EULERIAN CALCULUS FOR THE CONTRACTION IN THE WASSERSTEIN DISTANCE *

FELIX OTTO †‡ AND MICHAEL WESTDICKENBERG †§

Abstract. We consider the porous medium equation on a compact Riemannian manifold and give a new proof of the contraction of its semigroup in the Wasserstein distance. This proof is based on the insight that the porous medium equation does not increase the size of infinitesimal perturbations along gradient flow trajectories, and on an Eulerian formulation for the Wasserstein distance using smooth curves. Our approach avoids the existence result for optimal transport maps on Riemannian manifolds.

Key words. Porous medium equation; contraction; Wasserstein distance

AMS subject classifications. 58J65, 28D05, 60E15, 60G15

1. Introduction.

In this paper we consider the porous medium equation

$$\partial_t \rho - \Delta U(\rho) = 0 \tag{1.1}$$

and give an entirely "Eulerian" argument for the contraction of its semigroup in the Wasserstein distance. The argument is guided by the formal gradient flow structure of the porous medium equation proposed in [13].

More precisely, we choose as our state space $\mathcal{M}$ the space of probability measures $\rho(x)\,dx$, endowed with a suitable metric tensor g, see §2.2. The metric tensor g induces a distance on $(\mathcal{M}, g)$ that coincides with the Wasserstein distance $\mathcal{W}(\rho_0, \rho_1)$. Loosely speaking, this equivalence is a consequence of the Benamou–Brenier Eulerian formulation of the optimal transportation problem defining $\mathcal{W}(\rho_0, \rho_1)$ [4]. Then the porous medium equation is the gradient flow on $(\mathcal{M}, g)$ of the functional

$$E(\rho) = \int e(\rho)\,dx, \tag{1.2}$$

where the "osmotic pressure" $U(\rho)$ is related to the energy density $e(\rho)$ via

$$U(\rho) = \rho e'(\rho) - e(\rho) \quad \text{for } \rho \geqslant 0. \tag{1.3}$$

We notice that U is (strictly) monotone if and only if e is (strictly) convex, and that for strictly monotone U, (1.1) is of parabolic type. The contraction property for the porous medium semigroup then follows from the convexity of E on $(\mathcal{M}, g)$. The latter is a reformulation of McCann's displacement convexity [11].

This formal argument has been made rigorous in [13] using the fact that for any two points ρ_0, ρ_1 a shortest curve with respect to $\mathcal{W}$ exists. The existence of these shortest curves relies on Brenier's result [5] on the existence of a one-to-one optimal transport map $y = \Phi(x)$ between two measures $\rho_0(x)\,dx$ and $\rho_1(y)\,dy$. This can be

*Felix Otto and Michael Westdickenberg acknowledge partial support by EU network *Hyperbolic and kinetic equations* HRPN-CT-2002-00282 and by Sonderforschungsbereich 611 *Singular phenomena and scaling in mathematical models* at Bonn University.

†Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 10, 53115 Bonn, Germany.

‡*E-mail address*: `otto@iam.uni-bonn.de`.

§*E-mail address*: `mwest@iam.uni-bonn.de`.

seen as a Lagrangian approach. It is somewhat delicate since the optimal transport map Φ can be nonsmooth even if the densities ρ_0 and ρ_1 are smooth.

In this paper, we carry out a rigorous Eulerian approach based on the new insight

- that the porous medium equation does not increase the naturally defined action $\mathcal{A}(\rho)$ of *smooth* curves $[0,1] \ni s \mapsto \rho(s) \in \mathcal{M}$ (see Proposition 4.2);
- that the squared Wasserstein distance $\frac{1}{2}\mathcal{W}(\rho_0, \rho_1)^2$ is the infimum of $\mathcal{A}(\rho)$ over smooth curves connecting ρ_0 to ρ_1 (see Proposition 4.3).

Hence we can work in the "class of smooth objects". Alternatively, contraction estimates can also be derived in more elaborate frameworks based on metric space theory. We refer to the recent publications [6, 1] for further information.

Our approach allows to obtain the contraction property on a compact Riemannian manifold $\mathbb{M}^n$ (instead of $\mathbb{R}^n$) without additional effort. A sufficient condition is that the Ricci curvature of $\mathbb{M}^n$ be nonnegative. This is the well known Bakry–Emery criterion for the logarithmic Sobolev inequality [3] (which can be refined using Γ_2–calculus [2]). It turns out that contractivity of the semigroup for certain nonlinear evolutions equation is in fact equivalent to lower bounds for the Ricci curvature. This has been proved for the heat semigroup in [17] and more generally in [16]. Our Eulerian approach avoids the subtle existence result for optimal transport maps Φ on Riemannian manifolds by McCann [12].

2. Gradient flows. It is instructive to discuss our approach in the language of gradient flows. This heuristics will serve as a guideline for the rigorous argument.

2.1. Abstract framework. Let us quickly recall the mathematical structure required for a gradient flow. One first needs a smooth function $\mathcal{M} \ni \rho \mapsto E(\rho)$ on a differentiable manifold $\mathcal{M}$. The differential $\operatorname{diff} E$ of E is a co-tangent vector field:

$$\mathcal{M} \ni \rho \mapsto \operatorname{diff} E_{|\rho} \in T_\rho\mathcal{M}^*.$$

Therefore one also needs a metric tensor g on $\mathcal{M}$, i.e., a scalar product g_ρ on $T_\rho\mathcal{M}$ in every point $\rho \in \mathcal{M}$. This scalar product allows to identify co-tangent with tangent vectors, yielding the gradient vector field $\operatorname{grad} E$. The gradient flow of E on the Riemannian manifold $(\mathcal{M}, g)$ is then given by the dynamical system

$$\frac{d\rho}{dt} = -\operatorname{grad} E_{|\rho}. \tag{2.1}$$

For subsequent use, we shall reformulate (2.1). We recall that the differential $\operatorname{diff} E$ can be inferred from differentiating E along a curve $[0,1] \ni s \mapsto \rho(s) \in \mathcal{M}$:

$$\frac{d}{ds} E\big(\rho(s)\big) = \left\langle \operatorname{diff} E_{|\rho(s)},\, \frac{d\rho}{ds}(s) \right\rangle.$$

Then the gradient $\operatorname{grad} E$ is defined by the requirement that for any tangent vector field $[0,1] \ni s \mapsto \delta\rho(s) \in T_{\rho(s)}\mathcal{M}$ along the above curve we have

$$g_{\rho(s)}\big(\operatorname{grad} E_{|\rho(s)},\, \delta\rho(s)\big) = \big\langle \operatorname{diff} E_{|\rho(s)},\, \delta\rho(s)\big\rangle.$$

Now a trajectory $[0,\infty) \ni t \mapsto \rho(t) \in \mathcal{M}$ of (2.1) is characterized by the fact that for any tangent vector field $[0,\infty) \ni t \mapsto \delta\rho(t) \in T_{\rho(t)}\mathcal{M}$ one has

$$g_{\rho(t)}\left(\frac{d\rho}{dt}(t),\, \delta\rho(t)\right) + \big\langle \operatorname{diff} E_{|\rho(t)},\, \delta\rho(t)\big\rangle = 0 \quad \text{for all } t. \tag{2.2}$$

2.2. Heuristics: The porous medium equation as gradient flow. We are interested in the porous medium equation on a compact, connected Riemannian manifold $\mathbb{M}^n$ without boundary. We denote by $\cdot$ the metric tensor on $\mathbb{M}^n$, by ∇, $\nabla\cdot$ and $\Delta = \nabla\cdot\nabla$ the gradient, divergence and Laplacian on $\mathbb{M}^n$. Finally dx denotes the volume form on $\mathbb{M}^n$; without loss of generality we assume $\int_{\mathbb{M}^n} 1\,dx = 1$. The porous medium equation describes the evolution of a nonnegative density $\rho(t,x)$ on $\mathbb{M}^n$. It is given by the nonlinear diffusion equation

$$\partial_t \rho - \Delta U(\rho) = 0. \tag{2.3}$$

The porous medium equation preserves the total mass and we assume $\int_{\mathbb{M}^n} \rho\,dx = 1$ for definiteness. In view of this, our state space $\mathcal{M}$ is the space of all nonnegative functions $\rho\colon \mathbb{M}^n \to [0,\infty)$ with unit integral:

$$\int_{\mathbb{M}^n} \rho\,dx = 1. \tag{2.4}$$

We also may think of $\mathcal{M}$ as the space of probability measures $\rho(x)\,dx$ on $\mathbb{M}^n$. For convenience we will not distinguish in the following between functions and the measures they induce via the volume element dx defined on $\mathbb{M}^n$.

Following [13], we now introduce the metric tensor g on $\mathcal{M}$. Notice that in view of (2.4) we may think of infinitesimal perturbations $\delta\rho \in T_\rho\mathcal{M}$ of a state $\rho \in \mathcal{M}$ as functions $\delta\rho\colon \mathbb{M}^n \to \mathbb{R}$ with

$$\int_{\mathbb{M}^n} \delta\rho\,dx = 0. \tag{2.5}$$

For given $\rho \in \mathcal{M}$ we define the scalar product g_ρ on $T_\rho\mathcal{M}$ as

$$g_\rho\big(\delta\rho_0,\,\delta\rho_1\big) = \int_{\mathbb{M}^n} \nabla\phi_0\cdot\nabla\phi_1\,\rho\,dx, \tag{2.6}$$

where, up to additive constants, the functions $\phi_i\colon \mathbb{M}^n \to \mathbb{R}$ are defined by

$$\delta\rho_i - \nabla\cdot(\rho\nabla\phi_i) = 0. \tag{2.7}$$

Notice that (2.7) constitutes an elliptic equation with variable coefficient $\rho \geqslant 0$ for ϕ_i; (2.5) is necessary for the existence. If ρ is strictly positive and $\rho, \delta\rho_i$ are smooth, then (2.5) is also sufficient for the existence of a smooth solution ϕ_i. For later use we notice that $g_\rho(\delta\rho_0,\delta\rho_1)$ can be rewritten as

$$g_\rho\big(\delta\rho_0,\,\delta\rho_1\big) = -\int_{\mathbb{M}^n} \delta\rho_0\,\phi_1\,dx. \tag{2.8}$$

The quadratic part of the metric tensor can also be characterized variationally:

$$\tfrac{1}{2}g_\rho\big(\delta\rho,\,\delta\rho\big) = \sup_\phi \Big\{ -\int_{\mathbb{M}^n} \tfrac{1}{2}|\nabla\phi|^2\rho\,dx - \int_{\mathbb{M}^n} \delta\rho\,\phi\,dx \Big\}, \tag{2.9}$$

where the sup is taken over all smooth functions $\phi\colon \mathbb{M}^n \longrightarrow \mathbb{R}$. In view of (2.7), we may think of ϕ_i as the "velocity potential" that generates the infinitesimal change $\delta\rho_i$ of the density ρ.

We now formally argue that (2.3) is indeed the gradient flow of (1.2) on $(\mathcal{M}, g)$, reproducing the argument in [13]. We are given a nonnegative function $\rho = \rho(t,x)$ satisfying (2.4); we fix a time t. Let the function $\delta\rho$ of x be given with (2.5) and let ϕ be related to $\delta\rho$ by (2.7). Then we have on the one hand that

$$\begin{aligned}\big\langle \operatorname{diff} E_{|\rho},\, \delta\rho\big\rangle &= \int_{\mathbb{M}^n} e'(\rho)\,\delta\rho\,dx \\ &\overset{(2.7)}{=} -\int_{\mathbb{M}^n} e''(\rho)\nabla\rho\cdot\rho\nabla\phi\,dx \\ &\overset{(1.3)}{=} -\int_{\mathbb{M}^n} \nabla U(\rho)\cdot\nabla\phi\,dx \\ &= \int_{\mathbb{M}^n} \Delta U(\rho)\,\phi\,dx. \end{aligned} \tag{2.10}$$

On the other hand we have according to (2.8)

$$g_\rho\big(\partial_t\rho,\, \delta\rho\big) = -\int_{\mathbb{M}^n} \partial_t\rho\,\phi\,dx.$$

The combination of the last two identities gives, for any $\delta\rho$ satisfying (2.5),

$$g_\rho\big(\partial_t\rho,\, \delta\rho\big) + \big\langle \operatorname{diff} E_{|\rho},\, \delta\rho\big\rangle = -\int_{\mathbb{M}^n} \big(\partial_t\rho - \Delta U(\rho)\big)\,\phi\,dx.$$

In view of (2.2), this proves that indeed (2.3) is the gradient flow of (1.2) with respect to the metric tensor (2.6) defined on $\mathcal{M}$.

3. Convexity and contraction. In this section we discuss heuristically how the convexity of E on $(\mathcal{M}, g)$ implies contraction for the gradient flow.

3.1. Abstract framework. Recall that a function E on a Riemannian manifold $(\mathcal{M}, g)$ is convex if its Hessian Hess E is positive definite in any point $\rho \in \mathcal{M}$, i.e.,

$$g_\rho\Big(\delta\rho,\, \operatorname{Hess} E_{|\rho}\,\delta\rho\Big) \geqslant 0 \quad \text{for all } \delta\rho \in T_\rho\mathcal{M} \text{ and } \rho\in\mathcal{M}.$$

In an infinite-dimensional context, it is convenient to have alternative ways of probing convexity. We mention two possibilities:

- The standard way to probe convexity is by geodesics: If $[0,1] \ni s \mapsto \rho(s) \in \mathcal{M}$ is a geodesic, i.e., any curve for which

$$\frac{D}{ds}\frac{d\rho}{ds} = 0,$$

where $\frac{D}{ds}$ denotes the covariant derivative along $s \mapsto \rho(s)$, then we have

$$\frac{d^2}{ds^2}\,E\big(\rho(s)\big) \geqslant 0.$$

Indeed, this follows from the chain rule

$$\begin{aligned}\frac{d^2}{ds^2}\,E\big(\rho(s)\big) &= \frac{d}{ds}\,g_\rho\bigg(\frac{d\rho}{ds},\, \operatorname{grad} E_{|\rho}\bigg) \\ &= g_\rho\bigg(\frac{d\rho}{ds},\, \operatorname{Hess} E_{|\rho}\,\frac{d\rho}{ds}\bigg) + g_\rho\bigg(\frac{D}{ds}\frac{d\rho}{ds},\, \operatorname{grad} E_{|\rho}\bigg) \\ &= g_\rho\bigg(\frac{d\rho}{ds},\, \operatorname{Hess} E_{|\rho}\,\frac{d\rho}{ds}\bigg). \end{aligned} \tag{3.1}$$

- There is another way to probe convexity of E: For any gradient flow trajectory $[0,\infty) \ni t \mapsto \rho(t) \in \mathcal{M}$, i.e., any curve for which

$$\frac{d\rho}{dt} = -\operatorname{grad} E_{|\rho},$$

and any infinitesimal perturbation $[0,\infty) \ni t \mapsto \delta\rho(t) \in T_{\rho(t)}\mathcal{M}$ along this curve for which by the chain rule

$$\frac{D}{dt}\,\delta\rho = -\operatorname{Hess} E_{|\rho}\,\delta\rho, \tag{3.2}$$

we have that the size of this perturbation does not increase over time:

$$\frac{d}{dt}\,\tfrac{1}{2} g_\rho\big(\delta\rho,\,\delta\rho\big) \leqslant 0. \tag{3.3}$$

Indeed, this follows from

$$\frac{d}{dt}\,\tfrac{1}{2} g_\rho\big(\delta\rho,\,\delta\rho\big) = g_\rho\Big(\delta\rho,\,\frac{D}{dt}\,\delta\rho\Big) = -g_\rho\Big(\delta\rho,\,\operatorname{Hess} E_{|\rho}\,\delta\rho\Big). \tag{3.4}$$

The property (3.3) has a finite counterpart: Recall that the distance $\operatorname{dist}(\rho_0,\rho_1)$ between $\rho_0,\rho_1 \in \mathcal{M}$ induced by the metric tensor g is defined by

$$\tfrac{1}{2}\operatorname{dist}(\rho_0,\rho_1)^2 = \inf\left\{\mathcal{A}(\rho)\;\middle|\;[0,1]\ni s\mapsto\rho(s)\in\mathcal{M},\;\left\{\begin{array}{c}\rho(0,\cdot)=\rho_0\\ \rho(1,\cdot)=\rho_1\end{array}\right\}\right\}, \tag{3.5}$$

where $\mathcal{A}(\rho)$ is the natural action of a curve, i.e.,

$$\mathcal{A}(\rho) := \int_0^1 \tfrac{1}{2} g_\rho\Big(\frac{d\rho}{ds},\frac{d\rho}{ds}\Big)\,ds. \tag{3.6}$$

We now argue that (3.3) easily yields a global consequence of the convexity of E: The gradient flow of E is a contraction in dist. This means that for any two gradient flow trajectories $[0,\infty) \ni t \mapsto \rho_i(t) \in \mathcal{M}$, $i = 0,1$, i.e., any curves with

$$\frac{d\rho_i}{dt} = -\operatorname{grad} E_{|\rho_i},$$

we have

$$\operatorname{dist}(\rho_0,\rho_1) \quad \text{is nonincreasing in } t.$$

Indeed, by translational invariance in time, it is enough to show that

$$\tfrac{1}{2}\operatorname{dist}\big(\rho_0(t),\rho_1(t)\big)^2 \leqslant \tfrac{1}{2}\operatorname{dist}\big(\rho_0(0),\rho_1(0)\big)^2 \quad \text{for all } t \geqslant 0. \tag{3.7}$$

According to (3.5), for given $\epsilon > 0$, there exists a curve $[0,1] \ni s \mapsto \bar\rho(s) \in \mathcal{M}$ such that $\bar\rho(s=0) = \rho_0(t=0)$ and $\bar\rho(s=1) = \rho_1(t=0)$, with

$$\tfrac{1}{2}\operatorname{dist}\big(\rho_0(0),\rho_1(0)\big)^2 = \tfrac{1}{2}\operatorname{dist}\big(\bar\rho(0),\bar\rho(1)\big)^2 \geqslant \mathcal{A}(\bar\rho) - \epsilon. \tag{3.8}$$

Now for every $s \in [0,1]$ let $[0,\infty) \ni t \mapsto \rho(s,t) \in \mathcal{M}$ denote the solution of

$$\frac{d\rho(s,\cdot)}{dt} = -\operatorname{grad} E_{|\rho(s,\cdot)}, \tag{3.9}$$

with $\rho(s,0) = \bar\rho(s)$. Notice that then $\rho(0,t) = \rho_0(t)$ and $\rho(1,t) = \rho_1(t)$ so that

$$\tfrac{1}{2} \operatorname{dist}\big(\rho_0(t), \rho_1(t)\big)^2 \leqslant \mathcal{A}\big(\rho(\cdot,t)\big). \tag{3.10}$$

Taking the covariant derivative of (3.9) with respect to s yields

$$\frac{D}{\partial t}\frac{\partial\rho}{\partial s} = \frac{D}{\partial s}\frac{\partial\rho}{\partial t} = -\operatorname{Hess} E_{|\rho}\frac{\partial\rho}{\partial s}.$$

Thus we obtain from (3.3) applied to $\delta\rho = \frac{\partial\rho}{\partial s}$

$$\frac{\partial}{\partial t}\,\tfrac{1}{2} g_\rho\left(\frac{\partial\rho}{\partial s}, \frac{\partial\rho}{\partial s}\right) \leqslant 0.$$

Integration over s yields

$$\frac{d}{dt}\mathcal{A}\big(\rho(\cdot,t)\big) = \int_0^1 \frac{\partial}{\partial t}\,\tfrac{1}{2} g_\rho\left(\frac{\partial\rho}{\partial s}, \frac{\partial\rho}{\partial s}\right) ds \leqslant 0.$$

Integration over t yields

$$\mathcal{A}\big(\rho(\cdot,t)\big) \leqslant \mathcal{A}\big(\rho(\cdot,0)\big).$$

Together with (3.10) and (3.8) we therefore end up with

$$\tfrac{1}{2} \operatorname{dist}\big(\rho_0(t), \rho_1(t)\big)^2 \leqslant \tfrac{1}{2} \operatorname{dist}\big(\rho_0(0), \rho_1(0)\big)^2 + \epsilon,$$

and since $\epsilon > 0$ was arbitrary, (3.7) is proved.

Remark. It is possible to give an argument in favor of

$$\frac{\partial}{\partial t}\,\tfrac{1}{2} g_\rho\left(\frac{\partial\rho}{\partial s}, \frac{\partial\rho}{\partial s}\right) = -g_\rho\left(\frac{\partial\rho}{\partial s}, \operatorname{Hess} E_{|\rho}\frac{\partial\rho}{\partial s}\right) \tag{3.11}$$

that avoids using the covariant derivative altogether: Consider first a family of curves $[0,1] \ni s \mapsto \tilde\rho(s,t) \in \mathcal{M}$ for $t \in [0,\infty)$ such that $s \mapsto \tilde\rho(s,0)$ is a geodesic. Then

$$\frac{\partial}{\partial t}\,\tfrac{1}{2} g_{\tilde\rho}\left(\frac{\partial\tilde\rho}{\partial s}, \frac{\partial\tilde\rho}{\partial s}\right) = \frac{\partial}{\partial s}\, g_{\tilde\rho}\left(\frac{\partial\tilde\rho}{\partial s}, \frac{\partial\tilde\rho}{\partial t}\right) \quad \text{for } t = 0. \tag{3.12}$$

Indeed, given any function $[0,1] \ni s \mapsto \alpha(s) \in \mathbb{R}$ with $\alpha(0) = \alpha(1) = 0$ let

$$\hat\rho(s,t) := \tilde\rho\big(s, \alpha(s)t\big) \quad \text{for all } s, t.$$

Since $\hat\rho(0,t) = \tilde\rho(0,0)$ and $\hat\rho(1,t) = \tilde\rho(1,0)$, the definition of geodesic yields

$$0 = \frac{d}{dt}_{|t=0} \int_0^1 \tfrac{1}{2} g_{\hat\rho}\left(\frac{\partial\hat\rho}{\partial s}, \frac{\partial\hat\rho}{\partial s}\right) ds = \int_0^1 \frac{\partial}{\partial t}_{|t=0}\,\tfrac{1}{2} g_{\hat\rho}\left(\frac{\partial\hat\rho}{\partial s}, \frac{\partial\hat\rho}{\partial s}\right) ds. \tag{3.13}$$

On the other hand, we have $\frac{\partial\hat\rho}{\partial s}(s,t)=\frac{\partial\tilde\rho}{\partial s}(s,\alpha(s)t)+\alpha'(s)t\,\frac{\partial\tilde\rho}{\partial t}(s,\alpha(s)t)$ and therefore

$$\begin{aligned}\Big[\tfrac12 g_{\hat\rho}\Big(\frac{\partial\hat\rho}{\partial s},\frac{\partial\hat\rho}{\partial s}\Big)\Big](s,t)=&\Big[\tfrac12 g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial s}\Big)\Big](s,\alpha(s)t)\\&+\alpha'(s)t\Big[g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial t}\Big)\Big](s,\alpha(s)t)\\&+\tfrac12\big(\alpha'(s)t\big)^2\Big[g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial t},\frac{\partial\tilde\rho}{\partial t}\Big)\Big](s,\alpha(s)t).\end{aligned}$$

Using this identity in (3.13) then gives

$$\begin{aligned}0&=\int_0^1\alpha\,\frac{\partial}{\partial t}_{|t=0}\tfrac12 g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial s}\Big)\,ds+\int_0^1\alpha'\,g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial t}\Big)_{|t=0}\,ds\\&=\int_0^1\alpha\Big\{\frac{\partial}{\partial t}\,\tfrac12 g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial s}\Big)-\frac{\partial}{\partial s}\,g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial t}\Big)\Big\}_{|t=0}\,ds.\end{aligned}$$

This proves (3.12) because α was arbitrary. Consider now the family of gradient flows $\rho=\rho(s,\cdot)$ satisfying (3.9). For any $s_0\in[0,1]$ there exists a map $\tilde\rho$ such that

$$\left.\begin{cases}[0,1]\ni s\mapsto\tilde\rho(s,t)\text{ is a geodesic}\\ \tilde\rho(s_0,t)=\rho(s_0,t)\\ \frac{\partial\tilde\rho}{\partial s}(s_0,t)=\frac{\partial\rho}{\partial s}(s_0,t)\end{cases}\right\}\quad\text{for all }t\in[0,\infty).\tag{3.14}$$

At $s=s_0$ we then find

$$\begin{aligned}\frac{\partial}{\partial t}\,\tfrac12 g_\rho\Big(\frac{\partial\rho}{\partial s},\frac{\partial\rho}{\partial s}\Big)\overset{(3.14)}{=}\frac{\partial}{\partial t}\,\tfrac12 g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial s}\Big)&\overset{(3.12)}{=}\frac{\partial}{\partial s}\,g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\tilde\rho}{\partial t}\Big)\\&\overset{(3.14)}{=}\frac{\partial}{\partial s}\,g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\frac{\partial\rho}{\partial t}\Big)\\&\overset{(2.2)}{=}\frac{\partial}{\partial s}\Big[-\Big\langle\operatorname{diff}E_{|\rho},\frac{\partial\tilde\rho}{\partial s}\Big\rangle\Big]\\&\overset{(3.14)}{=}\frac{\partial}{\partial s}\Big[-\Big\langle\operatorname{diff}E_{|\tilde\rho},\frac{\partial\tilde\rho}{\partial s}\Big\rangle\Big]=-\frac{\partial^2}{\partial s^2}E(\tilde\rho).\end{aligned}$$

By definition of the Hessian, we have

$$\frac{\partial^2}{\partial s^2}E(\tilde\rho)=g_{\tilde\rho}\Big(\frac{\partial\tilde\rho}{\partial s},\operatorname{Hess}E_{|\tilde\rho}\frac{\partial\tilde\rho}{\partial s}\Big)\overset{(3.14)}{=}g_\rho\Big(\frac{\partial\rho}{\partial s},\operatorname{Hess}E_{|\rho}\frac{\partial\rho}{\partial s}\Big)$$

and (3.11) follows.

3.2. Heuristics: Convexity and induced metric. In this section we show heuristically how the abstract framework of the previous section yields a contraction property in the Wasserstein distance for the porous medium equation. This argument will be made rigorous in the remainder of the paper.

We recall the heuristic argument for the convexity of E on $(\mathcal{M},g)$ for which we probe the convexity along geodesics. Therefore we start by heuristically deriving the equation for geodesics, essentially reproducing [14]. An alternative heuristic derivation can be found in [13]. Notice first that within the abstract framework, the geodesic equation is the Euler–Lagrange equation (i.e., the first variation) of the action functional (3.6). We will take this venue. In view of (2.6), our action functional for a curve in $\mathcal{M}$, i.e., for a function $\rho\colon [0,1]\times\mathbb{M}^n \to [0,\infty)$ with $\int_{\mathbb{M}^n}\rho(s,x)\,dx = 1$ for all $s\in[0,1]$, takes the form

$$\mathcal{A}(\rho) = \iint_{[0,1]\times\mathbb{M}^n} \tfrac{1}{2}|\nabla\phi|^2\,\rho\,dx\,ds, \tag{3.15}$$

where the function $\phi\colon [0,1]\times\mathbb{M}^n \to \mathbb{R}$ is determined by

$$\partial_s\rho - \nabla\cdot(\rho\nabla\phi) = 0 \tag{3.16}$$

and plays the role of the tangent vector field along the curve. Like for the metric tensor itself, cf. (2.9), the action functional can be written variationally:

$$\begin{aligned}\mathcal{A}(\rho) &= \sup_\phi\Big\{-\iint_{[0,1]\times\mathbb{M}^n}\tfrac{1}{2}|\nabla\phi|^2\,\rho\,dx\,ds - \iint_{[0,1]\times\mathbb{M}^n}\phi\,\partial_s\rho\,dx\,ds\Big\}\\ &= \sup_\phi\Big\{-\iint_{[0,1]\times\mathbb{M}^n}\tfrac{1}{2}|\nabla\phi|^2\,\rho\,dx\,ds + \iint_{[0,1]\times\mathbb{M}^n}\partial_s\phi\,\rho\,dx\,ds\\ &\qquad + \int_{\mathbb{M}^n}\phi(0,x)\,\rho_0(x)\,dx - \int_{\mathbb{M}^n}\phi(1,x)\,\rho_1(x)\,dx\Big\},\end{aligned} \tag{3.17}$$

where the sup is taken over all smooth functions $\phi\colon [0,1]\times\mathbb{M}^n\to\mathbb{R}$. Here ρ_0,ρ_1 are the fixed end points of the curve, i.e., we have

$$\rho(0,\cdot) = \rho_0 \quad\text{and}\quad \rho(1,\cdot) = \rho_1. \tag{3.18}$$

To obtain the induced distance in $\mathcal{M}$, the expression (3.17) needs to be minimized over all functions $\rho\colon [0,1]\times\mathbb{M}^n\to[0,\infty)$ with $\int_{\mathbb{M}^n}\rho(\cdot,x)\,dx = 1$, see (3.5). In fact, we may think of minimizing (3.17) over all functions $\rho\colon [0,1]\times\mathbb{M}^n\to\mathbb{R}$ because (3.17) is $+\infty$ if (3.16) or (3.18) are violated. Maximizing in ϕ and minimizing in ρ amounts to a saddle-point problem. The first variation in ϕ is given by (3.16) & (3.18). The first variation in ρ is given by the Hamilton–Jacobi equation

$$\partial_s\phi - \tfrac{1}{2}|\nabla\phi|^2 = 0. \tag{3.19}$$

Hence the combination of the transport equation (3.16) and the Hamilton–Jacobi equation (3.19) form the geodesic equation. Note that the system (3.16) & (3.19) is of hyperbolic nature as a partial differential equation. The velocity $u = -\nabla\phi$ satisfies the "pressureless Euler equation"

$$\frac{D}{ds}u + Du\,u = 0,$$

and thus the flow $\partial_s\Phi = u\circ\Phi$ consists of geodesic trajectories, i.e., $\frac{D}{ds}\frac{\partial}{\partial s}\Phi = 0$. Notice that (3.16) states that $\rho(s,\cdot)$ is the push-forward of $\rho(s=0)$ under $\Phi(s,\cdot)$. This is what we call the Lagrangian approach. Geodesics in the sense of shortest curves were given a rigorous meaning for a Riemannian manifold $\mathbb{M}^n$ in [12].

Having identified the geodesic equation, we can probe the convexity of (1.2) along geodesics. This was first done in the Lagrangian framework in [11] and gave rise to the notion of displacement convexity. We reproduce the heuristic Eulerian argument from

[14]. Let $\rho\colon[0,1]\times\mathbb{M}^n\to[0,\infty)$ be a geodesic with tangent field $\phi\colon[0,1]\times\mathbb{M}^n\to\mathbb{R}$, i.e., let (3.16) & (3.19) be satisfied. As in (2.10), we find for the first derivative

$$\frac{dE}{ds}=\int_{\mathbb{M}^n}\Delta U(\rho)\,\phi\,dx=\int_{\mathbb{M}^n}U(\rho)\,\Delta\phi\,dx.$$

For the second derivative, we obtain

$$\begin{aligned}
\frac{d^2E}{ds^2} &= \int_{\mathbb{M}^n}\Big(U'(\rho)\,\partial_s\rho\,\Delta\phi+U(\rho)\,\Delta\partial_s\phi\Big)\,dx\\
&\overset{(3.16)}{\underset{(3.19)}{=}} \int_{\mathbb{M}^n}\Big(U'(\rho)\,\nabla\cdot(\rho\nabla\phi)\,\Delta\phi+U(\rho)\,\Delta\tfrac12|\nabla\phi|^2\Big)\,dx\\
&= \int_{\mathbb{M}^n}\Big(\rho U'(\rho)\,(\Delta\phi)^2+\nabla U(\rho)\cdot\nabla\phi\,\Delta\phi+U(\rho)\,\Delta\tfrac12|\nabla\phi|^2\Big)\,dx\\
&= \int_{\mathbb{M}^n}\Big(\rho U'(\rho)\,(\Delta\phi)^2+U(\rho)\Big(-\nabla\cdot(\nabla\phi\,\Delta\phi)+\Delta\tfrac12|\nabla\phi|^2\Big)\Big)\,dx\\
&= \int_{\mathbb{M}^n}\Big(\Big(\rho U'(\rho)-U(\rho)\Big)(\Delta\phi)^2+U(\rho)\Big(-\nabla\phi\cdot\nabla\Delta\phi+\Delta\tfrac12|\nabla\phi|^2\Big)\Big)\,dx.
\end{aligned}$$

We appeal to Bochner's formula (see [15]):

$$-\nabla\phi\cdot\nabla\Delta\phi+\Delta\tfrac12|\nabla\phi|^2=|\mathrm{D}^2\phi|^2+\nabla\phi\cdot\mathrm{Ric}\,\nabla\phi,$$

where $\mathrm{D}^2\phi$ denotes the Hessian of ϕ, $|A|^2$ stands for the trace of A^tA, and Ric denotes the Ricci curvature of $\mathbb{M}^n$. We thus obtain the formula

$$\frac{d^2E}{ds^2}=\int_{\mathbb{M}^n}\Big(\Big(\rho U'(\rho)-U(\rho)\Big)(\Delta\phi)^2+U(\rho)\Big(|\mathrm{D}^2\phi|^2+\nabla\phi\cdot\mathrm{Ric}\,\nabla\phi\Big)\Big)\,dx. \tag{3.20}$$

In view of (3.1), the right-hand side of (3.20) can be understood as the quadratic part of the Hessian of E in ρ in direction of the infinitesimal variation $\delta\rho=\nabla\cdot(\rho\nabla\phi)$. We notice that it is nonnegative for all functions $\rho\geqslant 0$ and ϕ, if and only if

$$\rho U'(\rho)\geqslant\big(1-\tfrac1n\big)U(\rho)\geqslant 0\quad\text{and}\quad\mathrm{Ric}(x)\geqslant 0\text{ for all }x\in\mathbb{M}^n$$

because $(\Delta\phi)^2\leqslant n|\mathrm{D}^2\phi|^2$. The convexity of E along geodesics in the Riemannian case $\mathbb{M}^n$ was given a rigorous meaning in [7].

To conclude it only remains to prove that (3.5) with (3.15) & (3.16) coincides with $\frac12\mathcal{W}^2(\rho_0,\rho_1)$. Recall that for $\rho_0,\rho_1\in\mathrm{Prob}(\mathbb{M}^n)$, $\mathcal{W}^2(\rho_0,\rho_1)$ is defined as

$$\inf\left\{\iint_{\mathbb{M}^n\times\mathbb{M}^n}d(x,y)^2\,d\pi(x,y)\;\middle|\;\begin{array}{l}\pi\in\mathrm{Prob}(\mathbb{M}^n\times\mathbb{M}^n),\\ \displaystyle\int_{\mathbb{M}^n}d\pi(\cdot,y)=\rho_0,\int_{\mathbb{M}^n}d\pi(x,\cdot)=\rho_1\end{array}\right\},$$

compare to [18]. Several heuristic arguments are possible here (cf. [13] and [14]). However, the rigorous proof we provide in the next section is no more diffcult than a heuristic one; therefore we refer to Proposition 4.3.

4. Rigorous Result: Contraction. We recall that $\mathbb{M}^n$ is a compact connected Riemannian manifold without boundary, with geodesic distance d and $\int_{\mathbb{M}^n} 1\,dx = 1$.

Here is our main result.

THEOREM 4.1 (Contraction estimate). *Assume that* $\rho U'(\rho) \geqslant \big(1-\frac{1}{n}\big)U(\rho) \geqslant 0$ *for all* $\rho \geqslant 0$, *and that* $\mathrm{Ric}(x) \geqslant 0$ *for all* $x \in \mathbb{M}^n$. *For nonnegative initial data* $\bar\rho_0, \bar\rho_1$ *with* $\int_{\mathbb{M}^n} \bar\rho_i\,dx = 1$ *consider solutions* ρ_i *of the porous medium equation*

$$\left.\begin{aligned} \partial_t\rho_i - \Delta U(\rho_i) &= 0 \\ \rho_i(t=0) &= \bar\rho_i \end{aligned}\right\} \quad \text{for } i = 0,1.$$

Then the Wasserstein distance of ρ_0 *and* ρ_1 *is nonincreasing in time, i.e.,*

$$\frac{d^+}{dt}\mathcal{W}^2(\rho_0,\rho_1) \leqslant 0. \tag{4.1}$$

Remark. We have seen in §3.2 that heuristically the convexity of E is equivalent to the conditions on U and Ric required in Theorem 4.1. We have seen in §3.1 that convexity of E is equivalent to the contractivity of the corresponding gradient flow. Hence we expect that the conditions on U and Ric are also necessary. This has been rigorously proven in [16]. Also in [16], the sufficiency of these conditions has been established using the Lagrangian approach mentioned in §1 which relies on [12].

The theorem will be a consequence of the following two propositions.

PROPOSITION 4.2. *Assume that* $\rho U'(\rho) \geqslant \big(1-\frac{1}{n}\big)U(\rho) \geqslant 0$ *for all* $\rho \geqslant 0$, *and that* $\mathrm{Ric}(x) \geqslant 0$ *for all* $x \in \mathbb{M}^n$. *Consider a family of smooth positive solutions of*

$$\partial_t\rho - \Delta U(\rho) = 0, \tag{4.2}$$

depending smoothly on the parameter $s \in [0,1]$. *For any* (s,t) *let* ϕ *be defined by*

$$\partial_s\rho - \nabla\cdot(\rho\nabla\phi) = 0.$$

Then the following holds

$$\frac{d}{dt}\iint_{[0,1]\times\mathbb{M}^n} |\nabla\phi|^2\rho\,dx\,ds \leqslant 0.$$

Remark. Proposition 4.2 is guided by the abstract observation of §3.1: Convexity can be probed by the gradient flow. More precisely, convexity expresses itself by the fact that the action of curves is reduced when the points along the curve are evolved by the gradient flow.

PROPOSITION 4.3. *Consider* $\rho_0\,dx, \rho_1\,dx \in \mathrm{Prob}(\mathbb{M}^n)$ *where* ρ_0, ρ_1 *are smooth and positive functions. Then the Wasserstein distance squared* $\frac{1}{2}\mathcal{W}^2(\rho_0,\rho_1)$ *equals*

$$\inf\left\{ \iint_{[0,1]\times\mathbb{M}^n} \tfrac{1}{2}|\nabla\phi|^2\rho\,dx\,ds \;\middle|\; \begin{aligned} &(\rho>0,\phi) \text{ smooth functions on } [0,1]\times\mathbb{M}^n, \\ &\partial_s\rho - \nabla\cdot(\rho\nabla\phi) = 0, \left\{ \begin{aligned} \rho(0,\cdot) &= \rho_0 \\ \rho(1,\cdot) &= \rho_1 \end{aligned} \right\} \end{aligned} \right\}. \tag{4.3}$$

Proof of Theorem 4.1. Assume first that the initial data are smooth and positive, and that U is linear for $\rho \notin [\alpha, 1/\alpha]$ with $\alpha > 0$ small. Then standard parabolic theory yields that solutions of the porous medium equation for smooth and positive initial data are also smooth and positive. By Proposition 4.3 we can, for any $\varepsilon > 0$, find smooth functions $\big(\bar\rho > 0, \bar\phi\big)$ on $[0,1] \times \mathbb{M}^n$, with

$$\partial_s \bar\rho - \nabla \cdot \big(\bar\rho \nabla \bar\phi\big) = 0, \quad \left\{ \begin{array}{l} \bar\rho(0,\cdot) = \bar\rho_0 \\ \bar\rho(1,\cdot) = \bar\rho_1 \end{array} \right\}$$

such that

$$\iint_{[0,1]\times\mathbb{M}^n} \big|\nabla\bar\phi\big|^2 \bar\rho \, dx\, ds \leqslant \mathcal{W}^2(\bar\rho_0, \bar\rho_1) + \varepsilon.$$

For any $s \in [0,1]$, let $\bar\rho(\cdot, s)$ evolve according to the porous medium equation. This yields a family ρ of solutions depending smoothly on s for which Proposition 4.2 applies. Using again the characterization of Proposition 4.3 then yields

$$\begin{aligned} \mathcal{W}^2\big(\rho_0(t), \rho_1(t)\big) &\leqslant \iint_{[0,1]\times\mathbb{M}^n} |\nabla\phi(t)|^2 \rho(t)\, dx\, ds \\ &\leqslant \iint_{[0,1]\times\mathbb{M}^n} \big|\nabla\bar\phi\big|^2 \bar\rho\, dx\, ds \leqslant \mathcal{W}^2(\bar\rho_0, \bar\rho_1) + \varepsilon \quad \text{for all } t > 0. \end{aligned} \tag{4.4}$$

Since $\varepsilon > 0$ was arbitrary, we obtain (4.1) in this case.

The general case follows by an approximation argument that we do not discuss in detail here. For general nonnegative initial data one can find sequences of smooth positive functions, converging strongly to the given $\bar\rho_0, \bar\rho_1$. Then standard theory for the porous medium equation yields that the solutions converge strongly in $L^1(\mathbb{M}^n)$, hence a posteriori also in the Wasserstein distance which metrizes the weak* topology of measures. Therefore the contraction estimate generalizes to this setting. Similarly, one can approximate a given U with $\rho U'(\rho) \geqslant \big(1 - \frac{1}{n}\big) U(\rho) \geqslant 0$ for all $\rho \geqslant 0$, by a sequence of functions that have the same property and are linear for small and large ρ, and that converge uniformly. Then again standard theory applies and allows to conclude. We refer to [13] where this program has been carried out in $\mathbb{R}^n$. ∎

Proof of Proposition 4.2. The following remark is at the core of Proposition 4.2:

Lemma 4.4. *Consider smooth functions $\big(\rho > 0, \delta\rho\big)$ on $[0,\infty) \times \mathbb{M}^n$ solving*

$$\left\{ \begin{array}{lll} \partial_t \rho & - \Delta U(\rho) & = 0, \\ \partial_t(\delta\rho) & - \Delta\Big(U'(\rho)\, \delta\rho\Big) & = 0. \end{array} \right. \tag{4.5}$$

For any t let ϕ be defined by

$$\delta\rho - \nabla \cdot \big(\rho \nabla\phi\big) = 0. \tag{4.6}$$

Then we have

$$\begin{aligned} &\frac{d}{dt} \int_{\mathbb{M}^n} \tfrac{1}{2} |\nabla\phi|^2 \rho\, dx \\ &\quad = -\int_{\mathbb{M}^n} \bigg(\Big(\rho U'(\rho) - U(\rho)\Big) (\Delta\phi)^2 + U(\rho) \Big(|\mathrm{D}^2\phi|^2 + \nabla\phi \cdot \mathrm{Ric}\, \nabla\phi \Big) \bigg)\, dx. \end{aligned} \tag{4.7}$$

Remark. Observe that the second equation in (4.5) describes the evolution of an infinitesimal perturbation $\delta\rho$ of ρ, see (3.2). Notice further that in view of (4.6), the left-hand side of (4.7) measures how the squared norm of $\delta\rho$ changes in time, cf. (3.3). Observe finally that the right-hand side expression of (4.7) coincides with what we expect to be—up to the sign—the Hessian, see (3.20). In this sense the formula (4.7) reproduces (3.4).

Proof. The left-hand side of (4.7) equals, after an integration by parts,

$$\frac{d}{dt}\int_{\mathbb{M}^n} \tfrac{1}{2}|\nabla\phi|^2\rho\,dx = \int_{\mathbb{M}^n} \Big(-\phi\,\nabla\cdot\big(\rho\,\partial_t\nabla\phi\big) + \tfrac{1}{2}|\nabla\phi|^2\,\partial_t\rho\Big)\,dx. \tag{4.8}$$

We express $-\nabla\cdot(\rho\,\partial_t\nabla\phi)$ in terms of ρ and ϕ. We find by differentiating (4.6)

$$\begin{aligned}
&-\nabla\cdot\big(\rho\,\partial_t\nabla\phi\big)\\
&\quad= -\partial_t(\delta\rho) + \nabla\cdot\big(\partial_t\rho\,\nabla\phi\big)\\
&\quad\overset{(4.5)}{=} -\Delta\Big(U'(\rho)\,\delta\rho\Big) + \nabla\cdot\Big(\Delta U(\rho)\,\nabla\phi\Big)\\
&\quad\overset{(4.6)}{=} -\Delta\Big(U'(\rho)\,\nabla\cdot(\rho\nabla\phi)\Big) + \nabla\cdot\Big(\Delta U(\rho)\,\nabla\phi\Big)\\
&\quad= -\Delta\Big(\big(\rho U'(\rho) - U(\rho)\big)\,\Delta\phi\Big) - \Delta\nabla\cdot\Big(U(\rho)\,\nabla\phi\Big) + \nabla\cdot\Big(\Delta U(\rho)\,\nabla\phi\Big).
\end{aligned}$$

Using this identity and (4.5) in (4.8) gives, after throwing all derivatives onto ϕ,

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{M}^n} \tfrac{1}{2}|\nabla\phi|^2\rho\,dx\\
&\quad= -\int_{\mathbb{M}^n} \bigg(\Big(\rho U'(\rho) - U(\rho)\Big)(\Delta\phi)^2 + U(\rho)\Big(-\nabla\Delta\phi\cdot\nabla\phi + \Delta\tfrac{1}{2}|\nabla\phi|^2\Big)\bigg)\,dx.
\end{aligned}$$

Then we use Bochner's formula

$$-\nabla\Delta\phi\cdot\nabla\phi + \Delta\tfrac{1}{2}|\nabla\phi|^2 = |\mathrm{D}^2\phi|^2 + \nabla\phi\cdot\mathrm{Ric}\,\nabla\phi$$

(see Proposition 3.3 of [15]) to conclude. ∎

Fix $s\in[0,1]$ and let $\delta\rho = \partial_s\rho(\cdot,s)$. Differentiating (4.2) with respect to s gives

$$\partial_t\big(\partial_s\rho\big) - \Delta\Big(U'(\rho)\,\partial_s\rho\Big) = 0.$$

Then Lemma 4.4 applies and yields

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{M}^n} \tfrac{1}{2}|\nabla\phi|^2\rho\,dx\\
&\quad= -\int_{\mathbb{M}^n} \bigg(\Big(\rho U'(\rho) - U(\rho)\Big)(\Delta\phi)^2 + U(\rho)\Big(|\mathrm{D}^2\phi|^2 + \nabla\phi\cdot\mathrm{Ric}\,\nabla\phi\Big)\bigg)\,dx.
\end{aligned} \tag{4.9}$$

Notice that $(\Delta\phi)^2 \leqslant n|\mathrm{D}^2\phi|^2$. By the assumption on U, we therefore get

$$\Big(\rho U'(\rho) - U(\rho)\Big)(\Delta\phi)^2 + U(\rho)|\mathrm{D}^2\phi|^2 \geqslant U(\rho)\Big(-\tfrac{1}{n}(\Delta\phi)^2 + |\mathrm{D}^2\phi|^2\Big) \geqslant 0.$$

Furthermore, we have $\nabla\phi\cdot\mathrm{Ric}\,\nabla\phi \geqslant 0$. This proves the proposition. ∎

Remark. The same reasoning also yields convergence rates: In fact, if

$$U(\rho)\,\xi\cdot\mathrm{Ric}(x)\xi \geqslant \lambda\rho|\xi|^2 \quad \text{for all } \rho\geqslant 0 \text{ and } (x,\xi)\in T\mathbb{M}^n \tag{4.10}$$

for a suitable constant $\lambda\in\mathbb{R}$, then (4.9) gives

$$\frac{d}{dt}\int_{\mathbb{M}^n}\tfrac12|\nabla\phi|^2\rho\,dx = -\int_{\mathbb{M}^n}U(\rho)\Big(\nabla\phi\cdot\mathrm{Ric}\,\nabla\phi\Big)\,dx \leqslant -2\lambda\int_{\mathbb{M}^n}\tfrac12|\nabla\phi|^2\rho\,dx.$$

We obtain exponential decay of $\int_{\mathbb{M}^n}\frac12|\nabla\phi|^2\rho\,dx$ with rate 2λ, thus of $\mathcal{W}^2(\rho_1,\rho_0)$, by (4.4). For the heat equation on the unit sphere, for example, condition (4.10) is satisfied with constant $\lambda=1$.

Proof of Proposition 4.3. We proceed in five steps.

Step 1: We first prove that $\frac12\mathcal{W}^2(\rho_0,\rho_1)\leqslant$ (4.3). Therefore assume that (ρ,ϕ) is admissible in (4.3). For abbreviation we introduce the velocity field $u:=-\nabla\phi$, such that $\partial_s\rho+\nabla\cdot(\rho u)=0$, and consider the flow induced by u:

$$\Phi\colon[0,1]\times\mathbb{M}^n\longrightarrow\mathbb{M}^n \quad\text{with}\quad \partial_s\Phi(s,x)=u\big(s,\Phi(s,x)\big),\ \Phi(0,x)=x, \tag{4.11}$$

for all $(s,x)\in[0,1]\times\mathbb{M}^n$. Then the measure $\rho(s,x)\,dx$ is the push-forward of the measure $\rho_0(x)\,dx$ under $\Phi(s,\cdot)$, i.e., we have for all smooth functions ζ on $\mathbb{M}^n$

$$\int_{\mathbb{M}^n}\zeta(x)\rho(s,x)\,dx=\int_{\mathbb{M}^n}\zeta\big(\Phi(s,x)\big)\rho_0(x)\,dx \quad\text{for all } s\in[0,1]. \tag{4.12}$$

Moreover, by definition of the geodesic distance d we have

$$d\big(x,\Phi(1,x)\big)^2\leqslant\int_0^1|\partial_s\Phi(s,x)|^2\,ds. \tag{4.13}$$

Let π be the nonnegative measure defined by

$$\iint_{\mathbb{M}^n\times\mathbb{M}^n}\zeta(x,y)\,d\pi(x,y)=\int_{\mathbb{M}^n}\zeta\big(x,\Phi(1,x)\big)\rho_0(x)\,dx \tag{4.14}$$

for all smooth functions ζ on $\mathbb{M}^n\times\mathbb{M}^n$. Thanks to (4.12), π is admissible in the definition of the Wasserstein distance $\mathcal{W}^2(\rho_0,\rho_1)$. Furthermore we have

$$\begin{aligned}
&\iint_{\mathbb{M}^n\times\mathbb{M}^n}d(x,y)^2\,d\pi(x,y)\\
&\overset{(4.14)}{=}\int_{\mathbb{M}^n}d\big(x,\Phi(1,x)\big)^2\rho_0(x)\,dx\\
&\overset{(4.13)}{\leqslant}\int_{\mathbb{M}^n}\bigg(\int_0^1|\partial_s\Phi(s,x)|^2\,ds\bigg)\rho_0(x)\,dx\\
&\overset{(4.11)}{=}\int_0^1\int_{\mathbb{M}^n}\big|u\big(s,\Phi(s,x)\big)\big|^2\rho_0(x)\,dx\,ds\\
&\overset{(4.12)}{=}\int_0^1\int_{\mathbb{M}^n}|u(s,x)|^2\rho(s,x)\,dx\,ds=\iint_{[0,1]\times\mathbb{M}^n}|\nabla\phi|^2\rho\,dx\,ds.
\end{aligned}$$

This proves our claim.

Step 2: Notice that any smooth vector field on $[0,1]\times\mathbb{M}^n$ can be identified with a pair (ρ,m), where ρ is a function on $[0,1]\times\mathbb{M}^n$ and m is an s–dependent vector field on $\mathbb{M}^n$ (such as $m=-\rho\nabla\phi$). We will now show that (4.3) equals

$$\inf\left\{ \iint_{[0,1]\times\mathbb{M}^n} \tfrac{1}{2}\rho^{-1}|m|^2\,dx\,ds \;\middle|\; \begin{array}{l} (\rho>0,m) \text{ smooth vector field on } [0,1]\times\mathbb{M}^n, \\ \partial_s\rho+\nabla\cdot m=0,\ \left\{\begin{array}{l}\rho(0,\cdot)=\rho_0\\ \rho(1,\cdot)=\rho_1\end{array}\right\} \end{array}\right\}. \tag{4.15}$$

That (4.15) does not exceed (4.3) is obvious. To prove the converse consider an admissible pair (ρ,m) in the sense of (4.15). By positivity of ρ we then find, for any $s\in[0,1]$, a smooth function ϕ on $\mathbb{M}^n$ solving the elliptic equation

$$\nabla\cdot\big(m+\rho\nabla\phi\big)=0 \quad \text{on } \mathbb{M}^n. \tag{4.16}$$

This ϕ depends smoothly on s because the pair (ρ,m) does. Since $\mathbb{M}^n$ has no boundary

$$\int_{\mathbb{M}^n}\big(m+\rho\nabla\phi\big)\cdot\nabla\phi\,dx=0.$$

Therefore by Cauchy–Schwarz

$$\begin{aligned}\int_{\mathbb{M}^n}\rho|\nabla\phi|^2\,dx &\overset{(4.16)}{=} \int_{\mathbb{M}^n}-m\cdot\nabla\phi\,dx \\ &\leqslant \left(\int_{\mathbb{M}^n}\rho^{-1}|m|^2\,dx\right)^{1/2}\left(\int_{\mathbb{M}^n}\rho|\nabla\phi|^2\,dx\right)^{1/2},\end{aligned}$$

and thus

$$\int_{\mathbb{M}^n}\tfrac{1}{2}\rho|\nabla\phi|^2\,dx \leqslant \int_{\mathbb{M}^n}\tfrac{1}{2}\rho^{-1}|m|^2\,dx.$$

Step 3: Now we generalize the functional (4.15) to a certain class of distributions and prove that then the inf is bounded by the Wasserstein distance, for any measures $\rho_0,\rho_1\in\operatorname{Prob}(\mathbb{M}^n)$. To achieve this, notice first that thanks to the Riemannian metric on $\mathbb{M}^n$, any smooth 1-form ω on $[0,1]\times\mathbb{M}^n$ can be identified with a pair (σ,ξ), where σ is a function on $[0,1]\times\mathbb{M}^n$ and ξ is an s–dependent vector field on $\mathbb{M}^n$, via

$$\big\langle\omega,\ (\rho,m)\big\rangle=\sigma\rho+\xi\cdot m$$

for all (smooth) vector fields (ρ,m). We write

$$\omega=\sigma\,ds+\xi\cdot dx.$$

The space of 1-forms can be topologized as usual in the theory of distributions, but we do not want to go into details and refer to [9, 10] instead. A linear functional on the space of smooth 1-forms is called a current. Any smooth vector field (ρ,m) defined on $[0,1]\times\mathbb{M}^n$ gives rise to a current T via

$$\Big\langle T,\ \sigma\,ds+\xi\cdot dx\Big\rangle := \iint_{[0,1]\times\mathbb{M}^n}\rho\sigma+m\cdot\xi\,dx\,ds. \tag{4.17}$$

But of course not all currents T can be represented in this form.

We consider currents defined on $[0,1]\times\mathbb{M}^n$ that satisfy

$$\Big\langle T,\ \partial_s\zeta\,ds+\nabla\zeta\cdot dx\Big\rangle=\int_{\mathbb{M}^n}\zeta(1,x)\,d\rho_1(x)-\int_{\mathbb{M}^n}\zeta(0,x)\,d\rho_0(x) \tag{4.18}$$

for all test functions ζ, for given $\rho_0,\rho_1\in\mathrm{Prob}(\mathbb{M}^n)$. If now T is of the form (4.17), then (4.18) is just the weak formulation of the continuity equation $\partial_s\rho+\nabla\cdot m=0$ with initial and final data ρ_0 and ρ_1. Following [4], we can generalize the action

$$\iint_{[0,1]\times\mathbb{M}^n}\tfrac{1}{2}\rho^{-1}|m|^2\,dx\,ds \tag{4.19}$$

as follows. For any current T with (4.18) we consider

$$\mathcal{A}(T):=\sup\Big\{\Big\langle T,\ \sigma\,ds+\xi\cdot dx\Big\rangle\ \Big|\ (\sigma,\xi)\text{ smooth vector field on }[0,1]\times\mathbb{M}^n\text{ with }\sigma+\tfrac{1}{2}|\xi|^2\leqslant 0\Big\}. \tag{4.20}$$

We claim that this $\mathcal{A}(T)$ coincides with (4.19) if T is of the form (4.17): Indeed, setting $\xi=\rho^{-1}m$ and $\sigma=-\frac{1}{2}|\xi|^2=-\frac{1}{2}\rho^{-2}|m|^2$ shows that (4.19) $\leqslant\mathcal{A}(T)$; and

$$\rho\sigma+m\cdot\xi\leqslant-\rho\,\tfrac{1}{2}|\xi|^2+m\cdot\xi\leqslant\tfrac{1}{2}\rho^{-1}|m|^2$$

for all admissible (σ,ξ) implies that $\mathcal{A}(T)\leqslant$ (4.19).

Step 4: Now we prove that $\frac{1}{2}\mathcal{W}^2(\rho_0,\rho_1)$ is bigger than or equal to

$$\inf\Big\{\mathcal{A}(T)\ \Big|\ T\text{ current on }[0,1]\times\mathbb{M}^n\text{ satisfying (4.18)}\Big\}. \tag{4.21}$$

Consider any transference plan $\pi\in\mathrm{Prob}(\mathbb{M}^n\times\mathbb{M}^n)$ that is admissible in the definition of $\mathcal{W}^2(\rho_0,\rho_1)$, and let $\Phi\colon[0,1]\times\mathbb{M}^n\times\mathbb{M}^n\longrightarrow\mathbb{M}^n$ be defined by

$$[0,1]\ni s\mapsto\Phi(s,x,y)\quad\text{is the shortest geodesic between }x\text{ and }y.$$

Then we have in particular

$$\int_0^1|\partial_s\Phi(s,x,y)|^2\,ds=d(x,y)^2\quad\text{and}\quad\left\{\begin{array}{l}\Phi(0,x,y)=x\\ \Phi(1,x,y)=y\end{array}\right\}. \tag{4.22}$$

We define a current T on $[0,1]\times\mathbb{M}^n$ as follows: For all 1-forms $\sigma\,ds+\xi\cdot dx$ let

$$\Big\langle T,\ \sigma\,ds+\xi\cdot dx\Big\rangle:=\iint\int_0^1\Big\{\sigma\big(s,\Phi(s,x,y)\big)+\xi\big(s,\Phi(s,x,x)\big)\cdot\partial_s\Phi(s,x,y)\Big\}\,ds\,d\pi(x,y).$$

This current satisfies the admissibility condition (4.18). Indeed we have

$$\begin{aligned}
&\Big\langle T,\ \partial_s\zeta\,ds + \nabla\zeta\cdot dx\Big\rangle \\
&\quad= \iint\int_0^1 \Big\{\partial_s\zeta\big(s,\Phi(s,x,y)\big) + \nabla\zeta\big(s,\Phi(s,x,y)\big)\cdot\partial_s\Phi(s,x,y)\Big\}\,ds\,d\pi(x,y) \\
&\quad= \iint\int_0^1 \frac{d}{ds}\Big\{\zeta\big(s,\Phi(s,x,y)\big)\Big\}\,ds\,d\pi(x,y) \\
&\quad= \iint \zeta\big(1,\Phi(1,x,y)\big)\,d\pi(x,y) - \iint \zeta\big(0,\Phi(0,x,y)\big)\,d\pi(x,y) \\
&\quad\overset{(4.22)}{=} \iint \zeta(1,y)\,d\pi(x,y) - \iint \zeta(0,x)\,d\pi(x,y) \\
&\quad= \int \zeta(1,y)\,d\rho_1(y) - \int \zeta(0,x)\,d\rho_0(x)
\end{aligned}$$

for all test functions ζ. Now we argue that $\mathcal{A}(T) \leqslant \iint \frac{1}{2}d(x,y)^2\,d\pi(x,y)$. Indeed we have for any vector field (σ,ξ) admissible in (4.20) that

$$\begin{aligned}
&\Big\langle T,\ \sigma\,ds + \xi\cdot dx\Big\rangle \\
&\quad= \iint\int_0^1 \Big\{\sigma\big(s,\Phi(s,x,y)\big) + \xi\big(s,\Phi(s,x,y)\big)\cdot\partial_s\Phi(s,x,y)\Big\}\,ds\,d\pi(x,y) \\
&\quad\leqslant \iint\int_0^1 \Big\{-\tfrac{1}{2}\big|\xi\big(s,\Phi(s,x,y)\big)\big|^2 + \xi\big(s,\Phi(s,x,y)\big)\cdot\partial_s\Phi(s,x,y)\Big\}\,ds\,d\pi(x,y) \\
&\quad\leqslant \iint\int_0^1 \tfrac{1}{2}|\partial_s\Phi(s,x,y)|^2\,ds\,d\pi(x,y) \\
&\quad\overset{(4.22)}{=} \iint \tfrac{1}{2}d(x,y)^2\,d\pi(x,y).
\end{aligned}$$

Step 5. To conclude the proof of the proposition it is then sufficient to show that the two inf in (4.15) and (4.21) coincide. This will follow from Proposition 5.1 below which shows that any current T satisfying the admissibility condition (4.18) for smooth and positive data ρ_0,ρ_1 can in fact be approximated by a current T_ε that is representable by a smooth vector field $(\rho_\varepsilon \geqslant 0, m_\varepsilon)$, in such a way that (4.18) still holds with T_ε in place of T and $\limsup_{\varepsilon\to 0}\mathcal{A}(T_\varepsilon) \leqslant \mathcal{A}(T)$.

The only detail that needs to be settled is (strict) positivity of ρ_ε. We argue as follows. Since $\rho_0,\rho_1 > 0$ and $\mathbb{M}^n$ is compact, there exists $0 < \delta < 1$ with $\rho_0,\rho_1 \geqslant \delta$. Recall that by assumption $\int_{\mathbb{M}^n} 1\,dx = 1$. Then we consider

$$\tilde\rho_0 := \frac{\rho_0 - \delta}{1-\delta} \quad\text{and}\quad \tilde\rho_1 := \frac{\rho_1 - \delta}{1-\delta}$$

which are in $\mathrm{Prob}(\mathbb{M}^n)$. Let $\tilde T$ be the current constructed in Step 4, based on an admissible transference plan π in the definition of $\mathcal{W}^2(\tilde\rho_0,\tilde\rho_1)$. As shown there

$$\mathcal{A}(\tilde T) \leqslant \tfrac{1}{2}\mathcal{W}^2(\tilde\rho_0,\tilde\rho_1). \tag{4.23}$$

We apply Proposition 5.1 to $\tilde T$. This gives an approximation $\tilde T_\varepsilon$ that satisfies (4.18) and is representable by smooth vector fields $(\tilde\rho_\varepsilon \geqslant 0, \tilde m_\varepsilon)$, such that

$$\limsup_{\varepsilon\to 0}\mathcal{A}(\tilde T_\varepsilon) \leqslant \mathcal{A}(\tilde T). \tag{4.24}$$

In view of the remark in Step 3, the admissibility condition (4.18) amounts to

$$\partial_s \tilde{\rho}_\varepsilon + \nabla \cdot \tilde{m}_\varepsilon = 0, \left\{ \begin{array}{l} \tilde{\rho}_\varepsilon(0,\cdot) = \tilde{\rho}_0 \\ \tilde{\rho}_\varepsilon(1,\cdot) = \tilde{\rho}_1 \end{array} \right\}.$$

Now notice that $\tilde{\rho}_0$ and $\tilde{\rho}_1$ are constructed in such a way that

$$(\rho_\varepsilon, m_\varepsilon) := \big((1-\delta)\tilde{\rho}_\varepsilon + \delta, (1-\delta)\tilde{m}_\varepsilon\big)$$

is admissible in (4.15) because $\rho_\varepsilon \geqslant \delta > 0$. We have $\rho_\varepsilon^{-1}|m_\varepsilon|^2 \leqslant (1-\delta)\,\tilde{\rho}_\varepsilon^{-1}|\tilde{m}_\varepsilon|^2$, and thus by the remark in Step 3

$$\iint_{[0,1]\times\mathbb{M}^n} \tfrac{1}{2}\rho_\varepsilon^{-1}|m_\varepsilon|^2 \,dx\,ds \leqslant (1-\delta) \iint_{[0,1]\times\mathbb{M}^n} \tfrac{1}{2}\tilde{\rho}_\varepsilon^{-1}|\tilde{m}_\varepsilon|^2 = (1-\delta)\,\mathcal{A}(\tilde{T}_\varepsilon). \quad (4.25)$$

In view of (4.23), (4.24) and (4.25) it remains to argue that

$$\mathcal{W}(\tilde{\rho}_0, \tilde{\rho}_1) \leqslant \mathcal{W}(\rho_0, \rho_1) + o(1) \quad \text{as } \delta \to 0.$$

By the triangle inequality for the Wasserstein distance (see Theorem 7.3 of [18]),

$$\mathcal{W}(\tilde{\rho}_0, \tilde{\rho}_1) \leqslant \mathcal{W}(\rho_0, \rho_1) + \mathcal{W}(\rho_0, \tilde{\rho}_0) + \mathcal{W}(\rho_1, \tilde{\rho}_1). \quad (4.26)$$

In order to conclude it suffices therefore to prove that the last two terms on the right-hand side of (4.26) can be made small by choosing δ appropriately. We consider the transference plan $\pi \in \mathrm{Prob}(\mathbb{M}^n \times \mathbb{M}^n)$ defined by

$$\begin{aligned} &\iint_{\mathbb{M}^n\times\mathbb{M}^n} \zeta(x,y)\,d\pi(x,y) \\ &\quad := \int_{\mathbb{M}^n} \zeta(x,x)\,\big(\rho_0(x) - \delta\big)\,dx + \tfrac{\delta}{1-\delta} \iint_{\mathbb{M}^n\times\mathbb{M}^n} \zeta(x,y)\,\big(\rho_0(y) - \delta\big)\,dx\,dy \end{aligned} \quad (4.27)$$

for all ζ. This π is admissible in the definition of $\mathcal{W}^2(\rho_0, \tilde{\rho}_0)$ because

$$\begin{aligned} &\iint_{\mathbb{M}^n\times\mathbb{M}^n} \zeta(x)\,d\pi(x,y) \\ &\quad = \int_{\mathbb{M}^n} \zeta(x)\,\big(\rho_0(x) - \delta\big)\,dx + \tfrac{\delta}{1-\delta} \int_{\mathbb{M}^n} \zeta(x)\,dx \int_{\mathbb{M}^n} \big(\rho_0(y) - \delta\big)\,dy \\ &\quad = \int_{\mathbb{M}^n} \zeta(x)\,\rho_0(x)\,dx \end{aligned}$$

and similarly

$$\begin{aligned} &\iint_{\mathbb{M}^n\times\mathbb{M}^n} \zeta(y)\,d\pi(x,y) \\ &\quad = \int_{\mathbb{M}^n} \zeta(y)\,\big(\rho_0(y) - \delta\big)\,dy + \tfrac{\delta}{1-\delta} \int_{\mathbb{M}^n} dx \int_{\mathbb{M}^n} \zeta(y)\,\big(\rho_0(y) - \delta\big)\,dy \\ &\quad = \int_{\mathbb{M}^n} \zeta(y)\,\frac{\rho_0(y) - \delta}{1-\delta}\,dy. \end{aligned}$$

Using $\zeta(x,y) := d(x,y)^2$ in (4.27) then yields

$$\mathcal{W}^2(\rho_0, \tilde{\rho}_0) \leqslant \delta\,\mathrm{diam}(\mathbb{M}^n)^2.$$

The same argument applies to $\mathcal{W}^2(\rho_1, \tilde{\rho}_1)$, thereby finishing the proof. $\square$

5. Approximation of currents. In this section we prove the approximation result for currents used in the proof of Proposition 4.3. Notice that the regularization of currents is well understood, see e.g. [8, 10]. Here we need to adopt the standard arguments somewhat in order to obtain convergence of the action functional (4.20).

PROPOSITION 5.1. *Let $\mathbb{M}^n$ be a compact connected Riemannian manifold without boundary. For given measures $\rho_0, \rho_1 \in \mathrm{Prob}(\mathbb{M}^n)$ consider a current T on $[0,1] \times \mathbb{M}^n$ with $\mathcal{A}(T) < \infty$ which satisfies the admissibility condition*

$$\Big\langle T,\ \partial_s\zeta\,ds + \nabla\zeta\cdot dx\Big\rangle = \int_{\mathbb{M}^n} \zeta(1,x)\,d\rho_1(x) - \int_{\mathbb{M}^n} \zeta(0,x)\,d\rho_0(x) \tag{5.1}$$

for all test functions ζ defined on $[0,1]\times\mathbb{M}^n$. Then we have

1. *There exists a family of currents $\{T_\varepsilon\}_{\varepsilon>0}$ representable in the form*

$$\Big\langle T_\varepsilon,\ \sigma\,ds + \xi\cdot dx\Big\rangle = \iint_{[0,1]\times\mathbb{M}^n} \rho_\varepsilon\sigma + m_\varepsilon\cdot\xi\,dx\,ds \tag{5.2}$$

   for suitable vector fields $(\rho_\varepsilon \geqslant 0, m_\varepsilon)$ which are smooth inside $(0,1)\times\mathbb{M}^n$. The admissibility condition (5.1) still holds with T_ε in place of T and

$$\limsup_{\varepsilon\to 0} \mathcal{A}(T_\varepsilon) \leqslant \mathcal{A}(T).$$

2. *If ρ_0 and ρ_1 are smooth functions, we may assume that the fields $(\rho_\varepsilon, m_\varepsilon)$ are smooth up to the boundary, and thus $\rho_\varepsilon(0,\cdot) = \rho_0$ and $\rho_\varepsilon(1,\cdot) = \rho_1$.*

Remark. Since the action functional $\mathcal{A}(T)$ is lower semicontinuous in the usual weak* topology of currents (see [10]), it even holds that $\lim_{\varepsilon\to 0}\mathcal{A}(T_\varepsilon) = \mathcal{A}(T)$.

Proof. We start with a remark on notation. Because of the action and the admissibility condition, the s– and x–variables have to be treated differently. However, it will often be convenient to lump s– and x–variables together; therefore we will write $\mathbf{x} = (s,x)$, $\boldsymbol{\xi} = (\sigma,\xi)$ and $\mathbf{m} = (\rho, m)$. As a rule, bold symbols always denote $(n+1)$–dimensional objects (vector fields, parameters, operators, sets).

The approximating currents T_ε are obtained by regularization of T. We proceed as usual (see [8, 10]): Since a current is a linear form on 1-forms, we regularize T by duality, i.e., by constructing a linear operator that regularizes 1-forms $\boldsymbol{\xi}\cdot d\mathbf{x}$. This must be done in such a way that exact 1-forms $\boldsymbol{\xi}\cdot d\mathbf{x} = \boldsymbol{\nabla}\zeta\cdot d\mathbf{x}$ turn into exact 1-forms since by assumption (5.1), T vanishes on exact 1-forms that are compactly supported in $(0,1)\times\mathbb{M}^n$. Recall that pulling-back a 1-form under a smooth map preserves exactness. Therefore we regularize $\boldsymbol{\xi}\cdot d\mathbf{x}$ as follows: We construct a family of diffeomorphisms $\{\boldsymbol{\Phi}(\mathbf{z},\cdot)\}_{\mathbf{z}}$ of $\mathbb{R}\times\mathbb{M}^n$, parametrized by $\mathbf{z}\in\mathbb{R}\times\mathbb{R}^n$, we then consider its pull-back $\boldsymbol{\Phi}(\mathbf{z},\cdot)^{\#}(\boldsymbol{\xi}\cdot d\mathbf{x})$ and average over $\mathbf{z}$.

In order to preserve the boundary condition (5.1), it is necessary that $\boldsymbol{\Phi}(\mathbf{z},\cdot)$ leaves the complement of $(0,1)\times\mathbb{M}^n$ invariant. On the other hand, in order to achieve the regularizing effect, it is important that $\mathbf{z}\in\mathbb{R}\times\mathbb{R}^n$ "acts transitively" on $(0,1)\times\mathbb{M}^n$. Because of topological reasons, this cannot be achieved globally by a single map $\boldsymbol{\Phi}$ in general. We have to work locally with several maps $\boldsymbol{\Phi}$, each of which is attached to some open set U of a suitable covering of $\mathbb{M}^n$.

More precisely, we consider a finite covering $\{U_i\}_{i=1}^N$ of $\mathbb{M}^n$ subordinate to some atlas, with $U_i \subset \mathbb{M}^n$ homeomorphic to the unit ball $B_1(0)\subset\mathbb{R}^n$ and $\phi_i\colon U_i \longrightarrow \mathbb{R}^n$

the corresponding smooth coordinate map. We may assume that each ϕ_i extends to a neighborhood of U_i and $B_1(0)$. Based on this map, we shall construct an operator $T \mapsto T^{U_i}_{\varepsilon_i}$ for $\varepsilon_i > 0$ with the following properties:

(a) The operator $T \mapsto T^{U_i}_{\varepsilon_i}$ regularizes in $(0,1) \times U_i$, i.e., $T^{U_i}_{\varepsilon_i}$ is representable in $(0,1) \times U_i$ by a smooth vector field $(\rho_{\varepsilon_i} \geqslant 0, m_{\varepsilon_i})$ as in (5.2).
(b) The operator $T \mapsto T^{U_i}_{\varepsilon_i}$ does not destroy smoothness, i.e., if T is representable by a smooth vector field in $(0,1) \times V$ with $V \subset \mathbb{M}^n$ open, then also $T^{U_i}_{\varepsilon_i}$ is representable by a smooth vector field in $(0,1) \times V$.
(c) The new current is admissible in the sense that (5.1) still holds with $T^{U_i}_{\varepsilon_i}$ in place of T, and we have upper semicontinuity of the action

$$\limsup_{\varepsilon_i \to 0} \mathcal{A}\big(T^{U_i}_{\varepsilon_i}\big) \leqslant \mathcal{A}(T).$$

Then the composition

$$T \;\mapsto\; T^{U_1}_{\varepsilon_1} \;\mapsto\; \big(T^{U_1}_{\varepsilon_1}\big)^{U_2}_{\varepsilon_2} \;\mapsto\; \cdots \;\mapsto\; \big(\cdots\big(T^{U_1}_{\varepsilon_1}\big)^{U_2}_{\varepsilon_2}\cdots\big)^{U_N}_{\varepsilon_N}$$

yields an approximate current T_ε with all the properties required by the proposition. In particular, we obtain a vector field $(\rho_\varepsilon \geqslant 0, m_\varepsilon)$ that represents T_ε in the sense of (5.2) and is smooth throughout $(0,1) \times \mathbb{M}^n$. It suffices to consider each operator $T \mapsto T^{U_i}_{\varepsilon_i}$ separately and check that (a)–(c) are satisfied. To simplify the notation, we will suppress the index i and do not indicate the dependence on U_i. The idea of regularizing a current defined on a manifold by composing several operators attached to local coordinate maps already appeared in [8].

We proceed in several steps.

Step 1: As mentioned before, the regularization $T \mapsto T_\varepsilon$ is based on a family of diffeomorphisms $\{\mathbf{\Phi}(\mathbf{z}, \cdot)\}_{\mathbf{z}}$ of $\mathbb{R} \times \mathbb{M}^n$, parametrized by $\mathbf{z} \in \mathbb{R} \times \mathbb{R}^n$ and attached to the open set $U \subset \mathbb{M}^n$. We would like these diffeomorphisms to leave the complement of $(0,1) \times U$ invariant, but we cannot impose this since in order to control the action we need that the first component of $\mathbf{\Phi}$ does not depend on x.

We shall use

- a smooth map $\mathbf{\Phi} = (\Phi_0, \Phi)\colon (\mathbb{R} \times \mathbb{R}^n) \times (\mathbb{R} \times \mathbb{M}^n) \longrightarrow \mathbb{R} \times \mathbb{M}^n$ (whose construction is postponed until Step 7) with the following properties:

$$\forall\, \mathbf{z} \in \mathbb{R} \times \mathbb{R}^n \quad \mathbf{\Phi}(\mathbf{z}, \cdot) \text{ is a diffeomorphism of } \mathbb{R} \times \mathbb{M}^n \text{ onto } \mathbb{R} \times \mathbb{M}^n, \tag{5.3a}$$

$$\forall\, \mathbf{x} \in \mathbb{R} \times \mathbb{M}^n \quad \mathbf{\Phi}(0, \mathbf{x}) = \mathbf{x}, \tag{5.3b}$$

$$\forall\, \mathbf{x} \in (0,1) \times U \quad \mathbf{\Phi}(\cdot, \mathbf{x}) \text{ is a diffeomorphism of } \mathbb{R} \times \mathbb{R}^n \text{ onto } (0,1) \times U, \tag{5.3c}$$

$$\Phi_0 \text{ does not depend on } x, \tag{5.3d}$$

$$\forall\, (\mathbf{z}, \mathbf{x}) \in (\mathbb{R} \times \mathbb{R}^n) \times \big((0,1) \times (\mathbb{M}^n - U)\big) \quad \Phi(\mathbf{z}, \mathbf{x}) = x, \tag{5.3e}$$

$$\forall\, (\mathbf{z}, \mathbf{x}) \in (\mathbb{R} \times \mathbb{R}^n) \times \big(\big(\mathbb{R} - (0,1)\big) \times \mathbb{M}^n\big) \quad \mathbf{\Phi}(\mathbf{z}, \mathbf{x}) = \mathbf{x}. \tag{5.3f}$$

We shall also need the following maps which exist by (5.3a) and (5.3c):

- the right inverse $\boldsymbol{\Theta} = (\Theta_0, \Theta)\colon (\mathbb{R}\times\mathbb{R}^n)\times(\mathbb{R}\times\mathbb{M}^n) \longrightarrow \mathbb{R}\times\mathbb{M}^n$ of $\boldsymbol{\Phi}$ which is characterized by

$$\forall\,(\mathbf{z},\mathbf{x}) \in (\mathbb{R}\times\mathbb{R}^n)\times(\mathbb{R}\times\mathbb{M}^n) \quad \boldsymbol{\Phi}\big(\mathbf{z},\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\big) = \mathbf{x}; \tag{5.4}$$

- the left inverse $\boldsymbol{\Psi} = (\Psi_0, \Psi)\colon \big((0,1)\times U\big)\times\big((0,1)\times U\big) \longrightarrow \mathbb{R}\times\mathbb{R}^n$ of $\boldsymbol{\Phi}$ which is characterized by

$$\forall\,(\mathbf{x},\mathbf{y}) \in \big((0,1)\times U\big)\times\big((0,1)\times U\big) \quad \boldsymbol{\Phi}\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}),\mathbf{x}\big) = \mathbf{y}. \tag{5.5}$$

For later reference we collect some properties: Let $D_1\boldsymbol{\Phi}$ and $D_2\boldsymbol{\Phi}$ denote the derivatives of $\boldsymbol{\Phi}$ with respect to the first resp. second variable. Then

$$\forall\,(\mathbf{z},\mathbf{x}) \in (\mathbb{R}\times\mathbb{R}^n)\times(\mathbb{R}\times\mathbb{M}^n) \quad D_2\boldsymbol{\Phi}(\mathbf{z},\mathbf{x}) \text{ has full rank,} \tag{5.6}$$

$$\forall\,(\mathbf{z},\mathbf{x}) \in (\mathbb{R}\times\mathbb{R}^n)\times\big((0,1)\times U\big) \quad D_1\boldsymbol{\Phi}(\mathbf{z},\mathbf{x}) \text{ has full rank,} \tag{5.7}$$

as a consequence of (5.3a) and (5.3c). From (5.4) we obtain

$$\begin{aligned} D_{\mathbf{z}}\boldsymbol{\Theta}(\mathbf{z},\mathbf{x}) &= -\Big(D_2\boldsymbol{\Phi}\big(\mathbf{z},\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\big)\Big)^{-1} D_1\boldsymbol{\Phi}\big(\mathbf{z},\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\big),\\ D_{\mathbf{x}}\boldsymbol{\Theta}(\mathbf{z},\mathbf{x}) &= \Big(D_2\boldsymbol{\Phi}\big(\mathbf{z},\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\big)\Big)^{-1}, \end{aligned} \tag{5.8}$$

which together with (5.6) implies that $\boldsymbol{\Theta}$ is smooth. Similarly, (5.5) yields

$$\begin{aligned} D_{\mathbf{y}}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}) &= \Big(D_1\boldsymbol{\Phi}\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}),\mathbf{x}\big)\Big)^{-1},\\ D_{\mathbf{x}}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}) &= -\Big(D_1\boldsymbol{\Phi}\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}),\mathbf{x}\big)\Big)^{-1} D_2\boldsymbol{\Phi}\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}),\mathbf{x}\big), \end{aligned} \tag{5.9}$$

so $\boldsymbol{\Psi}$ is smooth by (5.7). Moreover, we gather from (5.3b) that

$$D_2\boldsymbol{\Phi}(\mathbf{z},\cdot) = \mathrm{Id} + \mathcal{O}(|\mathbf{z}|) \quad \text{as } |\mathbf{z}| \to 0. \tag{5.10}$$

Finally, we notice that the properties above entail that

$$\forall\,\mathbf{y} \in (0,1)\times U \qquad \lim_{\substack{\mathbf{x}\in(0,1)\times U\\ \mathbf{x}\to\partial((0,1)\times U)}} |\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})| = +\infty. \tag{5.11}$$

We argue by contradiction. Indeed suppose that (5.11) fails. Then there exist a sequence $\{\mathbf{x}_\nu\}_\nu \subset (0,1)\times U$ and $\mathbf{x} \in \partial\big((0,1)\times U\big)$, $\mathbf{z} \in \mathbb{R}\times\mathbb{R}^n$ with

$$\lim_{\nu\to\infty} \mathbf{x}_\nu = \mathbf{x} \quad \text{and} \quad \lim_{\nu\to\infty} \boldsymbol{\Psi}(\mathbf{y},\mathbf{x}_\nu) = \mathbf{z}.$$

Passing to the limit in (5.5) yields by continuity of $\boldsymbol{\Phi}$ that

$$\boldsymbol{\Phi}(\mathbf{z},\mathbf{x}) = \mathbf{y} \;\in\; (0,1)\times U. \tag{5.12}$$

Now recall that $\mathbf{x} \in \partial\big((0,1)\times U\big) = \big((0,1)\times\partial U\big) \cup \big(\{0,1\}\times\bar{U}\big)$. If $\mathbf{x} \in (0,1)\times\partial U$, then (5.12) contradicts (5.3e); if $\mathbf{x} \in \{0,1\}\times\bar{U}$, then (5.12) contradicts (5.3f).

We now introduce our T_ε. We select a smooth nonnegative function k on $\mathbb{R}\times\mathbb{R}^n$ with compact support in $B_1(0)$ and $\iint_{\mathbb{R}\times\mathbb{R}^n} k(\mathbf{z})\,d\mathbf{z} = 1$. For $\varepsilon > 0$, we denote by

$k_\varepsilon(\mathbf{z}) = k(\mathbf{z}/\varepsilon)/\varepsilon^{n+1}$ the rescaled kernel. Given a smooth 1-form $\boldsymbol{\xi}\cdot d\mathbf{x}$ on $\mathbb{R}\times\mathbb{M}^n$ and $\mathbf{z}\in\mathbb{R}\times\mathbb{R}^n$ we consider its pull-back $\boldsymbol{\Phi}(\mathbf{z},\cdot)^\#(\boldsymbol{\xi}\cdot d\mathbf{x}) =: \boldsymbol{\xi}(\mathbf{z},\cdot)\cdot d\mathbf{x}$. Observe that in terms of the vector fields this means

$$\forall\ \mathbf{x}\in\mathbb{R}\times\mathbb{M}^n \quad \boldsymbol{\xi}(\mathbf{z},\mathbf{x}) = \big(D_\mathbf{x}\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)^t\,\boldsymbol{\xi}\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big), \tag{5.13}$$

where A^t denotes the transpose of A with respect to the metric on $\mathbb{R}\times\mathbb{M}^n$. Then we define the smeared out 1-form $\boldsymbol{\xi}_\varepsilon\cdot d\mathbf{x}$ by averaging $\boldsymbol{\xi}(\mathbf{z},\cdot)\cdot d\mathbf{x}$ over $\mathbf{z}$ with respect to k_ε. On the level of the vector fields this means

$$\forall\ \mathbf{x}\in\mathbb{R}\times\mathbb{M}^n \quad \boldsymbol{\xi}_\varepsilon(\mathbf{x}) = \iint_{\mathbb{R}\times\mathbb{R}^n}\boldsymbol{\xi}(\mathbf{z},\mathbf{x})\ \ k_\varepsilon(\mathbf{z})\,d\mathbf{z}. \tag{5.14}$$

Finally, we introduce T_ε by duality, i.e., for all 1-forms $\boldsymbol{\xi}\cdot d\mathbf{x}$ we put

$$\Big\langle T_\varepsilon,\ \boldsymbol{\xi}\cdot d\mathbf{x}\Big\rangle := \Big\langle T,\ \boldsymbol{\xi}_\varepsilon\cdot d\mathbf{x}\Big\rangle. \tag{5.15}$$

Step 2: We first argue that T_ε has a smooth representative in $(0,1)\times U$. In order to see this, we write (5.14) in form of

$$\boldsymbol{\xi}_\varepsilon(\mathbf{x}) = \iint_{(0,1)\times U} K_\varepsilon(\mathbf{y},\mathbf{x})\,\boldsymbol{\xi}(\mathbf{y})\,d\mathbf{y} \quad \text{for } \mathbf{x}\in(0,1)\times U. \tag{5.16}$$

Indeed, we shall see that (5.16) holds for the tensor field

$$K_\varepsilon(\mathbf{y},\mathbf{x}) = \left\{\begin{array}{ll} \Big(-\big(D_\mathbf{y}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\big)^{-1}D_\mathbf{x}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\Big)^t & k_\varepsilon\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\big)\ \det D_\mathbf{y}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}) \\ & \text{for } \mathbf{x}\in(0,1)\times U \\ 0 & \text{otherwise} \end{array}\right\}$$
$$\text{for all } \mathbf{y}\in\big(0,1\big)\times U. \tag{5.17}$$

Notice that $K_\varepsilon(\mathbf{y},\mathbf{x})$ is an endomorphism from the tangent space $T_\mathbf{y}(\mathbb{R}\times\mathbb{M}^n)$ into $T_\mathbf{x}(\mathbb{R}\times\mathbb{M}^n)$, and that $K_\varepsilon(\mathbf{y},\mathbf{x})$ is smooth in $(\mathbf{y},\mathbf{x})\in\big((0,1)\times U\big)\times(\mathbb{R}\times\mathbb{M}^n)$. Indeed, if $\mathbf{y}$ varies in a compact subset of $(0,1)\times U$ and $\mathbf{x}\in(0,1)\times U$ is close to $\partial\big((0,1)\times U\big)$, we learn from (5.11) that $k_\varepsilon\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\big)=0$ and thus $K_\varepsilon(\mathbf{y},\mathbf{x})=0$, because k_ε has bounded support. We check (5.16): For all $\mathbf{x}\in(0,1)\times U$

$$\begin{aligned}
\boldsymbol{\xi}_\varepsilon(\mathbf{x}) &\overset{(5.14)}{=} \iint_{\mathbb{R}\times\mathbb{R}^n}\boldsymbol{\xi}(\mathbf{z},\mathbf{x})\ \ k_\varepsilon(\mathbf{z})\,d\mathbf{z} \\
&\overset{(5.13)}{=} \iint_{\mathbb{R}\times\mathbb{R}^n}\big(D_\mathbf{x}\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)^t\,\boldsymbol{\xi}\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)\ \ k_\varepsilon(\mathbf{z})\,d\mathbf{z} \\
&\overset{(5.5)}{=} \iint_{(0,1)\times U}\Big(D_2\boldsymbol{\Phi}\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x}),\mathbf{x}\big)\Big)^t\,\boldsymbol{\xi}(\mathbf{y})\ \ k_\varepsilon\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\big)\ \det D_\mathbf{y}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\,d\mathbf{y} \\
&\overset{(5.9)}{=} \iint_{(0,1)\times U}\Big(-\big(D_\mathbf{y}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\big)^{-1}D_\mathbf{x}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\Big)^t\,\boldsymbol{\xi}(\mathbf{y})\ \ k_\varepsilon\big(\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\big) \\
&\qquad\qquad\times\det D_\mathbf{y}\boldsymbol{\Psi}(\mathbf{y},\mathbf{x})\,d\mathbf{y} \\
&\overset{(5.17)}{=} \iint_{(0,1)\times U} K_\varepsilon(\mathbf{y},\mathbf{x})\,\boldsymbol{\xi}(\mathbf{y})\,d\mathbf{y}.
\end{aligned}$$

We now argue that in $(0,1)\times U$, T_ε is represented by $\mathbf{m}_\varepsilon$ defined through

$$\mathbf{m}_\varepsilon(\mathbf{y})\cdot\boldsymbol{\xi} := \Big\langle T,\ \big(K_\varepsilon(\mathbf{y},\cdot)\,\boldsymbol{\xi}\big)\cdot d\mathbf{x}\Big\rangle \quad \text{for } \mathbf{y}\in(0,1)\times U,\ \boldsymbol{\xi}\in T_{\mathbf{y}}(\mathbb{R}\times\mathbb{M}^n). \tag{5.18}$$

Since $K_\varepsilon(\mathbf{y},\mathbf{x})$ is smooth in $(\mathbf{y},\mathbf{x})\in\big((0,1)\times U\big)\times(\mathbb{R}\times\mathbb{M}^n)$, $\mathbf{m}_\varepsilon$ is smooth in $\mathbf{y}\in(0,1)\times U$. We check that $\mathbf{m}_\varepsilon$ is indeed the representative of T_ε in $(0,1)\times U$. Let $\boldsymbol{\xi}$ be a smooth vector field compactly supported in $(0,1)\times U$. Then

$$\begin{aligned}
\Big\langle T_\varepsilon,\ \boldsymbol{\xi}\cdot d\mathbf{x}\Big\rangle &\overset{(5.15)}{=} \Big\langle T,\ \boldsymbol{\xi}_\varepsilon\cdot d\mathbf{x}\Big\rangle \\
&\overset{(5.16)}{=} \iint_{(0,1)\times U} \Big\langle T,\ \big(K_\varepsilon(\mathbf{y},\cdot)\,\boldsymbol{\xi}(\mathbf{y})\big)\cdot d\mathbf{x}\Big\rangle\, d\mathbf{y} \\
&\overset{(5.18)}{=} \iint_{(0,1)\times U} \mathbf{m}_\varepsilon(\mathbf{y})\cdot\boldsymbol{\xi}(\mathbf{y})\, d\mathbf{y}.
\end{aligned}$$

Step 3: We now prove that the operator $T\mapsto T_\varepsilon$ does not destroy smoothness. More precisely, we shall argue that for $V\subset\mathbb{M}^n$ open

$$\begin{aligned}
&T \text{ has a smooth representative in } (0,1)\times V \\
&\quad\Longrightarrow\quad T_\varepsilon \text{ has a smooth representative in } (0,1)\times(U\cup V)
\end{aligned} \tag{5.19}$$

and that

$$\begin{aligned}
&T \text{ has a smooth representative in a neighborhood of } \{0,1\}\times\mathbb{M}^n \\
&\quad\Longrightarrow\quad T_\varepsilon \text{ has the same property.}
\end{aligned} \tag{5.20}$$

To treat both situations simultaneously, we consider a set $\mathbf{V}$ that is relatively open in $[0,1]\times\mathbb{M}^n$ and in which T is represented by a smooth vector field $\mathbf{m}$ in the sense that for all smooth vector fields $\boldsymbol{\xi}$ compactly supported in $\mathbf{V}$

$$\Big\langle T,\ \boldsymbol{\xi}\cdot d\mathbf{x}\Big\rangle = \iint_{\mathbb{R}\times\mathbb{M}^n} \mathbf{m}\cdot\boldsymbol{\xi}\, d\mathbf{x}. \tag{5.21}$$

Then we claim that in the set

$$\mathbf{V}_\varepsilon := \bigcap_{\mathbf{z}\in\overline{B_\varepsilon(0)}} \boldsymbol{\Phi}(\mathbf{z},\mathbf{V}) \tag{5.22}$$

the regularized current T_ε is represented by

$$\mathbf{m}_\varepsilon(\mathbf{x}) := \left\{\begin{array}{ll} \displaystyle\iint_{\mathbb{R}\times\mathbb{R}^n} & \big(D_{\mathbf{x}}\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\big)^{-1}\mathbf{m}\big(\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\big)\ \ k_\varepsilon(\mathbf{z})\,\det D_{\mathbf{x}}\boldsymbol{\Theta}(\mathbf{z},\mathbf{x})\,d\mathbf{z} \\ & \qquad\text{for } \mathbf{x}\in\mathbf{V}_\varepsilon\cap\big((0,1)\times\mathbb{M}^n\big) \\ \mathbf{m}(\mathbf{x}) & \qquad\text{for } \mathbf{x}\in\mathbf{V}_\varepsilon\cap\big(\{0,1\}\times\mathbb{M}^n\big) \end{array}\right\}. \tag{5.23}$$

Notice first that $\mathbf{V}_\varepsilon$ is relatively open in $[0,1]\times\mathbb{M}^n$ since (5.3a) & (5.3f) give

$$\mathbf{V}_\varepsilon = \big([0,1]\times\mathbb{M}^n\big) - \boldsymbol{\Phi}\Big(\overline{B_\varepsilon(0)},\ \big([0,1]\times\mathbb{M}^n\big)-\mathbf{V}\Big).$$

According to (5.21), (5.22) & (5.4), the vector field $\mathbf{m}_\varepsilon$ is well defined. Moreover, $\mathbf{m}_\varepsilon$ inherits the smoothness of $\mathbf{m}$ separately in both subsets of $\mathbf{V}_\varepsilon$. Hence we only need to check that $\mathbf{m}_\varepsilon$ is regular throughout $\mathbf{V}_\varepsilon$. By smoothness of $\boldsymbol{\Theta}$ and (5.3f), the function $(s,x) \mapsto \boldsymbol{\Theta}\big(\mathbf{z},(s,x)\big)$ approaches the identity map as $s \to \{0,1\}$, uniformly in all derivatives and both in $\mathbf{z} \in \overline{B_\varepsilon(0)}$ and $x \in \mathbb{M}^n$. This implies in particular that $D_\mathbf{x}\boldsymbol{\Theta} \to \mathrm{Id}$ and $\det D_\mathbf{x}\boldsymbol{\Theta} \to 1$. Since by assumption $\mathbf{m}$ is smooth in $\mathbf{V}$, regularity of $\mathbf{m}_\varepsilon$ then follows easily by standard arguments. Therefore the operator $T \mapsto T_\varepsilon$ does not destroy smoothness in the above sense. We now check that $\mathbf{m}_\varepsilon$ is indeed the representative. Let a smooth vector field $\boldsymbol{\xi}$ be given that is compactly supported in $\mathbf{V}_\varepsilon$. Because of (5.22), $\boldsymbol{\xi}(\mathbf{z},\cdot)$ defined in (5.13) is compactly supported in $\mathbf{V}$ for all $\mathbf{z} \in \overline{B_\varepsilon(0)}$, and so $\boldsymbol{\xi}_\varepsilon$ is compactly supported in $\mathbf{V}$, by definition (5.14). We obtain

$$
\begin{aligned}
&\Big\langle T_\varepsilon,\ \boldsymbol{\xi}\cdot d\mathbf{x}\Big\rangle \\
&\overset{(5.15)}{=} \Big\langle T,\ \boldsymbol{\xi}_\varepsilon\cdot d\mathbf{x}\Big\rangle \\
&\overset{(5.21)}{=} \iint_{\mathbb{R}\times\mathbb{M}^n} \mathbf{m}\cdot\boldsymbol{\xi}_\varepsilon\, d\mathbf{x} \\
&\overset{(5.14)}{=} \iint_{\mathbb{R}\times\mathbb{R}^n} \iint_{\mathbb{R}\times\mathbb{M}^n} \mathbf{m}(\mathbf{x})\cdot\boldsymbol{\xi}(\mathbf{z},\mathbf{x})\, d\mathbf{x}\ \ k_\varepsilon(\mathbf{z})\, d\mathbf{z} \\
&\overset{(5.13)}{=} \iint_{\mathbb{R}\times\mathbb{R}^n} \iint_{\mathbb{R}\times\mathbb{M}^n} D_\mathbf{x}\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\,\mathbf{m}(\mathbf{x})\cdot\boldsymbol{\xi}\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)\, d\mathbf{x}\ \ k_\varepsilon(\mathbf{z})\, d\mathbf{z} \\
&\overset{(5.4)}{=} \iint_{\mathbb{R}\times\mathbb{R}^n} d\mathbf{z}\ k_\varepsilon(\mathbf{z}) \\
&\qquad\times \iint_{\mathbb{R}\times\mathbb{M}^n} D_\mathbf{x}\boldsymbol{\Phi}\big(\mathbf{z},\boldsymbol{\Theta}(\mathbf{z},\mathbf{y})\big)\,\mathbf{m}\big(\boldsymbol{\Theta}(\mathbf{z},\mathbf{y})\big)\cdot\boldsymbol{\xi}(\mathbf{y})\ \det D_\mathbf{y}\boldsymbol{\Theta}(\mathbf{z},\mathbf{y})\, d\mathbf{y} \\
&\overset{(5.8)}{=} \iint_{\mathbb{R}\times\mathbb{R}^n} d\mathbf{z}\ k_\varepsilon(\mathbf{z}) \\
&\qquad\times \iint_{\mathbb{R}\times\mathbb{M}^n} \big(D_\mathbf{y}\boldsymbol{\Theta}(\mathbf{z},\mathbf{y})\big)^{-1}\mathbf{m}\big(\boldsymbol{\Theta}(\mathbf{z},\mathbf{y})\big)\cdot\boldsymbol{\xi}(\mathbf{y})\ \det D_\mathbf{y}\boldsymbol{\Theta}(\mathbf{z},\mathbf{y})\, d\mathbf{y} \\
&\overset{(5.23)}{=} \iint_{\mathbf{V}_\varepsilon} \mathbf{m}_\varepsilon(\mathbf{y})\cdot\boldsymbol{\xi}(\mathbf{y})\, d\mathbf{y}.
\end{aligned}
$$

If now $\mathbf{V} = (0,1)\times V$ with $V \subset \mathbb{M}^n$ open, then (5.3d)–(5.3f) entail

$$
\mathbf{V}_\varepsilon - \big((0,1)\times U\big) = \mathbf{V} - \big((0,1)\times U\big).
$$

In particular, $\mathbf{V}_\varepsilon$ contains an open neighborhood of $\mathbf{V}\cap\big((0,1)\times\partial U\big)$. Therefore T_ε is smooth in $\mathbf{V}\cup\big((0,1)\times U\big)$. This establishes (5.19). Similarly, if $\mathbf{V} = [0,\alpha)\times\mathbb{M}^n$ for some $0<\alpha<1$, then $\mathbf{V}_\varepsilon = [0,\alpha')\times\mathbb{M}^n$ for some $0<\alpha'\leqslant\alpha$, by (5.3d) & (5.3f). Therefore T_ε is smooth up to the boundary $\{0\}\times\mathbb{M}^n$. The same argument applies to $\mathbf{V} = (1-\alpha,1]\times\mathbb{M}^n$. This establishes (5.20).

Step 4: We now argue that T_ε is admissible if T is, i.e., if

$$
\Big\langle T,\ \boldsymbol{\nabla}\zeta\cdot d\mathbf{x}\Big\rangle = \int_{\mathbb{M}^n} \zeta(1,x)\, d\rho_1(x) - \int_{\mathbb{M}^n} \zeta(0,x)\, d\rho_0(x) \tag{5.24}
$$

for all smooth functions ζ on $\mathbb{R}\times\mathbb{M}^n$. Consider the gradient field $\boldsymbol{\xi} := \boldsymbol{\nabla}\zeta$. We gather from (5.13) and the chain rule that

$$
\boldsymbol{\xi}(\mathbf{z},\mathbf{x}) = \boldsymbol{\nabla}_\mathbf{x}\zeta(\mathbf{z},\mathbf{x}) \quad \text{where} \quad \zeta(\mathbf{z},\mathbf{x}) := \zeta\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big).
$$

We thus infer from (5.14) that

$$\boldsymbol{\xi}_\varepsilon(\mathbf{x}) = \boldsymbol{\nabla}\zeta_\varepsilon(\mathbf{x}) \quad \text{where} \quad \zeta_\varepsilon(\mathbf{x}) := \iint_{\mathbb{R}\times\mathbb{R}^n} \zeta(\mathbf{z},\mathbf{x})\ \ k_\varepsilon(\mathbf{z})\, d\mathbf{z}. \tag{5.25}$$

Then (5.3f) implies that

$$\forall\ (\mathbf{z},\mathbf{x}) \in (\mathbb{R}\times\mathbb{R}^n) \times \big((\mathbb{R}-(0,1)) \times \mathbb{M}^n\big) \quad \left\{ \begin{array}{c} \zeta(\mathbf{z},\mathbf{x}) = \zeta(\mathbf{x}) \\ \text{and thus} \\ \zeta_\varepsilon(\mathbf{x}) = \zeta(\mathbf{x}) \end{array} \right\}. \tag{5.26}$$

Hence we obtain as desired

$$\begin{aligned} \Big\langle T_\varepsilon,\ \boldsymbol{\nabla}\zeta \cdot d\mathbf{x} \Big\rangle &\overset{(5.15)}{\underset{(5.25)}{=}} \Big\langle T,\ \boldsymbol{\nabla}\zeta_\varepsilon \cdot d\mathbf{x} \Big\rangle \\ &\overset{(5.24)}{=} \int_{\mathbb{M}^n} \zeta_\varepsilon(1,x)\, d\rho_1(x) - \int_{\mathbb{M}^n} \zeta_\varepsilon(0,x)\, d\rho_0(x) \\ &\overset{(5.26)}{=} \int_{\mathbb{M}^n} \zeta(1,x)\, d\rho_1(x) - \int_{\mathbb{M}^n} \zeta(0,x)\, d\rho_0(x). \end{aligned}$$

Step 5: Now we address the action estimate. We claim that for small ε

$$\mathcal{A}(T_\varepsilon) \leqslant \big(1 + \mathcal{O}(\varepsilon)\big)\, \mathcal{A}(T) + \mathcal{O}(\varepsilon) \tag{5.27}$$

with the modulus $\mathcal{O}(\varepsilon)$ only depending on $\boldsymbol{\Phi}$. Let $\boldsymbol{\xi} = (\sigma, \xi)$ be an admissible vector field in the definition of $\mathcal{A}$, i.e., for which

$$\sigma + \tfrac{1}{2}|\xi|^2 \leqslant 0. \tag{5.28}$$

Consider $\boldsymbol{\xi}(\mathbf{z},\mathbf{x}) = \big(\sigma(\mathbf{z},\mathbf{x}), \xi(\mathbf{z},\mathbf{x})\big)$ defined in (5.13). We will then show that the modified vector field

$$\begin{gathered} \Big(\lambda(\mathbf{z})\, \big(\sigma(\mathbf{z},\cdot) - \mu(\mathbf{z})\big),\ \lambda(\mathbf{z})\, \xi(\mathbf{z},\cdot)\Big) \ \text{ is admissible} \\ \text{for suitable constants } \lambda(\mathbf{z}) = 1 - \mathcal{O}(|\mathbf{z}|),\ \mu(\mathbf{z}) = O(|\mathbf{z}|). \end{gathered} \tag{5.29}$$

Indeed, the anisotropy condition (5.3d) on $\boldsymbol{\Phi} = (\Phi_0, \Phi)$ and (5.13) give

$$\begin{gathered} \sigma(\mathbf{z},\mathbf{x}) = \partial_s\Phi_0(\mathbf{z},s)\, \sigma\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big) + \partial_s\Phi(\mathbf{z},\mathbf{x}) \cdot \xi\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big), \\ \xi(\mathbf{z},\mathbf{x}) = \big(D_x\Phi(\mathbf{z},\mathbf{x})\big)^t\, \xi\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big). \end{gathered}$$

Because of (5.10) this yields the estimates

$$\begin{gathered} \sigma(\mathbf{z},\mathbf{x}) \leqslant \big(1 - \mathcal{O}(|\mathbf{z}|)\big)\, \sigma\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big) + \mathcal{O}(|\mathbf{z}|)\, \big|\xi\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)\big|, \\ |\xi(\mathbf{z},\mathbf{x})| \leqslant \big(1 + \mathcal{O}(|\mathbf{z}|)\big)\, \big|\xi\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)\big|. \end{gathered} \tag{5.30}$$

Using Young's inequality

$$2\lambda(\mathbf{z})\, \big|\xi\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)\big| \leqslant 1 + \lambda(\mathbf{z})^2\, \big|\xi\big(\boldsymbol{\Phi}(\mathbf{z},\mathbf{x})\big)\big|^2, \tag{5.31}$$

we notice that the latter implies

$$\begin{aligned}
&\lambda(\mathbf{z})\,\big(\sigma(\mathbf{z},\mathbf{x})-\mu(\mathbf{z})\big)+\tfrac{1}{2}|\lambda(\mathbf{z})\,\xi(\mathbf{z},\mathbf{x})|^2\\
&\overset{(5.30)}{\leqslant}\ \lambda(\mathbf{z})\,\Big(\big(1-\mathcal{O}(|\mathbf{z}|)\big)\,\sigma\big(\mathbf{\Phi}(\mathbf{z},\mathbf{x})\big)+\mathcal{O}(|\mathbf{z}|)\,\big|\xi\big(\mathbf{\Phi}(\mathbf{z},\mathbf{x})\big)\big|-\mu(\mathbf{z})\Big)\\
&\qquad+\lambda(\mathbf{z})^2\,\big(1+\mathcal{O}(|\mathbf{z}|)\big)^2\,\tfrac{1}{2}\big|\xi\big(\mathbf{\Phi}(\mathbf{z},\mathbf{x})\big)\big|^2\\
&\overset{(5.31)}{\leqslant}\ \lambda(\mathbf{z})\,\big(1-\mathcal{O}(|\mathbf{z}|)\big)\,\sigma\big(\mathbf{\Phi}(\mathbf{z},\mathbf{x})\big)+\lambda(\mathbf{z})^2\,\big(1+\mathcal{O}(|\mathbf{z}|)\big)\,\tfrac{1}{2}\big|\xi\big(\mathbf{\Phi}(\mathbf{z},\mathbf{x})\big)\big|^2\\
&\qquad+\mathcal{O}(|\mathbf{z}|)-\lambda(\mathbf{z})\,\mu(\mathbf{z})\\
&\overset{(5.28)}{\leqslant}\ \lambda(\mathbf{z})\,\Big(-\big(1-\mathcal{O}(|\mathbf{z}|)\big)+\big(1+\mathcal{O}(|\mathbf{z}|)\big)\,\lambda(\mathbf{z})\Big)\,\tfrac{1}{2}\big|\xi\big(\mathbf{\Phi}(\mathbf{z},\mathbf{x})\big)\big|^2\\
&\qquad+\mathcal{O}(|\mathbf{z}|)-\lambda(\mathbf{z})\,\mu(\mathbf{z})
\end{aligned}$$

which in turn yields (5.29).

Choosing $\zeta(s,x)=s$ in the admissibility condition (5.24) yields

$$\Big\langle T,\ 1\,ds\Big\rangle=1. \tag{5.32}$$

Thus we have by definition of $\mathcal{A}(T)$:

$$\begin{aligned}
&\Big\langle T,\ \sigma(\mathbf{z},\cdot)\,ds+\xi(\mathbf{z},\cdot)\cdot dx\Big\rangle\\
&\quad=\ \frac{1}{\lambda(\mathbf{z})}\Big\langle T,\ \lambda(\mathbf{z})\,\big(\sigma(\mathbf{z},\cdot)-\mu(\mathbf{z})\big)\,ds+\lambda(\mathbf{z})\,\xi(\mathbf{z},\cdot)\cdot dx\Big\rangle+\mu(\mathbf{z})\Big\langle T,\ 1\,ds\Big\rangle\\
&\overset{(5.29)}{\underset{(5.32)}{\leqslant}}\ \frac{1}{\lambda(\mathbf{z})}\,\mathcal{A}(T)+\mu(\mathbf{z})\\
&\overset{(5.29)}{=}\ \big(1+\mathcal{O}(|\mathbf{z}|)\big)\,\mathcal{A}(T)+\mathcal{O}(|\mathbf{z}|).
\end{aligned} \tag{5.33}$$

We therefore obtain as desired

$$\begin{aligned}
\Big\langle T_\varepsilon,\ \sigma\,ds+\xi\cdot dx\Big\rangle &\overset{(5.15)}{=}\Big\langle T,\ \sigma_\varepsilon\,ds+\xi_\varepsilon\cdot dx\Big\rangle\\
&\overset{(5.14)}{=}\iint_{\mathbb{R}\times\mathbb{R}^n}\Big\langle T,\ \sigma(\mathbf{z},\cdot)\,ds+\xi(\mathbf{z},\cdot)\cdot dx\Big\rangle\ k_\varepsilon(\mathbf{z})\,d\mathbf{z}\\
&\overset{(5.33)}{=}\big(1+\mathcal{O}(\varepsilon)\big)\mathcal{A}(T)+\mathcal{O}(\varepsilon).
\end{aligned}$$

Since (σ,ξ) was arbitrary with (5.28), this yields (5.27) by definition of $\mathcal{A}(T_\varepsilon)$.

Step 6: Let T_ε have a smooth representative $(\rho_\varepsilon,m_\varepsilon)$ in $(0,1)\times\mathbb{M}^n$ and satisfy $\mathcal{A}(T_\varepsilon)<\infty$. We now argue that $\rho_\varepsilon\geqslant 0$. More precisely, we shall show that

$$\Big\langle T_\varepsilon,\ \zeta\,ds\Big\rangle\geqslant 0\quad\text{for all smooth test function }\zeta\text{ with }\zeta\geqslant 0. \tag{5.34}$$

Indeed, for $n\in\mathbb{N}$ the vector field $\boldsymbol{\xi}=(-n\zeta,0)$ is admissible and yields

$$-n\Big\langle T_\varepsilon,\ \zeta\,ds\Big\rangle=\Big\langle T_\varepsilon,\ \boldsymbol{\xi}\cdot d\mathbf{x}\Big\rangle\leqslant\mathcal{A}(T_\varepsilon),$$

which gives (5.34) in the limit $n\to\infty$. By (5.2), this proves that $\rho_\varepsilon\geqslant 0$.

Step 7: It remains to construct the map $\mathbf{\Phi} = (\Phi_0, \Phi)$. This is done in a series of short steps. Starting point is the diffeomorphism $h_0\colon (0,1) \longrightarrow \mathbb{R}$ defined by

$$h_0(s) = \big(s - \tfrac{1}{2}\big) \exp\bigg(\frac{1}{s(1-s)}\bigg).$$

Next we introduce the map $\Phi_0\colon \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$:

$$\Phi_0(u,s) = \left\{ \begin{array}{ll} h_0^{-1}\big(h_0(s) + u\big) & \text{for } s \in (0,1) \\ s & \text{otherwise} \end{array} \right\}. \tag{5.35}$$

The properties of the exponential function imply that Φ_0 is smooth. In particular, we have by the inverse function theorem that

$$\partial_u \Phi_0(u,s) = \left\{ \begin{array}{ll} \Big(h_0'\big(\Phi_0(u,s)\big)\Big)^{-1} & \text{for } s \in (0,1) \\ 0 & \text{otherwise} \end{array} \right\},$$

$$\partial_s \Phi_0(u,s) = \left\{ \begin{array}{ll} \Big(h_0'\big(\Phi_0(u,s)\big)\Big)^{-1} h_0'(s) & \text{for } s \in (0,1) \\ 1 & \text{otherwise} \end{array} \right\}.$$

The map Φ_0 obviously has the homomorphism property

$$\begin{gathered} \forall\, u, u' \in \mathbb{R},\ s \in \mathbb{R} \quad \Phi_0(u + u', s) = \Phi_0\big(u, \Phi_0(u', s)\big), \\ \forall\, s \in \mathbb{R} \quad \Phi_0(0, s) = s. \end{gathered} \tag{5.36}$$

In a similar way, we introduce the diffeomorphism $\tilde{h}\colon B_1(0) \to \mathbb{R}^n$ defined by

$$\tilde{h}(\tilde{x}) = \tilde{x} \exp\bigg(\frac{1}{1 - |\tilde{x}|^2}\bigg),$$

and the map $\tilde{\Phi}\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ through

$$\tilde{\Phi}(z, \tilde{x}) = \left\{ \begin{array}{ll} \tilde{h}^{-1}\big(\tilde{h}(\tilde{x}) + z\big) & \text{for } \tilde{x} \in B_1(0) \\ \tilde{x} & \text{otherwise} \end{array} \right\}. \tag{5.37}$$

Again, the properties of the exponential function imply that $\tilde{\Phi}$ is smooth, and

$$D_z \tilde{\Phi}(z, \tilde{x}) = \left\{ \begin{array}{ll} \Big(D\tilde{h}\big(\tilde{\Phi}(z, \tilde{x})\big)\Big)^{-1} & \text{for } \tilde{x} \in B_1(0) \\ 0 & \text{otherwise} \end{array} \right\},$$

$$D_{\tilde{x}} \tilde{\Phi}(z, \tilde{x}) = \left\{ \begin{array}{ll} \Big(D\tilde{h}\big(\tilde{\Phi}(z, \tilde{x})\big)\Big)^{-1} D\tilde{h}(\tilde{x}) & \text{for } \tilde{x} \in B_1(0) \\ \mathrm{Id} & \text{otherwise} \end{array} \right\}.$$

As for Φ_0 above, the map $\tilde{\Phi}$ has the homomorphism property

$$\begin{gathered} \forall\, z, z' \in \mathbb{R}^n,\ \tilde{x} \in \mathbb{R}^n \quad \tilde{\Phi}(z + z', \tilde{x}) = \tilde{\Phi}\big(z, \tilde{\Phi}(z', \tilde{x})\big), \\ \forall\, \tilde{x} \in \mathbb{R}^n \quad \tilde{\Phi}(0, \tilde{x}) = \tilde{x}. \end{gathered} \tag{5.38}$$

Recall that $U \subset \mathbb{M}^n$ is an open subset homeomorphic to the ball $B_1(0) \subset \mathbb{R}^n$, with $\phi\colon U \longrightarrow \mathbb{R}^n$ a coordinate map. We may assume that ϕ extends to some neighborhood of U and $B_1(0)$. Then the composition

$$h(x) := \tilde{h}\big(\phi(x)\big) \quad \text{for all } x \in U$$

defines a diffeomorphism $h\colon U \longrightarrow \mathbb{R}^n$. In view of (5.37),

$$\Phi(z,x) = \left\{ \begin{array}{ll} \phi^{-1}\Big(\tilde{\Phi}\big(z,\phi(x)\big)\Big) & \text{for } x \in U \\ x & \text{otherwise} \end{array} \right\}$$

defines a smooth map $\Phi\colon \mathbb{R}^n \times \mathbb{M}^n \longrightarrow \mathbb{M}^n$, and it is immediately clear that the properties (5.37) and (5.38) are conserved, i.e., we have

$$\Phi(z,x) = \left\{ \begin{array}{ll} h^{-1}\big(h(x)+z\big) & \text{for } x \in U \\ x & \text{otherwise} \end{array} \right\} \tag{5.39}$$

and

$$\begin{gathered} \forall\, z, z' \in \mathbb{R}^n,\ x \in \mathbb{M}^n \quad \Phi(z+z',x) = \Phi\big(z,\Phi(z',x)\big), \\ \forall\, x \in \mathbb{M}^n \quad \Phi(0,x) = x. \end{gathered} \tag{5.40}$$

Finally, we consider the smooth function $\delta\colon \mathbb{R} \longrightarrow \mathbb{R}$ defined through

$$\delta(s) = \left\{ \begin{array}{ll} \exp\left(-\dfrac{1}{s(1-s)}\right) & \text{for } s \in (0,1) \\ 0 & \text{otherwise} \end{array} \right\}. \tag{5.41}$$

We now introduce the smooth map $\mathbf{\Phi}\colon (\mathbb{R}\times\mathbb{R}^n)\times(\mathbb{R}\times\mathbb{M}^n) \longrightarrow (\mathbb{R}\times\mathbb{M}^n)$:

$$\mathbf{\Phi}\big((u,z),(s,x)\big) := \Big(\Phi_0(u,s), \Phi\big(\delta(s)z,x\big)\Big).$$

Let us quickly check that $\mathbf{\Phi}$ has the required properties. We establish the diffeomorphism properties (5.3a) by explicitly giving the right inverse

$$\mathbf{\Theta}\big((u,z),(s,x)\big) = \Big(\Phi_0(-u,s), \Phi\big(-\delta\big(\Phi_0(-u,s)\big)z,x\big)\Big)$$

which defines a smooth map $\mathbf{\Theta}\colon (\mathbb{R}\times\mathbb{R}^n)\times(\mathbb{R}\times\mathbb{M}^n) \longrightarrow (\mathbb{R}\times\mathbb{M}^n)$. Then (5.36) & (5.40) imply (5.4) and (5.3b). The anisotropy property (5.3d) is clear by construction. The first invariance property (5.3e) can be read off from (5.39); the second invariance property (5.3f) follows from (5.35) for the Φ_0–component and from (5.41) combined with (5.40) for the Φ–component. The second diffeomorphism property (5.3c) follows from an explicit formula for the left inverse

$$\mathbf{\Psi}\big((t,y),(s,x)\big) = \left(h_0(t) - h_0(s), \frac{h(y)-h(x)}{\delta(s)}\right)$$

which defines a smooth map $\mathbf{\Psi}\colon \big((0,1)\times U\big)\times\big((0,1)\times U\big) \longrightarrow \mathbb{R}\times\mathbb{R}^n$ (recall that δ is positive on $(0,1)$). The identities (5.35) & (5.39) imply (5.5) as desired.

Step 8: In this last step, we prove that if ρ_0 and ρ_1 are smooth functions and if T is a current defined on $[0,1]\times\mathbb{M}^n$ satisfying the admissibility condition (5.1), then the regularization T_ε of T is representable by a vector field $(\rho_\varepsilon \geqslant 0, m_\varepsilon)$ that is smooth up to the boundary $\{0,1\} \times \mathbb{M}^n$. This can be achieved by first approximating T by an admissible current T_α that is representable by a smooth vector field in $[0,\alpha) \times \mathbb{M}^n$ and $(1-\alpha, 1] \times \mathbb{M}^n$ for suitable α, and then applying to T_α the regularization procedure described above. Since the operators $T \mapsto T_{\varepsilon_i}^{U_i}$ do not destroy smoothness in stripes around the boundary $\{0,1\} \times \mathbb{M}^n$ (see Step 3) the regularized current $(T_\alpha)_\varepsilon$ can be represented by a vector field that attains the data ρ_0 and ρ_1 smoothly as desired.

We proceed as follows: For $0 < \alpha < \frac{1}{2}$ we consider the map

$$\mathbf{\Phi}_\alpha(s,x) := \big(\alpha + (1-2\alpha)s, x\big) \quad \text{for } (s,x) \in [0,1] \times \mathbb{M}^n.$$

Given a smooth 1-form $\boldsymbol{\xi} \cdot d\mathbf{x}$, let $\mathbf{\Phi}_\alpha^{\#}(\boldsymbol{\xi} \cdot d\mathbf{x}) =: \boldsymbol{\xi}_\alpha \cdot d\mathbf{x}$ be its pull-back under the map $\mathbf{\Phi}_\alpha$. In terms of the vector field $\boldsymbol{\xi} = (\sigma, \xi)$ this means

$$\forall\, \mathbf{x} \in [0,1] \times \mathbb{M}^n \quad \boldsymbol{\xi}_\alpha(\mathbf{x}) = \big(D_{\mathbf{x}} \mathbf{\Phi}_\alpha(\mathbf{x})\big)^t \boldsymbol{\xi}\big(\mathbf{\Phi}_\alpha(\mathbf{x})\big), \tag{5.42}$$

hence $\boldsymbol{\xi}_\alpha = (\sigma_\alpha, \xi_\alpha)$ with

$$\sigma_\alpha(\mathbf{x}) = (1-2\alpha)\, \sigma\big(\mathbf{\Phi}_\alpha(\mathbf{x})\big) \quad \text{and} \quad \xi_\alpha(\mathbf{x}) = \xi\big(\mathbf{\Phi}_\alpha(\mathbf{x})\big). \tag{5.43}$$

We define the approximate current T_α by duality as

$$\begin{aligned} \Big\langle T_\alpha,\ \boldsymbol{\xi} \cdot d\mathbf{x} \Big\rangle := \Big\langle T,\ \boldsymbol{\xi}_\alpha \cdot d\mathbf{x} \Big\rangle &+ \iint_{[0,\alpha] \times \mathbb{M}^n} \sigma(s,x)\, \rho_0(x)\, dx\, ds \\ &+ \iint_{[1-\alpha,1] \times \mathbb{M}^n} \sigma(s,x)\, \rho_1(x)\, dx\, ds \end{aligned} \tag{5.44}$$

for all smooth 1-forms $\boldsymbol{\xi} \cdot d\mathbf{x} = \sigma\, ds + \xi \cdot dx$. This T_α is admissible: If ζ is a smooth function and $\boldsymbol{\xi} := \boldsymbol{\nabla}\zeta$ the gradient field, then (5.42) implies that

$$\boldsymbol{\xi}_\alpha(\mathbf{x}) = \boldsymbol{\nabla}\zeta_\alpha(\mathbf{x}) \quad \text{where} \quad \zeta_\alpha(\mathbf{x}) := \zeta\big(\mathbf{\Phi}_\alpha(\mathbf{x})\big),$$

and therefore

$$\begin{aligned} \Big\langle T_\alpha,\ \boldsymbol{\nabla}\zeta \cdot d\mathbf{x} \Big\rangle &= \int_{\mathbb{M}^n} \zeta(1-\alpha, x)\, \rho_1(x)\, dx + \iint_{[1-\alpha,1] \times \mathbb{M}^n} \partial_s \zeta(s,x)\, \rho_1(x)\, dx\, ds \\ &\quad - \int_{\mathbb{M}^n} \zeta(\alpha, x)\, \rho_0(x)\, dx + \iint_{[0,\alpha] \times \mathbb{M}^n} \partial_s \zeta(s,x)\, \rho_0(x)\, dx\, ds \\ &= \int_{\mathbb{M}^n} \zeta(1,x)\, \rho_1(x)\, dx - \int_{\mathbb{M}^n} \zeta(0,x)\, \rho_0(x)\, dx. \end{aligned}$$

It follows easily from (5.44) that T_α is represented by the smooth vector field $(\rho_0, 0)$ in the stripe $[0,\alpha) \times \mathbb{M}^n$ and by $(\rho_1, 0)$ in $(1-\alpha, 1] \times \mathbb{M}^n$. Now we consider the action: Notice first that if $\boldsymbol{\xi} = (\sigma, \xi)$ is admissible in the definition of $\mathcal{A}$, then also the modified vector field $(1-2\alpha)\, \boldsymbol{\xi}_\alpha$ is admissible since by (5.43)

$$(1-2\alpha)\sigma_\alpha + \tfrac{1}{2}|(1-2\alpha)\xi_\alpha|^2 = (1-2\alpha)^2 \Big(\sigma \circ \mathbf{\Phi}_\alpha + \tfrac{1}{2}|\xi \circ \mathbf{\Phi}_\alpha|^2\Big) \leqslant 0.$$

The extra integrals in (5.44) do not contribute to $\mathcal{A}(T_\alpha)$ because $\sigma \leqslant -\frac{1}{2}|\xi|^2 \leqslant 0$ and $\rho_0, \rho_1 \geqslant 0$. This yields as above $\mathcal{A}(T_\alpha) \leqslant (1-2\alpha)^{-1} \mathcal{A}(T)$, and thus

$$\limsup_{\alpha \to 0} \mathcal{A}(T_\alpha) \leqslant \mathcal{A}(T).$$

Now we regularize as before to conclude. $\square$

REFERENCES

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[2] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, vol. 10 of Panoramas et Synthèses, Société Mathématique de France, Paris, 2000.

[3] D. Bakry and M. Émery, *Diffusions hypercontractives*, in Séminaire de probabilités XIX, vol. 1123 of Lecture Notes in Math., Springer, Berlin, 1985.

[4] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.

[5] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.

[6] J. A. Carrillo, R. J. McCann, and C. Villani, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Mat. Iberoamericana, 19 (2003), pp. 971–1018.

[7] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math., 146 (2001), pp. 219–257.

[8] G. de Rham, *Variétés différentiables*, Hermann et Cie, Paris, 1955.

[9] H. Federer, *Geometric measure theory*, Springer, Berlin, 1969.

[10] G. Giaquinta, A. Modica, and J. Souček, *Cartesian currents in the calculus of variations*, Springer, Berlin, 1998.

[11] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math., 128 (1997), pp. 153–179.

[12] ———, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal., 11 (2001), pp. 589–608.

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

[14] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173 (2000), pp. 361–400.

[15] G. Petersen, *Riemannian Geometry*, Springer, New York, 1998.

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

[17] K.-Th. Sturm and M.-K. von Renesse, *Transport inequalities, gradient estimate, entropy and Ricci curvature*, SFB 611-Preprint, 80 (2003). Appears in: Comm. Pure Appl. Math.

[18] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, 2003.