Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Evolutionary variational inequalities on the Hellinger–Kantorovich and spherical Hellinger–Kantorovich spaces

Vaios Laschos[1], Alexander Mielke[1,2]

submitted: November 23, 2022

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: vaios.laschos@wias-berlin.de
alexander.mielke@wias-berlin.de

[2] Humboldt-Universität zu Berlin
Institut für Mathematik
Rudower Chaussee 25
12489 Berlin
Germany

No. 2973
Berlin 2022

2020 *Mathematics Subject Classification.* 28A33, 54E35, 49Q2, 49J35, 49J40, 49K35, 46G99.

Key words and phrases. Hellinger distance, Wasserstein distance, minimizing movement scheme, geodesic semiconvexity, evolutionary variational inequality.

V.L. was partially supported by DFG under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC 2046/1, project ID 390685689). A.M. was partially supported by DFG under Priority programm *Variational Methods for Predicting Complex Phenomena in Engineering Structures and Materials* (SPP 2256, project ID 441470105, subproject Mi 459/9-1) .

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Evolutionary variational inequalities on the Hellinger–Kantorovich and spherical Hellinger–Kantorovich spaces

Vaios Laschos, Alexander Mielke

Abstract

We study the minimizing movement scheme for families of geodesically semiconvex functionals defined on either the Hellinger–Kantorovich or the Spherical Hellinger–Kantorovich space. By exploiting some of the finer geometric properties of those spaces, we prove that the sequence of curves, which are produced by geodesically interpolating the points generated by the minimizing movement scheme, converges to curves that satisfy the Evolutionary Variational Inequality (EVI), when the time step goes to 0.

1 Introduction

Let X be a geodesic metric space and $\mathcal{M}(X)$ the space of all nonnegative and finite Borel measures on X. Independently in [LMS16, LMS18], and [GaM17, CP*15] the space $(\mathcal{M}(X), \mathsf{HK})$ was introduced and studied, where HK denotes the Hellinger-Kantorovich or Wasserstein-Fisher-Rao distance. In [LMS16, LMS18], it was proved that $(\mathcal{M}(X), \mathsf{HK})$ is a geodesic space itself and all geodesic curves were characterized. In [LaM19], the spherical Hellinger Kantorovich distance SHK was introduced and it was proved that the set of all probability measures $\mathcal{P}(X) = \left\{ \mu \in \mathcal{M}(X) \,\middle|\, \mu(X) = 1 \right\}$ endowed with SHK is also a geodesic metric space.

For the rest of the paper, we have that $X \subset \mathbb{R}^d$ and that X is a compact, convex set with nonempty interior. We introduce a family of entropy functionals, i.e.

$$\mathsf{E}(\mu) = \int_X E(\rho(x))\mathcal{L}^d(\mathrm{d}x) + E'_\infty \mathrm{d}\mu^s, \qquad \mu = \rho\mathcal{L}^d + \mu^s \ \text{ and } \ \mu^s \perp \mathcal{L}^d, \tag{1.1}$$

where $E'_\infty = \lim_{t\to\infty} \frac{E(t)}{t} = \lim_{t\to\infty} E'(t)$, and μ^s the singular part of μ with respect to $\mathcal{L}^d$, i.e. the Lebesgue measure restricted at X. In this paper we are going to study De Giorgi's *minimizing movement (MM) scheme*, also known as JKO scheme (after [JKO98]) in the case of the Wasserstein space,

$$\mu_1 = \inf_{\mu\in\mathcal{M}(X)} \left\{ \frac{\mathsf{HK}^2(\mu_0,\mu)}{2\tau} + \mathsf{E}(\mu) \right\} \qquad \mu_1 = \inf_{\mu\in\mathcal{P}(X)} \left\{ \frac{\mathsf{SHK}^2(\mu_0,\mu)}{2\tau} + \mathsf{E}(\mu) \right\}, \tag{1.2}$$

for the Hellinger-Kantorovich and Spherical Hellinger-Kantorovich space respectively. We are going to limit our exploration to cases where the functionals E are of the form (1.1), and satisfy the following basic convexity assumptions.

Assumption A

DOI 10.20347/WIAS.PREPRINT.2973 Berlin 2022

1 $E : \mathbb{R}^+ \to \mathbb{R}$ *is a convex function.*

2 E *is geodesically* λ*-convex for some* $\lambda \in \mathbb{R}$.

We note that by [AFP00, Theorem 5.2], the functional E is lower semicontinuous, and it is the relaxation of itself when is defined only on $(\mathcal{M}_{\mathsf{ac}}(X), \mathsf{HK})$ or on $(\mathcal{P}_{\mathsf{ac}}(X), \mathsf{SHK})$. We will also make the following extra assumptions, that is necessary only for the case of $(\mathcal{M}(X), \mathsf{HK})$. This extra assumption, will assure lower Lebesgue density bounds for the solutions of the MM scheme as long as the initial data μ_0 has a Lebesgue density that is bounded from below. As it will be shown later, for the case of $(\mathcal{P}(X), \mathsf{SHK})$, Assumption A is sufficient, due to the total mass remaining fixed.

Assumption B $\quad \exists\, c_{\text{low}} > 0 : \quad E'(c_{\text{low}}) < 0.$

The main goal is to show that geodesic interpolation of points that are iteratively generated by the scheme give rise to sequences of curves with good limiting properties. More specifically, we show that such sequences of curves converge, when τ converges to zero, to curves that satisfy the Evolutionary Variational Inequalities (EVI) for the metric gradient system $(\mathcal{M}(X), \mathsf{E}, \mathsf{HK})$ or $(\mathcal{P}(X), \mathsf{E}, \mathsf{SHK})$, respectively.

Before we proceed, we briefly remind the reader of the definition of EVI. It involves the upper right Dini derivative

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\zeta(t) := \limsup_{h\to 0^+} \frac{1}{h}\left(\zeta(t+h) - \zeta(t)\right).$$

Definition 1.1 (EVI solutions for metric gradient systems) *Let* $(\mathcal{X}, \mathsf{d})$ *be a metric space and* $\phi : \mathcal{X} \to (-\infty, \infty]$ *a lsc functional, then* $(\mathcal{X}, \mathsf{d}, \phi)$ *is called a* metric gradient system*. For* $T \in (0, \infty)$ *and* $\lambda \in \mathbb{R}$ *we say that a continuous curve* $\boldsymbol{x} : [0, T) \to \mathcal{X}$ *is an* EVI_λ solution *for the metric gradient system* $(\mathcal{X}, \mathsf{d}, \phi)$*, if* $\phi(\boldsymbol{x}(t)) < \infty$ *for all* $t \in (0, T)$ *and for every "observer"* $x_{\mathsf{ob}} \in \mathcal{X}$*, we have*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(t), x_{\mathsf{ob}}) + \frac{\lambda}{2}\mathsf{d}^2(\boldsymbol{x}(t), x_{\mathsf{ob}}) \leq \phi(x_{\mathsf{ob}}) - \phi(\boldsymbol{x}(t)) \quad \textit{for all } t \in (0, T). \tag{1.3}$$

If furthermore $T = \infty$*, we call* $\boldsymbol{x}$ *a* complete EVI_λ solution.

EVIs are used to provide a generalization of the definition of gradient flows in the more abstract setting of geodesic metric spaces, see [AGS05]. For a nice exposition on EVIs the reader is advised to follow the trilogy of papers [MuS20, MuS22]. Our main result (cf. Theorem 5.6) relies on some these results and reads as follows:

Main Result *Let* $X \subset \mathbb{R}^d$ *be a compact, convex set with nonempty interior. Furthermore, let* E *be of the form* (1.1) *and let it satisfy Assumption A.*

Then, for all $\mu_0 = \rho_0\mathcal{L}^d$ *with* $0 < \underline{\rho}_0 \leq \rho_0(x) \leq \overline{\rho}_0 < \infty$ *a.e. in* X*, the solutions of the MM scheme for the gradient system* $(\mathcal{P}(X), \mathsf{E}, \mathsf{SHK})$ *(see* $(1.2)_{\mathsf{SHK}}$*) converge to a complete solution* $\boldsymbol{\mu} : [0, \infty) \to \mathcal{P}(X)$ *of EVI*$_\lambda$*. Moreover, for all* $\mu_0 \in \overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{SHK}} \subset \mathcal{P}(X)$ *there exists a unique EVI solution emanating from* μ_0.

If in addition Assumption B is satisfied, then the minimizing scheme for the gradient system $(\mathcal{M}(X), \mathsf{E}, \mathsf{HK})$ *(see* $(1.2)_{\mathsf{HK}}$*) converges to a complete solution* $\boldsymbol{\mu} : [0, \infty) \to \mathcal{M}(X)$ *of the EVI*$_\lambda$*. Moreover, for each* $\mu_0 \in \overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{HK}} \subset \mathcal{M}(X)$ *there exists a unique EVI solution emanating from* μ_0*.*

A family of functionals E satisfying both Assumptions A and B on the Hellinger-Kantorovich space $(\mathcal{M}(X), \mathsf{E})$ is the following.

Example 1.2 (The case $(\mathcal{M}(X), \mathsf{HK}, \mathsf{E})$) *Consider $\mathsf{E}^\gamma_{\alpha,m}$ generated by $E^\gamma_{\alpha,m}(c) = \alpha c^m + \gamma c$ with $\alpha > 0,\ m > 1,$ and $\gamma < 0$. Then, according to [LMS22, Sec. 7] we know that E is geodesically λ-convex on $(\mathcal{M}(X), \mathsf{E})$ with $\lambda = 2\gamma$.*

Another example of functionals satisfying both Assumption A and B, on the Hellinger-Kantorovich space but also on the Spherical Hellinger Kantorovich space follows.

Example 1.3 (Both cases $(\mathcal{M}(X), \mathsf{HK}, \mathsf{E})$ and $(\mathcal{P}(X), \mathsf{SHK}, \mathsf{E})$) *Let $E(c) = -c^q$ and assume that either $d = 1$ and $q \in [1/3, 1/2]$ or $d = 2$ and $q = 1/2$, then [LMS22, Sec. 7] ensures that E is geodesically 0-convex on $(\mathcal{M}(X), \mathsf{HK})$, and Proposition 5.5 gives the same for $(\mathcal{P}(X), \mathsf{SHK})$.*

Open Question 1.4 *In [LMS22] the functionals E of the form* (1.1) *that are semiconvex were fully characterized. For the case of the Spherical Hellinger-Kantorovich space, only very few semiconvexity results are known, and these results are corollaries of general theorems that connect metric spaces with their spherical counterparts, see Proposition 5.5. A general characterization of geodesically semi-convex functionals on the Spherical Hellinger-Kantorovich space is still elusive. We leave this as an open question, and we welcome any suggestion for collaboration in this direction.*

Unlike with other definitions of gradient flows, the EVI approach guarantees some useful properties. One of the most important, is the asymptotic stability for sequences of curves that satisfy EVI (see [MuS20, Sec. 3.2]). More specifically, under very weak convergence assumptions for a sequence of functionals G_k, to some G_∞, we get for free that the sequence of solutions to the respective EVI converge to a solution of EVI with respect to the limit functional. One can easily show that the limit case EVI for $\mathsf{G}_m = \mathsf{E}^\gamma_{\alpha,m}$ for $m \to \infty$, where $\mathsf{E}^\gamma_{\alpha,m}$ as in Example 1.2, corresponds to the functional

$$\mathsf{E}^\gamma_\infty(\mu) = \begin{cases} \gamma\mu(X) & \text{for } \mu = \rho\mathcal{L} \quad \text{with} \quad \rho(x) \leq 1 \text{ a.e.}, \\ \infty & \text{otherwise}, \end{cases} \tag{1.4}$$

which was studied in [DiC20]. By this stability property of the EVI_λ, the existence of complete EVI_λ solutions follows for a bigger class of interesting functionals that satisfy a weaker version of Assumption B. A particularly interesting case covered by this weaker assumption is the density $E(c) = \alpha c^m$. Comparing with Example 1.2 we observe that the negative term γc, which controls the mass, is no longer needed here. More specifically we have the following extension.

Extension *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore, let E be of the form* (1.1) *and let it satisfy Assumption A. If $E'(0) = 0$, then for every $\mu_0 \in \overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{HK}}$ there exists a unique curve $\boldsymbol{\mu} : [0, \infty) \to \mathcal{M}(X)$ emanating from μ_0 that is a complete EVI$_\lambda$ solution. For $d \leq 2$, no extra assumption in $E'(0)$ is needed.*

One of the main reasons why gradient of entropy functionals on spaces of measures are studied is that quite often are proven to be solutions to well known partial differential equations (PDEs), as it was initially shown for the case of the Fokker-Planck and diffusion equations, which was to proven to be a gradient flow of the differential entropy on the Wasserstein space, see [JKO98, Ott01]. Further applications in [PQV14, DiC20] involve reaction-diffusion equations where the reaction term exactly correspond to the Hellinger part in HK.

To emphasize this connection we follow [LiM13, LMS16] and provide some formal arguments showing that these PDEs are gradient-flow equations for the metric gradient system $(\mathcal{M}(X), \mathsf{HK}, \mathsf{E})$ or $(\mathcal{P}(X), \mathsf{SHK}, \mathsf{E})$, respectively. For this we use the concept of the Onsager operator, which can be seen as the inverse of the formal Riemannian tensor associated with $\mathsf{HK}_{\alpha,\beta}$ or $\mathsf{SHK}_{\alpha,\beta}$, where we now allow for the general parameters $\alpha, \beta > 0$ such that $\mathsf{HK} = \mathsf{HK}_{1,4}$ and $\mathsf{SHK} = \mathsf{SHK}_{1,4}$. Assuming that $\mu = \rho \mathrm{d}x$ we define the Onsager operators $\mathbb{K}^{\mathsf{HK}}_{\alpha,\beta}$ and $\mathbb{K}^{\mathsf{SHK}}_{\alpha,\beta}$ via

$$\begin{aligned}\mathbb{K}^{\mathsf{HK}}_{\alpha,\beta}(\mu)\,\xi &= -\operatorname{div}(\alpha\rho\nabla\xi) + \beta\rho\xi,\\ \mathbb{K}^{\mathsf{SHK}}_{\alpha,\beta}(\rho)\,\xi &= -\operatorname{div}(\alpha\rho\nabla\xi) + \beta\rho\left(\xi - \textstyle\int_X \rho\xi \mathrm{d}x\right).\end{aligned}$$

Assuming further that $\mathcal{L}^d = \mathcal{L}|_X$ is the d-dimensional Lebesgue measure such that $\mathsf{E}(\rho\mathrm{d}x) = \int_X E(\rho)\mathrm{d}x$ we have the derivative $\mathrm{D}\mathsf{E}(\rho\mathrm{d}x) = E'(\rho)$ and the (formal) gradient-flow equations read

$$\begin{aligned}\dot\rho = -\mathbb{K}^{\mathsf{HK}}_{\alpha,\beta}(\rho\mathrm{d}x)\,\mathrm{D}\mathsf{E}(\rho\mathrm{d}x) \quad &= \alpha\operatorname{div}(\rho E''(\rho)\nabla\rho) - \beta\rho E'(\rho),\\ \dot\rho = -\mathbb{K}^{\mathsf{SHK}}_{\alpha,\beta}(\rho\mathrm{d}x)\,\mathrm{D}\mathsf{E}(\rho\mathrm{d}x) \quad &= \alpha\operatorname{div}(\rho E''(\rho)\nabla\rho) - \beta\rho\left(E'(\rho) - \textstyle\int_X \rho E'(\rho)\mathrm{d}x\right).\end{aligned}$$

The gradient flow with respect to SHK is constituted by three parts, namely a classical nonlinear diffusion term, a local reaction term, and a nonlocal that takes care of the mass constraint.

In [Fle20], the author provides a rigorous proof to the above heuristics, by showing that the curves generated by the JKO scheme converge to solutions of some known reaction-diffusion PDEs. In combination, the two results guarantee that solutions to these PDEs, satisfy an EVI with respect to those functionals on the Hellinger-Kantorovich and Spherical Hellinger-Kantorovich space.

Our paper is separated in five parts. In Section 2 we are going to provide various equivalent definition of the HK, and some useful Lemmas. In Section 3 we are going to to prove that if μ_0, has nice densities bounds then the minimizers μ_1, of both the JKO schemes in (1.2) has also nice densities bounds. Specifically for the Spherical Hellinger-Kantorovich space we recover a discrete maximal principle. In Section 4, we are going provide an abstract existence theorem for EVI solutions based on geometric properties of the underlying geodesic space such κ-concavity and the local-angle condition (LAC). Here we use a series of results from [MuS22] which are cited in full detailed but without proof, except for our abstract existence result in Theorem 4.9 that relies on the density of $\cup A_\kappa$, where A_κ are suitable subsets of $\mathfrak{X}$ in which the metric is κ-concave. Our abstract existence result extends the approach provided in [MuS20, MuS22] of proving EVI solutions for λ-convex functions to situations where κ-concavity of the squared distance is not globally true. We only need that κ-concavity holds only in suitable subsets A_κ instead of the whole space. We reckon that this localization approach is an interesting extension on its own.

In Section 5 we then show that the necessary κ-concavity and the LAC for $(\mathcal{M}(X), \mathsf{HK})$ and $(\mathcal{P}(X), \mathsf{SHK})$ can be obtained from the theory developed in [LaM19]. Combining this with the theory of geodesic convexity of [LMS22] then allows us to establish our main existence results for EVI solutions.

2 The metric spaces $(\mathcal{M}(X), \mathsf{HK})$ and $(\mathcal{P}(X), \mathsf{SHK})$

2.1 Notation and preliminaries

We will denote by $\mathcal{M}(X)$ the space of all nonnegative and finite Borel measures on X endowed with the weak topology induced by the duality with the continuous and bounded functions of $\mathrm{C}_b(X)$. The subset of measures with finite quadratic moment will be denoted by $\mathcal{M}_2(X)$. The spaces $\mathcal{P}(X)$ and $\mathcal{P}_2(X)$ are the corresponding subsets of probability measures. If $\mu \in \mathcal{M}(X)$ and $\boldsymbol{T} : X \to X$ is a Borel map, $\boldsymbol{T}_\sharp\mu$ will denote the push-forward measure on $\mathcal{M}(X)$, defined by

$$\boldsymbol{T}_\sharp\mu(B) := \mu(\boldsymbol{T}^{-1}(B)) \quad \text{for every Borel set } B \subset X. \tag{2.1}$$

We will often denote elements of $X \times X$ by (x_0, x_1) and the canonical projections by $\pi^i : (x_0, x_1) \to x_i$ for $i = 0, 1$. A transport plan on X is a measure $M_{01} \in \mathcal{M}(X{\times}X)$ with marginals $\mu_i := \pi^i_\sharp M_{01}$.

2.2 The logarithmic-entropy transport formulation

Here we first provide the definition of the $\mathsf{HK}(\mu_0, \mu_1)$ distance in terms of a minimization problem that balances a specific transport problem of measures $\sigma_0\mu_0$ and $\sigma_1\mu_1$ with the relative entropies of $\sigma_j\mu_j$ with respect to μ_j. For the characterization of the Hellinger–Kantorovich distance via the static Logarithmic-Entropy Transport (LET) formulation, we define the logarithmic entropy density $F : [0,\infty[\;\to [0,\infty[$ via $F(r) = r\log r - r + 1$ and the cost function $\ell : [0,\infty[\;\to [0,\infty]$ via $\ell(R) = -2\log(\cos(R))$ for $R < \frac{\pi}{2}$ and $\ell \equiv +\infty$ otherwise. For given measures μ_0, μ_1 the LET functional $\mathsf{LET}(\,t\,; \mu_0, \mu_1) : \mathcal{M}(X \times X) \to [0,\infty[$ reads

$$\mathsf{LET}(H_{01}; \mu_0, \mu_1) := \int_X F(\sigma_0)\mathrm{d}\mu_0 + \int_X F(\sigma_1)\mathrm{d}\mu_1 + \iint_{X\times X} \ell(\mathsf{d}_X(x, x_1))\mathrm{d}H_{01} \tag{2.2}$$

with $\eta_i := (\pi_i)_\sharp H_{01} = \sigma_i\mu_i \ll \mu_i$. With this, the equivalent formulation of the Hellinger–Kantorovich distance as entropy-transport problem reads as follows.

Theorem 2.1 (LET formulation, [LMS18, Sec. 5]) *For all $\mu_0, \mu_1 \in \mathcal{M}(X)$ we have*

$$\mathsf{HK}^2(\mu_0, \mu_1) = \min\big\{\mathsf{LET}(H_{01}; \mu_0, \mu_1) \,\big|\, H_{01} \in \mathcal{M}(X{\times}X),\ (\pi_i)_\sharp H_{01} \ll \mu_i\big\}. \tag{2.3}$$

An optimal transport plan H_{01}, which always exists, gives the effective transport of mass. Note, in particular, that only $\eta_i \ll \mu_i$ is required and the cost of a deviation of η_i from μ_i is given by the entropy functionals associated with F. Moreover, the cost function ℓ is finite in the case $\mathsf{d}_X(x_0, x_1) < \frac{\pi}{2}$, which highlights the sharp threshold between transport and pure absorption-generation mentioned earlier.

Amongst the many characterizations of HK discussed in [LMS18] there is one that connects HK with the classical Kantorovich-Wasserstein distance on the cone $\mathfrak{C}$ over the base space (X, d_X) with metric

$$\mathsf{d}_{\mathfrak{C}}^2(z_0, z_1) := r_0^2 + r_1^2 - 2r_0r_1\cos_\pi(\mathsf{d}_X(x_0, x_1)), \quad z_i = [x_i, r_i], \tag{2.4}$$

where as above $\cos_b(a) = \cos(\min\{b, a\})$. Measures in $\mathcal{M}(X)$ can be "lifted" to measures in $\mathcal{M}(\mathfrak{C})$, e.g. by considering the measure $\mu \otimes \delta_1$ for $\mu \in \mathcal{M}(X)$. Moreover, we can define the projection of measures in $\mathcal{M}_2(\mathfrak{C})$ onto measures in $\mathcal{M}(X)$ via

$$\mathfrak{P} : \begin{cases} \mathcal{M}_2(\mathfrak{C}) & \to & \mathcal{M}(X), \\ \lambda & \mapsto & \int_{r=0}^\infty r^2\, \lambda(\cdot, \mathrm{d}r). \end{cases}$$

For example, the lift $\lambda = m_0\delta_{\{\mathbf{0}\}} + \mu \otimes \frac{1}{r(\cdot)^2}\delta_{r(\cdot)}$, with $m_0 \geq 0$ and $r : \mathrm{supp}(\mu) \to]0,\infty[$ arbitrary, gives $\mathfrak{P}\lambda = \mu$. Now, the cone space formulation of the Hellinger–Kantorovich distance of two measures $\mu_0, \mu_1 \in \mathcal{M}(X)$ is given as follows.

Theorem 2.2 (Optimal transport formulation on the cone) *For $\mu_0, \mu_1 \in \mathcal{M}(X)$ we have*

$$\begin{aligned} \mathsf{HK}^2(\mu_0,\mu_1) &= \min\Big\{ \mathrm{W}^2_{\mathsf{d}_\mathfrak{C}}(\lambda_0,\lambda_1) \,\Big|\, \lambda_i \in \mathcal{P}_2(\mathfrak{C}),\ \mathfrak{P}\lambda_i = \mu_i \Big\} \\ &= \min\Big\{ \iint_{\mathfrak{C}\times\mathfrak{C}} \mathsf{d}^2_\mathfrak{C}(z_0,z_1)\mathrm{d}\Lambda_{01}(z_0,z_1) \,\Big|\, \pi^i_\sharp\Lambda_{01} = \lambda_i,\ \textit{and}\, \mathfrak{P}\lambda_i = \mu_i \Big\}. \end{aligned}$$

This result will be needed for proving K-semiconcavity in Theorem 5.1.

2.3 Dilation-Transportation

Definition 2.3 *[Dilation-transport system] A quintuple $(\nu, q_0, \boldsymbol{T}_0, q_1, \boldsymbol{T}_1)$ with $\nu \in \mathcal{M}(\mathcal{Y})$, $\boldsymbol{T}_i : \mathcal{Y} \to X$, and $q_i \in \mathrm{L}^2(Y;\nu)$ is called dilation-transportation system for (μ_0, μ_1), if*

$$(\boldsymbol{T}_i)_\sharp(q_i^2\nu) = \mu_i. \tag{2.5}$$

is satisfied. If the dilation-transportation system, has the form $(\mu_0, 1, \mathbf{I}, q, \boldsymbol{T})$, then we call $(q, \boldsymbol{T})$ a dilation-transportation couple from μ_0 to μ_1. If for the dilation-transportation system/couple we have

$$\mathsf{HK}^2(\mu_0,\mu_1) = \int_Y \big((q_0-q_1)^2 + 4q_0q_1\sin^2(|\boldsymbol{T}_0 - \boldsymbol{T}_1|/2 \wedge \pi/2)\big)\,\mathrm{d}\nu, \tag{2.6}$$

then we will call that an optimal *dilation-transportation system/couple from μ_0 to μ_1.*

Definition 2.4 *For $\mu_0, \mu_1 \in \mathcal{M}(X)$, we define the following sets:*

$$A_i' = \{x \in X : \mathrm{dist}(x, \mathrm{supp}(\mu_{1-i} < \pi/2))\}, \qquad A_i'' := X \setminus A_i'. \tag{2.7}$$

We also define the following measure

$$\mu_i'(t) := (\mu_i)_{|_{A_i'}}(t) = \mu_i(t \cap A_i'), \qquad \mu_i''(t) := (\mu_i)_{|_{A_i''}}(t) = \mu_i(t \cap A_i). \tag{2.8}$$

Definition 2.5 (Reduced couple) *A couple of measures $(\mu_0, \mu_1) \in \mathcal{M}(X)^2$ is called reduced if $\mu_0 = \mu_0', \mu_1 = \mu_1'$.*

For every couple (μ_0, μ_1), the couple (μ_0', μ_1') is always reduced. Now, we have the following theorem that is a simplified version of [LMS22, Corollary 3.5].

Theorem 2.6 *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Let $(\mu_0, \mu_1) \in \mathcal{M}(X)^2$, and $\mu_0 \simeq \mathcal{L}^d$. Then there exists an optimal dilation-transportation couple $(q, \boldsymbol{T})$ from μ_0 to μ_1, with $\|\boldsymbol{T}(x) - x\| < \pi/2\ \ \mathcal{L}^d$-a.e. If furthermore $\mu_1' \ll \mathcal{L}^d$, then $\widetilde{\boldsymbol{T}}$ is essentially injective.*

Remark 2.7 *A transport plan $\boldsymbol{T}$ as in Theorem 2.6 has a version that is fully injective. From now on, without loss of generality we will make the assumption that $\boldsymbol{T}$ is fully injective to simplify the arguments. Even more, a couple like that, will be called an injective optimal dilation-transportation couple from μ_0 to μ_1.*

In the next Lemma, we will show that if the couple $(q, \boldsymbol{T})$ is an injective optimal dilation-transportation couple from μ_0 to μ_1, then it acts as an injective optimal dilation-transportation couple from $(\mu_0)_{|A}$ to $(\mu_1)_{|\boldsymbol{T}(A)}$, for every measurable set A. Even more for any partition $\{A_i\}$ of X, the total dilation-transportation cost squared is equal to the sum of the squares of the dilation-transportation costs for each part of the partition. This straightforward lemma will be used in the next section, for the construction of measures that violate the minimum assumption for the MM scheme if the minimum candidate does not have nice density bounds. These construction will be achieved by cutting and gluing the potential candidate with other measures.

Lemma 2.8 *Let $(\mu_0, \mu_1) \in \mathcal{M}(X)^2$ with $\mu_0', \mu_1' \ll \mathcal{L}^d$ and $\mu_1'' = 0$. Let $(q, \boldsymbol{T})$ be an injective optimal dilation-transportation couple from μ_0 to μ_1 and let $A_i, i = 1, \dots, n$, be a partition of X. If μ_0^i is the restriction of μ_0 on A_i, and μ_1^i, the restriction of μ_1 on $\boldsymbol{T}(A_i)$, then we have:*

- $(q, \boldsymbol{T})$ *an injective optimal dilation-transportation couple from μ_0^i to μ_1^i.*
- $\sum_i \mu_1^i = \mu_1$.
- $\mathsf{HK}^2(\mu_0, \mu_1) = \sum_{i=1}^n \mathsf{HK}^2(\mu_0^i, \mu_1^i)$.

Proof. We have

$$\begin{aligned}
&\int_X \zeta(x)\mu_1^i(\mathrm{d}x) = \int_X \mathbb{I}_{\boldsymbol{T}(A^i)}(x)\zeta(x)\mu_1(\mathrm{d}x) = \int_X \mathbb{I}_{\boldsymbol{T}(A^i)}(\boldsymbol{T}(x))\zeta(\boldsymbol{T}(x))q^2(x)\mu_0(\mathrm{d}x)\\
&= \int_{(X\setminus A^i)} \mathbb{I}_{\boldsymbol{T}(A^i)}(\boldsymbol{T}(x))\zeta(\boldsymbol{T}(x))q^2(x)\mu_0(\mathrm{d}x) + \int_{A_i} \mathbb{I}_{\boldsymbol{T}(A^i)}(\boldsymbol{T}(x))\zeta(\boldsymbol{T}(x))q^2(x)\mu_0(\mathrm{d}x)\\
&= 0 + \int_{A_i} \zeta(\boldsymbol{T}(x))q^2(x)\mu_0(\mathrm{d}x) = \int_X \zeta(\boldsymbol{T}(x))q^2(x)\mu_0^i(\mathrm{d}x).
\end{aligned}$$

By summing over i, we also obtain $\sum \mu_1^i = \mu_1$. Since $\|\boldsymbol{T}(x) - x\| < \frac{\pi}{2}$, it holds that for every $i \in 1, \dots, n$, the couple (μ_0^i, μ_1^i) is reduced and we can construct an optimal dilation-transportation couple $(q^i, \boldsymbol{T}^i)$ for (μ_0^i, μ_1^i), with cost, $\mathsf{HK}^2(\mu_0^i, \mu_1^i)$. For every i, we have that $(q, \boldsymbol{T})$ is a dilation-transportation couple between μ_0^i, and μ_1^i. Therefore we have

$$\begin{aligned}
\mathsf{HK}^2(\mu_0^i, \mu_1^i) &= \int_X \left(1 + q^i(x)^2 - 2q^i(x)\cos(|x - \boldsymbol{T}^i(x)|)\right)\mu_0^i(\mathrm{d}x)\\
&\le \int_X \left(1 + q^2(x) - 2q(x)\cos(|x - \boldsymbol{T}(x)|)\right)\mu_0^i(\mathrm{d}x).
\end{aligned} \tag{2.9}$$

We now define $(\widetilde{q}, \widetilde{\boldsymbol{T}})$ by

$$\widetilde{\boldsymbol{T}}(x) = \boldsymbol{T}^i(x), \quad x \in A^i, \qquad \widetilde{q}(x) = q^i(x), \quad x \in A^i, \tag{2.10}$$

we have

$$\mu_1 = \sum_{i=1}^n \mu_1^i = \sum_{i=1}^n \widetilde{\boldsymbol{T}}_\#(\widetilde{q}^2\mu_0^i) = \widetilde{\boldsymbol{T}}_\#(\widetilde{q}^2\mu_0),$$

and therefore it is a dilation-transportation couple for (μ_0, μ_1) with total cost

$$\sum_{i=1}^n \mathsf{HK}^2(\mu_0^i, \mu_1^i) \le \sum_{i=1}^n \int_X \left(1 + q^2(x) - 2q(x)\cos\left(|x - \boldsymbol{T}(x)|\right)\right)\mu_0^i(\mathrm{d}x) = \mathsf{HK}^2(\mu_0, \mu_1) \tag{2.11}$$

If at least one of the estimates in 2.9 is strict then the above inequality is also strict which implies that $(\widetilde{q}, \widetilde{\boldsymbol{T}})$ is a dilation-transportation couple that has less cost than $(q, \boldsymbol{T})$, which contradicts the fact that the latter is optimal. From the above we have that (2.9) is an equality for every i, and therefore $(q, \boldsymbol{T})$ is an injective optimal dilation-transportation couple between μ_0^i, and μ_1^i, and $\sum_{i=1}^n \mathsf{HK}^2(\mu_0^i, \mu_1^i) = \mathsf{HK}^2(\mu_0, \mu_1)$. ■

We will also use the following results, which was discussed carefully in [LaM19].

Theorem 2.9 (Scaling property of HK) *For all $\mu_0, \mu_1 \in \mathcal{M}(X)$ and $t_0, t_1 \geq 0$ we have*

$$\mathsf{HK}^2(t_0^2\mu_0, t_1^2\mu_1) = t_0 t_1 \mathsf{HK}^2(\mu_0, \mu_1) + (t_0^2 - t_0 t_1)\mu_0(X) + (t_1^2 - t_0 t_1)\mu_1(X). \tag{2.12}$$

Even more, if H_{01} is an optimal plan for the LET formulation of $\mathsf{HK}(\mu_0, \mu_1)$, then $H_{01}^{t_0 t_1} = t_0 t_1 H_{01}$ is an optimal plan for $\mathsf{HK}(t_0^2\mu_0, t_1^2\mu_1)$. When $t_0 = t_1 = t$,

Choosing $t_0/t_1 = (\mu_1(X)/\mu_0(X))^{1/2}$ in (2.12) we obtain the lower bound

$$\mathsf{HK}^2(\mu_0, \mu_1) \geq \left(\sqrt{\mu_0(X)} - \sqrt{\mu_1(X)}\right)^2. \tag{2.13}$$

2.4 The Spherical HK

The Spherical Hellinger-Kantorovich space $(\mathcal{P}(X), \mathsf{SHK})$ was introduced in [LaM19], and the distance metric is related to the Hellinger-Kantorovich distance HK restricted to $\mathcal{P}(X) \times \mathcal{P}(X)$ through the formula by

$$\mathsf{SHK}(\nu_0, \nu_1) = \arccos\left(1 - \frac{1}{2}\mathsf{HK}^2(\nu_0, \nu_1)\right) = 2\arcsin\left(\frac{1}{2}\mathsf{HK}(\nu_0, \nu_1)\right).$$

The important point is that $(\mathcal{P}(X), \mathsf{SHK})$ is still a geodesic space in the sense of Definition 4.1. Moreover, it is shown [LaM19] how that all geodesics connection ν_0 and ν_1 in $(\mathcal{P}(X), \mathsf{SHK})$ can be obtained by the geodesics in $\mathcal{M}(X), \mathsf{HK})$.

3 Density bounds for the MM scheme

The purpose of this section is to show that the starting from a μ_0 with nice density bounds, the minimizer of the MM scheme (1.2) also enjoy similar bounds. This will later be used to retrieve concavity properties for the distance squared, along the geodesics that interpolate the points generated by the scheme. The approach is an extension of an idea by Felix Otto developed in [Ott96]. There, the author proved that for a specific class of functionals E, and for measures μ_0 that has density bounded from below by a number $c_{\min} > 0$, the one step minimizer of the MM scheme, $\mu_1 = \arg\min\left\{\frac{W_2(\mu_0,\mu)}{2\tau} + \mathsf{E}(\mu)\right\}$, has also the same property. The main argument was, that if the set of points with density smaller than c_{min} is not essential empty, then mass must have moved outside from it to an another set, resulting in density bigger than c_{min}. This would imply, that keeping some of the mass at place would not only have been cheaper with respect to the Wasserstein distance, but also would have resulted to a more “uniform"distribution of the density and therefor a smaller value of the functional, getting a contradiction to the assumption that μ_1 is a minimizer. Felix Otto, used this argument in [Ott96] to prove that the MM scheme for the p-density functionals, converge to solutions of the doubly degenerate diffusion equations. Our arguments are of similar nature, although we have to take into account that our setting also allows for destruction and creation of mass.

3.1 Motivation of density bounds

The ODE case already gives an indication that should be addressed.

Consider any convex domain $X \subset \mathbb{R}^d$ with $\mathcal{L}^d(X) = 1$. Moreover, we restrict our view to measures with spatially constant Lebesgue density, i.e. $\mu(t) = c(t)\mathcal{L}^d$ as special solutions for the gradient system $(\mathcal{M}(X), \mathsf{E}, \mathsf{HK})$. Clearly, the equation for the scalar c is

$$\dot{c} = -4\, c\, E'(c).$$

Because of the above choices we have $\mathsf{HK}(c_0\mathcal{L}^d, c_1\mathcal{L}^d) = \big(\sqrt{c_1}-\sqrt{c_0}\big)^2$, and the MM scheme reduces to

$$\frac{1}{2\tau}\,(\sqrt{c_1}-\sqrt{c_0})^2 + E(c_1) \rightsquigarrow \min_{c_1 \geq 0} \quad \longleftrightarrow \quad 1 - \sqrt{\frac{c_0}{c_1}} + 2\tau E'(c_1) = 0. \tag{3.1}$$

Assuming $E'(c_{\mathrm{low}}) \leq 0 \leq E'(c_{\mathrm{upp}})$ we obtain the following trivial observations for the MM scheme solutions:

(D1) $\quad c_0 \geq c_{\mathrm{upp}} \iff c_1 \leq c_0,$

(D2) $\quad c_0 \leq c_{\mathrm{low}} \iff c_1 \geq c_0.$

However, in the case inbetween, we obtain nontrivial estimates:

(D3) $\quad c_1 \leq \max\Big\{a,\ \dfrac{c_0}{(1{+}2\tau\min\{E'(a),0\})^2}\Big\}$ whenever $2\tau E'(a) > -1;$

(D4) $\quad c_1 \geq \min\Big\{b,\ \dfrac{c_0}{(1{+}2\tau\max\{E'(b),0\})^2}\Big\}$ for all $b \geq 0.$

To see that the upper estimate in (D3) holds we set $\mu_a := \min\{E'(a),0\}$ with $0 \geq \mu_a > -1/(2\tau)$ and assume that (D3) does not hold, i.e. (i) $c_1 > a$ and (ii) $c_1 > c_0/(1{+}2\tau\mu_a)^{-2}$. Then, by (i) and monotonicity of E' we have $0 \leq (1{+}2\tau\mu_a)^2 \leq (1{+}2\tau E'(c_1))^2$. Exploiting the Euler-Lagrange equation in (3.1) we continue

$$(1{+}2\tau\mu_a)^2 \leq (1{+}2\tau E'(c_1))^2 \overset{\text{EL eqn.}}{=} c_0/c_1 \overset{\text{(ii)}}{<} (1{+}2\tau\mu_a)^2,$$

which is the desired contradiction. The lower estimate in (D4) follows similarly.

We have considered the simple scalar case because it turns out that similar estimates hold for the densities in a true minimization step for $(\mathcal{M}(X), \mathsf{HK}, \mathsf{E})$, see (3.10) in Proposition 3.2 for the upper estimate (D3) and (3.18) in Proposition 3.4 for a simplified version of the lower estimate (D4) using $\max\{E'(c_{\mathsf{low}}),0\} = 0$ following from Assumption B.

The "spherical" case for $(\mathcal{M}(X), \mathsf{E}, \mathsf{SHK})$ is in fact much better, because no sign conditions for $E'(c)$ are needed. To see this we again consider the pure Spherical Hellinger space $(\mathcal{P}(X), \mathsf{E}, \mathsf{SHe})$ with $\mathsf{SHe}(\nu_0,\nu_1) = 2\arcsin\big(\mathsf{He}(\nu_0,\nu_1)/2\big)$ and $\mathsf{He}(\mu_0,\mu_1)^2 = \int_X \big[\big(\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu}\big)^{1/2} - \big(\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu}\big)^{1/2}\big]^2 \mathrm{d}\mu$ for any μ with $\mu_0{+}\mu_1 \ll \mu$. The corresponding gradient flow for for absolutely continuous measures $\nu(t) = c(,\cdot)\mathrm{d}x$ leads to

$$\dot{c}(t,x) = -4c(t,x)\Big(E'(c(t,x)) - \int_X c(t,y)E'(c(t,y))\,\mathrm{d}y\Big).$$

For this flow it can be shown that $t \mapsto \inf c(t,\cdot)$ is increasing and $t \mapsto \sup c(t,\cdot)$ is decreasing. For this we consider any smooth convex function $\varphi :]0,\infty[\to \mathbb{R}$ and observe

$$\frac{1}{4}\frac{\mathrm{d}}{\mathrm{d}t}\int_X \varphi(c(t,x))\,\mathrm{d}x = \frac{1}{4}\int_X \varphi'(c)\dot c\,\mathrm{d}x = \int_X E'(c)c\,\mathrm{d}x \int_X \varphi'(c)c\,\mathrm{d}x - \int_X E'(c)\varphi'(c)c\,\mathrm{d}x$$
$$= \int_0^\infty E'(c)\,\mathrm{d}F_t(c)\int_0^\infty \varphi'(c)\,\mathrm{d}F_t(c) - \int_0^\infty E'(c)\varphi'(c)\,\mathrm{d}F_t(c) \overset{*}{\le} 0,$$

where $F_t(b) := \int_X c(t,x)\mathbb{1}_{c(t,\cdot)\le b}(x)\,\mathrm{d}x$ with $0 \le F_t(c) \le F_t(\infty) = 1$. The estimate $\overset{*}{\le}$ is a well-known rearrangement estimate following from the monotonicities of $E'(c)$ and φ', see [HLP34, Ch. 10.13].

Given c_0 with $\underline{c} \le c_0(x) \le \overline{c}$ we set $c_* = \frac{1}{2}\big(\overline{c}+\underline{c}\big)$ and $\delta = \overline{c} - c_*$. For $p \ge 2$ we choose $\varphi(c) = |c-c_*|^p$. Using $\mathcal{L}^d(X) = 1$ again we find $\|c(t)-c_*\|_{\mathrm{L}^p} \le \|c_0-c_*\|_{\mathrm{L}^p} \le \delta$. In the limit $p \to \infty$ we are left with $\|c(t)-c_*\|_{\mathrm{L}^\infty} \le \delta$, which implies $c(t,x) \in [c_*-\delta, c_*+\delta] = [\underline{c},\overline{c}]$ as desired.

3.2 A single minimization step for HK

We will first prove a lemma that works for both spaces and their respective incremental minimization schemes

$$\mu_1 = \inf_{\mu\in\mathcal{M}(X)}\left\{\frac{\mathsf{HK}^2(\mu_0,\mu)}{2\tau} + \mathsf{E}(\mu)\right\}, \qquad \mu_1 = \inf_{\mu\in\mathcal{P}(X)}\left\{\frac{\mathsf{SHK}^2(\mu_0,\mu)}{2\tau} + \mathsf{E}(\mu)\right\}, \tag{3.2}$$

under the sole assumption that the function E, which generates E in (1.1), is convex. We will write $(3.2)_{\mathsf{HK}}$ and $(3.2)_{\mathsf{SHK}}$ to distinguish the two different incremental minimization schemes.

We will start by providing a Lemma that is a generalization of Otto's argument in [Ott96]. Let μ_1 the one step solution to either of the minimization schemes (3.2). According to the Lemma, mass can be transferred from a set A to some set $\boldsymbol{T}(A)$, only if it results in a situation where the density $\rho_1 = \frac{\mathrm{d}\mu_1}{\mathrm{d}\mathcal{L}^d}$ at the destination $\boldsymbol{T}(A)$, is less than the destiny in the origin A. In the case of the Wasserstein distance where mass can only be transported, this is enough to prove bounds for ρ_1, by studying the set where the density is below the minimum of ρ_0. However, for the case of SHK the arguments are a bit more involved, while for HK, we even need the extra Assumption (B) to get the lower bound.

Lemma 3.1 *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore let E be as in* (1.1) *with E being a convex real function. Let finally $\mu_0 \simeq \mathcal{L}^d$, and $\mu_1 \in \mathcal{M}(X)$ as in* $(3.2)_{\mathsf{HK}}$ *or*$(3.2)_{\mathsf{SHK}}$. *Then either $\mu_1 \equiv 0$ or $\mu_1 \simeq \mathcal{L}^d$ and for the injective optimal dilation-transportation couple $(q,\boldsymbol{T})$ from μ_0 to μ_1 provided by Theorem 2.6, we have*

$$\rho_1(\boldsymbol{T}(x)) \le \rho_1(x) \quad \big(\textit{alternatively } \rho_1(\boldsymbol{T}^{-1}(x)) \ge \rho_1(x)\big), \quad \textit{almost everywhere.} \tag{3.3}$$

Proof. We are going to prove the statement in two parts. First we are going to show that $\mu_1 \ll \mathcal{L}^d$, with $\rho_1(\boldsymbol{T}(x)) \le \rho_1(x)$. On the second part we are going to prove $\mathcal{L}^d \ll \mu_1$.

Step A: Proving $\mu_1 \ll \mathcal{L}^d$ and $\rho_1(\boldsymbol{T}(x)) \le \rho_1(x)$. Constructing the counterexample.

We note that since $\mu_0 \simeq \mathcal{L}^d$, it exits a transportation dilation couple from μ_0 to μ_1. In case $\mu_1^s \not\equiv 0$, it exists set B such that $\mu_1^s(B) \neq 0$, and $\mathcal{L}^d(B) = 0$. The identity

$$0 < \mu_1(B) = \int_B \mu_1(dx) = \int_{T^{-1}(B)} q^2(x)\mu_0(dx), \tag{3.4}$$

guarantees that for the set $A = \boldsymbol{T}^{-1}(B)$ we have that $\mu_0(A) \neq 0$ and therefore $\mathcal{L}^d(A) > 0$. Having in mind to generate a contradiction to the assumption, we get that in both cases where either $\mu_1^s \not\equiv 0$, or $\mu_1^s \equiv 0$, if the assumption is violated then there exists $a < b$ and a set A with $\mathcal{L}^d(A) > 0$ such that $\rho_1(x) < a < b < \rho_1(\boldsymbol{T}(x))$ for every $x \in A$. When $\mu_1^s(\boldsymbol{T}(A)) \neq 0$, while $\mathcal{L}^d(\boldsymbol{T}(A)) = 0$, we set $\rho_1(\boldsymbol{T}(x))$ equal to ∞. We define μ_1^* to be the restriction of μ_1 onto $\boldsymbol{T}(A)$, and μ_0^* the restriction of μ_0 onto A. For $0 < t < 1$, we define the measure $\mu_1^t = \mu_1 - t\mu_1^* + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*$, which satisfies $\mu_1^t(X) = \mu_1(X)$. Moreover, by assumption we have $A \cap \boldsymbol{T}(A) = \emptyset$ and can decompose μ_1^t as

$$\mu_1^t = \mu_1\big|_{X\setminus(A\cap\boldsymbol{T}(A))} + (1-t)\mu_1\big|_{\boldsymbol{T}(A)} + \Big(\mu_1\big|_A + t\frac{\mu_1(\boldsymbol{T}(A))}{\mu_0(A)}\mu_0\big|_A\Big).$$

Thus, nothing is changed on $X \setminus (A \cap \boldsymbol{T}(A))$, while mass is taken away on $\boldsymbol{T}(A)$ proportional to $\mu_1^* = \mu_1|_{\boldsymbol{T}(A)}$ and added on A proportional to $\mu_0^* = \mu_0|_A$.

We will prove that the HK distance between μ_0 and μ_1 is not smaller than the resulting cost for this new dilation-transportation couple and therefore not smaller than $\mathsf{HK}(\mu_0, \mu_1^t)$. At the same time, we will show that $\mathsf{E}(\mu_1^t) < \mathsf{E}(\mu_1)$, for small enough t, leading this way to a contradiction.

Step A.1: Proving that the constructed measure is closer to μ_0. By applying Lemma 2.8 for A and $X \setminus A$, we obtain By applying Lemma 2.8 for A and $X \setminus A$, we obtain

$$\begin{aligned}
\mathsf{HK}^2(\mu_0, \mu_1) &= \mathsf{HK}^2(\mu_0-\mu_0^*, \mu_1-\mu_1^*) + \mathsf{HK}^2(\mu_0^*, \mu_1^*) \\
&= \mathsf{HK}^2(\mu_0-\mu_0^*, \mu_1-\mu_1^*) + \mathsf{HK}^2\left(\mu_0^*, (1-t)\mu_1^* + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*\right) \\
&\qquad + \mathsf{HK}^2(\mu_0^*, \mu_1^*) - \mathsf{HK}^2\left(\mu_0^*, (1-t)\mu_1^* + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*\right) \\
&\overset{\text{subadd.}}{\geq} \mathsf{HK}^2(\mu_0, \mu_1^t) + \mathsf{HK}^2(\mu_0^*, \mu_1^*) - \mathsf{HK}^2\left(\mu_0^*, (1-t)\mu_1^* + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*\right) \\
&\overset{\text{subadd.}}{\geq} \mathsf{HK}^2(\mu_0, \mu_1^t) + \mathsf{HK}^2(\mu_0^*, \mu_1^*) - \mathsf{HK}^2\left((1-t)\mu_0^*, (1-t)\mu_1^*\right) - \mathsf{HK}^2\left(t\mu_0^*, t\frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*\right) \\
&\overset{(2.12)}{\geq} \mathsf{HK}^2(\mu_0, \mu_1^t) + t\left(\mathsf{HK}^2(\mu_0^*, \mu_1^*) - \mathsf{HK}^2\left(\mu_0^*, \frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*\right)\right) \geq \mathsf{HK}^2(\mu_0, \mu_1^t).
\end{aligned}$$

In the last estimate we used that starting from a measure μ_0, then among all measures with a given mass M, the measure $M\mu_0$ has the least distance. Indeed, we have

$$\begin{aligned}
\mathsf{HK}^2\left(\mu_0^*, \frac{\mu_1^*(X)}{\mu_0^*(X)}\mu_0^*\right) &\overset{(2.12)}{=} \sqrt{\frac{\mu_1^*(X)}{\mu_0^*(X)}}\mathsf{HK}^2(\mu_0^*, \mu_0^*) + \left(\frac{\mu_1^*(X)}{\mu_0^*(X)} - \sqrt{\frac{\mu_1^*(X)}{\mu_0^*(X)}}\right)\mu_0^*(X) \\
&\quad + \left(1-\sqrt{\frac{\mu_1^*(X)}{\mu_0^*(X)}}\right)\mu_0^*(X) = \left(\sqrt{\mu_1^*(X)} - \sqrt{\mu_0^*(X)}\right)^2 \overset{(2.13)}{\leq} \mathsf{HK}^2(\mu_0^*, \mu_1^*)
\end{aligned}$$

We will treat the case where μ_1 has a singular part and the case $\mu_1 \ll \mathcal{L}^d$ separately.

Step A.2.1: Entropy functional estimate; the case $\mu^s \equiv 0$

$$\begin{aligned}\mathsf{E}(\mu_1^t) - \mathsf{E}(\mu_1) &= \int_X \big(E(\rho_1^t(x)) - E(\rho_1(x))\big)\,\mathcal{L}^d(\mathrm{d}x)\\ &\leq \int_X E'(\rho_1^t(x))\,\big(\rho_1^t(x) - \rho_1(x)\big)\,\mathcal{L}^d(\mathrm{d}x)\\ &= \int_X \big(E'(\rho_1^t(x) - E'\left(\tfrac{a+b}{2}\right)\big)\,(\rho_1^t(x) - \rho_1(x))\mathcal{L}^d(\mathrm{d}x),\end{aligned} \tag{3.5}$$

where in the last term the constant $E'(\frac{a+b}{2})$ could be inserted because of $\int_X \rho^t \mathrm{d}\mathcal{L}^d = \mu^t(X) = \mu_1(X) = \int_X \rho_1 \mathrm{d}\mathcal{L}^d$. The integrand in the last term of (3.5) can be estimated as follows:

$$\begin{aligned}&\big(E'(\rho_1^t(x)) - E'\left(\tfrac{a+b}{2}\right)\big)\,(\rho_1^t(x) - \rho_1(x))\\ &= t\left(E'\left(\rho_1(x) + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\rho_0^*(x) - t\rho_1^*(x)\right) - E'\left(\frac{a+b}{2}\right)\right)\left(\frac{\mu_1^*(X)}{\mu_0^*(X)}\rho_0^*(x) - \rho_1^*(x)\right)\\ &\leq \begin{cases} t\left(E'\left(a + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\rho_0(x)\right) - E'\left(\frac{a+b}{2}\right)\right)\frac{\mu_1^*(X)}{\mu_0^*(X)}\rho_0(x) & \text{on } A,\\ t\left(E'\left(\frac{a+b}{2}\right) - E'\left(b - t\rho_1(x)\right)\right)(\rho_1(x)) & \text{on } \boldsymbol{T}(A).\end{cases}\end{aligned} \tag{3.6}$$

In both cases the factor multiplying t are negative for very small values of t. Hence, we have shown $\mathsf{E}(\mu^t) < E(\mu_1)$ which is the desired contradiction.

Step A.2.2: Entropy functional estimate; the case $\mu_1^s \not\equiv 0$. Let B such that $\mathcal{L}^d(B) = 0$ and $\mu_1^s(B^c) = 0$

$$\begin{aligned}\mathsf{E}(\mu_1^t) - \mathsf{E}(\mu_1) &= \int_B \big(E(\rho_1^t(x)) - E(\rho_1(x))\big)\,\mathcal{L}^d(\mathrm{d}x) + E'_\infty\mu_1^t(B^c) - E'_\infty\mu_1(B^c)\\ &\leq \int_B E'(\rho_1^t(x))\,\big(\rho_1^t(x) - \rho_1(x)\big)\,\mathcal{L}^d(\mathrm{d}x) + E'_\infty\mu_1^t(B^c) - E'_\infty\mu_1(B^c)\\ &= \int_B \big(E'(\rho_1^t(x)) - E'\left(\tfrac{a+b}{2}\right)\big)\,(\rho_1^t(x) - \rho_1(x))\mathcal{L}^d(\mathrm{d}x)\\ &\quad + \left(E'_\infty - E'\left(\frac{a+b}{2}\right)\right)(\mu_1^t(B^c) - \mu_1(B^c)),\end{aligned} \tag{3.7}$$

where in the last term the constant $E'(\frac{a+b}{2})$ could be inserted because of $\mu^t(X) = \mu_1(X)$. The integrated in the last term of (3.5) can be estimated as follows:

$$\begin{aligned}&\int_B \big(E'(\rho_1^t(x)) - E'\left(\tfrac{a+b}{2}\right)\big)\,(\rho_1^t(x) - \rho_1(x))\mathcal{L}^d(\mathrm{d}x)\\ &\leq \int_A \big(E'(\rho_1^t(x)) - E'\left(\tfrac{a+b}{2}\right)\big)\,(\rho_1^t(x) - \rho_1(x))\mathcal{L}^d(\mathrm{d}x)\\ &\leq \int_A t\left(E'\left(a + t\frac{\mu_1^*(X)}{\mu_0^*(X)}\rho_0(x)\right) - E'\left(\frac{a+b}{2}\right)\right)\frac{\mu_1^*(X)}{\mu_0^*(X)}\rho_0(x),\end{aligned} \tag{3.8}$$

while

$$\mu_1^t(B^c) - \mu_1(B^c) = (\mu_1^t(\boldsymbol{T}(A)) - \mu_1(\boldsymbol{T}(A))) = t\left(E'\left(\frac{a+b}{2}\right) - E'_\infty\right)\mu_1(\boldsymbol{T}(A)) \tag{3.9}$$

In both cases the factor multiplying t are negative for very small values of t. Hence, we have shown $\mathsf{E}(\mu^t) < E(\mu_1)$ which is the desired contradiction.

Step B: Proving $\mu_1 \equiv 0$ or $\mathcal{L}^d \ll \mu_1$. We will assume that $\mu_1 \not\equiv 0$ and that there exists $B = \{x : \rho_1(x) = 0\}$ with $\mathcal{L}^d(B) > 0$ to reach a contradiction.

Since $\mathcal{L}^d(B) > 0$, there exist $x_0 \in X$ and $r_0 \in (0, \frac{\pi}{4})$ such that $\mathcal{L}^d(B(x_0, r_0) \cap B) > 0$ and $B(x_0, r_0) \subset X$. By the assumption $\mu_1 \not\equiv 0$, the set $B^c = X \setminus B$ satisfies $\mathcal{L}^d(B^c) > 0$, and therefore there exist $x_1 \in X$ and $r_1 \in (0, \frac{\pi}{4})$ with $\mathcal{L}^p(B(x_1, r_1) \cap B^c) > 0$ and $B(x_1, r_1) \subset X$.

We set $r_\theta = (1-\theta)r_0 + \theta r_1$ and $x_\theta = (1-\theta)x_0 + \theta x_1$ and observe that the convexity of X implies $B(x_\theta, r_\theta) \subset \mathbb{R}^d$ for $\theta \in [0, 1]$. As μ_1 is absolutely continuous the functions $\beta(\theta) := \mathcal{L}^d(B(x_\theta, r_\theta) \cap B)$ and $\gamma(\theta) := \mathcal{L}^d(B(x_\theta, r_\theta) \cap B^c)$ are continuous and satisfy $\beta(\theta) + \gamma(\theta) = c_d r_\theta^d > 0$ for all $\theta \in [0, 1]$. With $\beta(0) > 0$ and $\gamma(1) > 0$ we conclude that there exists $\theta \in [0, 1]$ such that $\mathcal{L}^d(B(x_\theta, r_\theta) \cap B) > 0$ and $\mathcal{L}^d(B(x_\theta, r_\theta) \cap B^c) > 0$. For economy of notation, we denote $B_0 = B(x_\theta, r_\theta) \cap B$ and $B_1 = B(x_\theta, r_\theta) \cap B^c$.

From the assumption that $\mu_0 \simeq \mathcal{L}^d$, we have $\mu_0(B_0) > 0$. Also by definition of B^c and therefore of B_1, we have $\mu_1(B_1) > 0$. Furthermore it holds that $\sup_{x\in B_0, y\in B_1} |x-y| < 2r_\theta < \frac{\pi}{2}$ which means that μ_1 has positive value in a set that has distance less than $\pi/2$ from points in B_0. More specifically, it holds $B_0 \subset \operatorname{supp}(\mu_0')$, and $B_1 \subset \operatorname{supp}(\mu_1')$, where μ_0' and μ_1' are as in Definition 2.4.

However the optimality conditions in [LMS22], provide $\infty > \sigma_1(T(x)), \sigma_0(x) > 0$ on $B_0 \subset \operatorname{supp}(\mu_0')$ and therefore we know that mass in B_0 could not be destroyed, but instead it must be transferred from B_0 somewhere in B^c. This implies that $\boldsymbol{T}(x) \in B^c$ for all $x \in B_0$. However in Step A1, we proved $\rho_1(\boldsymbol{T}(x)) \leq \rho_1(x)$ almost surely, therefore

$$0 < \rho_1(\boldsymbol{T}(x)) \leq \rho_1(x) = 0 \text{ for almost every } \ x \in B_0.$$

This gives us

$$0 < \int_{B_0} \rho_1(\boldsymbol{T}(x))\mathcal{L}^d \leq \int_{B_0} \rho_1(x)\mathcal{L}^d = \mu_1(B_0) = 0$$

which is a contradiction.

The SHK case: The proof for SHK is exactly the same, because the constructed counterexample has the same mass as the original and because the distance SHK is larger than HK distance, namely $\mathsf{SHK}(\mu_0, \mu_1) = 2\arcsin\left(\frac{1}{2}\,\mathsf{HK}(\mu_0, \mu_1)\right) \geq \mathsf{HK}(\mu_0, \mu_1)$. ∎

We proceed with proving upper and lower bounds for the density ρ_1 of the measure μ_1 defined by (3.2)$_{\mathsf{HK}}$, given that μ_0 satisfies some density bounds of its own, i.e. $c_{\min} < \rho_0 < c_{\max}$. The upper bound we retrieve is the same as in (D3), and the proof is relatively straightforward. We assume that the density ρ_1 of the minimizer μ_1 in (3.2)$_{\mathsf{HK}}$ is bigger than the expected density given by (D3) in a set of positive Lebesgue measure. By applying Lemma 3.1, we infer excessive creation of mass in some set B. We conclude that, what we "gain" in HK distance by restricting the growth on B is more than what we "lose" for the entropy function. Therefore ending up with a measure that has total energy less than μ_1, which is a contradiction. A similar bound was retrieved in [DiC20] for a class of Entropy functionals that were studied in that paper.

Proposition 3.2 (Upper bound for incremental densities) *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore consider* E *as in* (1.1) *with E being convex. Let finally $\mu_0 = \rho_0\mathcal{L}^d$ with $\rho_0(x) \leq c_{\max}$ and $\mu_1 \in \mathcal{M}(X)$ as in* (3.2)$_{\mathsf{HK}}$. *Then, for all $c_{\rm upp} \geq 0$ and $\tau > 0$ with $E'(c_{\rm upp})\tau > -1/2$ we have*

$$\rho_1(x) \leq \max\Big\{c_{\rm upp}, \frac{c_{\max}}{\left(1 + 2\tau\min\{E'(c_{\rm upp}), 0\}\right)^2}\Big\}. \tag{3.10}$$

Proof. We start by setting

$$k_\epsilon := \frac{1}{1+2\tau\ \min\{E'(c_{\rm upp}),0\}}\ +\epsilon \geq 1+\epsilon.$$

Step 1: Construction of sets. In order to arrive a a contradiction we define the set

$$B := \big\{\, x\in X \;\big|\; \rho_1(x)\geq \xi_\epsilon \,\big\} \quad \text{with } \xi_\epsilon := \max\{(1+\epsilon)^2 c_{\rm upp}, k_\epsilon^2 c_{\max}\}$$

and assume $\mu_1(B)>0$. By Lemma 3.1, for almost every $x\in \mathbf{T^{-1}}(B)$, we have

$$\rho_1(\boldsymbol{T}^{-1}(x)) \geq \rho_1(x)\geq \xi_\epsilon,$$

which yields $x\in B$, and we conclude $\boldsymbol{T}^{-1}(B)\subset B$. By using this we find

$$\begin{aligned}\mu_1(B) &= \int_B \rho_1(x)\mathcal{L}^d(\mathrm{d}x) \geq \int_{\boldsymbol{T}^{-1}(B)} \rho_1(x)\mathcal{L}^d(\mathrm{d}x)\\ &\geq \int_{\boldsymbol{T}^{-1}(B)} k_\epsilon^2 c_{\max}\mathcal{L}^d(\mathrm{d}x) \geq \int_{\boldsymbol{T}^{-1}(B)} k_\epsilon^2 \rho_0(x)\mathcal{L}^d(\mathrm{d}x) = k_\epsilon^2\mu_0(\boldsymbol{T}^{-1}(B)),\end{aligned}$$

where in the last estimates we applied both $\rho_1(x)\geq k_\epsilon^2 c_{\max}$ and $c_{\max}\geq \rho_0(x)$. By combining the estimate above with the definition of the injective optimal transport couple we have,

$$k_\epsilon^2\mu_0(\boldsymbol{T}^{-1}(B)) \leq \mu_1(B) = \int_{\boldsymbol{T}^{-1}(B)} q^2\mu_0(\mathrm{d}x),$$

which leads to the existence of a set $A\subset \boldsymbol{T}^{-1}(B)$, for which $\mu_0(A)>0$ and $q^2(x)\geq k_\epsilon^2$ for a.a. $x\in A$.

Step 2: Comparison of HK distances. We denote by μ_0^* the restriction of μ_0 on A, and by μ_1^* the restriction of μ_1 on $\boldsymbol{T}(A)$. Since $(\boldsymbol{T},tq)$ is a dilation transportation couple for $\mu_0^*, t^2\mu_1^*$ for $t\in[0,1]$, we get

$$\mathsf{HK}^2\left(\mu_0^*, t^2\mu_1^*\right) \leq \int_A \big(1+(tq(x))^2-2tq(x)\cos\left(|\boldsymbol{T}(x){-}x|\right)\big)\,\mu_0^*(\mathrm{d}x), \tag{3.11}$$

from which we obtain

$$\begin{aligned}\mathsf{HK}^2\left(\mu_0^*,\mu_1^*\right)-\mathsf{HK}^2\left(\mu_0^*,t^2\mu_1^*\right) &\geq \int_X \big(1+(q(x))^2-2q(x)\cos\left(|\boldsymbol{T}(x){-}x|\right)\big)\,\mu_0^*(\mathrm{d}x)\\ &\quad -\int_X \big(1+(tq(x))^2-2tq(x)\cos\left(|\boldsymbol{T}(x){-}x|\right)\big)\,\mu_0^*(\mathrm{d}x)\\ &= \int_X (1{-}t^2)q^2(x)-2(1{-}t)q(x)\cos\left(|\boldsymbol{T}(x){-}x|\right)\mu_0^*(\mathrm{d}x)\\ &\geq \int_X \left[(1{-}t^2)q^2(x)-\frac{2(1{-}t)}{q(x)}q^2(x)\right]\mu_0^*(\mathrm{d}x) \geq (1{-}t)\left(1+t-\frac{2}{k_\epsilon}\right)\int_X q^2(x)\mu_0^*(\mathrm{d}x)\\ &= (1{-}t)\left(1+t-\frac{2}{k_\epsilon}\right)\mu_1^*(X),\end{aligned} \tag{3.12}$$

where in the last estimate, we applied $q(x)\geq k_\epsilon$. Now, for $0\leq t\leq 1$, we define the measure

$$\mu_1^t := \mu_1-\mu_1^*+t^2\mu_1^* = \mu_1+(t^2{-}1)\mu_1^* = \big(\mathbf{1}_{X\setminus\boldsymbol{T}(A)}+t^2\mathbf{1}_{\boldsymbol{T}(A)}\big)\,\rho_1\mathcal{L}^d.$$

Applying Lemma 2.8 for A and $X \setminus A$, we get

$$\begin{aligned}
&\mathsf{HK}^2(\mu_0,\mu_1) = \mathsf{HK}^2(\mu_0-\mu_0^*,\mu_1-\mu_1^*) + \mathsf{HK}^2\left(\mu_0^*,\mu_1^*\right) \\
&\geq \mathsf{HK}^2(\mu_0-\mu_0^*,\mu_1-\mu_1^*) + \mathsf{HK}^2\left(\mu_0^*,t^2\mu_1^*\right) + \mathsf{HK}^2\left(\mu_0^*,\mu_1^*\right) - \mathsf{HK}^2\left(\mu_0^*,t^2\mu_1^*\right) \\
&\geq \mathsf{HK}^2(\mu_0,\mu_1^t) + (1{-}t)\left(1+t-\frac{2}{k_\epsilon}\right)\mu_1^*(X),
\end{aligned} \tag{3.13}$$

where the last inequality is a result of the sub-additivity of the squared distance and (3.12). So, for the new measure $\mu_1^t = \mu_1 + (t^2{-}1)\mu_1^*$, we have

$$\frac{1}{2\tau}\left(\mathsf{HK}^2(\mu_0,\mu_1) - \mathsf{HK}^2(\mu_0,\mu_1^t)\right) \geq \frac{1{-}t}{2\tau}\left(1+t-\frac{2}{k_\epsilon}\right)\mu_1^*(X). \tag{3.14}$$

For later use we recall that

$$\mu_1^*(X) = \int_X q^2\mu_0^*(\mathrm{d}x) = \int_A q^2\mu_0(\mathrm{d}x) \geq \int_A k_\epsilon^2\mu_0(\mathrm{d}x) \geq (1{+}\epsilon)^2\mu(A) > 0. \tag{3.15}$$

Step 3: Comparison of entropies. On $\boldsymbol{T}(A)$ we have $\rho_1^t = t^2\rho_1 \geq t^2\xi_\epsilon$, and the convexity of E gives

$$\begin{aligned}
\mathsf{E}(\mu_1) - \mathsf{E}(\mu_1^t) &= \int_{\boldsymbol{T}(A)} \left(E(\rho_1(x)) - E(\rho_1^t(x))\right)\mathcal{L}^d(\mathrm{d}x) \\
&\geq \int_{\boldsymbol{T}(A)} E'(\rho_1^t)(\rho_1(x)-\rho_1^t(x))\mathcal{L}^d(\mathrm{d}x) \geq E'(t^2\xi_\epsilon)\int_{\boldsymbol{T}(A)}(\rho_1{-}\rho_1^t)\mathcal{L}^d(\mathrm{d}x).
\end{aligned}$$

For all $t \in \left((1{+}\epsilon)^{-1/2},1\right)$ we have $t^2\xi_\epsilon \geq c_{\mathrm{upp}}$, and therefore $E'(t^2\xi_\epsilon) \geq E'(c_{\mathrm{upp}})$ by the monotonicity of E'. With this we arrive at the lower bound

$$\begin{aligned}
\mathsf{E}(\mu_1) - \mathsf{E}(\mu_1^t) &\geq E'(c_{\mathrm{upp}})\int_{\boldsymbol{T}(A)}(\rho_1(x)-\rho_1^t(x))\mathcal{L}^d(\mathrm{d}x) \\
&= E'(c_{\mathrm{upp}})\int_{\boldsymbol{T}(A)}(1{-}t^2)\rho_1(x)\mathcal{L}^d(\mathrm{d}x) = (1{-}t^2)E'(c_{\mathrm{upp}})\mu_1^*(X).
\end{aligned} \tag{3.16}$$

Step 4: Minimization provides contradiction. Because μ_1 is a minimizer we reach a contradiction if we find a $t \in (0,1)$ such that

$$\eta(t) := \frac{1}{2\tau}\mathsf{HK}^2(\mu_0,\mu_1) + \mathsf{E}(\mu_1) - \left(\frac{1}{2\tau}\mathsf{HK}^2(\mu_0,\mu_1^t) + \mathsf{E}(\mu_1^t)\right) > 0.$$

Combining the estimates (3.14) and (3.16) we find, for $t \in \left((1{+}\epsilon)^{-1/2},1\right)$, the lower estimate

$$\eta(t) \geq \overline{\eta}(t)\mu_1^*(X) \quad \text{with } \overline{\eta}(t) := \frac{1{-}t}{2\tau}\left(1+t-\frac{2}{k_\epsilon}\right) + (1{-}t^2)E'(c_{\mathrm{upp}}).$$

Clearly, we have $\overline{\eta}(1) = 0$ and find

$$\overline{\eta}'(1) = -\frac{1}{\tau}\left(1+2\tau E'(c_{\mathrm{upp}}) - \frac{1}{k_\epsilon}\right) \leq -\frac{1}{\tau}\left(\frac{1}{k_\epsilon{-}\epsilon} - \frac{1}{k_\epsilon}\right) = -\frac{1}{\tau}\,\frac{\varepsilon}{(k_\epsilon{-}\epsilon)k_\epsilon} < 0.$$

Recalling $\mu_1^*(X) > 0$ from (3.15) we obtain $\eta(t) > 0$ for all $t < 1$ that are sufficiently close to $t = 1$. Thus, the assumption $\mu(B) > 0$ must have been false, and we conclude $\rho_1(x) \leq \xi_\epsilon$ a.e. in X. As $\epsilon > 0$ was arbitrary, the assertion is established. ■

Despite what one would expect from (D4) in Section 3.1, we were unable to retrieve a strong lower bound without assumption B. We expect that a proof, in case the results hold true, or a counterexample in case it does not, will provide some interesting insight on the Hellinger-Kantorovich space. Therefore we leave this as our second open question.

Open Question 3.3 *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore let* E *be as in* (1.1) *with convex E. Let finally $\mu_0 \simeq \mathcal{L}^d$ with $\rho_0(x) \geq c_{\min}$, and $\mu_1 \in \mathcal{M}(X)$ as in* (3.2)$_{\mathsf{HK}}$. *Is it true that*

$$\rho_1(x) \geq \max\Big\{c_{\text{low}},\ \frac{c_{\min}}{(1{+}2\tau\max\{E'(c_{\text{low}}),0\})^2}\Big\} \quad \textit{for all } c_{\text{low}} \geq 0\ ? \tag{3.17}$$

We are now proceeding with our second result, which is proved under Assumption B, i.e. there exists $c_{\text{low}} > 0$ such that $E'(c_{\text{low}}) < 0$. The proof, like most in this section, is achieved by contradiction. We assume that ρ_1 falls bellow $c_{\text{low}} > 0$ in some set, then by Lemma 3.1 we get that mass inside that set must have been destroyed. We argue that keeping part of that mass is not only cheaper for the HK distance, but also results in a smaller value for E.

Proposition 3.4 (Lower bound for incremental densities) *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore let* E *be as in* (1.1) *with convex E an satisfying Assumption B. Let finally $\mu_0 \simeq \mathcal{L}^d$, and $\mu_1 \in \mathcal{M}(X)$ as in* (3.2)$_{\mathsf{HK}}$. *Then, we have*

$$\rho_1(x) \geq \min\{c_{\text{low}}, c_{\min}\}. \tag{3.18}$$

Proof. We proceed as in the previous proposition by choosing $\epsilon > 0$ and setting

$$A = A^\epsilon = \left\{ x \in X \;\middle|\; \rho_1(x) \leq \frac{\min\{c_{\text{low}}, c_{\min}\}}{1+\epsilon} \right\}.$$

We assume that $\mathcal{L}^d(A) > 0$ with the intention to produce a contradiction.

By point Lemma 3.1, for $\mathcal{L}^d$ almost every $x \in A$, we have $\rho_1(\boldsymbol{T}(x)) \leq \rho_1(x)$, and therefore $\boldsymbol{T}(x) \in A$. We infer that $\mathcal{L}^d(\boldsymbol{T}(A)) \leq \mathcal{L}^d(A)$, and therefore

$$\begin{aligned} \mu_1(\boldsymbol{T}(A)) &= \int_{\boldsymbol{T}(A)} \rho_1(x)\mathcal{L}^d(\mathrm{d}x) \leq \frac{\min\{c_{\text{low}}, c_{\min}\}}{1+\epsilon}\mathcal{L}^d(\boldsymbol{T}(A)) \\ &\leq \frac{\min\{c_{\text{low}}, c_{\min}\}}{1+\epsilon}\mathcal{L}^d(A) \leq \frac{\mu_0(A)}{1+\epsilon}. \end{aligned} \tag{3.19}$$

Since $\mathcal{L}^d(A) > 0$ and therefore $\mu_0(A) > 0$, in order for (3.19) to hold, it must exist $A^* \subset A$ with $\mu_0(A^*)$, $\mathcal{L}^d(A^*) > 0$, such that $q^2(x) \leq \frac{1}{1+\epsilon}$ for every $x \in A^*$. Now we define μ_0^* the restriction of μ_0 on A^*, and μ_1^*, the restriction of μ_1 in $\boldsymbol{T}(A^*)$. For $0 \leq t \leq 1$, we define the measure $\mu_1^t = \mu_1 + t\mu_0^*$. Our next step is to show $\mathsf{HK}^2(\mu_0, \mu_1) \geq \mathsf{HK}^2(\mu_0, \mu_1^t)$ for $t \in (0, t_*)$ with $t_* := 1 - \mu_1^*(X)^2/\mu_0^*(X)^2$. We note that

$$\mu_1^*(X) = \mu_1^*(\boldsymbol{T}(A)) = \int_A q^2(x)\,\mu_0^*(\mathrm{d}x) \leq \frac{1}{1+\epsilon}\int_A \mu_0^*(\mathrm{d}x) = \mu^*(A) = \mu^*(X).$$

By (2.12) in Theorem 2.9 we have

$$
\begin{aligned}
&\mathsf{HK}^2((1{-}t)\mu_0^*,\mu_1^*)\\
&\overset{(2.12)}{=} \sqrt{1{-}t}\,\mathsf{HK}^2(\mu_0^*,\mu_1^*) + \big(1{-}t - \sqrt{1{-}t}\big)\,\mu_0^*(X) + \big(1-\sqrt{1{-}t}\big)\,\mu_1^*(X)\\
&\overset{(2.12)}{=} \sqrt{1{-}t}\,\mathsf{HK}^2(\mu_0^*,\mu_1^*) - \sqrt{1-t}\,\big(1-\sqrt{1-t}\big)\,\mu_0^*(X) + \big(1-\sqrt{1{-}t}\big)\,\mu_1^*(X)\\
&= \sqrt{1{-}t}\,\mathsf{HK}^2(\mu_0^*,\mu_1^*) \;-\; (1-\sqrt{1{-}t})\,\big(\sqrt{1{-}t}\,\mu_0^*(X) - \mu_1^*(X)\big)\\
&\overset{t<t_*}{\le} \sqrt{1{-}t}\,\mathsf{HK}^2(\mu_0^*,\mu_1^*) < \mathsf{HK}^2(\mu_0^*,\mu_1^*).
\end{aligned}
\tag{3.20}
$$

Combining this with Lemma 2.8 we find

$$
\begin{aligned}
&\mathsf{HK}^2(\mu_0,\mu_1) \overset{\text{Lem. 2.8}}{=} \mathsf{HK}^2(\mu_0{-}\mu_0^*,\mu_1{-}\mu_1^*) + \mathsf{HK}^2(\mu_0^*,\mu_1^*)\\
&\overset{(3.20)}{\ge} \mathsf{HK}^2(\mu_0{-}\mu_0^*,\mu_1{-}\mu_1^*) + \mathsf{HK}^2((1{-}t)\mu_0^*,\mu_1^*)\\
&= \mathsf{HK}^2(\mu_0{-}\mu_0^*,\mu_1{-}\mu_1^*) + \mathsf{HK}^2((1{-}t)\mu_0^*,\mu_1^*) + \mathsf{HK}^2(t\mu_0^*,t\mu_0^*) \ge \mathsf{HK}^2(\mu_0,\mu_1^t),
\end{aligned}
\tag{3.21}
$$

where the last estimate follows from subadditivity. Thus, the HK distance is estimated.

For the energies we use $\rho_1(x) \ge c_{\text{low}}$ for a.a. $x \in A^* \subset A$ and obtain the estimate

$$
\begin{aligned}
&\mathsf{E}(\mu_1^t) - \mathsf{E}(\mu) = \int_{A^*} E(\rho_1(x){+}t\rho_0(x)) - E(\rho_1(x))\mathcal{L}^d(\mathrm{d}x)\\
&\le \int_{A^*} E'(\rho_1(x))t\rho_0(x)\mathcal{L}^d(\mathrm{d}x) \le t\,E'(c_{\text{low}}) \int_{A^*} \rho_0(x)\mathcal{L}^d(\mathrm{d}x) = t\,E'(c_{\text{low}})\mu_0^*(X) < 0.
\end{aligned}
$$

Using this and (3.21) we obtain, for $t \in (0,t_*)$ the estimate

$$
\frac{1}{2\tau}\mathsf{HK}^2(\mu_0,\mu_1) + \mathsf{E}(\mu_1) - \left(\frac{1}{2\tau}\mathsf{HK}^2(\mu;\mu_1^t) + \mathsf{E}(\mu_1^t)\right) \ge -tE'(c_{\text{low}})\mu_0^*(X) > 0,
$$

which is a contradiction to the minimization property of μ_1. Hence, the assumption $\mathcal{L}^d(A) > 0$ is false, and the assertion (3.18) is established. ■

3.3 A single minimization step for SHK

We will proceed with the theorem for the spherical Hellinger Kantorovich describing the propagation of density bounds for the incremental minimization scheme for the gradient system $(\mathcal{P}(X), \mathsf{SHK}, \mathsf{E})$.

For retrieving density bounds for the case where the metric is the spherical Hellinger Kantorovich instead of the standard Hellinger Kantorovich, one does not need to impose any extra assumptions on the derivative of E. The purpose for this assumption in the previous section, was to discourage low values of the density. Since, the functional does not prefer very low density, and moving mass in order to create this density increases the cost of the transportation, one can, by contradictions discard a case like this. However for the case of the spherical Hellinger Kantorovich, one can utilize the fact that the total mass remain fixed in order to negate the existence of such a possibility. Indeed the main argument for the spherical Hellinger Kantorovich goes as follows. If the assumption for the lower bound of ρ_1 is violated, then the fact that the mass remains constant, guarantees the existence of two sets A, B of positive measure that lead to the contradiction. More specifically, for mass leaving A, we have growth (i.e. $q > 1$), with a resulting density at the target bigger than $c_{\min}$. At the same time for B, we have that the final density is strictly smaller than $c_{\min}$ and part of the mass that left B was destroyed, (i.e. $q < 1$). One can show that it is cheaper to reduce the growth of the mass leaving A and going $\boldsymbol{T}(A)$ resulting in less density in $\boldsymbol{T}(A)$, where at the same time for the mass leaving B we retain a portion at place instead of destroying it during the transportation, which again leads to a cheaper cost. This way we can construct a new measure that contradicts the optimality of μ_1.

Proposition 3.5 (Density bounds for SHK) *Let X be the closure of an open and convex subset of $\mathbb{R}^d$. Furthermore consider a finite doubling measure $\mathcal{L}^d \in \mathcal{M}(X)$ with $\mathcal{L}^d \simeq \mathcal{L}$ and let* E *be as in* (1.1) *with convex E. Let finally $\mu_0 = \rho\mathcal{L}^d$ with $c_{\min} \leq \rho_0(x) \leq c_{\max}$. and $\mu_1 \in \mathcal{P}(X)$ as in* (3.2)$_{\mathsf{SHK}}$. *Then, we have the density bounds*

$$c_{\min} \leq \rho_1(x) \leq c_{\max} \quad \textit{almost everywhere in } X. \tag{3.22}$$

Proof. Step 1: Existence of sets A, B that lead to contradiction. Let $(q, \boldsymbol{T})$ be an optimal dilation-transportation couple from μ_0 to μ_1. To prove the last statement we will need that if the set $\{x : \rho_1(x) < c_{\min}\}$ has positive measure, then we could find $a > 0$, and two sets A, B, such that

1. $\forall x \in A,\ q^2(x) > 1 + a$ and $\rho_1(\boldsymbol{T}(x)) > (1 + a)c_{\min}$.
2. $\forall x \in B,\ q^2(x) \leq 1 - a$ and $\rho_1(x) \leq (1 - a)\, c_{\min}$.

If this is true then we can construct a new measure $\mu_1^{t,s}$ with unit mass, but with less density at $\boldsymbol{T}(A)$ and more density at B, resulting in a lower value in the minimizing scheme than μ_1. This will violate the assumption that μ_1 is a minimizer.

Since $\{x : \rho_1(x) < c_{\min}\}$ has positive measure, it is straightforward to show that it exists $a_1 > 0$ such that for $\tilde{a} \leq a_1$ the set $\{x : \rho_1(x) \leq (1 - \tilde{a})\, c_{\min}\}$ has also positive measure. Now, by Lemma 3.1, it holds $\rho_1(\boldsymbol{T}(x)) \leq \rho_1(x)$, for $\mathcal{L}^d$ almost every $x \in X$, and therefore

$$A^{\tilde{a}} = \{x : \rho_1(\boldsymbol{T}(x))) \leq (1 - \tilde{a})\, c_{\min}\} \supset \{x : \rho_1(x) \leq (1 - \tilde{a})\, c_{\min}\} \tag{3.23}$$

has also positive measure. By using Lemma 3.1 once more, we obtain that $\rho_1(\boldsymbol{T}(\boldsymbol{T}(x))) \leq \rho_1(\boldsymbol{T}(x))$, from which, we can deduce that $\boldsymbol{T}(x) \in A^{\tilde{a}}$ for $\mathcal{L}^d$ almost every $x \in A^{\tilde{a}}$, and therefore $\mathcal{L}^d(\boldsymbol{T}(A^{\tilde{a}})) \leq \mathcal{L}^d(A^{\tilde{a}})$. Now that gives us

$$\begin{aligned} \mu_1(\boldsymbol{T}(A^{\tilde{a}})) = \int_{\boldsymbol{T}(A^{\tilde{a}})} \rho_1(y)\mathcal{L}^d(\mathrm{d}y) &\leq [(1 - \tilde{a})\, c_{\min}]\, \mathcal{L}^d(\boldsymbol{T}(A_0^{\tilde{a}})) \\ &\leq (1 - \tilde{a}) \left(c_{\min}\mathcal{L}^d(A_0^{\tilde{a}})\right) \leq (1 - \tilde{a})\, \mu(A^{\tilde{a}}). \end{aligned} \tag{3.24}$$

By refining (3.24), we can also obtain that

$$\mu_1(\boldsymbol{T}(A^0)) < \mu_0(A^0). \tag{3.25}$$

Indeed we have,

$$\begin{aligned} &\mu_1(\boldsymbol{T}(A^0)) = \int_{\boldsymbol{T}(A^0)} \rho_1(y)\, \mathcal{L}^d(\mathrm{d}y) = \int_{\boldsymbol{T}(A^0)\setminus \boldsymbol{T}(A_0^{a_1})} \rho_1(y)\mathcal{L}^d(\mathrm{d}y) + \int_{\boldsymbol{T}(A^{a_1})} \rho_1(y)\mathcal{L}^d(\mathrm{d}y) \\ &= \int_{\boldsymbol{T}(A\setminus A^{a_1})} \rho_1(y)\mathcal{L}^d(\mathrm{d}y) + \int_{\boldsymbol{T}(A^{a_1})} \rho_1(y)\mathcal{L}^d(\mathrm{d}y) \\ &\leq c_{\min}\mathcal{L}^d(\boldsymbol{T}(A^0 \setminus A^{a_1})) + (1 - a_1)\, c_{\min}\mathcal{L}^d(\boldsymbol{T}(A^{a_1})) \\ &< c_{\min} \left(\mathcal{L}^d(\boldsymbol{T}(A^0 \setminus A^{a_1})) + \mathcal{L}^d(\boldsymbol{T}(A^{a_1}))\right) = c_{\min}\mathcal{L}^d(\boldsymbol{T}(A^0)) \leq c_{\min}\mathcal{L}^d(A^0) \leq \mu_0(A^0). \end{aligned}$$

Now we set $B^{\tilde a} = \{x : \rho_1(\boldsymbol{T}(x)) > (1+\tilde a)\, c_{\min}\}$, which gives $B^{\tilde a} \cup A^{-\tilde a} = X$. Since (3.25) is true, we can find $a_2 > 0$, such that $\mu_1(\boldsymbol{T}(A^{-a_2})) < \mu_0(A^{-a_2})$. Now we have

$$\mu_0(X) = \mu_0(B^{a_2}) + \mu_0(B^{-a_2}),\ \text{and}\ \mu_1(X) = \mu_1(\boldsymbol{T}(X)) = \mu_1(\boldsymbol{T}(B^{a_2})) + \mu_1(\boldsymbol{T}(A^{-a_2})).$$

By combining the inequalities we can find $a_2 > a_3 > 0$ for which $\mu_1(\boldsymbol{T}(B^{a_2})) > (1+a_3)\, \mu_0(B^{a_2})$. From (3.24) we can conclude that there exists a subset of A^{a_2}, with positive $\mathcal{L}^d$ measure, for which

$$q^2(x) > (1+a_3) \quad \text{and} \quad \rho_1(\boldsymbol{T}(x)) > (1+a_2)\, c_{\min} > (1+a_3)\, c_{\min}.$$

To find B we repeat the same procedure that we used with $A^{\tilde a}$, but with the sets $\{x : \rho_1(x) < (1-\tilde a)\, c_{\min}\}$.

Step 2: Construction of counterexample. Now, we define μ_0^A the restriction of μ_0 on A, $\mu_1^{\boldsymbol{T}(A)}$ the restriction of μ_1 on $\boldsymbol{T}(A)$, μ_0^B the restriction of μ_0 on B, and $\mu_1^{\boldsymbol{T}(B)}$ the restriction of μ_1 on $\boldsymbol{T}(B)$. Let $\mu_1^{t,s}$, defined by

$$\mu_1^{t,s} = \mu_1 - s\mu_1^{\boldsymbol{T}(A)} + t\mu_0^B.$$

Using the splitting from Lemma 2.8 we obtain

$$\begin{aligned}
\mathsf{HK}^2(\mu_0,\mu_1) &= \mathsf{HK}^2(\mu_0 - \mu_0^A - \mu_0^B, \mu_1 - \mu_1^{\boldsymbol{T}(A)} - \mu_1^{\boldsymbol{T}(B)}) + \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + \mathsf{HK}^2(\mu_0^B, \mu_1^{\boldsymbol{T}(B)})\\
&\geq \mathsf{HK}^2(\mu_0 - \mu_0^A - \mu_0^{\boldsymbol{T}(A)}, \mu_1 - \mu_1^A - \mu_1^{\boldsymbol{T}(B)}) + \mathsf{HK}^2(\mu_0^A, (1{-}s)\mu_1^{\boldsymbol{T}(A)})\\
&\quad + \mathsf{HK}^2((1{-}t)\mu_0^{\boldsymbol{T}(B)}, \mu_1^{\boldsymbol{T}(B)}) + \mathsf{HK}^2(t\mu_0^B, t\mu_0^B)\\
&\quad + \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + \mathsf{HK}^2(\mu_0^B, \mu_1^{\boldsymbol{T}(B)}) - \mathsf{HK}^2(\mu_0^A, (1{-}s)\mu_1^{\boldsymbol{T}(A)}) - \mathsf{HK}^2((1{-}t)\mu_0^B, \mu_1^{\boldsymbol{T}(B)})\\
&\geq \mathsf{HK}^2(\mu_0, \mu_1^{t,s}) + \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) - \mathsf{HK}^2(\mu_0^A, (1{-}s)\mu_1^{\boldsymbol{T}(A)})\\
&\quad + \mathsf{HK}^2(\mu_0^B, \mu_1^{\boldsymbol{T}(B)}) - \mathsf{HK}^2((1{-}t)\mu_0^B, \mu_1^{\boldsymbol{T}(B)}),
\end{aligned}$$

where in the first estimate we used (3.20) and the last one is due to subadditivity.

By Theorem 2.9, for $0 < s < 1 - \frac{1}{(1+a)^2}$, we have

$$\begin{aligned}
&\mathsf{HK}^2(\mu_0^A, (1{-}s)\mu_1^{\boldsymbol{T}(A)})\\
&= \sqrt{1{-}s}\,\mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + \left(1 - \sqrt{1{-}s}\right)\mu_0^A(X) + \left(1{-}s - \sqrt{1{-}s}\right)\mu_1^{\boldsymbol{T}(A)}(X)\\
&\leq \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + (1 - \sqrt{1{-}s})\mu_1^{\boldsymbol{T}(A)}(X)\left(\frac{\mu_0^A(X)}{\mu_1^{\boldsymbol{T}(A)}(X)} - \sqrt{1{-}s}\right)\\
&\leq \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + (1 - \sqrt{1{-}s})\mu_1^{\boldsymbol{T}(A)}(X)\left(\frac{1}{1{+}a} - \sqrt{1{-}s}\right) < \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}).
\end{aligned}$$

Similarly for $t < a(1-a)$, we have

$$\begin{aligned}
\mathsf{HK}^2((1{-}t)\mu_0^{**}, \mu_1^{**}) &= \sqrt{1{-}t}\,\mathsf{HK}^2(\mu_0^*, \mu_1^*) + \left(1{-}t - \sqrt{1{-}t}\right)\mu_0^*(X) + \left(1 - \sqrt{1{-}t}\right)\mu_1^*(X)\\
&\leq \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + (1 - \sqrt{1{-}t})\mu_0^A(X)\left(\frac{\mu_1^{\boldsymbol{T}(A)}(X)}{\mu_0^A(X)} - \sqrt{1{-}t}\right)\\
&\leq \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}) + (1 - \sqrt{1{-}t})\left(1 - a - \sqrt{1{-}t}\right) < \mathsf{HK}^2(\mu_0^A, \mu_1^{\boldsymbol{T}(A)}).
\end{aligned}$$

Now if we take $0 < t_0 < a(1-a), 0 < s_0 < 1 - \frac{1}{(1+a)^2}$ such that $s\mu_1^{\boldsymbol{T}(A)} = t\mu_0^B$, we create a measure $\mu^{t,s}(X) = 1$, and $\mathsf{HK}^2(\mu_0,\mu_1) \geq \mathsf{HK}^2(\mu_0, \mu_1^{t,s})$ To see that $\mathsf{E}(\mu_1^{t,s}) < \mathsf{E}(\mu_1)$, we repeat the arguments from Lemma 3.1. ∎

3.4 Many minimization steps

Applying the MM scheme means to apply the minimization problems (3.2) iteratively. If we repeat the minimization problems, we will show that the density bounds are such that we keep good a priori bounds the depend only on the actual time $t = n\tau$ but not on the number of steps.

For the MM scheme of $(\mathcal{M}(X), \mathsf{HK}, \mathsf{E})$ we have derived the upper and lower bounds for ρ_1 that depend on

$$\underline{\rho}_0 := \operatorname{ess\,inf} \rho_0 \quad \text{and} \quad \overline{\rho}_0 = \operatorname{ess\,sup} \rho_0$$

in the form

$$\mathsf{HK}: \qquad \min\big\{\underline{\rho}_0, c_{\text{low}}\big\} \leq \rho_1(x) \leq \max\Big\{a, \frac{\overline{\rho}_0}{(1 + 2\tau \min\{E'(a), 0\})^2}\Big\}, \tag{3.26}$$

$$\mathsf{SHK}: \qquad \underline{\rho}_0 \leq \rho_1(x) \leq \overline{\rho}_0, \tag{3.27}$$

where $a > 0$ is arbitrary as long as $2\tau E'(a) > -1$, and $c_{\text{low}} > 0$ is from Assumption B, which is needed for (3.26) but not for (3.27).

Thus, when constructing μ_n^τ by the MM scheme (3.2) we easily can apply these bounds and obtain $\mu_n^\tau = \rho_n \mathcal{L}^d$ with the corresponding density bounds. If E' has a positive root $c_* > 0$, then changes sign, we immediately obtain a global bound for all iterates in the form

$$\min\{c_*, \underline{\rho}_0\} \leq \rho_k(x) \leq \max\{c_*, \overline{\rho}_0\} \quad \text{a.e. in } X.$$

Thus, the more difficult cases are when E is either strictly decreasing (i.e. $E'(c) < 0$ for all $c > 0$) of strictly increasing (i.e. $E'(c) > 0$ for all $c > 0$). Both these cases can occur and are relevant, e.g. for the choices $E(c) = -\sqrt{c}$ or $E(c) = c^2$.

Also in these general cases we are able to provide suitable upper and lower density bounds that only depend on $k\tau$, which is the original time in the gradient-flow equation.

Proposition 3.6 (General lower and upper density bounds for HK) *Assume that $E : [0, \infty) \to \mathbb{R}$ is lsc, convex. Assume $0 < \underline{\rho}_0 \leq \rho(x) \leq \overline{\rho}_0 < \infty$ and set and set*

$$\overline{S} := \inf\big\{\, E'(c) \,\big|\, c \geq \overline{\rho}_0 \,\big\} > -\infty.$$

Assume further $\tau\overline{S} \geq -1/4$, then for all $k \in \mathbb{N}$ we have $\mu_k^\tau = \rho_k^\tau \mathcal{L}^d$ where ρ_k^τ satisfies, for all $k \in \mathbb{N}$, the general density bounds

$$\rho_k^\tau(x) \leq \overline{\rho}_0\, \mathrm{e}^{8\max\{-\overline{S},0\}\, k\tau} \text{ a.e. in } X. \tag{3.28}$$

If furthermore Assumption B is satisfied then

$$\rho_k^\tau(x) \geq \min\{\underline{\rho}_0, c_{\text{low}}\} \tag{3.29}$$

Proof. The result follows essentially by iterating (3.26).

We are going to prove only the upper bound since the lower one is trivial. We construct a nondecreasing sequence (a_k) via $a_0 = \overline{\rho}_0$ and the recursion $a_k \mapsto a_{k+1} = a_k/(1{+}2\tau \min\{E'(a_k), 0\})^2 \geq a_k$. Using the upper estimate in (3.26) an an induction over k, we obtain for $\overline{\rho}_k := \operatorname{ess\,sup} \rho_k^\tau$, the estimate

$$\rho_k^\tau(x) \leq \overline{\rho}_k \leq a_k \quad \text{for all } k \in \mathbb{N}.$$

However, the monotonicities of (a_k) and of E' imply

$$\min\{E'(a_k),0\} \geq \min\{E'(a_k),0\} = \min\{\overline{S},0\} = .$$

Thus, the recursion for a_k implies

$$a_k \leq \frac{a_0}{\big(1{+}2\tau\min\{\overline{S},0\}\big)^{2k}} \overset{*}{\leq} a_0\,\big(1{-}4\tau\min\{\overline{S},0\}\big)^{2k} \leq \overline{\rho}_0\,\mathrm{e}^{8\max\{-\overline{S},0\}\,k\tau},$$

where in $\overset{*}{\leq}$ we used $\tau\overline{S} \geq -1/4$ and the estimate $(1{-}r)^{-2} \leq 1{+}2r$ for all $r \in [0,1/2]$. ■

The upper bound can be improved from exponential in $t = k\tau$ into quadratic, if we impose a suitable lower bound for E', namely $E'(c) \geq -e_*/\sqrt{c}$ for $c \geq c_*$ and some $e_* \geq 0$.

The following result will not be needed for the rest of the paper, however, it is instructive because the negative functionals E with $E(c) = -a_* c^q$ play a special role later (see Proposition 5.5), and there we have the restriction $q \in [\frac{d}{d+2},\frac{1}{2}]$, which exactly provides $E'(c) = -a_* q c^{q-1} \geq -a_* q/\sqrt{c}$ for $c \geq 1$ because of $q{-}1 \leq -\frac{1}{2}$.

Proposition 3.7 (Iteration of the upper bound for HK) *Assume that $E : [0,\infty) \to \mathbb{R}$ is lsc, convex, and satisfies $E'(c) \geq -e_*/\sqrt{c}$ for $c \geq c_*$. Then, $\overline{\rho}_k = \operatorname{ess\,sup} \rho_k^\tau$ satisfies*

$$\overline{\rho}_k \leq \Big(\sqrt{\max\{\overline{\rho}_0, c_*, 4\tau^2 e_*^2\}} + 4e_*\,k\tau\Big)^2 \quad \textit{for all } k \in \mathbb{N}, \tag{3.30}$$

Proof. We apply the upper bound for HK iteratively using $\overline{\rho}_k = \operatorname{ess\,sup}\rho_k$ with $\overline{\rho}_0 = c_{\max}$ by choosing suitable $a = a_k$. Clearly, the estimate is monotone in $\overline{\rho}_k$, hence we may replace $\overline{\rho}_0$ by $\max\{\overline{\rho}_0, c_*, 4\tau^2 e_*^2, c_{max}\}$. Thus, we obtain $\overline{\rho}_k \geq \overline{\rho}_0 \geq c_*$ and choosing $a_k = \overline{\rho}_k$ is admissible because $2\tau E'(a_k) \geq -2\tau e_*/\sqrt{\overline{\rho}_k} \geq -1/2 > -1$. Moreover, we obtain

$$\sqrt{\overline{\rho}_{k+1}} \leq \frac{\sqrt{\overline{\rho}_k}}{1+2\tau e_*/\sqrt{\overline{\rho}_k}} \leq \sqrt{\overline{\rho}_k}\,\Big(1+4\tau e_*/\sqrt{\overline{\rho}_k}\Big) = \sqrt{\overline{\rho}_k} + 4\tau e_*.$$

where we used the estimate $(1{-}\alpha)^{-1} \leq 1+2\alpha$ for $\alpha \in [0,1/2]$. Thus, the desired estimate (3.30) follows. ■

4 Existence for EVI using local κ-concavity

Following [Sav07, LaM19] we first give precise definitions of geodesics curves in a general metric space $(\mathfrak{X},\mathsf{d})$ and the local angle condition (LAC). Based on these fundamental concepts, [Sav11, MuS22] introduces the two geometry-descriptive functions $\langle\cdot,\cdot\rangle_{\mathrm{up}}$ and $\measuredangle(\cdot,\cdot)$ (cf. Definition 4.5) that allow us to quantify the relationship between two geodesics emanating from the same point. Next we discuss semiconvex and semiconcave functions in the sense of geodesic κ-concavity or geodesic λ-convexity. When the squared distance $\frac{1}{2}\mathsf{d}^2_{\mathfrak{X}}(t,x_{\mathsf{ob}})$ is κ-concave along a geodesic $\widehat{\boldsymbol{xy}}$ with respect to some observer x_{ob} the derivative of $t \mapsto \frac{1}{2}\mathsf{d}^2_{\mathfrak{X}}(\widehat{\boldsymbol{xy}}(t),x_{\mathsf{ob}})$ can be estimated in terms of the two quantities mentioned above. Finally, we discuss Evolutionary Variational Inequalities EVI$_\lambda$ for a metric gradient system $(\mathfrak{X},\mathsf{d},\phi)$ and show that solutions can be constructed via the minimizing movement scheme, if ϕ is strongly λ-convex with compact sublevels and the closure of $\mathrm{dom}(\phi)$ can be written as the closure of the union of sets A_κ, where $\frac{1}{2}\mathsf{d}_{\mathcal{X}}^2$ is κ-concave. See Theorem 4.9 for the exact statement.

The main arguments for our existence theory are based on results developed in [Sav07, Sav11] which was a prelude for the work in [MuS20] and [MuS22]. Hence, we provide the corresponding result here without our own proofs expecting a soon publication of the latter work. This contains in particular Proposition 4.7 and the estimates on the Minimizing Movement scheme in Section 4.4. We would like however to remark that we extend some of the results in [MuS22] by weakening the assumption that the functional ϕ must be universally K-concave for some $K > 0$. Instead, we assume that we have concave bounds K_n for a collection of nested sets A_n, whose closure is the domain of ϕ. Proofs that depend on this modified assumption are provided in both versions of the paper.

4.1 Geodesic spaces and the local angle condition (LAC)

We now provide some basic definitions for geodesics in metric spaces and some of their properties.

Definition 4.1 (Geodesics) *Let $(\mathcal{X}, \mathsf{d})$ be a metric space. A curve $\widehat{\boldsymbol{xy}} : [0,1] \to \mathcal{X}$ is called a* (constant-speed) geodesic joining x to y *if*

$$\widehat{\boldsymbol{xy}}(0) = x, \quad \widehat{\boldsymbol{xy}}(1) = y, \quad \textit{and}$$
$$\mathsf{d}(\widehat{\boldsymbol{xy}}(t_1), \widehat{\boldsymbol{xy}}(t_2)) = |t_2 - t_1| \mathsf{d}(x,y) \textit{ for all } t_1, t_2 \in [0,1].$$

We will denote the set of all such geodesics with $\mathrm{Geod}(x,y)$.

The metric space $(\mathcal{X}, \mathsf{d})$ is called a geodesic space*, if for all points $x,\ y \in \mathcal{X}$ the set $\mathrm{Geod}(x,y)$ is nonempty. The metric d is then called a* geodesic distance*.*

For our theory it will be important to introduce the concept of geodesic covers of an arbitrary set. It plays the role of the convexification in Banach spaces, however, our covering notion is not idempotent, i.e. $A^{\mathrm{Geod}} \subsetneq \big(A^{\mathrm{Geod}}\big)^{\mathrm{Geod}}$ is possible.

Definition 4.2 (Geodesic cover) *Let $(\mathcal{X}, \mathsf{d})$ be a metric space and A a subset of $\mathcal{X}$. We define the* geodesic cover A^{Geod} *of A by*

$$A^{\mathrm{Geod}} := \big\{ \widehat{\boldsymbol{xy}}(t) \,\big|\, t \in [0,1],\ \widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x,y),\ x, y \in A \big\}. \tag{4.1}$$

In the following we introduce the two geometric concepts in geodesic metric spaces called the *local angle condition* (LAC) and *κ-concavity*. These properties are going to be utilized in the sequel to prove that the curves occurring by geodesically interpolating the points produced by the minimizing scheme, converge to solutions of the EVI$_\lambda$, when the minimization step τ tends to zero. In order to introduce LAC, we first introduce the notions of comparison angle between three points and of local angle between two geodesics emanating from the same point.

Definition 4.3 (Comparison and local angles) *Let $(\mathcal{X}, \mathsf{d})$ be a metric space. For three points $x, y, z \in \mathcal{X}$ with $x \notin \{y, z\}$, we set*

$$A(x; y, z) := \frac{\mathsf{d}^2(x,y) + \mathsf{d}^2(x,z) - \mathsf{d}^2(y,z)}{2\mathsf{d}(x,y)\mathsf{d}(x,z)} \in [-1, 1]$$

and define the comparison angle $\measuredangle(x; y, z) \in [0, \pi]$ *with vertex* x *via*

$$\measuredangle(x; y, z) := \arccos(A(x; y, z)) \in [0, \pi].$$

Let $\widehat{\boldsymbol{xy}}$ *and* $\widehat{\boldsymbol{xz}}$ *be two geodesics in* $\mathcal{X}$ *emanating from point* $x = \widehat{\boldsymbol{xy}}(0) = \widehat{\boldsymbol{xz}}(0)$. *The* upper angle $\sphericalangle_{\rm up}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) \in [0, \pi]$ *and the* lower angle $\sphericalangle_{\rm lo}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) \in [0, \pi]$, *between* $\widehat{\boldsymbol{xy}}$ *and* $\widehat{\boldsymbol{xz}}$ *are defined by*

$$\begin{aligned} \sphericalangle_{\rm up}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) &:= \limsup_{s,t\downarrow 0} \measuredangle(x, \widehat{\boldsymbol{xy}}(s), \widehat{\boldsymbol{xz}}(t)), \\ \sphericalangle_{\rm lo}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) &:= \liminf_{s,t\downarrow 0} \measuredangle(x; \widehat{\boldsymbol{xy}}(s), \widehat{\boldsymbol{xz}}(t)). \end{aligned} \tag{4.2}$$

If $\sphericalangle_{\rm up}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) = \sphericalangle_{\rm lo}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}})$ *holds, we say that the* local angle *exists in the strict sense and write* $\sphericalangle(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}})$.

The local angle condition concerns three geodesics emanating from one point and states that the sum of the three local angles does not exceed 2π.

Definition 4.4 (Local angle condition) *We say that a point* x *of a geodesic metric space* $(\mathcal{X}, \mathsf{d})$ *satisfies* LAC*, if for any three geodesics* $\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}, \widehat{\boldsymbol{xw}}$ *emanating from* x, *we have*

$$\sphericalangle_{\rm up}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) + \sphericalangle_{\rm up}(\widehat{\boldsymbol{xz}}, \widehat{\boldsymbol{xw}}) + \sphericalangle_{\rm up}(\widehat{\boldsymbol{xw}}, \widehat{\boldsymbol{xy}}) \le 2\pi.$$

In geodesic spaces $(\mathcal{X}, \mathsf{d})$ the set of all geodesics emanating from a point x may be considered as a (nonlinear) surrogate of the tangent space defined in the case of manifolds. We now introduce a kind of scalar product between two geodesics emanating from one point as a generalization of the classical inner product on the tangent space. Moreover, we define the function $\mathbb{A}$ that measures how much two geodesic curves emanating from one point are exactly opposite to each other.

Definition 4.5 (Comparing two geodesics) *In the geodesic space* $(\mathcal{X}, \mathsf{d})$ *consider three points* $x, y, z \in \mathcal{X}$ *and two geodesics* $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$, $\widehat{\boldsymbol{xz}} \in \mathrm{Geod}(x, z)$ *and define*

$$\begin{aligned} \langle \widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}} \rangle_{\rm up} &:= \mathsf{d}(x, y)\mathsf{d}(x, z) \cos(\sphericalangle_{\rm up}(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}})) \\ &= \liminf_{s,t\downarrow 0} \frac{1}{2st} \big(\mathsf{d}^2(x, \widehat{\boldsymbol{xy}}(s)) + \mathsf{d}^2(x, \widehat{\boldsymbol{xz}}(t)) - \mathsf{d}^2(\widehat{\boldsymbol{xy}}(s), \widehat{\boldsymbol{xz}}(t))\big) \end{aligned} \tag{4.3}$$

$$\mathbb{A}^2(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) = \mathsf{d}^2(x, y) + \mathsf{d}^2(x, z) + 2\langle \widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}} \rangle_{\rm up} \ge 0. \tag{4.4}$$

The second form of $\langle \widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}} \rangle_{\rm up}$ given in (4.3) is easily derived from the above definition when taking into account that $\cos$ is decreasing on $[0, \pi]$, which turns the limsup in $\sphericalangle_{\rm up}$ into a liminf.

Considering a Hilbert space with scalar product $(x|y)$ and norm $\|z\|$ the geodesic curves are given via $\widehat{\boldsymbol{xy}} = (1{-}t)x + ty$, $\widehat{\boldsymbol{xz}} = (1{-}t)x + tz$ and we easily find

$$\langle \widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}} \rangle_{\rm up} = \big(y{-}x\big|z{-}x\big) \quad \text{and} \quad \mathbb{A}^2(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) = \big\|(y{-}x) - (z{-}x)\big\|^2,$$

i.e. we have $\mathbb{A}^2(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) = 0$ if and only if $x = \frac{1}{2}(y{+}z)$. For general geodesic spaces we may consider a geodesic $\widehat{\boldsymbol{yz}} \in \mathrm{Geod}(y, z)$, choose x as the midpoint $\widehat{\boldsymbol{yz}}(1/2)$, and set

$$\widehat{\boldsymbol{xy}}(t) = \widehat{\boldsymbol{yz}}\,((1{-}t)/2) \quad \text{and} \quad \widehat{\boldsymbol{xz}}(t) = \widehat{\boldsymbol{yz}}\,((1{+}t)/2) \quad \text{for } t \in [0, 1].$$

Then, $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$, $\widehat{\boldsymbol{xz}} \in \mathrm{Geod}(x, z)$ and $\mathbb{A}^2(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) = 0$.

4.2 Semiconvex and semiconcave functions

For a geodesic metric space $(\mathfrak{X}, \mathsf{d})$, we now provide the definition of κ-(semi)concavity and λ-(semi)convexity of a function $\phi : \mathfrak{X} \to (-\infty, \infty]$ along a geodesic $(\mathfrak{X}, \mathsf{d})$ or of a functional $F : \mathfrak{X} \times \mathfrak{X} \to (-\infty, \infty]$, along some geodesic with respect to some observer. For this, we recall that a function $f : [0, 1] \to (-\infty, \infty]$ is called κ-concave or λ-convex, if the mapping $t \mapsto f - \kappa t^2/2$ is concave or $t \mapsto f - \lambda t^2/2$ is convex, respectively. We emphasize that κ and λ can lie in all of $\mathbb{R}$. Subsequently, we will shortly say κ-concave and λ-convex.

For a κ-concave function $f : [0, 1] \to (-\infty, \infty]$ we have the inequality

$$f \ \kappa\text{-concave} \quad \Longrightarrow \quad \forall\, t \in [0, 1] : \ f(t) + \frac{1}{2}\,\kappa\, t(1-t) \geq (1-t)\, f(0) + t\, f(1). \tag{4.5}$$

For the following definition we recall that the elements of $\mathrm{Geod}(x, y)$ are constant-speed geodesics of length and speed $\mathsf{d}(x, y)$.

Definition 4.6 (κ-concavity/convexity) *Let $(\mathfrak{X}, \mathrm{d})$ be a geodesic space, A and B subsets of $\mathfrak{X}$, and $\kappa \in \mathbb{R}$.*

(A) A function $\phi : \mathfrak{X} \to (-\infty, \infty]$ is called κ-concave (convex) on A*, if for all $x, y \in A$ there exists an $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$ such that the function $t \mapsto \phi(\widehat{\boldsymbol{xy}}(t))$ is $\kappa\, \mathsf{d}^2_{\mathfrak{X}}(x, y)$-concave (convex). ϕ is called* strongly κ-concave (convex) on A*, if the previous condition holds for all $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$. If $A = \mathfrak{X}$, then we simply that ϕ is (strongly) κ-concave (convex).*

(B) For $x, y \in \mathfrak{X}$, we say that a functional $F : \mathfrak{X} \times \mathfrak{X} \to (-\infty, \infty]$ is κ-concave (convex) along a geodesic $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$ with respect to the observer $x_{\mathsf{ob}} \in \mathfrak{X}$, if $t \mapsto F(\widehat{\boldsymbol{xy}}(t), x_{\mathsf{ob}})$ is $\kappa\, \mathsf{d}^2(x, y)$-concave (convex). Furthermore, we say that F is κ-concave (convex) in A with respect to observers from B, if for every couple of points $x, y \in A$, there exists a geodesic $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$ such that for every $x_{\mathsf{ob}} \in B$ we have that F is κ-concave (convex) along $\widehat{\boldsymbol{xy}}$ with respect to x_{ob}. We finally say that F is strongly *κ-concave (convex) in A with respect to observers from B, if for every couple of points $x, y \in A$, all geodesics $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$, and all $x_{\mathsf{ob}} \in B$ the function F is κ-concave (convex) along $\widehat{\boldsymbol{xy}}$ with respect to x_{ob}.*

In the previous definition, the points $\widehat{\boldsymbol{xy}}(t)$ don't have to lie in A for $t \in (0, 1)$. Of course, we have $\widehat{\boldsymbol{xy}}(t) \in A^{\mathrm{Geod}}$ for all $t \in [0, 1]$.

A crucial step in the convergence theory of minimizing movement solutions is to exploit the κ-concavity of the squared distance with respect to suitable observers x_{ob}. The point is that one obtains an upper estimate of the upper right Dini derivative $\frac{\mathrm{d}^+}{\mathrm{d}t} \frac{1}{2} \mathsf{d}^2(\widehat{\boldsymbol{xy}}(t), x_{\mathsf{ob}})$ in terms of the two geometric quantities $\langle \cdot, \cdot \rangle_{\mathrm{up}}$ and $\triangle$ introduced in Definition 4.5. Here the upper right Dini derivative is defined via

$$\frac{\mathrm{d}^+}{\mathrm{d}t} \zeta(t) := \limsup_{h \downarrow 0} \frac{1}{h}(\zeta(t+h) - \zeta(t)).$$

For the proof of the following result we again refer to [Sav11, MuS22].

Proposition 4.7 (Differentiation of squared distance along geodesics) *If $(\mathfrak{X}, \mathrm{d})$ is a geodesic space and $x, y \in \mathfrak{X}$ such that $F = \frac{1}{2} \mathsf{d}^2_{\mathfrak{X}}$ is κ-concave along $\widehat{\boldsymbol{xy}} \in \mathrm{Geod}(x, y)$ with respect to an observer x_{ob}, then we have*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} \frac{1}{2} \mathsf{d}^2(\widehat{\boldsymbol{xy}}(t), x_{\mathsf{ob}}) \leq -\langle \widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xx_{\mathsf{ob}}}} \rangle_{\mathrm{up}} + \kappa \mathsf{d}(x, y) \mathsf{d}(\widehat{\boldsymbol{xy}}(t), x) \text{ a.e. on } [0, 1]. \tag{4.6}$$

If furthermore $(\mathcal{X}, \mathsf{d})$ *satisfies LAC at* x, *then for a.a.* $t \in [0,1]$ *we have*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}\mathsf{d}^2(\widehat{\boldsymbol{xy}}(t), x_{\mathsf{ob}}) \le \langle \widehat{\boldsymbol{xz}}, \widehat{\boldsymbol{xx_{\mathsf{ob}}}}\rangle_{\mathrm{up}} + \mathsf{d}(x, x_{\mathsf{ob}})\measuredangle(\widehat{\boldsymbol{xy}}, \widehat{\boldsymbol{xz}}) + t\kappa\, \mathsf{d}^2(x,y), \tag{4.7}$$

for every $z \in \mathcal{X}$ *and* $\widehat{\boldsymbol{xz}} \in \mathrm{Geod}(x,z)$.

4.3 EVI$_\lambda$ and construction of solutions

We first recall the standard definitions and results from [MuS20, Sec. 3] and then introduce our notations.

We have the following theorem that provides an alternative form of (1.1) that uses integration instead of differentiation. This form can, in a straightforward manner, be combined with the lower semicontinuity properties of the distance d and of ϕ, thus allowing to show that various limits of EVI$_\lambda$ solutions are again EVI$_\lambda$ solutions.

The two results given in the next theorem are taken from [MuS20, Thm. 3.3+3.5].

Proposition 4.8 (Characterizations and properties of EVI solutions)
(A) A curve $\boldsymbol{x} : [0,T) \to \mathcal{X}$ *satisfies EVI*$_\lambda$ *with respect to* ϕ, *if and only if for all* $x_{\mathsf{ob}} \in \mathrm{dom}(\phi)$ *the two maps* $t \mapsto \phi(\boldsymbol{x}(t))$ *and* $t \mapsto \mathsf{d}^2(\boldsymbol{x}(t), x_{\mathsf{ob}})$ *belong to* $\mathrm{L}^1_{\mathrm{loc}}((0,T))$ *and*

$$\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(t), x_{\mathsf{ob}}) - \frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(s), x_{\mathsf{ob}}) + \int_s^t \Big(\phi(\boldsymbol{x}(r)) + \frac{\lambda}{2}\mathsf{d}^2(\boldsymbol{x}(r), x_{\mathsf{ob}})\Big)\,\mathrm{d}r \le (t-s)\phi(x_{\mathsf{ob}}) \quad \text{for all } s,t \in (0,T) \text{ with } s<t. \tag{4.8}$$

(B) If $\boldsymbol{x} : [0,T) \to \mathcal{X}$ *is an EVI*$_\lambda$ *solution, then* $t \mapsto \phi(\boldsymbol{x}(t))$ *is non-increasing and hence continuous from the right (by lsc of* ϕ*).*

(C) If $\boldsymbol{x}^1, \boldsymbol{x}^2 : [0,\infty) \to \mathcal{X}$ *are EVI*$_\lambda$ *solutions, then we have*

$$\mathsf{d}(\boldsymbol{x}^1(t), \boldsymbol{x}^2(t)) \le \mathrm{e}^{-\lambda(t-s)}\mathsf{d}(\boldsymbol{x}^1(s), \boldsymbol{x}^2(s)) \quad \text{for all } s,t \in [0,T) \text{ with } s<t. \tag{4.9}$$

We are now able to formulate our main existence result for EVI$_\lambda$ solutions for metric gradients systems $(\mathcal{X}, \mathsf{d}, \phi)$ which relies on the geodesic structure of $(\mathcal{X}, \mathsf{d})$, the λ-convexity of the potential ϕ, and some "local κ-concavity" of $\frac{1}{2}\mathsf{d}^2_{\mathcal{X}}(t, x_{\mathsf{ob}})$. The construction of solutions will be done by the minimizing movement scheme and geodesic interpolation. For a given $x \in \mathcal{X}$ and a time step τ the discrete solutions $(x^\tau_n)_{n=0,1,\dots,N}$ of the minimization movement schemes are defined via

$$x^\tau_0 = x \quad \text{and} \quad x^\tau_n \in \arg\min\Big\{ \frac{1}{2\tau}\mathsf{d}^2_{\mathcal{X}}(x, x^\tau_{n-1}) + \phi(x) \;\Big|\; x \in \mathcal{X}\Big\} \text{ for } n \in \mathbb{N}. \tag{4.10}$$

As all our ϕ are λ-convex for some $\lambda \in \mathbb{R}$, the functional $x \mapsto \frac{1}{2\tau}\mathsf{d}^2_{\mathcal{X}}(x, x^\tau_{n-1}) + \phi(x)$ is quadratically bounded from below for all $\tau > 0$ with $\lambda + 1/\tau > 0$. Thus under suitable assumptions on ϕ minimizers exist for sufficiently small τ.

Theorem 4.9 (Existence of EVI$_\lambda$ solutions) *Let* $(\mathcal{X}, \mathsf{d})$ *be a geodesic metric space and* $\phi : \mathcal{X} \to (-\infty, \infty]$ *a strongly* λ*-convex functional with compact sublevels. Further assume that there exists a nested sequence of sets* $A_\kappa \subset \mathrm{dom}(|\partial\phi|)$ *with* $\overline{\cup A_\kappa} = \overline{\mathrm{dom}(\phi)}$, *for which the statements (A1) to (A3) are true:*

(A1) *For every $\kappa \in \mathbb{N}$, we have that $\frac{1}{2}\mathsf{d}^2 : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ is strongly κ-concave in A_κ with observers in A_κ^{Geod}, and satisfies LAC for every point in A_κ.*

(A2) *For all $x_0 \in \cup_{\kappa\in\mathbb{N}} A_\kappa$ there exists κ_0 such that for all $\kappa > \kappa_0$ there exists $T(x_0,\kappa) > 0$ such that for all time steps $\tau > 0$ the n-step minimization scheme $(x_k^\tau)_{k=0,\ldots,n}$ remains in A_κ as long as $n < T/\tau + 1$.*

(A3) *For every $x \in \overline{\mathrm{dom}(\phi)}$, there exists a sequence $x_m \in \cup A_\kappa$ with $x_m \to x$ and $\phi(x_m) \to \phi(x)$.*

Then, for every $x_0 \in \cup A_\kappa$ there exists a unique EVI$_\lambda$ solution $\boldsymbol{x} : [0, T_\infty(x_0)) \to \mathcal{X}$ with $\boldsymbol{x}(0) = x_0$, where $T_\infty(x_0) = \lim_{\kappa\to\infty} T(x_0,\kappa)$ (w.l.o.g. $\kappa \mapsto T(x_0,\kappa)$ is non-decreasing).

If additionally, assumption (A2) holds for all $x_0 \in \cup A_\kappa$ with $T_\infty(x_0) = \infty$, then for all $\widehat{x}_0 \in \overline{\mathrm{dom}(\phi)}$ there exists a complete EVI$_\lambda$ solution $\boldsymbol{x}$ with $\boldsymbol{x}(0) = \widehat{x}_0$.

The proof will be completed in Section 4.5 after the necessary a priori estimates for the discrete minimizing movement solutions are collected next.

4.4 Estimates for the MM scheme

In this subsection we state a few of the results from [Sav11, MuS22] quite explicitly, especially to emphasize the dependence on the semiconvexity parameter λ of ϕ and the semiconcavity parameter κ of $\frac{1}{2}\mathsf{d}^2$. This concerns our Lemma 4.10, Proposition 4.11, Corollary 4.13, and Lemma 4.14.

We first state some very basic estimates that hold true for the discrete solutions of the MM scheme. Using geodesic interpolation, λ-convexity of ϕ and κ-concavity of the squared distance, one then obtains sharper estimates that allows us to control the distance between different approximants.

These estimates will be used later prove the existence of curves that satisfy EVI$_\lambda$, but also to bound the distance of the EVI$_\lambda$ satisfying curve from the approximating curves occurring by geodesically interpolating the points of the minimizing scheme. Before we proceed we are going to define the metric slope $|\partial\phi|$ of ϕ at a point $x \in \mathcal{X}$ via

$$|\partial\phi|(x) = \begin{cases} +\infty & \text{for } x \notin \mathrm{dom}(\phi), \\ 0 & \text{for } x \in \mathrm{dom}(\phi) \text{ is isolated}, \\ \limsup_{y\to x} \frac{\max\{0,\phi(x)-\phi(y)\}}{d(x,y)} & \text{otherwise}. \end{cases} \tag{4.11}$$

Lemma 4.10 (Euler equation for discrete solutions) *For a given $x_{n-1}^\tau \in \mathcal{X}$ and $\tau > 0$, assume that $x_n^\tau \in \mathcal{X}$ is a solution to the minimizing movement scheme* (4.10)*. Then, for every observer $x_{\mathsf{ob}} \in \mathcal{X}$ and all geodesics $\overrightarrow{\boldsymbol{x_n^\tau x_{n-1}^\tau}} \in \mathrm{Geod}(x_n^\tau, x_{n-1}^\tau)$ and $\overrightarrow{\boldsymbol{x_n^\tau x_{\mathsf{ob}}}} \in \mathrm{Geod}(x_n^\tau, x_{\mathsf{ob}})$, we have*

$$\frac{1}{\tau}\langle \overrightarrow{\boldsymbol{x_n^\tau x_{n-1}^\tau}}, \overrightarrow{\boldsymbol{x_n^\tau x_{\mathsf{ob}}}} \rangle_{\mathrm{up}} + \frac{\lambda}{2}\mathsf{d}^2(x_n^\tau, x_{\mathsf{ob}}) + \phi(x_n^\tau) \le \phi(x_{\mathsf{ob}}), \tag{4.12}$$

$$\text{(a)}\ \ |\partial\phi(x_n^\tau)| \le \frac{\mathsf{d}(x_{n-1}^\tau, x_n^\tau)}{\tau}, \quad \text{(b)}\ \ (1{+}\lambda\tau)\frac{\mathsf{d}(x_{n-1}^\tau, x_n^\tau)}{\tau} \le |\partial\phi|(x_{n-1}^\tau), \quad \text{(c)}\ \ (1{+}\lambda\tau)|\partial\phi|(x_n^\tau) \le |\partial\phi|(x_{n-1}^\tau). \tag{4.13}$$

We now provide an estimate of how much a geodesic interpolation of a minimizing scheme deviates from being an EVI$_{\lambda^*}$ solution for the modified $\lambda^* = 2\min\{0,\lambda\} - 2 \leq -2$, where λ is such that ϕ is λ-convex. We note and highlight that λ-convexity of ϕ with $\lambda > 0$ will not be helpful here. The following technical estimate will be used later to prove that interpolating curves $\boldsymbol{x}^\tau$ converge to some curve $\boldsymbol{x}$ when the time steps τ converges to 0. One important feature of the following result is that the error terms Δ_n^τ are independent of the observer point x_{ob}.

Proposition 4.11 (Discrete error estimates) *Let $\lambda \leq 0$, $\tau > 0$ with $1{+}\lambda\tau > 0$, and $x_0^\tau \in A_{\kappa_0}$ be fixed that the discrete solution $\{x_n^\tau\}_{n\in\mathbb{N}}$ of (4.10) satisfies the following: For $T > 0$, let $\kappa \in \mathbb{N}$ be such that $x_n^\tau \in A_\kappa$, for all $n \in \mathbb{N}_0$ with $n < \frac{T}{\tau} + 1$. Then, all geodesic interpolators $\boldsymbol{x}^\tau : [0,\infty) \to \mathfrak{X}$, given by*

$$\boldsymbol{x}^\tau(t) = \overset{\frown}{\boldsymbol{x_n^\tau x_{n+1}^\tau}}((t{-}n\tau)/\tau) \text{ for } t \in [n\tau,(n{+}1)\tau] \text{ with } \overset{\frown}{\boldsymbol{x_n^\tau x_{n+1}^\tau}} \in \mathrm{Geod}(x_n^\tau, x_{n+1}^\tau), \tag{4.14}$$

satisfies, for all $x_{\mathsf{ob}} \in A_\kappa^{\mathrm{Geod}}$ and almost all $t \in [0,T]$, the estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}^\tau(t), x_{\mathsf{ob}}) + \frac{\lambda^*}{2}\mathsf{d}^2(\boldsymbol{x}^\tau(t), x_{\mathsf{ob}}) \leq \phi(x_{\mathsf{ob}}) - \phi(\boldsymbol{x}^\tau(t)) + \Delta^{\tau,\kappa}(t), \tag{4.15}$$

with $\lambda^* = 2\lambda - 2$ and $\Delta^\tau(t) = \Delta_n^\tau$ for $t \in [n\tau,(n{+}1)\tau)$, where

$$\Delta_n^{\tau,\kappa} = \begin{cases} (1{-}2\lambda)\mathsf{d}^2(x_0^\tau, x_1^\tau) + (1 + (1{+}\lambda\tau)^{-1})\,|\partial\phi|^2(x_0^\tau) & \text{for } n = 0, \\ (1{-}2\lambda{+}\kappa/\tau)\,\mathsf{d}^2(x_n^\tau, x_{n+1}^\tau) + \frac{1}{\tau^2}\measuredangle^2\Big(\overset{\frown}{\boldsymbol{x_n^\tau x_{n-1}^\tau}}, \overset{\frown}{\boldsymbol{x_n^\tau x_{n+1}^\tau}}\Big) & \text{for } n \in \mathbb{N}. \end{cases} \tag{4.16}$$

The next result exploits the strength of the EVI formulation with an arbitrary observer. If we have two approximate solutions $\boldsymbol{x}_i$ to the EVI, where the error term $\Delta_i(t)$ does not depend on the observer, then we obtain a control on the distances between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.

Lemma 4.12 (Distance between approximate EVI solutions) *Let Δ_i, $i = 1,2$, be nonnegative real functions in $L^1([0,T])$. Let also $\boldsymbol{x}_i \in \mathrm{AC}_{\mathrm{loc}}([0,\infty))$, $i = 1,2$, be two locally absolutely continuous functions satisfying EVI$_\lambda$ for some $\lambda^* < 0$ in the form*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}_i(t), x_{\mathsf{ob}}) + \frac{\lambda^*}{2}\mathsf{d}^2(\boldsymbol{x}_i(t), x_{\mathsf{ob}}) \leq \phi(x_{\mathsf{ob}}) - \phi(\boldsymbol{x}_i(t)) + \Delta_i(t), \tag{4.17}$$

for a.a. $t \in [0,T]$ and every $x_{\mathsf{ob}} \in \boldsymbol{x}_{3-i}([0,T])$ for $i = 1,2$. Then we have the estimate

$$\sup_{t\in[0,T]} \mathrm{e}^{\lambda^* t}\mathsf{d}(\boldsymbol{x}_1(t), \boldsymbol{x}_2(t)) \leq \mathsf{d}(\boldsymbol{x}_1(0), \boldsymbol{x}_2(0)) + \big\|2\mathrm{e}^{2\lambda^* t}(\Delta_1{+}\Delta_2)\big\|_{L^1[0,T]}^{1/2} \tag{4.18}$$

Proof. For $s,t \in [0,T]$ we define the function $q(s,t) := \frac{1}{2}\mathsf{d}(\boldsymbol{x}_1(s), \boldsymbol{x}_2(t))$. Applying [AGS05, Lem. 4.3.4] we obtain the differential inequality

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}_1(t), \boldsymbol{x}_2(t)) &\leq \limsup_{h\downarrow 0}\frac{q(t,t) - q(t{-}h,t)}{h} + \limsup_{h\downarrow 0}\frac{q(t,t) - q(t,t{+}h)}{h} \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}_1(s), \boldsymbol{x}_2(t))|_{s=t} + \frac{\mathrm{d}}{\mathrm{d}r}\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}_1(t), \boldsymbol{x}_2(r))|_{r=t} \\ &\leq -\lambda^*\mathsf{d}^2(\boldsymbol{x}_1(t), \boldsymbol{x}_2(t)) + \Delta_1(t){+}\Delta_2(t), \end{aligned} \tag{4.19}$$

for almost every $t > 0$, where we used (4.17) for $\boldsymbol{x}_i$ with observer $x_{\mathsf{ob}} = \boldsymbol{x}_{3-i}$. Multiplying the inequality by $2\mathrm{e}^{2\lambda^* t}$ and integrating in time yields the desired estimate. ∎

The following result specifies this estimate by looking at the solutions obtained as geodesic interpolants from MM schemes with two different time steps $\tau > 0$ and $\sigma > 0$.

Corollary 4.13 (Comparison of MM solutions) *Let τ, σ, two time steps and let $\boldsymbol{x}^\tau$ and $\boldsymbol{x}^\sigma$ be two piecewise geodesic interpolants defined in* (4.14) *for initial conditions $x_0^\tau, x_0^\sigma \in A_{\kappa_0}$, respectively. Let ϕ be λ-convex with $\lambda \leq 0$ and set $\lambda^* = 2\lambda - 2$. Then we have*

$$\sup_{t\in[0,T]} e^{\lambda^* t}\mathsf{d}(\boldsymbol{x}^\tau(t), \boldsymbol{x}^\sigma(t)) \leq \mathsf{d}(x_0^\tau, x_0^\sigma) + \left\|2e^{2\lambda^* t}(\Delta^\tau{+}\Delta^\sigma)\right\|^{1/2}_{L^1[0,T]}, \tag{4.20}$$

where Δ^τ and Δ^σ are defined via (4.16) *in Proposition 4.11.*

Thus, it remains to control the error functions Δ^τ. The main problem is to control the terms $\frac{1}{\tau}\measuredangle^2\left(\overrightarrow{\boldsymbol{x}_n^\tau\boldsymbol{x}_{n-1}^\tau}, \overrightarrow{\boldsymbol{x}_n^\tau\boldsymbol{x}_{n+1}^\tau}\right)$, which control the change of the "directions and length" of the connecting geodesic interpolants. For this we use the improved incremental energy estimate (4.12) which allows us to invoke a telescope sum by inserting $x_{\mathsf{ob}} = x_{n+1}^\tau$.

Lemma 4.14 (Controlling the incremental errors) *Let $\lambda \leq 0$, $\tau \in (0,1)$ with $\tau\lambda > -1/2$, $x_0^\tau \in A_{\kappa_0}$, and let x_n^τ be defined iteratively by the MM scheme. For $T > 0$ and $\kappa \geq \kappa_0$ assume $x_n^\tau \in A_\kappa$ as long as $n < T/\tau + 1$. Then, the error function Δ^τ defined in* (4.16) *in Proposition 4.11 satisfies the estimate*

$$\|\mathrm{e}^{2\lambda^* t}\Delta^\tau\|_{L^1([0,T])} \leq \tau\,(4 + \tau\kappa)\,|\partial\phi|^2(x_0^\tau), \tag{4.21}$$

where we recall that $\lambda^* = 2\lambda - 2$.

4.5 Proof of the main abstract result in Theorem 4.9

Having prepared the above the preliminary estimates for the solutions x_n^τ of the minimizing movement scheme, we are now ready to give the proof of the abstract existence result. The important point in the proof is that the value κ of the κ-concavity of d^2 is occurring only in a few places that are well controlled. In particular, it is needed only on the time-discrete level (see e.g. (4.6) and (4.16)), but it disappears in the EVI formulation.

Proof of Theorem 4.9.
Let λ such that ϕ is geodesically λ-convex, then ϕ is also geodesically $\min\{\lambda, 0\}$-convex. As before we set $\lambda^* = 2\min\{\lambda, 0\} - 2$.

Step 1. Limit passage on approximate solutions $\boldsymbol{x}^{\tau_k}$. We now exploit the assumptions (A1) to (A3) of Theorem 4.9. For a given initial point $x_0 \in \cup_{\kappa\in\mathbb{N}} A_\kappa$ there exist $\kappa_0 \in \mathbb{N}$ such that $T(x_0, \kappa) > 0$ for all $\kappa > \kappa_0$. Fixing a $\kappa_* > \kappa_0$ we define time steps $\tau_k = T/k$ for $k \in \mathbb{N}$, where $T = T(x_0, \kappa_*)$ from assumption (A2). Hence, we have $x_n^{\tau_k} \in A_{\kappa_*}$ as long as $n < T/\tau_k + 1$. By construction of the geodesic interpolators in (4.14) the function $\boldsymbol{x}^{\tau_k} : [0,T] \to \mathfrak{X}$ satisfies $\boldsymbol{x}^{\tau_k}(t) \in A_{\kappa_*}^{\mathrm{Geod}}$ for all $t \in [0,T]$.

Thus, we are able to apply Corollary 4.13 and Lemma 4.14 and obtain (for $\tau_k, \tau_{k'} \leq 1$)

$$\sup_{t\in[0,T]} \mathsf{d}(\boldsymbol{x}^{\tau_k}(t), \boldsymbol{x}^{\tau_{k'}}(t)) \leq \mathrm{e}^{-\lambda^* T} 2\left((\tau_k{+}\tau_{k'})(4{+}\kappa_*)\right)^{1/2}|\partial\phi|(x_0). \tag{4.22}$$

Therefore the curves $\boldsymbol{x}^{\tau_k} : [0,T] \to \mathfrak{X}$ converge uniformly in the compact and hence complete sublevel $\left\{\, y \in \mathfrak{X} \;\middle|\; \phi(y) \leq \phi(x_0) \,\right\}$ to a continuous limiting curve $\boldsymbol{x} : [0,T] \to \mathfrak{X}$ with $\boldsymbol{x}(0) = x_0$.

Step 2. $\boldsymbol{x}$ is the unique EVI$_{\lambda^*}$ solution. We now return to the approximate EVI formulation (4.15) for the interpolants $\boldsymbol{x}^{\tau_k}$.

For a general observer $x_{\mathsf{ob}} \in \overline{\mathrm{dom}(\phi)}$ we can choose a sequence $(y_m)_{m\in\mathbb{N}}$ with

$$y_m \in A_{\kappa_m} \subset \mathfrak{X}, \quad y_m \overset{m\to\infty}{\longrightarrow} x_{\mathsf{ob}}, \quad \phi(y_m) \overset{m\to\infty}{\longrightarrow} \phi(x_{\mathsf{ob}}).$$

Without loss of generality we may assume $\kappa_* \leq \kappa_m$. Choosing $x_{\mathsf{ob}} = y_m \in A_{\kappa_m} \subset A^{\mathrm{Geod}}_{\kappa_m}$ in (4.15) for $\boldsymbol{x}^{\tau_k}$ and integrating over the interval (s,t) we find

$$\begin{aligned} &\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}^{\tau_k}(t), y_m) - \frac{1}{2}\mathsf{d}^2(\boldsymbol{x}^{\tau_k}(s), y_m) + \int_s^t \left(\phi(\boldsymbol{x}^{\tau_k}(r)) + \frac{\lambda^*}{2}\mathsf{d}^2(\boldsymbol{x}^{\tau_k}(r), y_m)\right) \mathrm{d}r \\ &\leq (t{-}s)\phi(y_m) + (t{-}s)\mathrm{e}^{-2\lambda^* t}\tau_k(4{+}\tau_k\kappa_m)\,|\partial\phi|^2(x_0). \end{aligned} \tag{4.23}$$

Keeping m fixed, taking $k\to\infty$, and using the lower semicontinuity of ϕ, we obtain

$$\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(t), x_m) - \frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(s), x_m) + \int_s^t \left(\phi(\boldsymbol{x}(r)) + \frac{\lambda^*}{2}\mathsf{d}^2(\boldsymbol{x}(r), x_m)\right) \mathrm{d}r \leq (t{-}s)\phi(x_m),$$

Note that κ_m has disappeared because of $\tau_k \to 0$ for $k\to\infty$. Now $m\to\infty$ yields

$$\frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(t), x_{\mathsf{ob}}) - \frac{1}{2}\mathsf{d}^2(\boldsymbol{x}(s), x_{\mathsf{ob}}) + \int_s^t \left(\phi(\boldsymbol{x}(r)) + \frac{\lambda^*}{2}\mathsf{d}^2(\boldsymbol{x}(r), x_{\mathsf{ob}})\right) \mathrm{d}r \leq (t{-}s)\phi(x_{\mathsf{ob}}),$$

where we have convergence on the left-hand side and use the lsc of ϕ on the right-hand side. As the inequality trivially holds for $x_{\mathsf{ob}} \in \mathfrak{X}\setminus\overline{\mathrm{dom}(\phi)}$, we have shown that $\boldsymbol{x} : [0,T]$ is an EVI$_{\lambda^*}$ solution. The uniqueness follows by applying Corollary 4.13 with $\Delta^\tau = \Delta^\sigma = 0$.

Step 3. Extension to $t \in [0, T_\infty(x_0))$. In the previous step the solution $\boldsymbol{x} : [0, T(x_0,\kappa_*)] \to X$ was well-defined and unique. However, $\kappa_* > \kappa_0(x_0)$ was arbitrary. Hence, we can extend the solution uniquely to any interval $[0, T(x_0,\kappa)]$ with $\kappa > \kappa_0$. Taking the limit $\kappa\to\infty$ we obtain a unique solution on $[0, T_\infty(x_0)) \subset \cup_{\kappa>\kappa_0}[0, T(x_0,\kappa)]$.

Step 4. Complete EVI flow on $\overline{\mathrm{dom}(\phi)}$. We now further assume $T_\infty(x_0) = \infty$ for all $x_0 \in \cup_\kappa A_\kappa$. We now consider an arbitrary $x_0 \in \overline{\mathrm{dom}(\phi)}$. Since $\overline{\cup A_\kappa} = \overline{\mathrm{dom}(\phi)}$, there exists a sequence $(x_0^m)_{m\in\mathbb{N}}$ with $x_0^m \in A_{\kappa_m} \subset \cup_{\kappa\in\mathbb{N}} A_\kappa$ and $x^m \overset{m\to\infty}{\longrightarrow} x$. Define $\boldsymbol{x}^m : [0,\infty) \to X$ to be the unique EVI$_{\lambda^*}$ solution starting in x_0^m. By (4.9) in Proposition 4.8(C), we have

$$\mathrm{e}^{\lambda^* t}\mathsf{d}(\boldsymbol{x}^m(t), \boldsymbol{x}^{m'}(t)) \leq \mathsf{d}(\boldsymbol{x}_0^m, \boldsymbol{x}_0^{m'}) \quad \text{for all } t \geq 0.$$

Thus, for all $T > 0$ the sequence $\boldsymbol{x}^m$ is Cauchy in the space $C([0,T];\mathfrak{X})$. Therefore, it converges locally uniform to a limit $\boldsymbol{x} : [0,\infty) \to \mathfrak{X}$, which satisfies the initial condition $\boldsymbol{x}(0) = x_0 = \lim x_0^m$. Since each curve $\boldsymbol{x}^m$ satisfies the integrated form (4.8) of EVI$_{\lambda^*}$, the lower semicontinuity of ϕ guarantees that the limit curve $\boldsymbol{x}$ is again an EVI solution.

Step 5. Correcting λ^* back to λ. Above we have constructed EVI$_{\lambda^*}$ solutions, but our functional ϕ is geodesically λ-convex and $\lambda > \lambda^*$. To recover the correct λ, we can apply [MuS20, Cor. 3.12] because we know that $\boldsymbol{x}$ is an EVI$_{\lambda^*}$ solution for $(\mathfrak{X}, \mathsf{d}, \phi)$ and that ϕ is λ-convex with $\lambda \geq \lambda^*$. Hence, $\boldsymbol{x}$ is also EVI$_\lambda$ solution. ∎

5 Semiconcavity and EVI flows for $(\mathcal{M}(X), \mathsf{HK}, \mathsf{E})$ and $(\mathcal{P}(X), \mathsf{SHK}, \mathsf{E})$

We now combine the theory developed in the previous two sections, namely the existence result for EVI flows provided in Theorem 4.9 with the semiconcavity results established in [LaM19, Sec. 4].

5.1 Semiconcavity of $\frac{1}{2}\mathsf{HK}^2$ and $\frac{1}{2}\mathsf{SHK}^2$

In order to apply Theorem 4.9 in the case of $\mathsf{HK}, \mathsf{SHK}$, we need to provide some semiconcavity results. More specifically, we need to prove that point (A1) is satisfied for a sequence of sets A_κ. Before we proceed, we will define the following two collections of sets. For $\delta \in (0,1)$ we define the set

$$\mathcal{M}_\delta(X) = \left\{ \mu \in \mathcal{M}(X) : \mu \ll \mathcal{L}^d,\ \delta \leq \frac{d\mu}{d\mathcal{L}^d}(x) \leq \frac{1}{\delta},\ \text{for } \mathcal{L}^d\text{-a.e. } x \in X \right\}. \tag{5.1}$$

For positive numbers $\mathsf{d}_1, \mathsf{d}_2$, we also define

$$\widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X) = \left\{ \mu \in \mathcal{M}(X) : \forall x \in X :\ \mathsf{d}_2 \leq \frac{\mu\left(B\left(x, \mathsf{d}_1\right)\right)}{\mathcal{L}^d(B\left(x, \mathsf{d}_1\right))} \leq \frac{1}{\mathsf{d}_2} \right\}. \tag{5.2}$$

It is straightforward to see that for all $\mathsf{d}_1 > 0$ it holds $\mathcal{M}_\delta(X) \subset \widetilde{\mathcal{M}}_{\mathsf{d}_1,\delta}(X)$. Furthermore all elements in $\mathcal{M}_\delta(X)$ have total mass bounded by $\frac{1}{\delta}\mathcal{L}^d(X)$.

In [LaM19, Thm. 4.8], it was stated and proved that for a set $X \subset \mathbb{R}^d$ that is compact, convex and with nonempty interior, there exists $\kappa(\delta) \in \mathbb{R}$, such that $(\mathcal{M}(X), \mathsf{HK})$ is κ-concave on $\mathcal{M}_\delta(X)$. We clarify at this point, that in practice Theorem 4.8 was stated for more general metric spaces and for reference measures ν that are doubling. However for simplification we are going to recall any theorems or lemmas we need from [LaM19] directly adapted to to our setting, avoiding all the extra generality related to doubling measures and abstract metric spaces. We remind the reader, that for a compact, convex set $X \subset \mathbb{R}^d$ with nonempty interior, the Lebesgue measure is doubling and the Euclidean distance is 2-concave. Although [LaM19, Thm. 4.8] was stated in this weaker form, the given proof provides a stronger result, namely the following:

Theorem 5.1 (K-concavity for $(\mathcal{M}(X), \mathsf{HK})$) *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Then, there exists $\kappa(\delta) \in \mathbb{R}$, such that $(\mathcal{M}(X), \mathsf{HK})$ is κ-concave on $\mathcal{M}_\delta(X)$, with respect to observers in $\mathcal{M}_\delta^{\mathrm{Geod}}(X)$.*

At the moment of writing [LaM19], we were not aware that this version will be useful, however now this property is exactly the assumption (A1) in our Theorem 4.9. Therefore we will recall some lemmas from there and provide a short proof of Theorem 5.1.

Lemma 5.2 ([LaM19, Lem. 4.9]) *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. There exists $0 < C_{\min} \leq C_{\max}$ such that for every $\mu_0, \mu_1 \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ and any optimal plan H_{01} for $\mathsf{LET}_{\boldsymbol{d}}(\,\cdot\,; \mu_0, \mu_1)$ we have*

$$C_{\min} \leq \sigma_i(x_i) \leq C_{\max}, \quad \eta_i\text{-a.e.} \tag{5.3}$$

where $\eta_i = \pi^i_\# H_{01} = \sigma_i \mu_i$ for $i = 0, 1$. Furthermore, any transportation happens in distances strictly less than some $\frac{\pi}{2}$, i.e. there exists $\mathfrak{D} < \frac{\pi}{2}$ that depends only on $\mathsf{d}_1, \mathsf{d}_2$, such that $\mathsf{d}_X(x_0, x_1) \leq \mathfrak{D}$ for H_{01} almost every (x_0, x_1).

Lemma 5.3 ([LaM19, Lem. 4.10]) *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior and $\mathcal{M}_\delta(X)$ be as in (5.1). Then, for each $\delta > 0$ there exist $\mathsf{d}_1 \in (0, \frac{\pi}{2})$ and $\mathsf{d}_2 > 0$ such that any constant-speed geodesic $\mu_{\mathbf{01}}$ connecting μ_0 to μ_1 with $\mu_0, \mu_1 \in \mathcal{M}_\delta(X)$ satisfies $\boldsymbol{\mu}_{01}(t) \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ for all $t \in [0, 1]$.*

From Lemma 5.3 we obtain

$$\mathcal{M}_\delta(X) \subset \mathcal{M}_\delta{}^{\mathrm{Geod}}(X) \subset \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X).$$

For the proof of [LaM19, Thm. 4.8] also the following result is used.

Lemma 5.4 ([LaM19, Lem. 4.11]) *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior and $\widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ be as in* (5.2)*. Then, there exist $R_{\min}, R_{\max} > 0$ that depend on $\mathsf{d}_1, \mathsf{d}_2$, such that for μ_0, μ_1 with $\boldsymbol{\mu}_{01}(t) \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ and $\mu_2 \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ we can find measures $\lambda_0, \lambda_1, \lambda_2, \lambda_t \in \mathcal{P}_2(\mathfrak{C}[R_{\min}, R_{\max}]$ with*

$$\mathfrak{P}\lambda_i = \mu_i, \quad \mathfrak{P}\lambda_t = \boldsymbol{\mu}_{01}(t), \quad \mathrm{W}_{\mathsf{d}_\mathfrak{C}}(\lambda_i, \lambda_t) = \mathsf{HK}(\mu_i, \boldsymbol{\mu}_{01}(t)) \quad \textit{for } i = 0, 1, 2.$$

As the reader can see in both, Lemma 5.4, and in the actual proof of [LaM19, Thm. 4.8], the observer does not need to be in $\mathcal{M}_\delta$, but it suffices that it lies in $\widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$, which contains $\mathcal{M}_\delta^{\mathrm{Geod}}(X)$, therefore the same proof follows through.

Proof of Theorem 5.1. By Lemma 5.3 there exists $0 < \mathsf{d}_1 < \frac{\pi}{2}$ and $0 < \mathsf{d}_2$ such that every geodesic $\boldsymbol{\mu}_{01}$ connecting $\mu_0, \mu_1 \in \mathcal{M}_\delta(X)$ satisfies $\boldsymbol{\mu}_{01}(t) \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ for all $t \in [0, 1]$. We also have $\mu_2 \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X) \supset \mathcal{M}_\delta{}^{\mathrm{Geod}}(X)$.

We would like to utilize the equivalent definitions of K-concavity provided in [LaM19, Cor. 2.24(iii)] where a function $f : [0, 1] \to \mathbb{R}$ is K-concave if for every $t_1, t_2 \in [0, 1]$ with $t_1 < t_2$ the mapping $\tilde{f}_i^{[t_1,t_2]}(t) = f_i\left(t_1 + t(t_2 - t_1)\right)$ satisfies

$$\tilde{f}_i^{[t_1,t_2]}(t) + Kt(1-t)(t_2-t_1)^2 \geq (1-t)\tilde{f}_i^{[t_1,t_2]}(0) + t\tilde{f}_i^{[t_1,t_2]}(1) \text{ for all } t \in [0, 1]. \tag{5.4}$$

In that direction, we take $\tilde{\mu}_0 = \boldsymbol{\mu}_{01}(t_1), \tilde{\mu}_1 = \boldsymbol{\mu}_{01}(t_2)$ for $t_1, t_2 \in [0, 1]$, and $\tilde{\boldsymbol{\mu}}_{01}(t) = \boldsymbol{\mu}_{01}(t(t_2-t_1)+t_1)$. By Lemma 5.4, there exists $R_{\min}, R_{\max}$ that depend on $\mathsf{d}_1, \mathsf{d}_2$, and therefore on δ, such that for every $\tilde{\mu}_0, \tilde{\mu}_1, \tilde{\mu}_2 \in \widetilde{\mathcal{M}}_{\mathsf{d}_1,\mathsf{d}_2}(X)$ and $0 < t < 1$ we can find measures $\lambda_0, \lambda_1, \lambda_2, \lambda_t \in \mathcal{P}_2$ $(\mathfrak{C}[R_{\min}, R_{\max}])$ with

$$\mathfrak{P}\lambda_i = \tilde{\mu}_i, \quad \mathfrak{P}\lambda_t = \tilde{\boldsymbol{\mu}}_{01}(t), \text{ and } \mathrm{W}_{\mathsf{d}_\mathfrak{C}}(\lambda_i, \lambda_t) = \mathsf{HK}(\tilde{\mu}_i, \tilde{\boldsymbol{\mu}}_{01}(t)),\ i = 0, 1, 2, \tag{5.5}$$

see Theorem 2.2. Using the geodesic property of $\tilde{\boldsymbol{\mu}}_{01}$ yields

$$\begin{aligned} \mathrm{W}_{\mathsf{d}_\mathfrak{C}}(\lambda_0, \lambda_t) + \mathrm{W}_{\mathsf{d}_\mathfrak{C}}(\lambda_1, \lambda_t) &= \mathsf{HK}(\mu_0, \tilde{\boldsymbol{\mu}}_{01}(t)) + \mathsf{HK}(\mu_1, \tilde{\boldsymbol{\mu}}_{01}(t)) \\ &= \mathsf{HK}(\tilde{\mu}_0, \tilde{\mu}_1) \leq \mathrm{W}_{\mathsf{d}_\mathfrak{C}}(\lambda_0, \lambda_1). \end{aligned}$$

Hence, it is straightforward to see that there exists a geodesic $\boldsymbol{\lambda}_{01}$ connecting λ_0, λ_1, such that $\boldsymbol{\lambda}_{01}(t) = \lambda_t$. Furthermore, by [Lis06, Thm. 6] there is a plan $\boldsymbol{\Lambda}_{0\to 1}$ on the geodesics such that $\Lambda_{ts} := (e_t, e_s)_\sharp \boldsymbol{\Lambda}_{0\to 1}$ is an optimal plan between $\boldsymbol{\lambda}(t)$ and $\boldsymbol{\lambda}(s)$. Now, by using a gluing lemma, we can find a plan $\boldsymbol{\Lambda}_{2t}^{0\to 1}$ in $\mathcal{P}((C[0, 1]; \mathfrak{C}) \times \mathfrak{C})$, such that $\Lambda_{01} = (e_0, e_1)_\sharp \left(\pi_\sharp^{0\to 1} \boldsymbol{\Lambda}_{2t}^{0\to 1}\right)$, and $(e_t(\pi^{0\to 1}) \times I)_\sharp \boldsymbol{\Lambda}_{2t}^{0\to 1}$ is an optimal plan for $\mathrm{W}_{\mathsf{d}_\mathfrak{C}}(\lambda_2, \boldsymbol{\lambda}_{01}(t))$. Finally by applying the last part of Lemma 5.2, we get the existence of a $\mathfrak{D} < \frac{\pi}{2}$ such that $|x_2 - x_t| < \mathfrak{D}$ for $(e_t(\pi^{0\to 1}) \times I)_\sharp \boldsymbol{\Lambda}_{2t}^{0\to 1}$ almost every (z_2, z_t), similarly $|x_0 - x_1| < \mathfrak{D}$ for Λ_{01} almost every $[z_0, z_1]$. Therefore, for $\boldsymbol{\Lambda}_{2t}^{0\to 1}$ almost every $(z_2, \boldsymbol{z}(\cdot, z_0, z_1))$, where $\boldsymbol{z}(\cdot, z_0, z_1)$ is a geodesic connecting z_0, z_1, we have

$$x_0, x_1, x_2, \bar{\boldsymbol{x}}(t, z_0, z_1) \in B\left(\bar{\boldsymbol{x}}(t, z_0, z_1), d\right).$$

By [LaM19, Prop. 2.27] we get a K' such that

$$\mathsf{d}_{\mathfrak{C}}^2(z_2, \boldsymbol{z}(t, z_0, z_1)) + K't(1-t)\mathsf{d}_{\mathfrak{C}}^2(z_0, z_1) \geq (1-t)\mathsf{d}_{\mathfrak{C}}^2(z_2, z_0) + t\,\mathsf{d}_{\mathfrak{C}}^2(z_2, z_1), \tag{5.6}$$

for $\boldsymbol{\Lambda}_{2t}^{0\to 1}$ almost every $(z_2, \boldsymbol{z}(\cdot, z_0, z_1))$. By integrating with respect to $\boldsymbol{\Lambda}_{2t}^{0\to 1}$, we find

$$\mathrm{W}_{\mathsf{d}_{\mathfrak{C}}}^2(\lambda_2, \boldsymbol{\lambda}_{01}(t)) + K't(1-t)\mathrm{W}_{\mathsf{d}_{\mathfrak{C}}}^2(\lambda_0, \lambda_1) \geq (1-t)\mathrm{W}_{\mathsf{d}_{\mathfrak{C}}}^2(\lambda_2, \lambda_0) + t\,\mathrm{W}_{\mathsf{d}_{\mathfrak{C}}}^2(\lambda_2, \lambda_1). \tag{5.7}$$

Using (5.5) we find the desired semiconcavity, and Theorem 5.1 is proved. ■

5.2 Geodesic semiconvexity of functionals on HK and SHK

In [LMS22] the question of geodesic λ-convexity of functionals E with reference measure $\mathcal{L}^d$ on a d-dimensional domain are discussed in detail. It is shown that E defined as in (1.1) in terms of a lsc and convex density functions E with $E(0) = 0$ is λ-convex on $(\mathcal{M}(X), \mathsf{HK})$ if and only if the auxiliary function

$$N_{E,\lambda} : (0,\infty)^2 \to \mathbb{R} \cup \{\infty\};\ (\rho,\gamma) \mapsto \left(\frac{\rho}{\gamma}\right)^d E\left(\frac{\gamma^{2+d}}{\rho^d}\right) - \frac{\lambda}{2}\gamma^2$$

satisfies the following two conditions:

$$\begin{aligned} &N_{E,\lambda} : (0,\infty)^2 \to \mathbb{R} \cup \{\infty\} \ \text{ is convex and} \\ &\rho \mapsto (d-1)\, N_{E,\lambda}(\rho,\gamma) \ \text{ is non-increasing.} \end{aligned} \tag{5.8}$$

It is shown that the density functions E of the form

$$E(c) = \alpha_0 c + \alpha_1 c^{p_1} + \cdots + \alpha_m c^{p_m}$$

lead to geodesically $2\alpha_0$-convex E if $\alpha_0 \in \mathbb{R}$ and $\alpha_i \geq 0$ and $p_i > 1$ for $i = 1, \ldots, m$. Moreover in dimensions $d \in \{1, 2\}$ the density function $E(c) = -\beta c^q$ with $\beta \geq 0$ and $q \in [d/(d+2), 1/2]$ lead to geodesically convex functionals E.

So far there doesn't seem to be a theory for semiconvexity on $\mathcal{P}(X), \mathsf{SHK})$ which can be used to provide examples. In Appendix A we establish the following nontrivial class of examples.

Proposition 5.5 *Consider dimension $d \in \{1, 2\}$ and a bounded convex domain $X \subset \mathbb{R}^d$ that is the closure of an open set. Then, the functional E_q defined via*

$$\mathsf{E}_q(\mu) = -\int_X \rho^q \,\mathrm{d}x \quad \textit{for } \mu = \rho\,\mathrm{d}x + \mu^\perp$$

is geodesically convex on $(\mathcal{P}(X), \mathsf{SHK})$ if $q \in [d/(d+2), 1/2]$.

With this result we are sure that the following main existence result for EVI flows on $(\mathcal{P}(X), \mathsf{SHK}, \mathsf{E})$ provides at least the solutions to a small, but nontrivial family of nonlinear partial differential equations.

5.3 The Main Result

In this section we collect the results from the previous sections and provide the proof of our main result, which we repeat here for convenience.

Theorem 5.6 *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore, let* E *be a functional defined as in* (1.1) *that satisfies Assumption A.*

Then, for all $\mu_0 = \rho_0\mathcal{L}^d \in \mathcal{P}(X)$ with $0 < \underline{\rho}_0 \leq \rho_0(x) \leq \overline{\rho}_0 < \infty$ a.e. in X, the solutions of the MM scheme $(1.2)_{\mathsf{SHK}}$ *converge to a complete solution $\boldsymbol{\mu} : [0,\infty) \to \mathcal{P}(X)$ of EVI$_\lambda$ for $(\mathcal{P}(X), \mathsf{E}, \mathsf{SHK})$. Moreover, for all $\mu_0 \in \overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{SHK}} \subset \mathcal{P}(X)$ there exists a unique EVI solution emanating from μ_0.*

If in addition Assumption B is satisfied, then for all $\mu_0 = \rho_0\mathcal{L}^d \in \mathcal{M}(X)$ with $0 < \underline{\rho}_0 \leq \rho_0(x) \leq \overline{\rho}_0 < \infty$ a.e. in X the MM scheme $(1.2)_{\mathsf{HK}}$ *converges to a curve $\boldsymbol{\mu} : [0,\infty) \to \mathcal{M}(X)$ that satisfies the EVI$_\lambda$ for $(\mathcal{M}(X), \mathsf{E}, \mathsf{HK})$. Moreover, for each $\mu_0 \in \overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{SHK}} subset \mathcal{M}(X)$ there exists a unique EVI solution emanating from μ_0.*

Proof. We start from $\mu_0 = \rho_0\mathcal{L}^d$ with $0 < \underline{\rho}_0 \leq \rho_0(x) \leq \overline{\rho}_0 < \infty$. By Propositions 3.5 (for SHK) and 3.7 (for HK) we know that for every T there exists a $\tau_0 > 0$, such that for all $\tau \leq \tau_0$ and n with $n\tau < T$, the solutions μ_n of the MM scheme satisfy $\mu_n = \rho_n\mathcal{L}^d$ where $0 < \delta \leq \rho_n(x) \leq 1/\delta < \infty$. But this means that $\mu_n \in \overline{\mathcal{M}}_\delta(X)$ for all n with $n\tau < T$. Therefore, by taking $A_\kappa = \overline{\mathcal{M}}_\delta(X)$ and using Theorem 5.1 we see that all assumptions of Theorem 4.9 are satisfied, and the convergence of the MM scheme for density restricted initial data μ_0 follows.

For general initial data μ_0 in the closure $\overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{SHK}} \subset \mathcal{P}(X)$ or $\overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{HK}} \subset \mathcal{M}(X)$ of the domain of the functional E we can choose approximations $\mu_0^m = \rho_0^m\mathcal{L}^d$ satisfying $\rho_0^m \in [1/m, m]$ a.e. in X. The associated EVI solutions μ_m are complete and converge to the desired, complete EVI solution emanating from μ_0 by applying (4.18). ■

We conclude with a corollary that allows us to extend the existence of complete EVIs when Assumption B does not hold.

Corollary 5.7 *Let $X \subset \mathbb{R}^d$ be a compact, convex set with nonempty interior. Furthermore, let* E *be a functional as in* (1.1) *satisfying Assumption A. If $E'(0) = 0$, then for every $\mu_0 \in \overline{\mathrm{dom}(\mathsf{E})}^{\mathsf{HK}}$ there exists a unique curve $\boldsymbol{\mu} : [0,\infty) \to \mathcal{M}(X)$ emanating from μ_0 that is a complete EVI$_\lambda$ solution for $(\mathcal{M}(X), \mathsf{E}, \mathsf{HK})$. If $d \leq 2$, then the the same holds even without the extra assumption $E'(0) = 0$.*

Proof. For dimension $d \leq 2$, we define a new functional E_δ given by $E_\delta(c) = E(c) - \delta c^{1/2}$. By the results in [LMS22, Sec. 7] we know that E_δ still satisfies Assumption A if E does so. Moreover, the new functional satisfies Assumption B, i.e. there exists $c_{\mathrm{low}} > 0$ with $E_\delta'(c_{\mathrm{low}}) < 0$. Therefore, for every $\delta > 0$ Theorem 5.6 guarantees that we have unique EVI solution μ^δ for $(\mathcal{M}(X), \mathsf{E}_\delta, \mathsf{HK})$ emanating from every μ_0. Now by taking the limit $\delta \to 0$ we have Γ-convergence of E_δ to E, and therefore we obtain the convergence of μ^δ to a complete EVI solution μ for $(\mathcal{M}(X), \mathsf{E}, \mathsf{HK})$ by [MuS20, Sec. 3.2].

For the case of $d > 2$, we define $E_\delta(c) = E(c) - \delta c$. If E satisfies $E'(0) = 0$, then E_δ satisfies Assumption B for every $\delta > 0$. Now, the result follows with the same arguments as above. ■

A Transfer of λ-convexity between HK and SHK

We rely on the interpretation of $(\mathcal{M}(X), \mathsf{HK})$ as a cone of over $(\mathcal{P}(X), \mathsf{SHK})$ that was developed in [LaM19]. But first we consider a general geodesic space $(\mathcal{X}, \mathsf{d})$ and the associated cone $(\mathcal{C}, D)$, which take the places of $(\mathcal{P}(X), \mathsf{SHK})$ and $\mathcal{M}(X), \mathsf{HK})$, respectively.

We provide a general result showing that under suitable conditions the geodesic convexity of a p-homogeneous functional F on $(\mathcal{C}, D)$, the restriction of F to $(\mathcal{X}, \mathsf{d})$ is again geodesically convex.

Proposition A.1 (Transfer for negative, homogeneous functionals) *Assume that* $\mathsf{F} : \mathcal{C} \to [-\infty, 0]$ *is geodesically convex on* $(\mathcal{C}, D)$*, and that it is* p*-homogeneous for some* $p \geq 1/2$*, i.e.*

$$\mathsf{F}([x, r]) = r^p \mathsf{F}([x, 1]) \text{ for all } [x, r] \in \mathcal{C}.$$

Moreover, assume $\mathsf{d}(x, y) < \pi$ *for all* $x, y \in \mathcal{X}$*.*

Then, $\mathsf{E}(x) := \mathsf{F}([x, 1]) \in [-\infty, 0]$ *is geodesically convex on* $(\mathcal{X}, \mathsf{d})$*.*

Proof. We use the fact that all geodesics $x : [0, 1] \to \mathcal{X}$ connecting x and x_1 are given by the geodesics $z : [0, 1] \mapsto [\overline{x}(t), r(t)] \in \mathcal{C}$ by a simple reparametrization, see [LaM19, Thm. 2.7], namely, setting $\delta = d(x, x_1) \in]0, \pi[$ we have

$$x(t) = \overline{x}(\beta_\delta(t)) \text{ with } \beta_\delta(t) := \frac{\sin(t\delta)}{\sin(t\delta) + \sin((1-t)\delta)},$$

where $\beta_\delta(0) = 0$ and $\beta_\delta(1) = 1$. Moreover, $r(t) = 1 - t(1-t)D([x, 1], [x_1, 1])^2$ with $D([x, 1], [x_1, 1]) = (2(1-\cos\delta))^{1/2}$ can be rewritten via

$$r(\beta_\delta(t)) = r_\delta(t) := \frac{\sin(\delta)}{\sin(t\delta) + \sin((1-t)\delta)} \in [1/2, 1],$$

where $r(0) = r(1) = 1$. With this we obtain

$$\begin{aligned}
\mathsf{E}(x(t)) &= \mathsf{F}([x(t), 1]) && \text{(definition of E)} \\
&= \mathsf{F}([\overline{x}(\beta_\delta(t)), 1]) && \text{(reparametrization)} \\
&= \frac{1}{r_\delta(t)^p} \mathsf{F}([\overline{x}(\beta_\delta(t)), r_\delta(t)]) && \text{(p-homogeneity of F)} \\
&\leq \frac{1}{r_\delta(t)^p} \, \big((1-\beta_\delta(t))\, \mathsf{F}([\overline{x}(0), r(0)]) + \beta_\delta(t) \mathsf{F}([\overline{x}(1), r(1)])\big) && \text{(geodesic cvx of F on $\mathcal{C}$)} \\
&= \frac{1-\beta_\delta(t)}{r_\delta(t)^p} \, \mathsf{E}(x) + \frac{\beta_\delta(t)}{r_\delta(t)^p} \, \mathsf{E}(x_1) && \text{(definition of E)}
\end{aligned}$$

Because of $\mathsf{E}(x_j) \leq 0$ it suffices to show the two estimates

$$\frac{1-\beta_\delta(t)}{r_\delta(t)^p} \geq 1 - t \quad \text{and} \quad \frac{\beta_\delta(t)}{r_\delta(t)^p} \geq t \qquad \text{for all } t \in [0, 1] \text{ and } \delta \in [0, \pi]. \tag{A.1}$$

For the second estimate and the case $p \geq 1$ we use the explicit form and obtain

$$\frac{\beta_\delta(t)}{r_\delta(t)^p} = \frac{1}{r_\delta(t)^{p-1}} \, \frac{\sin(t\delta)}{\sin\delta} \geq t \quad \text{for all } t \in [0, 1],$$

where we used $r_\beta(t) \leq 1$ and that $t \mapsto \sin(t\delta)$ is concave on $[0,1]$ because of $\delta \in]0,\pi[$. Thus, the result certainly holds for $p \geq 1$.

However, it also holds for $p \in [1/2, 1]$ by the following arguments. Define the function

$$Q_p(t,\delta) = \frac{\sin(t\delta)}{t\,(\sin\delta)^p}\,(\sin(t\delta) + \sin((1-t)\delta))^{p-1}.$$

It suffices to show $Q_p(t,\delta) \geq 1$ for $t \in]0,1[$ and $\delta \in]0,\pi[$. Clearly, we have $Q_p(1,\delta) = 1$ and we find

$$\partial_t Q_p(1,\delta) = -1 + p\frac{\delta\cos\delta}{\sin\delta} + (1-p)\,\frac{\delta}{\sin\delta}.$$

For $p \in [1/2, 1]$ one can show that $\partial_t Q_p(t,\delta) \leq 0$ which implies $Q_p(t,\delta) \geq Q_p(1,\delta) = 1$ which is the desired second estimate in (A.1).

To see why $p \geq 1/2$ is necessary, we use $\lim_{\delta\to\pi^-} (\sin\delta\,\partial_t Q_p(1,\delta)) = \pi(1-2p)$. Thus, for $p < 1/2$ we have $\partial_t Q_p(t,\delta) > 0$ for $\delta \approx \pi$, which implies $Q_p(t,\delta) < 1$.

The first estimate in (A.1) follows similarly, namely by changing t to $1-t$. This proves the result. ∎

We now consider $(\mathcal{M}(X), \mathsf{HK})$ as the cone over $(\mathcal{P}(X), \mathsf{SHK})$ for some convex and compact $X \subset \mathbb{R}^d$. We first observe that [LaM19, Thm. 3.4] guarantees $\mathsf{SHK}(\nu_0,\nu_1) \leq \pi/2$ such that the condition $\mathsf{d}(x,y) < \pi$ is automatically satisfied.

From [LMS22] we know by that the functionals

$$\mathsf{E}_q(\mu) = \int_X \varrho(x)^q\,\mathrm{d}x \quad \text{for } \mu = \varrho\,\mathrm{d}x$$

are geodesically 0-convex on $(\mathcal{M}(X), \mathsf{HK})$ whenever $q > 1$. Moreover, in the sense of cones we have $\mu = r^2\varrho\mathcal{L}^d$ giving p-homogeneity with $p = 2q$, namely

$$\mathsf{E}_q(\mu) = \mathsf{E}_q(r^2\varrho\mathcal{L}^d) = \mathsf{F}_q([\varrho\mathcal{L}^d, r]) = r^{2q}\mathsf{F}_q([\varrho\mathcal{L}^d, 1]) = r^{2q}\mathsf{E}_q(\varrho\mathcal{L}^d).$$

However, the above result is not applicable because of $\mathsf{F}_q(\mu) \geq 0$.

Note also that the special case $q = 1$ leads to the mass functional $\mathsf{E}_\mathrm{M}(\mu) = \mathsf{E}_1(\mu) = \int_X \mu(\mathrm{d}x) = \mu(X)$ which is geodesically 2-convex for HK (as well as geodesically 2-concave). However, its spherical restriction is obviously constant, hence it is geodesically 0-convex and 0-concave. This means that we have a drop in the convexity, namely

$$0 = \Lambda_{\mathsf{SHK}} \lneqq \Lambda_{\mathsf{HK}} = 2.$$

However, the above result can be applied in the case of functionals of the form

$$\mathsf{E}_q(\mu) = -\int_X \varrho(x)^q\,\mathrm{d}x \quad \text{for } q \in (0,1) \text{ and } \mu = \varrho\mathcal{L}^d + \mu^\perp, \tag{A.2}$$

where $\mu^\perp$ is singular with respect to $\mathcal{L}^d$. This leads to the

Proof of Proposition 5.5. For $\nu_0, \nu_1 \in \mathcal{P}(X)$ we first observe $\mathsf{HK}^2(\nu_0,\nu_1) \leq \nu_0(X) + \nu_1(X)$ which implies $\mathsf{SHK}(\nu_0,\nu_1) = 2\arcsin(\frac{1}{2}\mathsf{HK}(\nu_0,\nu_1)) \leq 2\arcsin(\frac{1}{2}\sqrt{2}) = \pi/2 < \pi$.

It is shown in [LMS22] that E_q is geodesically 0-convex under the following conditions:

$$q \in [1/3, 1/2] \text{ and } d = 1 \qquad \text{or} \qquad q = 1/2 \text{ and } d = 2.$$

Moreover, we obviously have $\mathsf{E}_q(\mu) \leq 0$ and $\mathsf{E}_q(r^2\varrho\mathcal{L}^d) = \mathsf{F}_q([\varrho\mathcal{L}^d, r]) = r^{2q}\mathsf{F}_q([\varrho\mathcal{L}^d, 1]) = r^{2q}\mathsf{E}_q(\varrho\mathcal{L}^d)$. Using $q \geq 1/3$ we have p-homogeneity with $p = 2q \geq 2/3$. Hence, all assumptions of Proposition A.1 are satisfied, and the geodesic convexity of E_q restricted to $(\mathcal{P}(X), \mathsf{SHK})$ follows. ∎

Acknowledgments. The authors are grateful to Giuseppe Savaré for fruitful and stimulating discussions and for sharing the preprint [Sav11]. V. Laschos was supported by DFG, through the project EXC-2046/1 within Germany´s Excellence Strategy – The Berlin Mathematics Research Center MATH+ (project ID: 390685689). A. Mielke was partially supported by DFG through the project Mie 459/9-1 within the Priority Program SPP 2256 *Variational Methods for Predicting Complex Phenomena in Engineering Structures and Materials* (project ID: 441470105).

References

[AFP00] L. Ambrosio, N. Fusco, and D. Pallara: *Functions of bounded variation and free discontinuity problems*. Clarendon Press, Oxford, 2000.

[AGS05] 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.

[CP*15] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard: *An interpolating distance between optimal transport and Fisher–Rao metrics*. Found. Comput. Math. **18**:1 (2015) 1–44.

[DiC20] S. Di Marino and L. Chizat: *A tumor growth model of Hele-Shaw type as a gradient flow*. ESAIM Control Optim. Calc. Var. **26**:103 (2020) 1–38.

[Fle20] F. Fleißner: *A minimizing movement approach to a class of scalar reaction-diffusion equations*. arXiv:2002.04496 (2020) .

[GaM17] T. O. Gallouët and L. Monsaingeon: *A JKO splitting scheme for Kantovorich-Fisher-Rao gradient flows*. SIAM J. Math. Analysis **49**:2 (2017) 1100–1130.

[HLP34] G. H. Hardy, J. E. Littlewood, and G. Pólya: *Inequalities*. Cambridge University Press, 1934.

[JKO98] R. Jordan, D. Kinderlehrer, and F. Otto: *The variational formulation of the Fokker-Planck equation*. SIAM J. Math. Analysis **29**:1 (1998) 1–17.

[LaM19] V. Laschos and A. Mielke: *Geometric properties of cones with applications on the Hellinger–Kantorovich space, and a new distance on the space of probability measures*. J. Funct. Analysis **276**:11 (2019) 3529–3576.

[LiM13] M. Liero and A. Mielke: *Gradient structures and geodesic convexity for reaction-diffusion systems*. Phil. Trans. Royal Soc. A **371**:2005 (2013) 20120346, 28.

[Lis06] S. Lisini. Characterization of absolutely continuous curves in Wasserstein spaces. Calc. Var. Partial Differ. Eqns. **28**:1 (2006) 85–120.

[LMS16] M. Liero, A. Mielke, and G. Savaré: *Optimal transport in competition with reaction – the Hellinger–Kantorovich distance and geodesic curves*. SIAM J. Math. Analysis **48**:4 (2016) 2869–2911.

[LMS18] ———: *Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures*. Invent. math. **211** (2018) 969–1117.

[LMS22] ———: *Fine properties of geodesics and geodesic λ-convexity for the Hellinger–Kantorovich distance*. arXiv2208.14299v2 (2022).

[MuS20] M. Muratori and G. Savaré: *Gradient flows and evolution variational inequalities in metric spaces. I: structural properties*. J. Funct. Analysis **278**:4 (2020) 108347.

[MuS22] M. Muratori and G. Savaré: *Gradient flows and evolution variational inequalities in metric spaces. II: variational convergence and III: generation results*. In preparation, 2022.

[Ott96] F. Otto, *Double degenerate diffusion equations as steepest descent*, Preprint no. 480, SFB 256, University of Bonn, 1996.

[Ott01] ———: *The geometry of dissipative evolution equations: the porous medium equation*. Comm. Partial Diff. Eqns. **26** (2001) 101–174.

[PQV14] B. Perthame, F. Quirós, and J. L. Vázquez: *The Hele-Shaw asymptotics for mechanical models of tumor growth*. Arch. Rational Mech. Anal. **212**:1 (2014) 93–127.

[Sav07] G. Savaré, *Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds*, C. R. Math. Acad. Sci. Paris **345**:3 (2007) 151–154.

[Sav11] G. Savaré, *Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds*. Extended version of [Sav07]. Private Communication, 2011.