

## GRADIENT FLOWS FOR SEMICONVEX FUNCTIONS ON METRIC MEASURE SPACES – EXISTENCE, UNIQUENESS, AND LIPSCHITZ CONTINUITY

KARL-THEODOR STURM

(Communicated by Guofang Wei)

ABSTRACT. Given any continuous, lower bounded, and  $\kappa$ -convex function  $V$  on a metric measure space  $(X, d, m)$  which is infinitesimally Hilbertian and satisfies some synthetic lower bound for the Ricci curvature in the sense of Lott-Sturm-Villani, we prove existence and uniqueness for the (downward) gradient flow for  $V$ . Moreover, we prove Lipschitz continuity of the flow w.r.t. the starting point

$$d(x_t, x'_t) \leq e^{-\kappa t} d(x_0, x'_0).$$

Throughout this paper, let  $(X, d)$  be a locally compact geodesic space and let  $m$  be a locally finite Borel measure with full topological support. We always assume that the metric measure space  $(X, d, m)$  satisfies the  $\text{RCD}(K, \infty)$ -condition for some finite number  $K \in \mathbb{R}$ . Recall that this means that  $\text{Ent}(\cdot|m)$ , the Boltzmann entropy w.r.t.  $m$ , is weakly  $K$ -convex on the geodesic spaces  $(\mathcal{P}_2(X), W_2)$  and that  $\mathcal{E}$ , the Cheeger energy on  $(X, d, m)$ , is a quadratic functional on  $L^2(X, m)$ , [9], [10], [5], [3], [4]. Among many others, it implies that  $\int_X e^{-C' d^2(x, x')} m(dx) < \infty$  for some  $C' \in \mathbb{R}$ ,  $x' \in X$ . Note that the latter in turn implies  $\text{Ent}(\mu|m) \geq -C'' - C' W_2^2(\mu, \delta_{x'})$  for each  $\mu \in \mathcal{P}_2(X)$ .

Here and henceforth  $W_2$  will denote the 2-Kantorovich–Wasserstein distance and  $\mathcal{P}_2(X)$  the 2-Wasserstein space consisting of those Borel probability measures  $\mu$  on  $X$  which satisfy  $W_2(\mu, \delta_{x'}) < \infty$  (for some, hence all,  $x' \in X$ ). Recall that  $\text{Ent}(\mu|m) = \int \rho \log \rho dm$  provided  $\mu = \rho m$  is absolutely continuous w.r.t.  $m$  and  $\int \rho(\log \rho)_+ dm < \infty$ . Otherwise  $\text{Ent}(\mu|m) = +\infty$ . Also recall that  $\mathcal{E}(u) = \liminf \left\{ \int_X |Du_n|^2 dm : u_n \in \text{Lip}(X), u_n \rightarrow u \text{ in } L^2(X, m) \right\}$  with  $|Du|(x) = \limsup_{y \rightarrow x} \frac{|u(x) - u(y)|}{d(x, y)}$ .

Moreover, let a function  $V : X \rightarrow (-\infty, +\infty]$  be given which is continuous (w.r.t. the topology of the extended real line) and bounded from below by  $-C_0 - C_1 d^2(\cdot, x_0)$  for some  $C_0, C_1 \in \mathbb{R}$ ,  $x_0 \in X$ . Put  $X_0 = \{V < \infty\}$  and  $X' = \overline{X_0}$ .

---

Received by the editors August 9, 2017, and, in revised form, December 13, 2017.

2010 *Mathematics Subject Classification*. Primary 49Q15, 53C21, 58E35.

The author gratefully acknowledges support by the European Union through the ERC Advanced Grant “Metric measure spaces and Ricci curvature - analytic, geometric, and probabilistic challenges” (“RicciBounds”) as well as support by the German Research Foundation through the Hausdorff Center for Mathematics and the Collaborative Research Center 1060 “The Mathematics of Emergent Effects”.

**Definition.**

- (i) Given a number  $\kappa \in \mathbb{R}$ , we say that  $V$  is weakly  $\kappa$ -convex if for each  $x_0, x_1 \in X$  there exists a geodesic  $\gamma : [0, 1] \rightarrow X$  connecting them (i.e.,  $\gamma_0 = x_0, \gamma_1 = x_1$ ) such that

$$(1) \quad V(\gamma_t) \leq (1 - t)V(\gamma_0) + tV(\gamma_1) - \kappa t(1 - t)|\dot{\gamma}|^2/2$$

for each  $t \in [0, 1]$ . If (1) holds for all geodesics  $\gamma$  in  $X$  and all  $t \in [0, 1]$ , then  $V$  is called strongly  $\kappa$ -convex.

- (ii) Given  $x_0 \in X$ , an  $\text{EVI}_\kappa$ -gradient flow for  $V$  starting in  $x_0$  is an absolutely continuous curve  $(x_t)_{t \geq 0}$  in  $X$  emanating in  $x_0$  such that for all  $z \in X_0$  and a.e.  $t > 0$ ,

$$(2) \quad \frac{d}{dt} \frac{1}{2} d^2(x_t, z) + \kappa d^2(x_t, z) \leq V(z) - V(x_t).$$

Note that here and in the sequel—as common in metric geometry—a geodesic is always minimizing and parametrized by arc length.

The main result of this paper is the following.

**Theorem.** *If  $V$  is weakly  $\kappa$ -convex, then for each  $x_0 \in X'$  there exists a unique  $\text{EVI}_\kappa$ -gradient flow for  $V$  starting in  $x_0$ .*

It is well known—and easy to see—that the existence of an  $\text{EVI}_\kappa$ -gradient flow for  $V$  starting in  $x_0$  implies its uniqueness and the “contraction” property (or “expansion” bound)

$$(3) \quad d(x_t, x'_t) \leq e^{-\kappa t} d(x_0, x'_0)$$

for each pair of points  $x_0, x'_0 \in X'$  and associated flows  $(x_t)_{t \geq 0}$  and  $(x'_t)_{t \geq 0}$ . Moreover, every  $\text{EVI}_\kappa$ -gradient flow is also a gradient flow in the sense of energy dissipation

$$(4) \quad V(x_t) = V(x_s) - \frac{1}{2} \int_s^t (|\dot{x}_r|^2 + |\nabla^- V|^2(x_r)) dr$$

for all  $0 \leq s \leq t$ , where as usual  $|\nabla^- V|(x) := \limsup_{y \rightarrow x} \frac{V(x) - V(y)_+}{d(x, y)}$ . In particular,  $t \mapsto V(x_t)$  is nonincreasing.

**Corollary 1.** *For each  $x_0 \in X'$  there exists a unique curve  $(x_t)_{t \geq 0}$  such that  $\mu_t = \delta_{x_t}$  is the  $\text{EVI}_\kappa$ -gradient flow for  $S$  starting in  $\mu_0 = \delta_{x_0}$ .*

*The curve  $(x_t)_{t \geq 0}$  is the unique  $\text{EVI}_\kappa$ -gradient flow for  $V$  starting in  $x_0$ .*

*Proof.* Put  $\mu_0 = \delta_{x_0}$ . Then  $\text{Var}_2(\mu_0) = 0$  and thus, according to the previous lemma, also  $\text{Var}_2(\mu_t) = 0$ . This implies that each of the measures  $\mu_t$  is supported by a single point, say  $x_t$ .

Moreover, the fact that the curve  $(\delta_{x_t})_{t \geq 0}$  in  $\mathcal{P}_2(X')$  is the  $\text{EVI}_\kappa$ -gradient flow for  $S$  starting in  $\mu_0 = \delta_{x_0}$  immediately implies that the curve  $(x_t)_{t \geq 0}$  in  $X'$  is the (unique)  $\text{EVI}_\kappa$ -gradient flow for  $V$  starting in  $x_0$ . (To verify (2), just apply (7) to Diracs  $\nu = \delta_z$ .) □

**Corollary 2.** *The following properties are equivalent:*

- (i)  $V$  is weakly  $\kappa$ -convex;
- (ii)  $V$  is strongly  $\kappa$ -convex;

- (iii)  $V$  is  $\kappa$ -convex in EVI-sense, i.e., for each  $x_0 \in X'$  there exists a curve  $(x_t)_{t \geq 0}$  in  $X'$  such that for all  $z \in X_0$  and all  $t > 0$ ,

$$\frac{d}{dt} \frac{1}{2} d^2(x_t, z) + \kappa d^2(x_t, z) \leq V(z) - V(x_t).$$

Note that the implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) hold true in any geodesic space; see [4].

*Remark.* Our main result extends previous results in various respects.

In the setting of Alexandrov spaces (i.e., metric spaces with synthetic lower bounds for the sectional curvature), the existence, uniqueness, and Lipschitz continuity of gradient flows for semiconvex functions has been derived in an unpublished manuscript by Perelman and Petrunin [7]. The proof has been worked out and published by Plaut [8] and Lytchak [6]. This result requires no local compactness of the state space. The function  $V$ , however, is always required to be bounded and Lipschitz continuous.

In the setting of Hilbert spaces (and to some extent also for metric spaces), a far reaching theory of gradient flows for lower semicontinuous functions (with suitable lower growth bounds and compact sublevel sets) has been developed by Brezis as well as by De Giorgi and his school; see in particular the monograph [1] by Ambrosio-Gigli-Savaré. In this context, however, the state space  $X$  has to be flat (or of nonpositive curvature).

In the setting of metric measure spaces with synthetic Ricci bounds, currently no direct proof—without passing to the Wasserstein space but instead using more detailed properties of the Hessian of  $V$ ; cf. [11]—is known to the author.

The Theorem will be proven according to the following line of argumentation in the subsequent lemmata:

- There exists an  $\text{EVI}_{\kappa_n}$ -gradient flow for the functional  $S_n(\mu) = \frac{1}{n} \text{Ent}(\mu|m) + \int V d\mu$  on  $\mathcal{P}_2(X')$  with  $\kappa_n = \frac{1}{n}K + \kappa$ .
- In the limit  $n \rightarrow \infty$  this yields an  $\text{EVI}_{\kappa}$ -gradient flow for the functional  $S(\mu) = \int V d\mu$  on  $\mathcal{P}_2(X')$ .
- The  $\text{EVI}_{\kappa}$ -gradient flow for  $S$  is nondiffusive. In particular, the flow starting in a Dirac mass will not spread out: it is a Dirac mass in a moving point.

**Lemma 1.** *There exists an  $\text{EVI}_{\kappa_n}$ -gradient flow for the functional  $S_n(\mu) = \frac{1}{n} \text{Ent}(\mu|m) + \int V d\mu$  on  $\mathcal{P}_2(X')$  with  $\kappa_n = \frac{1}{n}K + \kappa$ .*

*Proof.* (i) Let us begin with some general remarks. Since  $X$  is geodesic and locally compact, it is “proper” in the sense that all closed balls are compact. By means of Prokhorov’s theorem this implies that  $W_2$ -bounded subsets of  $\mathcal{P}_2(X)$  are relatively compact (w.r.t. the usual topology of weak convergence of measures).

(ii) As a first step, we replace the space  $X$  by  $X'$  and the measure  $m$  by  $m' := 1_{X_0}m$ . Weak  $\kappa$ -convexity of the function  $V$  implies weak convexity of the set  $X_0$ , i.e., for each pair of points  $x_0, x_1 \in X_0$  there exists a connecting geodesic which stays in  $X_0$  (and minimizes the distance in  $X$ ).

Together with the fact that  $\text{RCD}(K, \infty)$  implies strong  $K$ -convexity of the Boltzmann entropy  $\text{Ent}(\cdot|m)$  on  $\mathcal{P}_2(X)$  ([3], Prop. 2.23), it yields weak  $K$ -convexity of the Boltzmann entropy  $\text{Ent}(\cdot|m')$  on  $\mathcal{P}_2(X')$ . Moreover, the Cheeger energy for  $(X', d, m')$  is still quadratic. (The relaxed gradients w.r.t.  $m$  and  $m'$ , resp., are local operators; they coincide on  $X_0$ .) Thus  $(X', d, m')$  satisfies the condition  $\text{RCD}(K, \infty)$ ; cf. [3], Thm. 4.19 (iii).

(iii) Next we consider the weighted mms  $(X', d, m'_1)$  with  $m'_1 = e^{-V} m'$ . Since  $V$  is continuous and bounded from below by  $-C_0 - C_1 d^2(\cdot, x_0)$ , it follows that the functional  $S : \mu \mapsto \int V d\mu$  on  $\mathcal{P}_2(X')$  is lower semicontinuous and bounded from below by  $-C_0 - C_1 W_2^2(\cdot, \delta_{x_0})$ . Indeed, if  $\mu_n \rightarrow \mu$  in  $\mathcal{P}_2(X')$ , then  $W_2(\mu_n, \delta_{x_0}) \rightarrow W_2(\mu, \delta_{x_0})$  and

$$\liminf_{n \rightarrow \infty} \int \tilde{V} d\mu_n \geq \liminf_{n \rightarrow \infty} \int \tilde{V}_i d\mu_n = \int \tilde{V}_i d\mu$$

for each  $i \in \mathbb{N}$ , where  $\tilde{V} := V + C_1 \cdot d^2(\cdot, x_0)$  and  $\tilde{V}_i := \tilde{V} \wedge i$  (which is bounded and continuous). Therefore,  $\liminf_{n \rightarrow \infty} \int \tilde{V} d\mu_n \geq \int \tilde{V} d\mu$ , and thus in turn  $\liminf_{n \rightarrow \infty} \int V d\mu_n \geq \int V d\mu$ .

The weak  $\kappa$ -convexity of  $V$  on  $X$  implies that  $S$  is weakly  $\kappa$ -convex on  $\mathcal{P}_2(X')$ . Indeed, if  $\mu_0 = \sum_{i=1}^\infty \lambda_i \cdot \delta_{x_i}$  and  $\mu_1 = \sum_{i=1}^\infty \lambda_i \cdot \delta_{y_i}$  for some  $\lambda_i \geq 0$  and  $x_i, y_i \in X'$  are given in such a way that  $q = \sum_{i=1}^\infty \lambda_i \cdot \delta_{(x_i, y_i)}$  is an optimal coupling of them and if  $z_i$  is a  $t$ -intermediate point of  $x_i, y_i$  (i.e.,  $\frac{1}{t}d(x_i, z_i) = \frac{1}{1-t}d(y_i, z_i) = d(x_i, y_i)$ ) with  $V(z_i) \leq (1-t)V(x_i) + tV(y_i) - \frac{\kappa}{2}t(1-t)d^2(x_i, y_i)$  for each  $i \in \mathbb{N}$ , then  $\mu_t := \sum_{i=1}^\infty \lambda_i \cdot \delta_{z_i}$  is a  $t$ -intermediate point of  $\mu_0, \mu_1$  and

$$S(\mu_t) \leq (1-t)S(\mu_0) + tS(\mu_1) - \frac{\kappa}{2}t(1-t)W_2^2(\mu_0, \mu_1).$$

For arbitrary discrete measures  $\mu_0 = \sum_{i=1}^\infty \alpha_i \cdot \delta_{x_i}$  and  $\mu_1 = \sum_{j=1}^\infty \beta_j \cdot \delta_{y_j}$ , each optimal coupling of them can be written as  $q = \sum_{i,j=1}^\infty \lambda_{ij} \cdot \delta_{(x_i, y_j)}$  for suitable  $\lambda_{ij} \geq 0$  with  $\sum_{k=1}^\infty \lambda_{ik} = \alpha_i$  and  $\sum_{k=1}^\infty \lambda_{kj} = \beta_j$ . Replacing  $\alpha_i \delta_{x_i}$  and  $\beta_j \delta_{y_j}$  by  $\sum_{k=1}^\infty \lambda_{ik} \delta_{x_i}$  and  $\sum_{k=1}^\infty \lambda_{kj} \delta_{y_j}$ , resp., and relabeling the indices, we may assume without restriction that  $\mu_0, \mu_1$  and  $q$  are always given as above.

Given general  $\mu_0, \mu_1 \in \mathcal{P}_2(X')$ , choose  $\mu_s^n \in \mathcal{P}_2(X')$  for  $s \in \{0, 1\}$  of the above type with  $\mu_s^n \rightarrow \mu_s$  and  $\int V d\mu_s^n \rightarrow \int V d\mu_s$  as  $n \rightarrow \infty$ . One way to achieve this is to fix for each  $n \in \mathbb{N}$  a partition  $X' = \bigcup_i X_i^n$  by sets of diameter less than  $1/n$ , to select points  $x_i^n \in \bar{X}_i^n$  with  $V(x_i^n) = \inf_{y \in \bar{X}_i^n} V(y)$ , and to put  $\mu_s^n = \sum_i \mu_s(X_i^n) \delta_{x_i^n}$ . Then obviously  $W_2(\mu_s^n, \mu_s) \leq 1/n$  and  $\int V d\mu_s^n \leq \int V d\mu_s$ . Thus the claim follows by lower semicontinuity of  $\int V d\mu$ .

Then for given  $t \in (0, 1)$  choose  $t$ -intermediate points  $\mu_t^n$  of  $\mu_0^n, \mu_1^n$  satisfying the above inequality for  $\kappa$ -convexity. Properness of  $X$  and  $W_2^2$ -boundedness of  $(\mu_t^n)_{n \in \mathbb{N}}$  implies that for each  $t \in (0, 1)$  there exists a converging subsequence, again denoted by  $\mu_t^n, n \in \mathbb{N}$ , of  $t$ -intermediate points of  $\mu_0^n, \mu_1^n$  with limit point  $\mu_t$  which then must be a  $t$ -intermediate point of  $\mu_0, \mu_1$ . Lower semicontinuity of  $S$  implies

$$\begin{aligned} S(\mu_t) &\leq \liminf_{n \rightarrow \infty} S(\mu_t^n) \\ &\leq \liminf_{n \rightarrow \infty} \left[ (1-t)S(\mu_0^n) + tS(\mu_1^n) - \frac{\kappa}{2}t(1-t)W_2^2(\mu_0^n, \mu_1^n) \right] \\ &= \left[ (1-t)S(\mu_0) + tS(\mu_1) - \frac{\kappa}{2}t(1-t)W_2^2(\mu_0, \mu_1) \right]. \end{aligned}$$

This is the weak  $\kappa$ -convexity of  $S$  on  $\mathcal{P}_2(X')$ .

Since  $\text{Ent}(\mu|m') + \int V d\mu = \text{Ent}(\mu|e^{-V} m')$  and since  $\text{Ent}(\mu|m')$  is strongly  $K$ -convex on  $\mathcal{P}_2(X')$ , this also yields the weak  $(K + \kappa)$ -convexity of  $\text{Ent}(\mu|e^{-V} m')$  on  $\mathcal{P}_2(X')$  or, in other words, the  $\text{CD}(K + \kappa, \infty)$ -condition for  $(X', d, m'_1)$ .

(iv) The next step will be to prove that the Cheeger energy for  $(X', d, m'_1)$  is quadratic or, equivalently, that the relaxed gradients in  $(X', d, m'_1)$  are quadratic.

The assumption on the quadratic lower bound for  $V$  guarantees the requested integrability condition for  $m'_1$ . Moreover, the local boundedness of  $V$  guarantees that on each closed ball in  $X_0$ , the measures  $m'$  and  $m'_1$  are equivalent with Radon–Nikodym density bounded from above and below (away from 0). Thus according to [2], Lemma 4.11, the relaxed gradients w.r.t.  $m'$  and w.r.t.  $m'_1$  coincide. Since the former already was proven to be quadratic, this finally proves the  $\text{RCD}(K + \kappa, \infty)$ -condition for  $(X', d, m'_1)$ .

(v) These observations will now be applied with  $n \cdot V$  in the place of  $V$ , which then yields that the space  $(X', d, e^{-nV} m')$  satisfies the  $\text{RCD}(\tilde{\kappa}_n, \infty)$ -condition with  $\tilde{\kappa}_n = K + n\kappa$ , [3], Prop. 6.19. The latter in turn implies the existence of an  $\text{EVI}_{\tilde{\kappa}_n}$ -gradient flow for the functional  $\tilde{S}_n(\mu) = \text{Ent}(\mu|e^{-nV} m') = \text{Ent}(\mu|m') + n \int V d\mu$  on  $\mathcal{P}_2(X')$ . Replacing  $\tilde{S}_n$  by  $S_n = \frac{1}{n} \tilde{S}_n$  and rescaling time by the factor  $n$  yields the claim.  $\square$

**Lemma 2.** *There exists an  $\text{EVI}_{\kappa}$ -gradient flow for the functional  $S(\mu) = \int V d\mu$  on  $\mathcal{P}_2(X')$ .*

*Proof.* (i) Let  $(\mu_t^n)_{t>0}$  denote the  $\text{EVI}_{\kappa_n}$ -gradient flow for the functional  $S_n(\mu) = \frac{1}{n} \text{Ent}(\mu|m) + \int V d\mu$  on  $\mathcal{P}_2(X')$  with  $\kappa_n = \frac{1}{n}K + \kappa$ . It is unique and well defined for each starting point  $\mu_0 \in \mathcal{P}_2(X')$ . Property (2) ( $\forall t > 0$ )—with  $S_n$  and  $W_2$  in place of  $V$  and  $d$ —is obviously equivalent to the property

$$(5) \quad \frac{1}{2} W_2^2(\mu_t^n, \nu) e^{2\kappa_n t} - \frac{1}{2} W_2^2(\mu_s^n, \nu) e^{2\kappa_n s} \leq \frac{1}{2\kappa_n} (e^{2\kappa_n t} - e^{2\kappa_n s}) [S_n(\nu) - S_n(\mu_t^n)]$$

( $\forall 0 \leq s < t$ ). In the case  $\kappa = 0$ , the prefactor  $\frac{1}{2\kappa} (e^{2\kappa t} - e^{2\kappa s})$  has to be replaced by  $(t - s)$ .

(ii) Note that our assumptions on  $m$  and  $V$  imply that  $S_n(\nu) \leq S_1(\nu) + C'' + C' W_2^2(\nu, \delta_{x'}) \leq C^*$  whenever  $S_1(\nu) < \infty$  and that  $-S_n(\mu_t^n) \leq C'' + C' W_2^2(\mu_t^n, \delta_{x'}) + C_0 + C_1 W_2^2(\mu_t^n, \delta_{x_0})$ . The differential version of (5) (or, in other words, (2) with  $S_n$  and  $W_2$ ) allows us to estimate uniformly in  $n$

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} W_2^2(\mu_t^n, \nu) &\leq -\kappa W_2^2(\mu_t^n, \nu) + S_n(\nu) - S_n(\mu_t^n) \\ &\leq C + c \cdot \frac{1}{2} W_2^2(\mu_t^n, \nu), \end{aligned}$$

which leads to the uniform growth bound

$$\frac{1}{2} W_2^2(\mu_t^n, \nu) \leq e^{ct} \cdot \frac{1}{2} W_2^2(\mu_0, \nu) + (e^{ct} - 1) \cdot C/c$$

with constants  $C$  and  $c$  which may depend on  $\nu$  but not on  $n$ . Thus for each  $t \in (0, 1)$ , the sequence  $(\mu_t^n)_{n \in \mathbb{N}}$  is  $W_2$ -bounded. Local compactness of  $W_2$ -bounded subsets of  $\mathcal{P}_2(X')$  implies that there exists some limit point  $\mu_t$  such that  $\mu_t^n \rightarrow \mu_t$  weakly—at least after passing to a subsequence. A diagonal argument allows us to choose this subsequence in such a way that convergence holds for all rational  $t \geq 0$ .

(iii) The quadratic lower bound (in terms of  $W_2$ ) for  $\mu \mapsto \text{Ent}(\mu|m)$  and the lower semicontinuity  $\mu \mapsto \int V d\mu$  imply that  $\liminf_{n \rightarrow \infty} S_n(\mu_t^n) \geq S(\mu_t)$ . Moreover,  $\lim_{n \rightarrow \infty} S_n(\nu) = S(\nu)$  provided  $\text{Ent}(\nu) < \infty$ . Thus for all such  $\nu$  and for all rational  $0 \leq s < t$ ,

$$(6) \quad \frac{1}{2} W_2^2(\mu_t, \nu) e^{2\kappa t} - \frac{1}{2} W_2^2(\mu_s, \nu) e^{2\kappa s} \leq \frac{1}{2\kappa} (e^{2\kappa t} - e^{2\kappa s}) [S(\nu) - S(\mu_t)].$$

Applying this estimate with  $\nu = \mu_s$  we obtain local Lipschitz continuity in  $t \in (0, \infty) \cap \mathbb{Q}$  of the flow  $(\mu_t)_t$  which allows us to extend it to all  $t$ , still satisfying the previous estimate. Thus, in particular,

$$(7) \quad \frac{d}{dt} \frac{1}{2} W_2^2(\mu_t, \nu) + \kappa W_2^2(\mu_t, \nu) \leq S(\nu) - S(\mu_t)$$

for a.e.  $t \geq 0$ .

(iv) The next step is to extend (6) (and in turn therefore also (7)) to all  $\nu \in \mathcal{P}_2(X')$ , i.e., to get rid of the constraint  $\text{Ent}(\nu|m) < \infty$ . That is, given  $\nu \in \mathcal{P}_2(X')$  we have to find a sequence of  $\nu_n \in \mathcal{P}_2(X')$  with  $\text{Ent}(\nu_n|m') < \infty$  such that  $W_2(\nu_n, \nu) \rightarrow 0$  and  $\int V d\nu_n \rightarrow \int V d\nu$  as  $n \rightarrow \infty$ .

One easily verifies that it suffices to prove this approximation result for functions  $V$  which are uniformly bounded from below. Indeed, assume that for each  $i \in \mathbb{N}$  there exists a sequence of measures  $\nu_{i,n} \in \mathcal{P}_2(X')$  with  $\text{Ent}(\nu_{i,n}|m') < \infty$  such that  $W_2(\nu_{i,n}, \nu) \leq \frac{1}{n}$  and  $\int V_i d\nu_{i,n} \leq \int V_i d\nu + \frac{1}{n}$  for all  $n$ , where  $V_i := \max\{V, -i\}$ . Then the sequence  $\nu_{n,n}$  will do the job for  $V$ :

$$\liminf_{n \rightarrow \infty} \int V d\nu_{n,n} \leq \liminf_{n \rightarrow \infty} \left[ \int V_n d\nu_{n,n} \right] \leq \liminf_{n \rightarrow \infty} \left[ \int V_n d\nu + \frac{1}{n} \right] = \int V d\nu.$$

Similarly, a simple monotone convergence argument shows that it suffices to prove the approximation results for measures  $\nu$  which are compactly supported in  $X_0$  since the assertions (6) and (7) are void if  $\nu$  is not supported in  $X_0$ .

Thus, let us now assume that  $\nu$  has compact support  $K$  in  $X_0$  and that  $V \geq 0$ . Fix  $\epsilon > 0$  such that  $K' := \overline{B}_\epsilon(K)$  is compact in  $X_0$ . Let  $\bar{\nu}_t, t > 0$ , be the gradient flow in  $\mathcal{P}_2(X')$  for the entropy  $\text{Ent}(\cdot|m')$  starting in  $\nu$  or, in other words, the heat distribution on  $X'$  after time  $t$  for the initial distribution  $\nu$ . Put  $\nu_t(\cdot) = \frac{1}{\alpha_t} \bar{\nu}_t(K' \cap \cdot)$ , where  $\alpha_t = \bar{\nu}_t(K')$ . Then continuity of  $t \mapsto \bar{\nu}_t$  in  $\mathcal{P}_2(X')$  implies  $\alpha_t \rightarrow 1$  as  $t \rightarrow 0$ . Thus

$$\limsup_{t \rightarrow 0} \int V d\nu_t = \limsup_{t \rightarrow 0} \frac{1}{\alpha_t} \int_{K'} V d\bar{\nu}_t \leq \int_{K'} V d\nu = \int V d\nu$$

since  $\bar{\nu}_t \rightarrow \nu$  weakly and since  $V$  is bounded and continuous on  $K'$ . Moreover, obviously  $\text{Ent}(\nu_t|m') < \infty$  for all  $t > 0$  and  $W_2(\nu_t, \nu) \rightarrow 0$  as  $t \rightarrow 0$ .

(v) Assume that the accumulation points of  $(\mu_t^n)_{n \in \mathbb{N}}$  are not unique, say  $\tilde{\mu}_t$  is another accumulation point (again for simplicity along one fixed subsequence for all rational  $t$ ). Applying the previous estimate (6) twice (first with  $\tilde{\mu}_s$  in place of  $\nu$ , then with  $\mu_t$  in place of  $\nu$  and  $\tilde{\mu}_s, \tilde{\mu}_t$  in place of  $\mu_s, \mu_t$ ) yields

$$\begin{aligned} & \frac{1}{2} W_2^2(\tilde{\mu}_t, \mu_t) e^{2\kappa t} + \frac{1}{2} W_2^2(\tilde{\mu}_s, \mu_t) [e^{2\kappa t} - e^{2\kappa s}] - \frac{1}{2} W_2^2(\mu_s, \tilde{\mu}_s) e^{2\kappa s} \\ & \leq \frac{1}{2\kappa} (e^{2\kappa t} - e^{2\kappa s}) [S(\tilde{\mu}_s) - S(\tilde{\mu}_t)], \end{aligned}$$

and thus in the limit  $t - s \rightarrow 0$ ,

$$\frac{d}{dt} \frac{1}{2} W_2^2(\mu_t, \tilde{\mu}_t) + \kappa W_2^2(\mu_t, \tilde{\mu}_t) \leq 0$$

for a.e.  $t$ . This finally implies  $W_2(\mu_t, \mu'_t) = 0$  and hence proves the uniqueness of the accumulation points.

(vi) Now let us consider two flows starting in different points, say  $\mu_0, \mu'_0 \in \mathcal{P}_2(X')$ . Again from (7) we may deduce  $\frac{d}{dt} \frac{1}{2} W_2^2(\mu_t, \mu'_t) + \kappa W_2^2(\mu_t, \mu'_t) \leq 0$  which

immediately yields the  $W_2$ -expansion bound

$$(8) \quad W_2(\mu_t, \mu'_t) \leq e^{-\kappa t} W_2(\mu_0, \mu'_0).$$

□

**Lemma 3.** *For each  $\mu_0 \in \mathcal{P}_2(X')$  the  $EVI_\kappa$ -flow for  $S$  satisfies*

$$\text{Var}_2(\mu_t) \leq e^{-2\kappa t} \text{Var}_2(\mu_0)$$

for all  $t > 0$ , where  $\text{Var}_2(\nu) := \iint d^2(x, y) \nu(dx)\nu(dy)$ .

*Proof.* The  $EVI_\kappa$ -property (6) for  $(\mu_t)_{t \geq 0}$  applied to  $\nu = \delta_y$  states

$$(9) \quad \frac{1}{2} \int d^2(x, y) d\mu_t(x) e^{2\kappa t} - \frac{1}{2} \int d^2(x, y) d\mu_s(x) e^{2\kappa s} \leq \frac{1}{2\kappa} (e^{2\kappa t} - e^{2\kappa s}) [V(y) - S(\mu_t)].$$

Integrating this w.r.t.  $\mu_t(dy)$  or  $\mu_s(dy)$ , resp., yields

$$(10) \quad \frac{1}{2} \text{Var}_2(\mu_t) e^{2\kappa t} - \frac{1}{2} \iint d^2(x, y) d\mu_s(x) d\mu_t(x) e^{2\kappa s} \leq 0$$

and

$$(11) \quad \frac{1}{2} \iint d^2(x, y) d\mu_s(x) d\mu_t(x) e^{2\kappa t} - \frac{1}{2} \text{Var}_2(\mu_s) e^{2\kappa s} \leq \frac{1}{2\kappa} (e^{2\kappa t} - e^{2\kappa s}) [S(\mu_s) - S(\mu_t)].$$

Adding up these two inequalities and letting  $s \rightarrow t$ , we obtain

$$\frac{d}{dt} \frac{1}{2} \text{Var}_2(\mu_t) e^{2\kappa t} \leq 0$$

since  $S(\mu_s) \rightarrow S(\mu_t)$ . This proves the claim. □

### APPENDIX 1

Let us finally add some observations which are not directly related to our main result but which might be of independent interest.

Denote the unique  $EVI_\kappa$ -flow for  $V$  starting in  $x_0 \in X'$  by  $x_t = \Phi_t(x_0), t > 0$ . Then for each  $t > 0$  the map  $\Phi_t : X' \rightarrow X'$  is Lipschitz continuous (with Lipschitz constant  $e^{-\kappa t}$ ). The family of maps  $\Phi_t, t > 0$ , is a semigroup w.r.t. composition:  $\Phi_t \circ \Phi_s = \Phi_{t+s}$ .

**Corollary 3.** *For each  $\mu_0 \in \mathcal{P}_2(X')$  the (unique)  $EVI_\kappa$ -flow for  $S$  starting in  $\mu_0$  is obtained via push forward as*

$$(12) \quad \mu_t = (\Phi_t)_* \mu_0.$$

That is,  $\int f(x) d\mu_t(x) = \int f(\Phi_t(x)) d\mu_0(x)$  for each bounded Borel function  $f$  on  $X$ .

This is a straightforward consequence of the following lemma which states that the  $EVI_\kappa$ -flow for  $S$  is  $\sigma$ -additive in  $\mu$ .

**Lemma 4.** *Assume that  $\mu_0 = \sum_{i=1}^n \lambda_i \mu_0^i$  for some probability measures  $\mu_0^i$  and positive numbers  $\lambda_i$  with  $\sum_i \lambda_i = 1$ , where  $n \in \mathbb{N}$  or  $n = \infty$ . For each  $i$ , let  $(\mu_t^i)_{t > 0}$  be the  $EVI_\kappa$ -flow for  $S$  starting in  $\mu_0^i$ . Then*

$$(13) \quad \mu_t = \sum_{i=1}^n \lambda_i \mu_t^i$$

is the  $EVI_\kappa$ -flow starting in  $\mu_0$ .

*Proof.* Let us define  $\mu_t$  as in (13), and for fixed  $t$  let  $q_t$  be an optimal coupling of  $\mu_t$  and  $\nu$ . Observe that  $\mu_t^i = f_t^i \mu_t$  for suitable bounded nonnegative functions  $f_t^i$  on  $X$ . Put  $dq_t^i(x, y) = f_t^i(x) dq_t(x, y)$  and  $\nu_t^i(\cdot) = q_t^i(X, \cdot)$ . Then  $\nu$  and  $q_t$  can be represented as

$$\nu = \sum_{i=1}^n \lambda_i \nu^i, \quad q_t = \sum_{i=1}^n \lambda_i q_t^i,$$

where for each  $i$  the probability measure  $q_t^i$  is an optimal coupling of  $\mu_t^i$  and  $\nu^i$  (since subcouplings of optimal couplings are optimal). Thus

$$(14) \quad W_2^2(\mu_t, \nu) = \sum_i \lambda_i W_2^2(\mu_t^i, \nu^i).$$

On the other hand, for each  $s > t$ ,

$$(15) \quad W_2^2(\mu_s, \nu) \leq \sum_i \lambda_i W_2^2(\mu_s^i, \nu^i).$$

(One should not expect equality since in general the optimal decomposition of  $\nu$  induced by the decomposition  $\mu_t = \sum_{i=1}^n \lambda_i \mu_t^i$  will not be optimal for the coupling with  $\mu_s = \sum_{i=1}^n \lambda_i \mu_s^i$ .) The  $\text{EVI}_\kappa$ -property for  $(\mu_s^i)_{s>t}$  (with observation point  $\nu^i$ ) implies

$$(16) \quad \lim_{s \searrow t} \frac{1}{2(s-t)} \left[ W_2^2(\mu_s^i, \nu^i) - W_2^2(\mu_t^i, \nu^i) \right] + \kappa W_2^2(\mu_t^i, \nu^i) \leq S(\nu^i) - S(\mu_t^i).$$

Adding up the last three inequalities yields

$$\begin{aligned} & \limsup_{s \searrow t} \frac{1}{2(s-t)} \left[ W_2^2(\mu_s, \nu) - W_2^2(\mu_t, \nu) \right] + \kappa W_2^2(\mu_t, \nu) \\ & \leq \sum_i \lambda_i \left[ \lim_{s \searrow t} \frac{1}{2(s-t)} \left( W_2^2(\mu_s^i, \nu^i) - W_2^2(\mu_t^i, \nu^i) \right) + \kappa W_2^2(\mu_t^i, \nu^i) \right] \\ & \leq \sum_i \lambda_i \left[ S(\nu^i) - S(\mu_t^i) \right] \\ & = \left[ S(\nu_t) - S(\mu_t) \right]. \end{aligned}$$

For the last equality, we made use of the additivity of  $S$ . □

### APPENDIX 2

Finally, let us consider  $\text{EVI}$ -gradient flows in a more general setting. Let  $(X, d)$  be an arbitrary complete geodesic space and let  $V : X \rightarrow (-\infty, +\infty]$  be a continuous function, finite on a dense subset  $X_0$  of  $X$ , and let a family of maps  $\Phi_t : X \rightarrow X$  be given such that for each  $x \in X$  the curve  $(\Phi_t(x))_{t \geq 0}$  is the unique  $\text{EVI}_\kappa$ -gradient flow for  $V$  (for some fixed  $\kappa \in \mathbb{R}$ ). This gradient flow comes with the “natural” parametrization

$$\frac{d}{dt} V(\Phi_t) = - |\nabla^- V(\Phi_t)|^2 = - |\dot{\Phi}_t|^2.$$

We will now have a closer look at a reparametrization  $(\Psi_a(x))_{a \in \mathbb{R}}$  of this flow with  $\frac{\partial}{\partial a} V(\Psi_a) = 1$ . For  $a \in \mathbb{R}$  and  $x \in X$  we put

$$T_a(x) = \inf \{ t \geq 0 : V(\Phi_t(x)) \leq a \}$$

(with  $\inf \emptyset := \infty$ ), and if  $T_a(x) < \infty$ ,

$$\Psi_a(x) = \Phi_{T_a(x)}(x).$$



The continuity of  $x \mapsto V(x)$  and  $t \mapsto \Phi_t(x)$  for  $t > 0$  implies that  $V(\Psi_a(x)) = a$  provided  $0 < T_a(x) < \infty$ .

**Proposition.** *For all  $x, x' \in X$  and all  $a \in \mathbb{R}$  with  $T_a(x), T_a(x') < \infty$ ,*

$$(17) \quad d\left(\Psi_a(x), \Psi_a(x')\right) \leq \exp\left(-\kappa \frac{T_a(x) + T_a(x')}{2}\right) \cdot d(x, x').$$

*Proof.* Given  $x, x' \in X$  and  $a \in \mathbb{R}$  put  $T = T_a(x) \wedge T_a(x')$  and  $S = |T_a(x) - T_a(x')|$ . Assume without restriction that  $T_a(x) \leq T_a(x')$ . Then  $T_a(x) = T$ ,  $T_a(x') = T + S$  and  $\Psi_a(x) = \Phi_T(x)$ ,  $\Psi_a(x') = \Phi_{T+S}(x')$ . Applying the basic contraction property (3) with  $t = T$  yields

$$(18) \quad d(\Phi_T(x), \Phi_T(x')) \leq e^{-\kappa T} d(x, x').$$

Next we apply (2) to the flow  $(\Phi_{T+t}(x'))_{t>0}$  with observation point  $z = \Phi_T(x)$ . It states that

$$\frac{d}{dt} \frac{1}{2} d^2(\Phi_T(x), \Phi_{T+t}(x')) + \kappa d^2(\Phi_T(x), \Phi_{T+t}(x')) \leq V(\Phi_T(x)) - V(\Phi_{T+t}(x')) \leq 0$$

for  $t \in [0, S]$ , which immediately yields

$$d(\Phi_T(x), \Phi_{T+S}(x')) \leq e^{-\kappa S/2} d(\Phi_T(x), \Phi_T(x')).$$

Together with (18) we thus obtain

$$d(\Phi_T(x), \Phi_{T+S}(x')) \leq e^{-\kappa(T+S/2)} d(x, x') = \exp\left(-\frac{\kappa}{2}(T_a(x) + T_a(x'))\right) \cdot d(x, x').$$

This is the claim.  $\square$

#### ACKNOWLEDGMENTS

The author gratefully acknowledges the valuable comments of Matthias Erbar, Nicola Gigli, Martin Kell, and Christian Ketterer on early drafts of this paper.

#### REFERENCES

- [1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005. MR2129498
- [2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, *Invent. Math.* **195** (2014), no. 2, 289–391. MR3152751
- [3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, *Duke Math. J.* **163** (2014), no. 7, 1405–1490. MR3205729
- [4] Matthias Erbar, Kazumasa Kuwada, and Karl-Theodor Sturm, *On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces*, *Invent. Math.* **201** (2015), no. 3, 993–1071. MR3385639
- [5] John Lott and Cédric Villani, *Ricci curvature for metric-measure spaces via optimal transport*, *Ann. of Math. (2)* **169** (2009), no. 3, 903–991. MR2480619
- [6] A. Lytchak, *Open map theorem for metric spaces*, *Algebra i Analiz* **17** (2005), no. 3, 139–159; English transl., *St. Petersburg Math. J.* **17** (2006), no. 3, 477–491. MR2167848
- [7] Perelman, G. and Petrunin, A., *Quasigeodesics and gradient curves in Alexandrov spaces*, unpublished preprint (1995), <http://www.math.psu.edu/petrunin/papers/papers.html>
- [8] Conrad Plaut, *Metric spaces of curvature  $\geq k$* , *Handbook of geometric topology*, North-Holland, Amsterdam, 2002, pp. 819–898. MR1886682
- [9] Karl-Theodor Sturm, *On the geometry of metric measure spaces. I*, *Acta Math.* **196** (2006), no. 1, 65–131. MR2237206
- [10] Karl-Theodor Sturm, *On the geometry of metric measure spaces. II*, *Acta Math.* **196** (2006), no. 1, 133–177. MR2237207

- [11] Karl-Theodor Sturm, *Ricci Tensor for Diffusion Operators and Curvature-Dimension Inequalities under Conformal Transformations and Time Changes*. ArXiv:1401.0687

INSTITUTE FOR APPLIED MATHEMATICS, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115  
BONN, GERMANY

*Email address:* `sturm@uni-bonn.de`