard Dirichlet form we have dealt with, the only reason why $\mathbb{V}'_{\mathcal{E}}$ and $\mathbb{D}'_{\mathcal{E}}$ do not coincide with $\mathbb{V}'$ and $\mathbb{D}'$, respectively, is that in the compact case the kernel of the Dirichlet energy functional $\mathcal{E} : \mathbb{H} \to [0, +\infty]$ coincides with the set of constant functions, hence is nontrivial. In fact $\mathbb{V}'_{\mathcal{E}}$ and $\mathbb{D}'_{\mathcal{E}}$ turn out to be identified as those elements of $\mathbb{V}'$ and $\mathbb{D}'$, respectively, that vanish on constant functions. On the contrary, in the noncompact case there holds

$$\mathcal{E}(f) = 0 \text{ and } f \in \mathbb{H} \quad \Longrightarrow \quad f = 0$$

provided $\mathcal{V}(\mathbb{M}^n) = \infty$, which is always true if (H2) is satisfied.

5.4. **Optimality for small times.** In what follows, even if the discussion could in principle be made more general, we will restrict ourselves to $\mathbb{M}^n = \mathbb{H}^n_K$, that is the n-dimensional hyperbolic space of constant sectional curvature $\mathrm{Sec} = -K$. The key starting point to show optimality is the next delicate result, inspired by [30, Proposition 6].

Lemma 5.6. *Let $K > 0$, $x \in \mathbb{H}^n_K$ and v be a unit tangent vector of $T_x \mathbb{H}^n_K$. Let $r, \delta > 0$. Denote by $v^\perp \subset T_x \mathbb{H}^n_K$ the orthogonal subspace to v and set $\mathrm{E} := \exp_x v^\perp \subset \mathbb{H}^n_K$. Let $w \in v^\perp$ be another unit tangent vector. Consider the point $y := \exp_x \delta v$ and set $w' := I^x_y(w)$, where $I^x_y : T_x \mathbb{H}^n_K \to T_y \mathbb{H}^n_K$ stands for the parallel-transport map along the geodesic $t \mapsto \exp_x tv$. Then*

$$\mathsf{d}(\exp_y rw', \mathrm{E}) = \delta \left(1 + \frac{K}{2} r^2 + O(r^3 + \delta r^2) \right) \qquad \text{as } (r, \delta) \to 0. \tag{5.28}$$

More in general, if $u' \in T_y \mathbb{H}^n_K$ is a unit tangent vector, then

$$\mathsf{d}(\exp_y ru', \mathrm{E}) = \delta \left(1 + \frac{K}{2} r^2 \sin^2\alpha(u', I^x_y(v)) + O(r^3 + \delta r^2) \right)$$
$$+ r \cos\alpha(u', I^x_y(v)) + O(r^3) \qquad \text{as } (r, \delta) \to 0, \tag{5.29}$$

where $\alpha(\cdot, \cdot) \in [0, \pi]$ denotes the angle between unit vectors in $T_y \mathbb{H}^n_K$. In all the above identities, the remainder terms $O(\cdot)$ can be considered independent of the chosen tangent vectors.

Proof. The expansion of formula (5.28) is exactly what is proved in [30, Section 8]. Consider now a general unit tangent vector $u' \in T_y \mathbb{H}^n_K$. Let us denote by $P_v(ru')$ and $P_{v^\perp}(ru')$ the projections, in the tangent space $T_y \mathbb{H}^n_K$, of the vector ru' on the subspace generated by $I^x_y(v)$ and on its orthogonal subspace $I^x_y(v^\perp)$, respectively. Clearly, we have:

$$\left| P_v(ru') \right| = \left| r \cos\alpha(u', I^x_y(v)) \right|, \qquad \left| P_{v^\perp}(ru') \right| = \left| r \sin\alpha(u', I^x_y(v)) \right| \tag{5.30}$$

and

$$P_v(ru') \perp P_{v^\perp}(ru'), \qquad P_v(ru') + P_{v^\perp}(ru') = ru'. \tag{5.31}$$

In agreement with [18], we put

$$\exp_y(P_v(ru'), P_{v^\perp}(ru')) := \exp_{\exp_y P_v(ru')}\left[I^y_{\exp_y P_v(ru')}\big(P_{v^\perp}(ru')\big)\right].$$

Thanks to (5.30) and (5.31), we can apply (5.28) with y replaced by $\exp_y P_v(ru')$ and rw' replaced by the vector $I^y_{\exp_y P_v(ru')}(P_{v^\perp}(ru'))$ (hence δ replaced by $\delta + r\cos\alpha(u', I^x_y(v))$ and r replaced by $|r\sin\alpha(u', I^x_y(v))|$), which yields

$$\mathsf{d}\big(\exp_y(P_v(ru'), P_{v^\perp}(ru')), \mathrm{E}\big)$$
$$= \delta\left(1 + \frac{K}{2}r^2\sin^2\alpha(u', I^x_y(v)) + O(r^3 + \delta r^2)\right) + r\cos\alpha(u', I^x_y(v)) + O(r^3). \tag{5.32}$$

In order to establish (5.29), first of all we take advantage of the triangle inequality, so as to obtain

$$\big|\mathsf{d}(\exp_y ru', \mathrm{E}) - \mathsf{d}\big(\exp_y(P_v(ru'), P_{v^\perp}(ru')), \mathrm{E}\big)\big| \leq \mathsf{d}\big(\exp_y ru', \exp_y(P_v(ru'), P_{v^\perp}(ru'))\big). \tag{5.33}$$

Still in agreement with [18], we denote by $h_y(P_v(ru'), P_{v^\perp}(ru'))$ the unique vector of $T_y\mathbb{H}^n_K$ such that

$$\exp_y\big(h_y(P_v(ru'), P_{v^\perp}(ru'))\big) = \exp_y(P_v(ru'), P_{v^\perp}(ru'));$$

on the other hand, by virtue of [18, formula (3)] there holds

$$\big|h_y(P_v(ru'), P_{v^\perp}(ru')) - ru'\big| = O(r^3),$$

so that

$$\mathsf{d}\big(\exp_y ru', \exp_y(P_v(ru'), P_{v^\perp}(ru'))\big) = O(r^3) \tag{5.34}$$

upon recalling the well-known fact that the Riemannian distance locally can be replaced by the Euclidean distance up an error of order $O(r^3)$ (see e.g. [41, formula (14.1)]). Estimate (5.29) then follows from (5.32), (5.33) and (5.34). $\qquad\square$

Taking advantage of Lemma 5.6, we are able to prove a lower bound for the Wasserstein distance between suitable radially-symmetric probability densities in $\mathbb{H}^n_K$.

Lemma 5.7. *Let $K > 0$ and $\{\rho^\epsilon\}_{\epsilon\in(0,1)}$ be a family of (continuous) radially-symmetric probability densities in $\mathbb{H}^n_K$, i.e. each $\rho^\epsilon : [0, +\infty) \mapsto [0, +\infty)$ satisfies*

$$\frac{|\mathbb{S}^{n-1}|}{K^{\frac{n-1}{2}}}\int_0^{+\infty}\rho^\epsilon(r)\sinh\left(\sqrt{K}r\right)^{n-1}\mathrm{d}r = 1 \qquad \forall\epsilon\in(0,1). \tag{5.35}$$

Suppose in addition that there exist some $\theta\in(0,1)$ and constants $C_1, C_2 > 0$ (independent of ϵ) such that

$$\frac{C_1}{\epsilon^n}\chi_{[0,\theta\epsilon]}(r) \leq \rho^\epsilon(r) \leq \frac{C_2}{\epsilon^n}\chi_{[0,\epsilon]}(r) \qquad \forall\epsilon\in(0,1), \quad \forall r\geq 0. \tag{5.36}$$

Let $x, y\in\mathbb{H}^n_K$ with $\mathsf{d}(x,y) =: \delta > 0$ and consider the probability measures μ^ϵ_x and μ^ϵ_y obtained by centering ρ^ϵ at x and y, respectively. That is, put $\mu^\epsilon_x := \rho^\epsilon(\mathsf{d}(\cdot, x))\mathcal{V}\in\mathscr{P}(\mathbb{H}^n_K)$ and $\mu^\epsilon_y := \rho^\epsilon(\mathsf{d}(\cdot, y))\mathcal{V}\in\mathscr{P}(\mathbb{H}^n_K)$. Then there exist constants $\bar\delta = \bar\delta(n, K, C_1, C_2, \theta) > 0$ and $\kappa = \kappa(n, C_1, C_2, \theta) > 0$ such that, if $\delta\in(0, \bar\delta)$,

$$\mathcal{W}_2(\mu^\epsilon_x, \mu^\epsilon_y) \geq \delta\left(1 + \kappa K\epsilon^2\right) \qquad \forall\epsilon\in(0, \bar\epsilon), \tag{5.37}$$

where $\bar\epsilon = \bar\epsilon(\delta, n, K, C_1, C_2, \theta)\in(0,1)$.

Proof. For simplicity we assume $K = 1$ and set $\mathbb{H}^n := \mathbb{H}^n_1$, since the modifications in order to deal with a general $K > 0$ are inessential. So, let $v\in T_x\mathbb{H}^n$ be the unit vector such that $\exp_x\delta v = y$. Let $i : \mathbb{R}^n \to T_x\mathbb{H}^n$ be an isometric isomorphism that preserves orientation. As in Lemma 5.6, we denote by I^x_y the parallel-transport map between $T_x\mathbb{H}^n$ and $T_y\mathbb{H}^n$ along the geodesic $t\mapsto\exp_x tv$. We then define the maps $\varphi_x : \mathbb{R}^n\to\mathbb{H}^n$ and $\varphi_y : \mathbb{R}^n\to\mathbb{H}^n$ as follows:

$$\varphi_x := \exp_x\circ\, i, \qquad \varphi_y := \exp_y\circ\, I^x_y\circ i.$$

First of all, we normalize ρ^ϵ in such a way that it is a probability measure on $\mathbb{R}^n$, namely we set

$$\rho_E^\epsilon(r) := h(\epsilon)\,\rho^\epsilon(r) \qquad \forall r \geq 0$$

with

$$h(\epsilon) := \frac{1}{1 - |\mathbb{S}^{n-1}| \int_0^\epsilon \rho^\epsilon(r)\,(\sinh(r)^{n-1} - r^{n-1})\,\mathrm{d}r} = 1 + O(\epsilon^2)\,, \tag{5.38}$$

where we used (5.35) and (5.36). Hence we put $\mu_E^\epsilon := \rho_E^\epsilon(|\cdot|)\mathscr{L}^n$, the symbol $\mathscr{L}^n$ standing for the Lebesgue measure on $\mathbb{R}^n$. Now we push forward the probability measure μ_E^ϵ on $\mathbb{H}^n$ by means of the maps φ_x and φ_y:

$$\hat{\mu}_x^\epsilon := (\varphi_x)_\sharp\,\mu_E^\epsilon\,, \qquad \hat{\mu}_y^\epsilon := (\varphi_y)_\sharp\,\mu_E^\epsilon\,. \tag{5.39}$$

It is possible to show that $\hat{\mu}_x^\epsilon$ and $\hat{\mu}_y^\epsilon$ are absolutely continuous w.r.t. to μ_x^ϵ and μ_y^ϵ, respectively, in a quantitative way; more precisely, there exist bounded functions $f_x^\epsilon : \mathbb{H}^n \to \mathbb{R}$ and $f_y^\epsilon : \mathbb{H}^n \to \mathbb{R}$ such that

$$\mathrm{d}\mu_x^\epsilon = \left(1 + \epsilon^2 f_x^\epsilon\right) \mathrm{d}\hat{\mu}_x^\epsilon\,, \qquad \mathrm{d}\mu_y^\epsilon = \left(1 + \epsilon^2 f_y^\epsilon\right) \mathrm{d}\hat{\mu}_y^\epsilon \tag{5.40}$$

and

$$\int_{\mathbb{H}^n} f_x^\epsilon\,\mathrm{d}\hat{\mu}_x^\epsilon = \int_{\mathbb{H}^n} f_y^\epsilon\,\mathrm{d}\hat{\mu}_y^\epsilon = 0\,. \tag{5.41}$$

Indeed, by construction φ_x and φ_y preserve radial lengths and angles. As a consequence, both $\hat{\mu}_x^\epsilon$ and $\hat{\mu}_y^\epsilon$ are represented on $\mathbb{H}^n$ by the same radial density $\hat{\rho}^\epsilon$ via the relation

$$\hat{\rho}^\epsilon(r)\,\sinh(r)^{n-1} = \rho_E^\epsilon(r)\,r^{n-1} = h(\epsilon)\,\rho^\epsilon(r)\,r^{n-1} \qquad \forall r \in (0,\epsilon)\,,$$

whence

$$\rho^\epsilon(r) = \frac{\sinh(r)^{n-1}}{h(\epsilon)\,r^{n-1}}\,\hat{\rho}^\epsilon(r) = \left(1 + \epsilon^2 \frac{\sinh(r)^{n-1} - h(\epsilon)\,r^{n-1}}{\epsilon^2\,h(\epsilon)\,r^{n-1}}\right) \hat{\rho}^\epsilon(r) =: \left(1 + \epsilon^2 f^\epsilon(r)\right) \hat{\rho}^\epsilon(r)$$
$$\forall r \in (0,\epsilon)$$

and therefore (5.40) holds with $f_x^\epsilon(\cdot) = f^\epsilon(\mathsf{d}(\cdot,x))$ and $f_y^\epsilon(\cdot) = f^\epsilon(\mathsf{d}(\cdot,y))$. Note that, in view of (5.38) and a standard Taylor expansion of $\sinh(r)$, the function f^ϵ is uniformly bounded by a constant that depends only on n and C_2. On the other hand, identity (5.41) just follows by the fact that μ_x^ϵ, $\hat{\mu}_x^\epsilon$, μ_y^ϵ, $\hat{\mu}_y^\epsilon$ are all probability measures.

Let E_0 and E_1 be the two disjoint, open, connected components in $\mathbb{H}^n$ separated by E, the latter set being defined as in Lemma 5.6. Assume for convenience that E_1 contains the point y. In order to prove (5.37), as in [30, Section 8] we choose the following 1-Lipschitz function $g : \mathbb{H}^n \to \mathbb{R}$:

$$g(z) := \begin{cases} \mathsf{d}(z,\mathrm{E}) & \text{if } z \in \mathrm{E}_1\,, \\ -\mathsf{d}(z,\mathrm{E}) & \text{otherwise}\,. \end{cases}$$

Upon recalling the duality formula (3.12) along with (3.11) and (5.40), we obtain:

$$\begin{aligned} \mathcal{W}_2(\mu_x^\epsilon, \mu_y^\epsilon) &\geq \mathcal{W}_1(\mu_x^\epsilon, \mu_y^\epsilon) \\ &\geq \int_{\mathbb{H}^n} g(z)\left(1 + \epsilon^2 f_y^\epsilon(z)\right) \mathrm{d}\hat{\mu}_y^\epsilon(z) - \int_{\mathbb{H}^n} g(z)\left(1 + \epsilon^2 f_x^\epsilon(z)\right) \mathrm{d}\hat{\mu}_x^\epsilon(z)\,. \end{aligned} \tag{5.42}$$

Since μ_x^ϵ is represented by a radially-symmetric density about x and $\mathbb{H}^n$ also has a radially-symmetric structure (about any point), by the definition of g it is not difficult to check that in fact

$$\int_{\mathbb{H}^n} g(z)\,\mathrm{d}\mu_x^\epsilon(z) = \int_{\mathbb{H}^n} g(z)\left(1 + \epsilon^2 f_x^\epsilon(z)\right) \mathrm{d}\hat{\mu}_x^\epsilon(z) = 0\,, \tag{5.43}$$

therefore we can focus on the first integral. By virtue of (5.39), we have:

$$\int_{\mathbb{H}^n} g(z)\left(1 + \epsilon^2 f_y^\epsilon(z)\right) \mathrm{d}\hat{\mu}_y^\epsilon(z) = \int_{\mathbb{R}^n} g(\varphi_y(q))\left(1 + \epsilon^2 f_y^\epsilon(\varphi_y(q))\right) \mathrm{d}\mu_E^\epsilon(q)\,; \tag{5.44}$$

on the other hand, thanks to (5.29) and the fact that μ_E^ϵ is supported in the Euclidean ball B_ϵ centered at the origin, we can write

$$\int_{\mathbb{R}^n} g(\varphi_y(q)) \left(1 + \epsilon^2 f_y^\epsilon(\varphi_y(q))\right) \mathrm{d}\mu_E^\epsilon(q)$$
$$= \int_{B_\epsilon} \left[\delta\left(1 + \frac{|q|^2 - (q \cdot p_v)^2}{2} + O(|q|^3 + \delta|q|^2)\right) + q \cdot p_v + O(|q|^3)\right] \left(1 + \epsilon^2 f^\epsilon(|q|)\right) \rho_E^\epsilon(|q|) \, \mathrm{d}q \, ,$$
(5.45)

where $p_v := i^{-1}(v)$. Clearly, by symmetry, the middle term involving $q \cdot p_v$ vanishes when integrated against any radial density. Hence, thanks to (5.36) (still the right-hand inequality) and (5.41), from (5.45) we can infer that

$$\int_{\mathbb{R}^n} g(\varphi_y(q)) \left(1 + \epsilon^2 f_y^\epsilon(\varphi_y(q))\right) \mathrm{d}\mu_E^\epsilon(q) = \delta\left[1 + \frac{n-1}{2n} \int_{B_\epsilon} |q|^2 \rho_E^\epsilon(|q|) \, \mathrm{d}q + O\left(\epsilon^3 + \delta\epsilon^2\right)\right] + O(\epsilon^3) \, .$$

In view of the left-hand inequality in (5.36), there exists a constant $\kappa > 0$ as in the statement such that

$$\int_{\mathbb{R}^n} g(\varphi_y(q)) \left(1 + \epsilon^2 f_y^\epsilon(\varphi_y(q))\right) \mathrm{d}\mu_E^\epsilon(q) \geq \delta\left[1 + 3\kappa\epsilon^2 + O\left(\epsilon^3 + \delta\epsilon^2\right)\right] + O(\epsilon^3) \, . \tag{5.46}$$

Upon collecting (5.42), (5.43), (5.44) and (5.46), the thesis follows by choosing $\bar{\delta}$ so small that $\left|O(\delta\epsilon^2)\right| \leq \kappa\epsilon^2$ for all $\delta \in (0, \bar{\delta})$ and $\bar{\epsilon}$ so small that $\left|\delta O(\epsilon^3)\right| + \left|O(\epsilon^3)\right| \leq \kappa\delta\epsilon^2$ for all $\epsilon \in (0, \bar{\epsilon})$ and all $\delta \in (0, \bar{\delta})$.   □

Proof of Theorem 2.5. Let $M = 1$. Thanks to [38, Theorem 1.1], we know that $\rho(\cdot, t)$ and $\hat{\rho}(\cdot, t)$ are represented by the same radial density centered at x and y, respectively. That is, $\rho(\cdot, t) = \tilde{\rho}(\mathsf{d}(\cdot, x), t)$ and $\hat{\rho}(\cdot, t) = \tilde{\rho}(\mathsf{d}(\cdot, y), t)$ for a suitable continuous, bounded, radially-nonincreasing family of densities $(r, t) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \tilde{\rho}(r, t)$. First of all we observe that, since $\mathbb{H}_K^n$ is a Cartan-Hadamard manifold, $\tilde{\rho}(r, t)$ lies below the *Euclidean* Barenblatt solution $\tilde{\rho}_E(r, t)$, see [21, Remark 2.12] and [38, Introduction]. This means that there exist constants $D = D(n, m) > 0$ and $k = k(n, m) > 0$ such that

$$\tilde{\rho}(r, t) \leq t^{-\frac{n}{2+n(m-1)}} \left(D - k\, r^2\, t^{-\frac{2}{2+n(m-1)}}\right)_+^{m-1} =: \tilde{\rho}_E(r, t) \qquad \forall (r, t) \in \mathbb{R}^+ \times \mathbb{R}^+ \, . \tag{5.47}$$

In particular,

$$\tilde{\rho}(r, t) \leq \frac{D^{m-1}}{t^{\frac{n}{2+n(m-1)}}} \chi_{[0, A(t)]}(r) \quad \forall (r, t) \in \mathbb{R}^+ \times \mathbb{R}^+ \, , \qquad A(t) := \sqrt{\frac{D}{k}}\, t^{\frac{1}{2+n(m-1)}} \, . \tag{5.48}$$

Now let

$$I(t) := \inf_{r \in \left[0, \frac{A(t)}{2}\right]} \tilde{\rho}(r, t) \qquad \forall t > 0 \, .$$

By mass conservation, (5.47) and the fact that $\tilde{\rho}(\cdot, t)$ is nonincreasing, we can deduce the following:

$$\frac{1}{|\mathbb{S}^{n-1}|} = K^{-\frac{n-1}{2}} \int_0^{\frac{A(t)}{2}} \tilde{\rho}(r, t) \sinh\left(\sqrt{K}r\right)^{n-1} \mathrm{d}r + K^{-\frac{n-1}{2}} \int_{\frac{A(t)}{2}}^{A(t)} \tilde{\rho}(r, t) \sinh\left(\sqrt{K}r\right)^{n-1} \mathrm{d}r$$
$$\leq K^{-\frac{n-1}{2}} \int_0^{\frac{A(t)}{2}} \tilde{\rho}_E(r, t) \sinh\left(\sqrt{K}r\right)^{n-1} \mathrm{d}r + K^{-\frac{n-1}{2}} I(t) \int_{\frac{A(t)}{2}}^{A(t)} \sinh\left(\sqrt{K}r\right)^{n-1} \mathrm{d}r \tag{5.49}$$
$$= \left[\frac{\lambda}{|\mathbb{S}^{n-1}|} + I(t)\, C\, t^{\frac{n}{2+n(m-1)}}\right] \left[1 + O\left(t^{\frac{2}{2+n(m-1)}}\right)\right] \, ,$$

where

$$\lambda := \left|\mathbb{S}^{n-1}\right| \int_0^{\frac{1}{2}\sqrt{\frac{D}{k}}} \tilde{\rho}_E(r,1)\, r^{n-1}\, \mathrm{d}r < 1\,, \qquad C := \int_{\frac{1}{2}\sqrt{\frac{D}{k}}}^{\sqrt{\frac{D}{k}}} r^{n-1}\, \mathrm{d}r > 0\,.$$

Note that in the last passage we have exploited the scaling properties of $\tilde{\rho}_E$. From (5.49) and the definition of $I(t)$, it is therefore apparent that there exist constants $D_1 = D_1(n,m) > 0$ and $t_1 = t_1(n,K,m) > 0$ such that

$$\tilde{\rho}(r,t) \geq \frac{D_1}{t^{\frac{n}{2+n(m-1)}}}\, \chi_{\left[0,\frac{A(t)}{2}\right]}(r) \qquad \forall (r,t) \in \mathbb{R}^+ \times (0,t_1)\,. \tag{5.50}$$

Hence, in order to estimate $\mathcal{W}_2(\rho(t),\hat{\rho}(t))$ from below, we are in position to apply Lemma 5.7. Indeed, if we set $\epsilon \equiv A(t)$ and $\rho^\epsilon \equiv \tilde{\rho}(\cdot,t)$, then by virtue of (5.48) and (5.50) we can claim that (5.36) is satisfied with $\theta = 1/2$ and suitable positive constants C_1, C_2 depending only on n and m, provided $\epsilon < A(t_1)$ (condition (5.36) is required to hold for $\epsilon \in (0,1)$ only for convenience). Estimate (2.7) for $M = 1$ is just (5.37), upon exploiting the above relation between t and ϵ, along with the trivial identity $\mathcal{W}_2(\delta_x,\delta_y) = \mathsf{d}(x,y)$.

In order to deal with a general mass $M > 0$, it is enough to notice that $M\rho(tM^{m-1})$ and $M\hat{\rho}(tM^{m-1})$ are still solutions of (2.1) starting from $M\delta_x$ and $M\delta_y$, respectively (recall that $\mathcal{W}_2^2$ is proportional to the mass). $\qquad\square$

References

[1] L. Ambrosio, N. Gigli, G. Savaré, "Gradient Flows in Metric Spaces and in the Space of Probability Measures". Second Edition. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.

[2] L. Ambrosio, N. Gigli, G. Savaré, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math. 195 (2014), 289–391.

[3] L. Ambrosio, N. Gigli, G. Savaré, *Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds*, Ann. Probab. 43 (2015), 339–404.

[4] L. Ambrosio, A. Mondino, G. Savaré, *Nonlinear diffusion equations and curvature conditions in metric measure spaces*, Mem. Amer. Math. Soc, Vol. 262, 1270 (2019).

[5] D. Bakry, T. Coulhon, M. Ledoux, L. Saloff-Coste, *Sobolev inequalities in disguise*, Indiana Univ. Math. J. 44 (1995), 1033–1074.

[6] D. Bakry, M. Émery, *Diffusions hypercontractives (French) [Hypercontractive diffusions]*, Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., 1123, Springer, Berlin, 1985.

[7] D. Bakry, I. Gentil, M. Ledoux, "Analysis and Geometry of Markov Diffusion Operators". Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 348. Springer, Cham, 2014.

[8] D. Bianchi, A.G. Setti, *Laplacian cut-offs, porous and fast diffusion on manifolds and other applications*, Calc. Var. Partial Differential Equations 57 (2018), Art. 4, 33 pp.

[9] F. Bolley, J.A. Carrillo, *Nonlinear diffusion: geodesic convexity is equivalent to Wasserstein contraction*, Comm. Partial Differential Equations 39 (2014), 1860–1869.

[10] M. Bonforte, G. Grillo, J.L. Vázquez, *Fast diffusion flow on manifolds of nonpositive curvature*, J. Evol. Equ. 8 (2008), 99–128.

[11] J.A. Carrillo, R.J. McCann, C. Villani, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Ration. Mech. Anal. 179 (2006), 217–263.

[12] T. Coulhon, D. Hauer, *Regularisation effects of nonlinear semigroups*, to appear in BCAM Springer Briefs, preprint arXiv: https://arxiv.org/abs/1604.08737.

[13] S. Daneri, G. Savaré, *Eulerian calculus for the displacement convexity in the Wasserstein distance*, SIAM J. Math. Anal. 40 (2008), 1104–1122.

[14] M. Erbar, *The heat equation on manifolds as a gradient flow in the Wasserstein space*, Ann. Inst. Henri Poincaré Probab. Stat. 46 (2010), 1–23.

[15] R.L. Foote, *Regularity of the distance function*, Proc. Amer. Math. Soc. 92 (1984), 153–155.

[16] A.R. Fotache, M. Muratori, *Smoothing effects for the filtration equation with different powers*, J. Differential Equations 263 (2017), 3291–3326.

[17] N. Fournier, B. Perthame, *Monge-Kantorovich distance for PDEs: the coupling method*, to appear in EMS Surveys in Mathematical Sciences, preprint arXiv: https://arxiv.org/abs/1903.11349.

[18] A.V. Gavrilov, *The double exponential map and covariant derivation*, Siberian Math. J. 48 (2007), 56–61.

[19] G. Grillo, M. Muratori, M.M. Porzio, *Porous media equations with two weights: smoothing and decay properties of energy solutions via Poincaré inequalities*, Discrete Contin. Dyn. Syst. 33 (2013), 3599–3640.

[20] G. Grillo, M. Muratori, F. Punzo, *The porous medium equation with large initial data on negatively curved Riemannian manifolds*, J. Math. Pures Appl. 113 (2018), 195–226.

[21] G. Grillo, M. Muratori, F. Punzo, *The porous medium equation with measure data on negatively curved Riemannian manifolds*, J. Eur. Math. Soc. (JEMS) 20 (2018), 2769–2812.

[22] G. Grillo, M. Muratori, J.L. Vázquez, *The porous medium equation on Riemannian manifolds with negative curvature. The large-time behaviour*, Adv. Math. 314 (2017), 328–377.

[23] G. Grillo, M. Muratori, J.L. Vázquez, *The porous medium equation on Riemannian manifolds with negative curvature: the superquadratic case*, Math. Ann. 373 (2019), 119–153.

[24] E. Hebey, "Nonlinear Analysis on Manifolds: Sobolev Spaces and Inequalities", Courant Lecture Notes in Mathematics, 5. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.

[25] R. Jordan, D. Kinderlehrer, F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. 29 (1998), 1–17.

[26] H. Li, G. Toscani, *Long-time asymptotics of kinetic models of granular flows*, Arch. Ration. Mech. Anal. 172 (2004), 407–428.

[27] S. Lisini, E. Mainini, A. Segatti, *A gradient flow approach to the porous medium equation with fractional pressure*, Arch. Ration. Mech. Anal. 227 (2018), 567–606.

[28] J.M. Lee, "Introduction to Smooth Manifolds". Second Edition. Graduate Texts in Mathematics, 218. Springer, New York, 2013.

[29] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations 26 (2001), 101–174.

[30] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. 256 (2009), 810–864.

[31] S. Ohta, A. Takatsu, *Displacement convexity of generalized relative entropies*, Adv. Math. 228 (2011), 1742–1787.

[32] F. Otto, M. Westdickenberg, *Eulerian calculus for the contraction in the Wasserstein distance*, SIAM J. Math. Anal. 37 (2005), 1227–1255.

[33] M.-K. von Renesse, K.-T. Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math. 58 (2005), 923–940.

[34] R.S. Strichartz, *Analysis of the Laplacian on the complete Riemannian manifold*, J. Funct. Anal. 52 (1983), 48–79.

[35] K.-T. Sturm, *Convex functionals of probability measures and nonlinear diffusions on manifolds*, J. Math. Pures Appl. 84 (2005), 149–168.

[36] N.Th. Varopoulos, *Small time Gaussian estimates of heat diffusion kernels. I. The semigroup technique*, Bull. Sci. Math. 113 (1989), 253–277.

[37] J.L. Vázquez, *Asymptotic behaviour for the heat equation in hyperbolic space*, to appear in Comm. Analysis and Geometry, preprint arXiv: https://arxiv.org/abs/1811.09034.

[38] J.L. Vázquez, *Fundamental solution and long time behavior of the porous medium equation in hyperbolic space*, J. Math. Pures Appl. 104 (2015), 454–484.

[39] J.L. Vázquez, "Smoothing and Decay Estimates for Nonlinear Diffusion Equations. Equations of Porous Medium Type", Oxford Lecture Series in Mathematics and its Applications, 33. Oxford University Press, Oxford, 2006.

[40] J.L. Vázquez, "The Porous Medium Equation. Mathematical Theory", Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, Oxford, 2007.

[41] C. Villani, "Optimal Transport, Old and New", Springer Verlag, Grundlehren der mathematischen Wissenschaften, 2008.

Nicolò De Ponti: Scuola Internazionale Superiore di Studi Avanzati (SISSA), Via Bonomea 265, 34136 Trieste (Italy)

E-mail address: ndeponti@sissa.it

Matteo Muratori: Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo Da Vinci 32, 20133 Milano (Italy)

E-mail address: matteo.muratori@polimi.it

Carlo Orrieri: Dipartimento di Matematica, Università degli Studi di Pavia, Via Ferrata 5, 27100 Pavia (Italy)

E-mail address: carlo.orrieri@unipv.it