EVOLUTION VARIATIONAL INEQUALITIES WITH GENERAL COSTS

PIERRE-CYRIL AUBIN-FRANKOWSKI, GIACOMO ENRICO SODINI, AND ULISSE STEFANELLI

Abstract. We extend the theory of gradient flows beyond metric spaces by studying evolution variational inequalities (EVIs) driven by general cost functions c, including Bregman and entropic transport divergences. We establish several properties of the resulting flows, including stability and energy identities. Using novel notions of convexity related to costs c, we prove that EVI flows are the limit of splitting schemes, providing assumptions for both implicit and explicit iterations.

Contents

1. Introduction

Context. The theory of gradient flows in metric spaces [AGS08], in particular in spaces of measures, is a cornerstone in the analysis of evolutionary systems. A reference tool in this context are minimizing movements, relating evolution to incremental minimization. The minimizing-movements approach is however not restricted to metric spaces. Indeed, it has been already applied out of the metric setting, including the case of Bregman divergences [Br67, eq.(1.4)] and their mirror flows [NY83]. These appear in entropic regularizations in optimal transport, see [Le14] and [PC19] for reviews. When leaving the metric setting, one is confronted with the question of which of the different formulations of gradient-flow evolution, namely, EDI, EDE, or EVI, as discussed in [AG13, San17] and in Section 2.3, should be considered.

Main contributions. In this article, we investigate the generalization of Evolution Variational Inqualities (EVIs) to general costs on general sets X. This provides a partial extension of the theory of [MS20] beyond the case of the square distance d^2 on a complete metric space X. Our notion of gradient flow takes the form

$$\frac{\mathrm{d}^+}{\mathrm{d}t} c(x, x_t) + \lambda c(x, x_t) \le \phi(x) - \phi(x_t) \quad \forall t > 0, x \in \mathscr{D}(\phi), \tag{1.1}$$

Date: May 2, 2025.

2020 *Mathematics Subject Classification.* 49J40, 37L05, 49J27, 49J52.

Key words and phrases. Gradient flows, Evolution variational inequality (EVI), Bregman divergence, Sinkhorn divergence, Kullback–Leibler divergence, Minimizing movements, Variational methods, Cross-convexity.

where X is a set, $c : X \times X \to [0, +\infty)$ is a cost function, $\phi : X \to (-\infty, +\infty]$ with domain $\mathscr{D}(\phi) = \{x \in X : \phi(x) < +\infty\}$. We call (1.1) EVI and we say that a curve $x : (0, +\infty) \to X$ is a EVI solution and write $x(\cdot) \in \mathrm{EVI}_\lambda(X, c, \phi)$ whenever (1.1) holds for some $\lambda \in \mathbb{R}$. We show in Theorem 2.4 that the differential formulation (1.1) has several equivalent integral expressions. For symmetric costs $c(x, y) = c(y, x)$, in Theorem 2.10 we establish that the set $\mathrm{EVI}_\lambda(X, c, \phi)$ of solutions still enjoys some of properties which hold in metric spaces. For example, the EVI is λ-contractive and one can prove an energy identity of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(x_{t+}) = -\lim_{h\downarrow 0}\frac{2c(x_t, x_{t+h})}{h^2} =: -|\dot{x}_{t+}|_c^2, \tag{1.2}$$

where $|\dot{x}_{t+}|_c$ plays the role of the metric derivative. Symmetry and nonnegativity of the cost are crucially used in order to obtain this. On the other hand, no notion of local slope is needed. In fact, although expression (1.1) is global in nature as it holds for all test points x, we discuss in Section 2.3 a local formulation of the form $\nabla_{2,1}c(x_t, x_t)\dot{x}_t \in \partial\phi(x_t)$, where $\nabla_{2,1}$ is the mixed-Hessian and $\partial\phi$ is the Fréchet subdifferential.

For geodesic metric spaces (X, d) and $c = d^2/2$, the EVI (1.1) has been related to some notion of convexity of ϕ w.r.t. d^2, see [MS20, Section 3.3]. Moreover, in metric spaces one may prove existence of solutions to (1.1) via minimizing movements and the implicit Euler scheme.

For general sets X and general costs c, a corresponding notion of convexity has been recently brought to evidence in [LAF23] by the first-named author. This notion takes the name of *cross-convexity* and delivers an extension of usual convexity. For $c = d^2/2$, cross-convexity takes the form of a discrete EVI, as in [AGS08, Corollary 4.1.3]. For c being a Bregman divergence, cross-convexity corresponds to the so-called *three-point inequality* in mirror descent, see, e.g., [CT93, Lemma 3.2]. In essence, cross-convexity can be interpreted as a compatibility property between energy and cost, see Section 3.2. In the cross-convex setting, existence of a solution $x(\cdot)$ to (1.1) for a given initial point x_0 can be obtained by taking the limit $\tau \to 0$ in the following alternating minimization (see (3.2)–(3.3)) based on the c/τ-transform $f^{c/\tau}(y) := \sup_{x'\in X}[f(x') - \frac{c(x',y)}{\tau}]$,

$$y_{i+1}^\tau \in \operatorname*{argmin}_{y\in X}\left\{\frac{c(x_i^\tau, y)}{\tau} + g(x_i^\tau) + f^{c/\tau}(y)\right\}, \tag{1.3}$$

$$x_{i+1}^\tau \in \operatorname*{argmin}_{x\in X}\left\{\frac{c(x, y_{i+1}^\tau)}{\tau} + g(x) + f^{c/\tau}(y_{i+1}^\tau)\right\}. \tag{1.4}$$

These iterations correspond to a splitting scheme over $\phi = f + g$, with one explicit step on f and one implicit on g. Schemes of this form were extensively studied in [LAF23]. Under some cross-convexity requirements (see Assumption 3.5), for c symmetric, a lower-bounded ϕ, and τ small enough, we show in Theorem 3.7 the following error estimate between the discrete-time scheme and its continuous-time limit:

$$c\left(\bar{x}_t^\tau, x_t\right) \le 2\tau(\phi(x_0) - \inf\phi) \quad \forall t \ge 0, \tag{1.5}$$

where $\bar{x}^\tau(\cdot)$ is the piece-wise constant in time interpolant of the discrete values $\{x_i^\tau\}$ on the uniform partition $\{i\tau\}$.

We now list three examples of costs c of interest.

Example 1.1. (distances, d^p) Let (X, d) be a complete metric space and $c = d^p$ for $p \ge 1$. Then, for $f = 0$, (1.4) is the usual implicit Euler discretization. Flows with asymmetric distances were also considered in [RMS08, CRZ09, OZ23].

Similarly to [AGS08], a key example in the following is $X = \mathcal{P}(\mathbb{X})$, that is the set of non-negative Borel measures over the measurable space $\mathbb{X}$ with total mass 1. Our analysis covers in particular that of [AGS08] if the cost c is the Wasserstein-2 distance. Two other prominent costs, on which we will showcase our assumptions and theorems, are the Kullback–Leibler and the Sinkhorn divergences.

Example 1.2. (Kullback–Leibler divergence, KL) The *Kullback–Leibler divergence* or *relative entropy* over probability measures is

$$\mathrm{KL}(\mu|\bar{\mu}) = \begin{cases} \int_{\mathbb{X}} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\bar{\mu}}(x)\right) \mathrm{d}\mu(x) & \text{if } \mu \ll \bar{\mu}, \\ +\infty & \text{otherwise}, \end{cases}$$

where, for any $\mu, \nu \in \mathcal{P}(\mathbb{X})$, we write $\mu \ll \nu$ when μ is absolutely continuous w.r.t ν, i.e., when it admits a Radon–Nikodym derivative $\mathrm{d}\mu/\mathrm{d}\nu$. The KL is a Bregman divergence of the entropy $\mathrm{KL}(\cdot|\rho)$ where ρ is interpreted as a reference positive measure (the Lebesgue measure on $\mathbb{X} \subset \mathbb{R}^d$, for instance). Note that KL is not symmetric and does not satisfy the triangle inequality.

Example 1.3. (Sinkhorn divergence, S_ϵ) Assume $\mathbb{X} \subset \mathbb{R}^n$ to be compact. Fix $\epsilon > 0$ and take a ground cost $c_{\mathbb{X}} \in C^1(\mathbb{X}\times\mathbb{X}; \mathbb{R})$ such that $\exp(-c_{\mathbb{X}}/\epsilon)$ is a positive definite and universal reproducing kernel, see [FSV+18]. Then, the *entropic optimal transport (EOT) dissimilarity* is defined as

$$\mathrm{OT}_\epsilon(\mu,\nu) = \min_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{X}\times\mathbb{X}} c_{\mathbb{X}}(x,y) d\pi(x,y) + \epsilon\, \mathrm{KL}(\pi|\mu\otimes\nu) = \varepsilon\, \mathrm{KL}(\pi|e^{-c_{\mathbb{X}}/\varepsilon}\mu\otimes\nu), \tag{1.6}$$

where we denote by $\Pi(\mu,\nu)$ the set of couplings having first marginal μ and second marginal ν. As $\mathrm{OT}_\epsilon(\mu,\mu) \neq 0$ in general [FSV+18, Section 1.2], the *Sinkhorn divergence* was defined in [GPC18] as

$$\mathrm{S}_\epsilon(\mu,\nu) = \mathrm{OT}_\epsilon(\mu,\nu) - \frac{1}{2}\,\mathrm{OT}_\epsilon(\mu,\mu) - \frac{1}{2}\,\mathrm{OT}_\epsilon(\nu,\nu) \tag{1.7}$$

which indeed fulfills $\mathrm{S}_\epsilon(\mu,\mu) = 0$. We refer to the introduction of [CDPS17] for a thorough presentation on OT_ϵ and to that of [LLM+24, p. 5] for S_ϵ. A self-contained discussion on the limiting behaviours for $\epsilon \to 0$ or $\epsilon \to \infty$ can be found in [NS23]. As evidenced in [LLM+24, Section 7.1], neither S_ϵ nor $\sqrt{\mathrm{S}_\epsilon}$ satisfy the triangle inequality, even though S_ϵ is symmetric and metrizes the convergence in law [FSV+18, Theorem 1].

Other natural examples of costs include doubly nonlinear evolution equations in a vector space X, $\Psi : X \to \mathbb{R}$, and $c_\tau(x,y) = \tau\Psi((x-y)/\tau)$, see [RMS08]. In our framework, we can only deal with homogeneous potentials Ψ. In fact, existence of formulations of the form (1.1) for Ψ not homogeneous do not presently seem available.

Related work. Reference works on EVIs in metric spaces are the three great expositions in [AGS08, AG13, MS20]. Notably, our theory is designed to recover some of these metric results in the case of (X,d) being a complete metric space and $c = d^2/2$. The continuous setting of Section 2 is somewhat closer to [MS20, Section 3], whereas the discrete setting of Section 3 takes inspiration from [AG13, Section 3.2.4], especially in its use of dyadic partitions.

Several extensions of the canonical gradient flows in metric settings have already been considered. In particular, [OZ23] deals with asymmetric distances and EDEs on Finsler manifolds, and [Cra17] studies EVIs when λd^2 in (1.1) is replaced by $\lambda\omega(d^2)$ for $\omega : \mathbb{R}_+ \to \mathbb{R}_+$, as motivated by the celebrated Osgood's criterion. Beyond metric spaces, [RS24] developed a theory for minimizing movements under so-called *action* costs $a(\tau,x,y)$. This theory actually covers our choice $c(x,y)/\tau$, stopping short of discussing the EVI formulation, nonetheless.

Compactness and completeness are the two alternative tools which are used to ascertain the existence of minimizing-movement limiting curves. Preferring one over the other naturally brings to alternative arguments, see e.g. [RW24]. Specifically, [MS20] is exclusively using completeness whereas [OZ23] is based on compactness. In our analysis, we elaborate on both approaches, by providing corresponding sets of assumptions.

In the metric setting, a discrete implicit-implicit splitting schemes based on the discrete EVI has been advanced in [CM11, Eq. (1.1)]. An explicit Euler scheme to deal with evolutions in the space of probability measures has been studied in [CSS23], where also the notion of EVI solution is used. To our best knowledge, we provide here the first theory for both implicit and explicit schemes in this EVI context and in such generality.

Plan of the paper. The paper is organized as follows. **Section 2** is devoted to formulate the EVI with general costs: in Section 2.1 we discuss the topological framework we adopt, and in particular the interplay between the cost function and the underlying topology. In Section 2.2, we present the equivalent formulations of the EVI and discuss the properties of EVI solutions. Section 2.3 discusses possible local formulations of the EVI, as well as how the EVI in the metric case leads to the metric slope, and hence to the EDE and EDI. In **Section 3**, we discuss the existence of EVI solutions for $\phi = f + g$ via an alternating minimization scheme. This is done first in Section 3.3 in the specific case of the implicit scheme when $f = 0$, and then in full generality in Section 3.4. Prior to this, sufficient compatibility conditions between the energy and cost, extending those from the metric setting, are discussed in Section 3.2.

2. Evolution Variational Inequalities with general cost c

Basic notation. In the following, we set $\mathbb{R}_+ = (0, +\infty)$. We use the shorthand l.s.c. for lower semicontinuous. Given the real function $u : \mathbb{R}_+ \to \mathbb{R}$, we indicate its *right derivative* at $t > 0$ (whenever existing) as $u'(t+)$. The *right Dini derivatives* are denoted by

$$\frac{\mathrm{d}^+}{\mathrm{d}t} u(t) = \limsup_{s \downarrow t} \frac{u(s) - u(t)}{s - t}, \quad \frac{\mathrm{d}}{\mathrm{d}t^+} u(t) = \liminf_{s \downarrow t} \frac{u(s) - u(t)}{s - t}.$$

Letting X be a nonempty set, we indicate a trajectory $x : \mathbb{R}_+ \to X$ with the symbol $x(\cdot)$ or $(x_t)_{t>0}$. The point at time $t > 0$ on the trajectory is denoted by x_t. The image of the map x over $I \subset \mathbb{R}_+$ is indicated by $\{x_t\}_{t \in I}$. Given a function $\phi : X \to (-\infty, +\infty]$, we set $\mathscr{D}(\phi) = \{x \in X \,:\, \phi(x) < +\infty\}$ to be its *domain* and we say that ϕ is *proper* if $\mathscr{D}(\phi) \neq \emptyset$.

Given a *directed set* $\mathcal{A}$ (i.e., a nonempty set with a preorder $\leq$ such that every pair of elements has an upper bound), we recall that a *net* in X is a mapping $x : \mathcal{A} \to X$, which we also indicate as $(x_a)_{a \in \mathcal{A}}$. We say that $(y_b)_{b \in \mathcal{B}}$ is a subnet of $(x_a)_{a \in \mathcal{A}}$ if there exists a nondecreasing *final* function $J : \mathcal{B} \to \mathcal{A}$ such that $y_b = x_{J(b)}$ for all $b \in \mathcal{B}$. We use the symbol $\sigma-\lim_a x_a$ to indicate the limit of $(x_a)_{a \in \mathcal{A}}$ with respect to the topology σ. We also write $x_\alpha \xrightarrow{\sigma} x$.

Costs. We assume to be given a *cost* $c : X \times X \to \mathbb{R}$. We say that the cost is *dissipative* if

$$c \geq 0 \quad \text{and} \quad c(x, x') = 0 \text{ if and only if } x = x'. \tag{Diss}$$

This in particular entails the dissipativity of both discrete-time and continuous-time evolutions driven by $\phi = f + g$. More precisely, in Section 3, for splitting schemes on $\phi = f + g$, nonnegativity is used to weave the two iterations together. For the explicit scheme, i.e. $g = 0$, nonnegativity is not necessary since c-concavity ensures that $\phi = f$ decreases. Costs fulfilling (Diss) that are symmetric have some surprising connections with notions of tropical monotonicity [AFG24, Proposition 4.1]. As also made clear in Section 3, we often handle specific expressions, in which differences of cost appear, so that replacing the lower bound 0 in (Diss) by any other constant is admissible.

2.1. Conditions for compatibility of the topology σ with (c, ϕ).

To discuss convergence and continuity we need to introduce a topology σ on the space X. This topology σ will be required to be compatible with the cost c and the energy ϕ. We start by specifying such compatibility requirement in Definition 2.1, providing later some sufficient conditions for such compatibility to hold.

Definition 2.1 (Compatible topology)**.** *Given c satisfying (Diss) and $\phi : X \to (-\infty, +\infty]$, we say that a Hausdorff topology σ on X is* compatible with the pair (c, ϕ) *if*

(1) *c-convergent nets in sublevels of ϕ are σ-convergent: if $(x_\alpha)_{\alpha \in \mathcal{A}} \subset \{\phi \leq r\}$ for some $r \in \mathbb{R}$, $x \in X$, and $c(x_\alpha, x) \to 0$ or $c(x, x_\alpha) \to 0$, then $x_\alpha \xrightarrow{\sigma} x$.*
(2) *ϕ is σ-lower semicontinuous and c is jointly σ-lower semicontinuous.*

We say that σ is forward-Cauchy-compatible with the pair (c, ϕ) *if furthermore*

(3) *c-forward-Cauchy sequences in sublevels of ϕ are σ-convergent: if $(x_n)_n \subset \{\phi \leq r\} \subset X$ for some $r \in \mathbb{R}$ satisfies*

$$\forall \varepsilon > 0 \, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } c(x_n, x_m) < \varepsilon \quad \forall n \geq m > N_\varepsilon,$$

then there exists $x \in X$ *such that* $x_n \xrightarrow{\sigma} x$.

Note that we do not require c to be symmetric. When this is the case, we simply refer to c-Cauchy sequences, omitting the term *forward*. As the topology σ is not assumed to be metrizable, in the following we deal with nets instead of sequences.

Example 1.1 (continued). Let $\omega : [0, +\infty) \to [0, +\infty)$ be a nonnegative, strictly increasing function such that $\omega(0) = 0$. Let (X, d) be a complete metric space with induced topology σ_d, and let $c := \omega(d)$. If ϕ is σ_d-lower semicontinuous, then σ_d is compatible with (c, ϕ) and c is jointly σ_d-continuous. In [CRZ09, RMS08, OZ23] asymmetric distances are also considered: assume (X, d) is an asymmetric metric space (that is, d satisfies all the axioms of a distance but symmetry) and let σ be the topology induced by the forward balls (cf. [OZ23, Definition 2.1], [CRZ09, Definition 2.1]). Then, the assumptions 4.2, 4.3, 4.7 of [CRZ09] on σ and ϕ are fulfilled in our setting, that is, σ is forward-Cauchy compatible with (c, ϕ).

Example 1.2 (continued). For $X \subset \mathcal{P}(\mathbb{X})$ with $\mathbb{X}$ being a Polish space and $c = \mathrm{KL}$, σ can be chosen to be the weak topology on measures which is dominated by the strong topology induced by the TV distance. By Pinsker's inequality, valid on measurable spaces [vEH14, Theorem 31], KL dominates the TV distance, whence (1) is satisfied. Moreover, it is known that KL is jointly σ-lower semicontinuous [vEH14, Theorem 19], so that (2) holds for all ϕ which are σ-l.s.c. To ensure finite values for c, we can take X to be a σ-closed subset of $\mathcal{P}(\mathbb{X})$ such that $\mathrm{KL}(x|y) < +\infty$ for all $x, y \in X$, e.g., fixing a measure ρ and setting $\mu \in X$ if and only if $\mu = f\rho$ with $f \in [a, b]$ ρ-a.s. with $0 < a < b < +\infty$. In the latter case, X is naturally identified with a subset of $L^1(\rho)$ and one could consider on X the $L^1(\rho)$-topology σ_ρ which is stronger than σ. Consequently, KL is jointly σ_ρ-continuous on X, so that σ_ρ is compatible with (c, ϕ) for any σ_ρ-lower semicontinuous functional $\phi : X \to (-\infty, +\infty]$.

Example 1.3 (continued). For $X = \mathcal{P}(\mathbb{X})$ with compact $\mathbb{X}$, $c = \mathrm{S}_\epsilon$ and σ the weak topology on measures, by [FSV⁺18, Theorem 1], S_ϵ is jointly σ-continuous, so (2) holds whenever ϕ is σ-l.s.c. Moreover, S_ϵ metrizes the weak convergence, hence (1) is satisfied. Since $X = \mathcal{P}(\mathbb{X})$ is σ-compact, owing to Lemma 2.2 below one has that forward-Cauchy-compatibility also holds.

We now present some sufficient conditions for the compatibility of σ with the pair (c, ϕ). The first condition applies to the case of a function ϕ with σ-compact sublevels.

Lemma 2.2 (ϕ with σ-compact sublevels)**.** *Let c be a cost on X satisfying* (Diss)*, let σ be a Hausdorff topology on X such that c is jointly σ-lower semicontinuous, and such that the sublevels of $\phi : X \to (-\infty, +\infty]$ are σ-compact. Then, σ is forward-Cauchy-compatible with (c, ϕ).*

Proof. Since σ is Hausdorff, the sublevels of ϕ are closed, so that ϕ is σ-lower semicontinuous and property (2) of Definition 2.1 is satisfied.

Let us show that property (1) holds: let $(x_\alpha)_{\alpha \in \mathcal{A}} \subset X$ be a net in a sublevel of ϕ and $x \in X$ be such that $c(x, x_\alpha) \to 0$. Let $(x_\beta)_{\beta \in \mathcal{B}}$ be any subnet of $(x_\alpha)_{\alpha \in \mathcal{A}}$ which is σ-converging to a point $y \in X$. By joint σ-lower semicontinuity of c, we have

$$0 = \liminf_{\beta \in \mathcal{B}} c(x, x_\beta) \geq c(x, y)$$

which implies that $x = y$ by (Diss). This, together with the σ-compactness of sublevels of ϕ, shows that the net $(x_\alpha)_{\alpha \in \mathcal{A}}$ σ-converges to x.

In order to check the c-forward-Cauchy property (3) of Definition 2.1, let $(x_n)_n \subset \{\phi \leq r\} \subset X$ for some $r \in \mathbb{R}$ and suppose that $(x_n)_n$ is c-forward-Cauchy. As the sublevel $\{\phi \leq r\}$ is σ-compact, we can find a σ-convergent subnet $(x_{J(\lambda)})_{\lambda \in \Lambda}$, where Λ is a directed set and $J : \Lambda \to \mathbb{N}$ is a final monotone function, namely, $\sigma\!-\lim_\lambda x_{J(\lambda)} = x$ for some $x \in X$. Fix $\varepsilon > 0$ and let $N_\varepsilon \in \mathbb{N}$ be such that $c(x_n, x_m) < \varepsilon$ for every $n \geq m \geq N_\varepsilon$. Since J is a final function, we can find some $\lambda_{\varepsilon,m} \in \Lambda$ such that $\lambda \geq m \geq N_\varepsilon$ for every $\lambda \geq \lambda_{\varepsilon,m}$. Thus we have

$$c(x_{J(\lambda)}, x_m) < \varepsilon \quad \text{for every } m \geq N_\varepsilon,\ \lambda \geq \lambda_{\varepsilon,m}.$$

Passing first to the $\liminf_\lambda$ and using the joint σ-lower semicontinuity of c, we get that

$$c(x, x_m) \leq \varepsilon \quad \text{for every } m \geq N_\varepsilon,$$

which means that x_n c-converges to x. By the first part of this proof, we deduce that x_n σ-converges to x, which proves the assertion. $\square$

Alternatively to checking that a given topology is compatible, one can use the cost c to construct one. Given a symmetric cost c satisfying (Diss), referred to as a *semimetric* in part of the literature, see [BP22, Section 2], one can define a topology σ_c on X by prescribing that a set $A \subset X$ is *open* (hence, $A \in \sigma_c$) if for every $x \in A$ there exists $r > 0$ such that

$$B_c(x, r) := \{y \in X : c(x, y) < r\} \subset A.$$

Notice that in general σ_c is not Hausdorff, $B_c(x, r) \notin \sigma_c$, c-convergent sequences do not have the c-Cauchy property, and c is not jointly σ_c-continuous. However, if c is a *regular semimetric* in the sense of Lemma 2.3 below, these patologies do not appear.

Lemma 2.3 (c regular semimetric)**.** *Suppose that c is a symmetric cost satisfying* (Diss) *and that it is additionally* regular*, i.e.*

$$\lim_{r\to 0} \sup_{y\in X} \sup_{z,w\in B_c(y,r)} c(z, w) = 0, \tag{2.1}$$

or, equivalently, that $\Phi_c : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ defined as

$$\Phi_c(r_1, r_2) := \sup\{c(x, y) \,|\, \exists\, p \in X : c(p, x) \le r_1,\ c(p, y) \le r_2\} \tag{2.2}$$

is continuous at $(0, 0)$. Moreover, assume that (X, c) is c-complete (i.e., c-Cauchy sequences are c-convergent) and ϕ is σ_c-lower semicontinuous. Then, σ_c is compatible with (c, ϕ) and c is jointly σ_c-continuous.

Proof. The equivalence between (2.1) and (2.2) can be found in [BP17, Lemma 1]. By [CJT18, Theorem 3.2], the regularity and completeness of c imply the existence of a complete metric ϱ on X which is uniformly equivalent to c, i.e.,

(1) for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\varrho(x, y) < \varepsilon$ whenever $c(x, y) < \delta$;
(2) for every $\varepsilon > 0$ there exists $\delta > 0$ such that $c(x, y) < \varepsilon$ whenever $\varrho(x, y) < \delta$.

This gives in particular that σ_c is induced by ϱ so that σ_c is Hausdorff, ϱ is σ_c-continuous (hence also c is), and that c-convergent sequences are ϱ-convergent, hence also σ_c-convergent. $\square$

2.2. **Equivalent formulations and main properties of the EVI.** In this section, we introduce the notion of EVI solution and discuss some of its properties. To start with, we present a result on equivalent characterizations of trajectories.

Theorem 2.4 (Equivalent EVI definitions)**.** *Let $\lambda \in \mathbb{R}$, $\phi : X \to (-\infty, +\infty]$, c satisfy* (Diss)*, and the topology σ be compatible with the pair (c, ϕ). Given the trajectory $x(\cdot) : \mathbb{R}_+ \to X$, for either $\lambda \ge 0$ or $\lambda < 0$ and c additionally separately σ-continuous, the following are equivalent:*

i) *$x(\cdot)$ is σ-continuous, $x_t \in \mathscr{D}(\phi)$ for all $t > 0$, and $x(\cdot)$ satisfies the* EVI in differential form*:*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} c(x, x_t) + \lambda c(x, x_t) \le \phi(x) - \phi(x_t) \quad \forall t > 0, x \in \mathscr{D}(\phi). \tag{2.3}$$

ii) *For all $x \in \mathscr{D}(\phi)$, the maps $t \mapsto \phi(x_t)$ and $t \mapsto c(x, x_t)$ belong to $L^1_{\rm loc}(\mathbb{R}_+)$, $(s, t) \mapsto c(x_s, x_t)$ is Lebesgue measurable on $\mathbb{R}_+ \times \mathbb{R}_+$, and $x(\cdot)$ satisfies the* EVI in integral form

$$c(x, x_t) - c(x, x_s) + \lambda \int_s^t c(x, x_r)\,\mathrm{d}r \le (t - s)\phi(x) - \int_s^t \phi(x_r)\,\mathrm{d}r \quad \forall\, 0 < s \le t. \tag{2.4}$$

iii) *$x(\cdot)$ satisfies the* EVI in exponential-integral form

$$e^{\lambda(t-s)} c(x, x_t) - c(x, x_s) \le E_\lambda(t - s)\,(\phi(x) - \phi(x_t)) \quad \forall\, 0 < s \le t, x \in \mathscr{D}(\phi), \tag{2.5}$$

where $E_\lambda(t) := \int_0^t e^{\lambda r}\,\mathrm{d}r = \begin{cases} \frac{e^{\lambda t}-1}{\lambda} & \text{for } \lambda \ne 0 \\ t & \text{for } \lambda = 0. \end{cases}$

Furthermore, if $x(\cdot)$ satisfies any of the above, then $t \mapsto \phi(x_t)$ is lower semicontinuous and non-increasing, and we have the oriented local Lipschitzianity relation

$$0 \le c(x_s, x_t) \le E_{-\lambda}(t - s)(\phi(x_{s_0}) - \phi(x_{t_0})) \quad \forall\, 0 < s_0 \le s \le t \le t_0. \tag{2.6}$$

Proof. The proof follows the blueprint of [MS20, Theorem 3.3], the main difference being that here we do not assume that c is jointly σ-continuous, but merely jointly σ-l.s.c. and we do not use the triangle inequality. This calls for some adaptation of the argument. Along this proof, we use the short-hand notation $\tilde{\phi} = \phi \circ x$.

i) $\Rightarrow$ ii): Fix $s > 0$ and let $\zeta(t) := c(x, x_t)$ and $\eta(t) := \frac{\lambda}{2} c(x, x_t) + \phi(x_t) - \phi(x)$. Both functions are lower semicontinuous, since ϕ, $c(x, \cdot)$, and $\lambda c(x, \cdot)$ are such, and $x(\cdot)$ is σ-continuous. Then, [MS20, Lemma A.1] gives that $\zeta, \eta \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ and $t \mapsto \zeta(t) + \int_s^t \eta(r)\,\mathrm{d}r$ is nonincreasing on $[s, +\infty)$ which is precisely (2.4). The lower semicontinuity, whence the measurability of $(s,t) \mapsto c(x_s, x_t)$ follows from the σ-continuity of $x(\cdot)$ and from the joint σ-lower semicontinuity of c.

iii) $\Rightarrow$ i): Inequality (2.5) in particular implies that $\phi(x_t) < +\infty$, so that we can choose $x = x_s \in \mathscr{D}(\phi)$ in (2.5). Since $c(x_s, x_s) = 0$ by (Diss), this gives

$$0 \leq \frac{e^{\lambda(t-s)}}{E_\lambda(t-s)} c(x_s, x_t) \leq \phi(x_s) - \phi(x_t). \tag{2.7}$$

This proves that $\tilde{\phi}$ is nonincreasing, hence also locally bounded in $\mathbb{R}_+$. As E_λ is continuous and $E_\lambda(0) = 0$, we can take limits in (2.5), obtaining for all $t_0 > 0$ and $x \in \mathscr{D}(\phi)$

$$\limsup_{t \downarrow t_0} c(x, x_t) \leq c(x, x_{t_0}) \leq \liminf_{s \uparrow t_0} c(x, x_s).$$

The second inequality entails that $t \mapsto c(x, x_t)$ is left σ-l.s.c. Evaluating at $x = x_{t_0}$ gives $\limsup_{t \downarrow t_0} c(x_{t_0}, x_t) = 0$, hence that $\lim_{t \downarrow t_0} c(x_{t_0}, x_t) = 0$. The compatibility with σ in Definition 2.1-(1) gives that $x(\cdot)$ is right σ-continuous, whence $t \mapsto c(x, x_t)$ is also right l.s.c., hence l.s.c. Since $\tilde{\phi}$ is nonincreasing, $x(\cdot)$ is right σ-continuous and ϕ is σ-l.s.c., we have that $\tilde{\phi}$ is l.s.c. As the lim sup of a sum is larger than the sum of a lim sup and a lim inf, we moreover have that

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=t_0} (e^{\lambda(t-t_0)} c(x, x_t)) \geq \frac{\mathrm{d}^+}{\mathrm{d}t}\Big|_{t=t_0} c(x, x_t) + \liminf_{h \downarrow 0} \lambda c(x, x_{t_0+h}) \quad \forall t_0 > 0.$$

Note that we have $\liminf_{h \downarrow 0} \lambda c(x, x_{t_0+h}) \geq \lambda c(x, x_{t_0})$ both for $\lambda \geq 0$, as c is jointly σ-l.s.c., and $\lambda < 0$, as c is assumed to be separately σ-continuous in this case. By dividing by $(t-s)$ in (2.5) and taking the limit $t \downarrow s$, we hence obtain (2.3).

To conclude, we have to show that $x(\cdot)$ is left σ-continuous. Letting $t > 0$, we use the nonincreasingness of $\tilde{\phi}$ and (2.7) to obtain for $s \in [t/2, t]$

$$0 \leq c(x_s, x_t) \leq \frac{E_\lambda(t-s)}{e^{\lambda(t-s)}} (\phi(x_s) - \phi(x_t)) \leq \frac{E_\lambda(t-s)}{e^{\lambda(t-s)}} (\phi(x_{t/2}) - \phi(x_t)). \tag{2.8}$$

Taking the lim sup as $s \uparrow t$, we deduce $\lim_{s \uparrow t} c(x_s, x_t) = 0$ so that, by the compatibility with σ in Definition 2.1-(1), we get that $x_s \xrightarrow{\sigma} x_t$ and $x(\cdot)$ is left σ-continuous at t.

ii) $\Rightarrow$ i): First of all, by taking limits in (2.4), for all $t_0 > 0$ and $x \in \mathscr{D}(\phi)$, we have that $\limsup_{t \downarrow t_0} c(x, x_t) \leq c(x, x_{t_0})$. Reasoning as in the previous implication, we deduce that $x(\cdot)$ is right σ-continuous. Now we show that $\tilde{\phi}$ is nonincreasing, first when restricted to its Lebesgue points, and then in general.

Let $0 < t_0 < t_1$ be Lebesgue points for $\tilde{\phi}$. For a.e. $s \in [t_0, t_1]$ we write (2.4) for $t = s + h$ and $x = x_s$. Integrating w.r.t. $s \in [t_0, t_1]$ we obtain

$$\begin{aligned} &\int_{t_0}^{t_1} c(x_s, x_{s+h})\,\mathrm{d}s + \lambda \int_{t_0}^{t_1} \int_s^{s+h} c(x_s, x_r)\,\mathrm{d}r\,\mathrm{d}s \\ &\qquad \leq \int_{t_0}^{t_1} \int_s^{s+h} \left(\tilde{\phi}(s) - \tilde{\phi}(r)\right) \mathrm{d}r\,\mathrm{d}s =: h^2 \eta(h). \end{aligned} \tag{2.9}$$

Now, let us work on the r.h.s. to relate it to $\tilde\phi(t_0)-\tilde\phi(t_1)$. We write

$$\begin{aligned}\int_{t_0}^{t_1}\int_s^{s+h}\Big(\tilde\phi(s)-\tilde\phi(r)\Big)\,\mathrm{d}r\,\mathrm{d}s &= \int_{t_0}^{t_1}\int_0^{h}\Big(\tilde\phi(s)-\tilde\phi(s+r)\Big)\,\mathrm{d}r\,\mathrm{d}s\\
&= \int_0^h\int_0^r\Big(\tilde\phi(t_0+\xi)-\tilde\phi(t_1+\xi)\Big)\,\mathrm{d}\xi\,\mathrm{d}r\\
&= \int_0^h\int_0^h\Big(\tilde\phi(t_0+\xi)-\tilde\phi(t_1+\xi)\Big)\,\chi_{[0,r]}(\xi)\,\mathrm{d}\xi\,\mathrm{d}r\\
&= \int_0^h\Big(\tilde\phi(t_0+\xi)-\tilde\phi(t_1+\xi)\Big)\,(h-\xi)\,\mathrm{d}\xi\\
&= h^2\int_0^1\Big(\tilde\phi(t_0+h\xi)-\tilde\phi(t_1+h\xi)\Big)\,(1-\xi)\,\mathrm{d}\xi. \end{aligned}\tag{2.10}$$

We set $u(h):=\int_{t_0}^{t_1}c(x_s,x_{s+h})\,\mathrm{d}s$ so that (2.9) yields

$$u(h)+\lambda\int_0^h u(r)\,\mathrm{d}r\le h^2\eta(h)\quad\forall h>0.$$

An application of Gronwall's lemma entails

$$h^{-2}u(h)\le e^{h\lambda^-}\sup_{\delta\in[0,h]}\eta(\delta)\tag{2.11}$$

with $\lambda^-:=\max\{0,-\lambda\}$ so that, passing to the $\limsup_{h\downarrow0}$ in (2.9), we get

$$0\overset{0\le c}{\le}\limsup_{h\downarrow0}\int_{t_0}^{t_1}\frac{c(x_s,x_{s+h})}{h^2}\,\mathrm{d}s\le\lim_{h\downarrow0}\eta(h)=\frac12(\tilde\phi(t_0)-\tilde\phi(t_1)),\tag{2.12}$$

where the last equality follows by, e.g., [SS05, Theorem 2.1] and the fact that t_0 and t_1 are Lebesgue points of $\tilde\phi$. Hence, $\tilde\phi$ is nonincreasing when restricted to its Lebesgue points. However, since ϕ is l.s.c.and $x(\cdot)$ is right σ-continuous, for every Lebesgue point t_0 and any $t>t_0$ we have that $\tilde\phi(t_0)\ge\tilde\phi(t)$. We deduce in particular that $x_t\in\mathscr{D}(\phi)$ for every $t>0$ and that, if s,t and $h>0$ are given in such a way that that $0<s<s+h<t$, then $\tilde\phi(t)\le\tilde\phi(t_0)$ for every Lebesgue point t_0 of $\tilde\phi$ in $(s,s+h)$. This in particular, gives that $\tilde\phi(t)\le\frac1h\int_0^h\tilde\phi(s+r)\,\mathrm{d}r$.

We now fix $0<s<t$ and take any $h>0$ such that $s+h<t$. By applying (2.4) to $x=x_s$ and $t=s+h$, discarding the nonnegative term $c(x_s,x_{s+h})$, and using the above bound for $\tilde\phi(t)$ we get

$$\tilde\phi(t)-|\lambda|\int_0^1 c(x_s,x_{s+rh})\,\mathrm{d}r\le\tilde\phi(s)\quad\forall\,0<s<s+h<t.$$

Passing to $\limsup_{h\downarrow0}$ and using the joint σ-lower semicontinuity of c and the σ-right continuity of $x(\cdot)$, an application of Fatou's lemma (recall that c is bounded from below) gives $\tilde\phi(t)\le\tilde\phi(s)$ for every $0<s<t$, as desired.

Using that $\tilde\phi$ is nonincreasing, we show next that $x(\cdot)$ is left σ-continuous. Choosing $x=x_s$ in (2.4) and setting $v(t):=c(x_s,x_t)$ we obtain

$$v(t)\le|\lambda|\int_s^t v(r)\,\mathrm{d}r+\int_s^t(\tilde\phi(s)-\tilde\phi(r))\,\mathrm{d}r\quad\forall\,0<s\le t.\tag{2.13}$$

Now fix any $t>0$ and take any $0<s_0<t<t_0$. By applying Gronwall's lemma to (2.13) and integrating by parts we deduce

$$\begin{aligned}v(t)&\le\int_s^t(\tilde\phi(s)-\tilde\phi(r))\,\mathrm{d}r+\int_s^t\int_s^r(\tilde\phi(s)-\tilde\phi(u))\,\mathrm{d}u\,|\lambda|e^{|\lambda|(t-r)}\,\mathrm{d}r\\
&\le\int_s^t(\tilde\phi(s_0)-\tilde\phi(t_0))\,\mathrm{d}r+\int_s^t\int_s^r(\tilde\phi(s_0)-\tilde\phi(t_0))\,\mathrm{d}u\,|\lambda|e^{|\lambda|(t-r)}\,\mathrm{d}r\\
&=(\tilde\phi(s_0)-\tilde\phi(t_0))\int_0^{t-s}e^{|\lambda|r}\,\mathrm{d}r\end{aligned}$$

for every $s_0 < s < t$. Taking the $\limsup_{s\uparrow t}$ in the above inequality we deduce that $\limsup_{s\uparrow t} c(x_s, x_t) \leq 0$, thus giving, via the usual compatibility conditions, that $x(\cdot)$ is left σ-continuous.

Furthermore, since $\tilde{\phi}$ is l.s.c., we have

$$\phi(x_s) \leq \liminf_{h\downarrow 0} \tilde{\phi}(s+h) \leq \liminf_{h\downarrow 0} \frac{1}{h} \int_s^{s+h} \tilde{\phi}(r)\,\mathrm{d}r. \tag{2.14}$$

So, when dividing (2.4) by $(t-s)$ and taking the lim sup for $t \downarrow s$, the r.h.s. can be bounded from above by $\phi(x) - \phi(x_s)$ and we obtain (2.3) using Fatou's lemma for $\lambda \geq 0$ and the separate σ-continuity of c and the dominated convergence theorem for $\lambda < 0$.

i) $\Rightarrow$ iii): Multiplying (2.3) by $e^{\lambda t}$ gives

$$\begin{aligned}
\frac{\mathrm{d}^+}{\mathrm{d}t}(e^{\lambda t} c(x, x_t)) &= \limsup_{h\downarrow 0} \frac{e^{\lambda(t+h)} c(x, x_{t+h}) - e^{\lambda t} c(x, x_t)}{h} \\
&= \limsup_{h\downarrow 0} \left[e^{\lambda(t+h)} \frac{c(x, x_{t+h}) - c(x, x_t)}{h} + c(x, x_t) \frac{e^{\lambda(t+h)} - e^{\lambda t}}{h} \right] \\
&\leq \limsup_{h\downarrow 0} \left[e^{\lambda(t+h)} \frac{c(x, x_{t+h}) - c(x, x_t)}{h} \right] + \limsup_{h\downarrow 0} \left[c(x, x_t) \frac{e^{\lambda(t+h)} - e^{\lambda t}}{h} \right] \\
&= e^{\lambda t} \frac{\mathrm{d}^+}{\mathrm{d}t} c(x, x_t) + \lambda e^{\lambda t} c(x, x_t) \leq e^{\lambda t}(\phi(x) - \phi(x_t)).
\end{aligned}$$

Integrating the latter through [MS20, Lemma A.1] and using the fact that $\tilde{\phi}$ does not increase, which is shown in ii) $\Leftrightarrow$ i), this proves that (2.5) holds. □

Remark 2.5 (Separate and joint lower semicontinuity). It would also be possible to formulate and prove Theorem 2.4 by assuming the cost c to be merely separately σ-lower semicontinuous in the case $\lambda \geq 0$, discarding the joint measurability of $(s,t) \mapsto c(x_s, x_t)$ in ii). Indeed, the only point where this is used is in (2.9) when writing a double integral, and the latter can be dropped using that $\lambda c \geq 0$ when $\lambda \geq 0$. In the case $\lambda < 0$, one should nevertheless assume c to be additionally jointly σ-continuous, in order for the full proof to work.

Remark 2.6 (On the EVI formulations and the estimates). Arguably, the differential (2.3) and integral (2.4) formulations of the EVI are the natural generalization of the classical formulations of the metric case in [MS20, Theorem 3.3]. The one with the exponential (2.5) is nevertheless very useful to achieve estimates. Its drawback is that it misses the energy identity by a factor 2, as easily seen when taking $\lambda = 0$ and $x = x_s$ in (2.5), or noticing that for $x = x_t$ the l.h.s. is negative. As a corrective, for symmetric costs, the missing term is expressed in (2.18) below. Notice also that the formulation in (2.5) does not depend on the topology σ.

Definition 2.7 (EVI solution)**.** *Let $\phi : X \to (-\infty, +\infty]$, and c satisfy* (Diss)*. A trajectory $x : \mathbb{R}_+ \to X$ satisfying* (2.5) *in Theorem 2.4 for some $\lambda \in \mathbb{R}$ is called a* EVI solution *for the triplet (X, c, ϕ) and we write $x(\cdot) \in \mathrm{EVI}_\lambda(X, c, \phi)$. Equivalently, $x(\cdot) \in \mathrm{EVI}_\lambda(X, c, \phi)$ if it satisfies either* (2.3) *or* (2.4) *for a topology σ on X which is compatible with the pair (c, ϕ) (with c additionally separately σ-continuous, if $\lambda < 0$).*

Remark 2.8 (Definition up to time $t = 0$). The equivalences of Theorem 2.4 holds also for a curve $x : [0, +\infty) \to X$ down to time $t = 0$, provided that the cost c is additionally separately σ-continuous (and not just jointly σ-lower semicontinuous). Indeed, in this case relation ii) written for $0 < s \leq t$ is equivalent to ii) written for $0 \leq s \leq t$ which is in turn equivalent to iii) for $0 \leq s \leq t$. To conclude, it is enough to observe that the implication iii) $\Rightarrow$ i) works exactly as above upon replacing $t_0 > 0$ with $t_0 \geq 0$.

On the basis of these observations, if c is additionally separately σ-continuous, one can extend the definition of EVI solutions to curves $x : [0, +\infty) \to X$ and, without introducing new notation, use the same symbol $\mathrm{EVI}_\lambda(X, c, \phi)$. In particular, notice that every curve $(x_t)_{t\geq 0} \in \mathrm{EVI}_\lambda(X, c, \phi)$ is σ-continuous in $[0, +\infty)$.

We now introduce a concept analogue to the metric derivative in the case $c = d^2/2$.

Definition 2.9 (c-cost derivative)**.** *Let c be a symmetric cost satisfying* (Diss) *and let $x : [0,+\infty) \to X$. We say that $x(\cdot)$ has a* c-cost derivative *at time $t \geq 0$ if*

$$|\dot{x}_{t+}|_c^2 := \lim_{h \downarrow 0} \frac{2c(x_t, x_{t+h})}{h^2} \tag{2.15}$$

exists and is finite.

Note that by using EVI it is easy to give a lower (resp. upper) bound to $c(x_t, x_{t+h})$ (resp. $c(x_{t+h}, x_t)$). The symmetry requirement in the above definition allows us to conclude that the two are indeed equal. Symmetry will also be crucial in the proof of Theorem 2.10, e.g., to prove the contraction estimate (2.16), which is obtained summing two EVIs with the curves involved playing different roles.

The following theorem is in the spirit of [MS20, Theorem 3.5], and its forerunner [DS14, Theorem 6.9], a distinctive aspect being however the fact that in the current general setting, the metric slope does not appear.

Theorem 2.10 (Properties of EVI solutions)**.** *Let $\lambda \in \mathbb{R}$, $\phi : X \to (-\infty, +\infty]$, c be a symmetric cost satisfying* (Diss)*, σ be compatible with the pair (c, ϕ), and let $x : \mathbb{R}_+ \to X$ belong to $\mathrm{EVI}_\lambda(X, c, \phi)$. Then the following holds.*

- **λ-contraction.** *Let $\tilde{x}(\cdot) \in \mathrm{EVI}_\lambda(X, c, \phi)$. Then,*
$$c(x_t, \tilde{x}_t) \leq e^{-2\lambda(t-s)} c(x_s, \tilde{x}_s) \quad \forall\, 0 < s \leq t. \tag{2.16}$$
- **Energy identity and regularizing effect.** *For every $t > 0$ the following right limits exist and are equal:*
$$\frac{\mathrm{d}}{\mathrm{d}t} \phi(x_{t+}) = -\lim_{h \downarrow 0} \frac{2c(x_t, x_{t+h})}{h^2} = -|\dot{x}_{t+}|_c^2. \tag{2.17}$$
Moreover the map $t \mapsto e^{2\lambda t} |x_{t+}|_c^2$ is nonincreasing and the map $t \mapsto \phi(x_t)$ is (locally semi-, if $\lambda < 0$) convex in $\mathbb{R}_+$. As a consequence, $t \mapsto \phi(x_t)$ is differentiable outside a countable set $\mathcal{T} \subset \mathbb{R}_+$.
- **A priori estimates***. For every $x \in \mathscr{D}(\phi)$, we have the following refinement of* (2.5)*:*
$$e^{\lambda(t-s)} c(x, x_t) - c(x, x_s) + \frac{(E_\lambda(t-s))^2}{2} |\dot{x}_{t+}|_c^2 \leq E_\lambda(t-s) \left(\phi(x) - \phi(x_t)\right) \quad \forall\, 0 < s \leq t. \tag{2.18}$$
- **Asymptotic behaviour as $t \to +\infty$.** *Assume that ϕ is bounded from below. If $\lambda > 0$ and either i) σ is Cauchy-compatible with (c, ϕ) or ii) the sublevel sets ϕ are σ-compact, then ϕ has a unique minimum point $\bar{x}$ and*
$$\begin{aligned} &\lambda c(\bar{x}, x_t) \leq \phi(x_t) - \phi(\bar{x}) \leq \frac{\lambda c(\bar{x}, x_{t_0})}{e^{\lambda(t-t_0)} - 1} \text{ and} \\ &c(\bar{x}, x_t) \leq c(\bar{x}, x_{t_0}) e^{-2\lambda(t-t_0)} \quad \forall\, 0 < t_0 < t. \end{aligned} \tag{2.19}$$
If $\lambda = 0$ and there exists a minimizer $\bar{x}$ of ϕ, then $t \mapsto c(\bar{x}, x_t)$ is nonincreasing and
$$\phi(x_t) - \phi(\bar{x}) \leq \frac{c(\bar{x}, x_{t_0})}{t - t_0} \quad \forall\, 0 < t_0 < t. \tag{2.20}$$
For $\lambda \geq 0$ and $\bar{x}$ a minimizer of ϕ, we furthermore have, for every $0 < t_0 < t$, that
$$2\lambda(\phi(x_t) - \phi(\bar{x})) \leq |\dot{x}_{t+}|_c^2, \quad |\dot{x}_{t+}|_c \leq |\dot{x}_{t_0+}|_c e^{-\lambda(t-t_0)}, \quad |\dot{x}_{t+}|_c \leq \frac{\sqrt{2c(\bar{x}, x_{t_0})}}{E_\lambda(t - t_0)}. \tag{2.21}$$
If $\lambda = 0$, ϕ has σ-compact sublevel sets, and c is jointly σ-continuous, then x_t σ-converges to a minimizer of ϕ as $t \to +\infty$.

Remark 2.11 (Time $t = 0$). Following up on Remark 2.8, we briefly discuss the possibility of extending some of the results of Theorem 2.10 to the case of a EVI solution $x : [0, +\infty) \to X$ defined at time $t = 0$, as well. Recall that, in this case, the cost function c is assumed to be additionally separately σ-continuous in order for the three formulations in Theorem 2.4 (written down to time $t = 0$) to be equivalent. We have the following:

(1) (λ-contractivity and uniqueness) If $\tilde{x} : [0, +\infty) \to X$ belongs to $\mathrm{EVI}_\lambda(X, c, \phi)$ and c is jointly σ-continuous at $(x_0, \tilde{x}_0)$, the contractivity estimate (2.16) holds also for $0 \le s \le t$. In particular, if $x_0 \in \mathscr{D}(\phi)$ and c is jointly σ-continuous at (x_0, x_0), there exists at most one EVI solution starting from x_0.
(2) (Stability w.r.t. the initial condition) If $x^n : [0, +\infty) \to X$ belong to $\mathrm{EVI}_\lambda(X, c, \phi)$, $x_0^n \xrightarrow{\sigma} x_0$, and c is jointly σ-continuous at the points (x_0^n, x_0) for every $n \in \mathbb{N}$, then $x_t^n \xrightarrow{\sigma} x_t$ for all $t \ge 0$. This directly follows from (2.16) written for $x^n(\cdot)$ and $x(\cdot)$ at times $0 = s \le t$.

Proof of Theorem 2.10. **λ-contraction.** Fix $0 < s \le t$. We write (2.5) for x_t with test point $x = \tilde{x}_t$, we multiply it by $\mathrm{e}^{\lambda(t-s)}$ and, using the symmetry of c, we sum it to (2.5) written for $\tilde{x}_t$ with test point $x = x_s$. The terms $e^{\lambda(t-s)} c(\tilde{x}_t, x_s)$ compensate by symmetry and we obtain

$$\begin{aligned}
&\mathrm{e}^{2\lambda(t-s)} c(\tilde{x}_t, x_t) - c(x_s, \tilde{x}_s) \\
&\quad \le \mathrm{E}_\lambda(t-s)\left(\phi(x_s) - \phi(x_t)\right) + \mathrm{E}_\lambda(t-s)\left(e^{\lambda(t-s)} - 1\right)\left(\phi(\tilde{x}_t) - \phi(x_t)\right)
\end{aligned}$$

for every $0 < s \le t < +\infty$. Reversing the roles of $x(\cdot)$ and $\tilde{x}(\cdot)$, summing up and multiplying by $\mathrm{e}^{2\lambda s}$ we get

$$2\mathrm{e}^{2\lambda t} c(x_t, \tilde{x}_t) - 2\mathrm{e}^{2\lambda s} c(x_s, \tilde{x}_s) \le \mathrm{e}^{2\lambda s} \mathrm{E}_\lambda(t-s)\left(\phi(x_s) - \phi(x_t) + \phi(\tilde{x}_s) - \phi(\tilde{x}_t)\right)$$

for every $0 \le s \le t < +\infty$. We now fix $s > 0$, divide by $t - s > 0$, and pass to the $\limsup$ as $t \downarrow s$. Observing that $\lim_{r\downarrow 0} \mathrm{E}_\lambda(r)/r = 1$ and using that $\tilde{\phi}$ is σ-l.s.c. so that the r.h.s. is nonpositive in the limit, we obtain

$$\frac{\mathrm{d}^+}{\mathrm{d}s} \mathrm{e}^{2\lambda s} c(x_s, \tilde{x}_s) \le 0 \qquad \text{for every } s > 0.$$

An application of [MS20, Lemma A.1] gives (2.16).

Energy identity and Regularizing effect. We start from (2.11) in the proof of the implication ii) $\Rightarrow$ i) in Theorem 2.4: recalling the definition of η and u in (2.9) and below (2.10), respectively, and using again the notation $\tilde{\phi} = \phi \circ x$, one has that for every $0 < t_0 < t_1$ it holds

$$\begin{aligned}
\int_{t_0}^{t_1} \frac{c(x_s, x_{s+h})}{h^2}\,\mathrm{d}s &\le e^{h\lambda^-} \sup_{0\le\delta\le h} \int_0^1 \left(\tilde{\phi}(t_0 + \delta\xi) - \tilde{\phi}(t_1 + \delta\xi)\right)(1 - \xi)\,\mathrm{d}\xi \\
&\le e^{h\lambda^-} \sup_{0\le\delta\le h} \int_0^1 \left(\tilde{\phi}(t_0) - \tilde{\phi}(t_1 + \delta\xi)\right)(1 - \xi)\,\mathrm{d}\xi,
\end{aligned}$$

where we used that $\tilde{\phi}$ is nonincreasing. Note that (2.11) is written for Lebesgue points $0 < t_0 < t_1$ of $\tilde{\phi}$ but it holds in general (the Lebesgue property is used only later in (2.12)). Passing to the $\limsup$ as $h \downarrow 0$ we obtain

$$\begin{aligned}
&\limsup_{h\downarrow 0} \int_{t_0}^{t_1} \frac{c(x_s, x_{s+h})}{h^2}\,\mathrm{d}s \\
&\quad \le \limsup_{h\downarrow 0} \int_0^1 \left(\tilde{\phi}(t_0) - \tilde{\phi}(t_1 + h\xi)\right)(1 - \xi)\,\mathrm{d}\xi \le \frac{1}{2}(\tilde{\phi}(t_0) - \tilde{\phi}(t_1)),
\end{aligned} \tag{2.22}$$

where the last inequality follows from Fatou's lemma and the local boundedness of $\tilde{\phi}$. We divide (2.22) by $t_1 - t_0$ and we take a $\liminf$ as $t_1 \downarrow t_0$, obtaining

$$\begin{aligned}
-\frac{\mathrm{d}^+}{\mathrm{d}t_0}\phi(x_{t_0}) &= \liminf_{t_1\downarrow t_0} \frac{\tilde{\phi}(t_0) - \tilde{\phi}(t_1)}{t_1 - t_0} \ge \liminf_{t_1\downarrow t_0} \limsup_{h\downarrow 0} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \frac{2c(x_s, x_{s+h})}{h^2}\,\mathrm{d}s \\
&\ge \limsup_{h\downarrow 0} \liminf_{t_1\downarrow t_0} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \frac{2c(x_s, x_{s+h})}{h^2}\,\mathrm{d}s \ge \limsup_{h\downarrow 0} \frac{2c(x_{t_0}, x_{t_0+h})}{h^2}
\end{aligned}$$

where we again used Fatou's lemma in the last inequality, as well as the joint σ- lower semicontinuity of c. This shows that

$$-\frac{\mathrm{d}^+}{\mathrm{d}t_0}\phi(x_{t_0}) \ge \limsup_{h\downarrow 0} \frac{2c(x_{t_0}, x_{t_0+h})}{h^2} \qquad \text{for every } t_0 > 0. \tag{2.23}$$

To show the converse inequality, we write (2.4) for $x = x_{t+h}$ and times t and $t+h$, and we use (2.6) to get

$$\begin{aligned}\frac{c(x_{t+h},x_t)}{h^2} &\geq \int_0^1 \frac{\phi(x_{t+hr})-\phi(x_{t+h})}{h}\,\mathrm{d}r - |\lambda|\int_0^1 \frac{c(x_{t+h},x_{t+rh})}{h}\,\mathrm{d}r\\ &\geq \int_0^1 \frac{\phi(x_{t+hr})-\phi(x_{t+h})}{h}\,\mathrm{d}r - |\lambda|(\phi(x_t)-\phi(x_{t+h}))\int_0^1 \frac{e^{-\lambda h(1-r)}-1}{-\lambda h}\,\mathrm{d}r\\ &= \int_0^1 \frac{\phi(x_{t+hr})-\phi(x_{t+h})}{h}\,\mathrm{d}r - \frac{|\lambda|}{2}(\phi(x_t)-\phi(x_{t+h}))(1+o(1))\end{aligned}$$

as $h\downarrow 0$. We can thus pass to the lim sup as $h\downarrow 0$ and, using the inequality $\limsup(a_n+b_n)\geq \limsup a_n + \liminf b_n$, discarding thus through lower semicontinuity the nonnegative term in $|\lambda|$, we obtain

$$\begin{aligned}\limsup_{h\downarrow 0}\frac{c(x_t,x_{t+h})}{h^2} &\geq \limsup_{h\downarrow 0}\int_0^1 \frac{\phi(x_{t+hr})-\phi(x_t)+\phi(x_t)-\phi(x_{t+h})}{h}\,\mathrm{d}r\\ &\geq \limsup_{h\downarrow 0}\frac{\phi(x_t)-\phi(x_{t+h})}{h} + \liminf_{h\downarrow 0}\int_0^1 \frac{\phi(x_{t+hr})-\phi(x_t)}{hr}r\,\mathrm{d}r\\ &\geq -\frac{\mathrm{d}}{\mathrm{d}t^+}\phi(x_t) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t^+}\phi(x_t) = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t^+}\phi(x_t)\\ &\geq -\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\phi(x_t) \geq \limsup_{h\downarrow 0}\frac{c(x_t,x_{t+h})}{h^2},\end{aligned}\tag{2.24}$$

where we used (2.23) in the last inequality and, to pass from the second to the third line, we have used the following inequality

$$\liminf_{h\downarrow 0}\int_0^1 \frac{\phi(x_{t+hr})-\phi(x_t)}{hr}r\,\mathrm{d}r \geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t^+}\phi(x_t).\tag{2.25}$$

To prove (2.25), we cannot use Fatou's lemma since the integrand is nonpositive. Instead let us denote the r.h.s. by $2\kappa\in[-\infty,0]$. If $\kappa=-\infty$, there is nothing to prove, so suppose $\kappa>-\infty$. For every $\varepsilon>0$, by the very definition of lim inf, we find $\delta_\varepsilon>0$ such that

$$\frac{\phi(x_{t+h'})-\phi(x_t)}{h'}\geq \kappa-\varepsilon \qquad \text{for every } h'\in(0,\delta_\varepsilon).$$

In particular

$$\frac{\phi(x_{t+hr})-\phi(x_t)}{hr}r\geq r(\kappa-\varepsilon) \qquad \text{for every } h\in(0,\delta_\varepsilon),\ r\in(0,1),$$

so that

$$\int_0^1 \frac{\phi(x_{t+hr})-\phi(x_t)}{hr}r\,\mathrm{d}r \geq \frac{1}{2}(\kappa-\varepsilon) \qquad \text{for every } h\in(0,\delta_\varepsilon).$$

Passing to the lim inf as $h\downarrow 0$ and then to the limit as $\varepsilon\downarrow 0$, gives (2.25).

The chain of inequalities (2.24) implies in particular that the right Dini derivative of $t\mapsto\phi(x_t)$ exists at every point $t>0$ and coincides with $-\limsup_{h\downarrow 0}\frac{2c(x_t,x_{t+h})}{h^2}$. Now that the existence of the right Dini derivative of $t\mapsto\phi(x_t)$ has been established, the same computations that led to (2.24), repeated with lim inf instead of lim sup (and thus using the inequality $\liminf(a_n+b_n)\geq \liminf a_n+\liminf b_n$) yield the existence of the limit $\lim_{h\downarrow 0}\frac{c(x_t,x_{t+h})}{h^2}$ for every $t>0$, thus establishing (2.17).

By (2.16) and the fact that $t\mapsto x_{t+h}$ belongs to $\mathrm{EVI}_\lambda(X,c,\phi)$ for every $h>0$, we deduce that $t\mapsto e^{2\lambda t}|\dot x_{t+}|_c^2$ is nonincreasing in $\mathbb{R}_+$. The equality in (2.17) gives that $t\mapsto e^{2\lambda t}\frac{\mathrm{d}}{\mathrm{d}t}\phi(x_{t+})$ is nondecreasing. As a consequence, the right derivative of the map

$$t\mapsto e^{2\lambda t}\phi(x_t)-2\lambda\int_0^t e^{2\lambda s}\phi(x_s)\,\mathrm{d}s$$

is nondecreasing. This in particular implies that $t \mapsto \phi(x_t)$ is convex on $\mathbb{R}_+$ if $\lambda \geq 0$ and locally semi-convex on $\mathbb{R}_+$ for $\lambda < 0$.

A priori estimates. To show (2.18) let us start by noting that, by taking derivatives in τ and using the definition of E_λ, for $\tau \geq 0$ we have

$$\frac{(E_\lambda(\tau))^2}{2} e^{-2\lambda\tau} = \frac{(E_{-\lambda}(\tau))^2}{2} = \frac{1}{2}\left(\int_0^\tau e^{-\lambda r}\,\mathrm{d}r\right)^2 = \int_0^\tau E_{-\lambda}(r) e^{-\lambda r}\,\mathrm{d}r.$$

Hence, using the fact that of $t \mapsto e^{\lambda t}|\dot{x}_{t+}|_c$ is nonincreasing on $\mathbb{R}_+$, the energy identity (2.17), the almost everywhere differentiability of $t \mapsto \phi(x_t)$, and by integrating by parts, we obtain

$$\begin{aligned}
\frac{(E_\lambda(t-s))^2}{2}|\dot{x}_{t+}|_c^2 &\leq \int_0^{t-s} E_{-\lambda}(r)e^{-\lambda r}e^{2\lambda r}|\dot{x}_{(s+r)+}|_c^2\,\mathrm{d}r \\
&\overset{(2.17)}{=} -\int_0^{t-s} E_{-\lambda}(r)e^{\lambda r}\left(\frac{\mathrm{d}}{\mathrm{d}r}\phi(x_{(s+r)})\right)\mathrm{d}r \\
&= \int_0^{t-s} e^{\lambda r}\phi(x_{(s+r)})\,\mathrm{d}r - E_{-\lambda}(t-s)e^{\lambda(t-s)}\phi(x_t) \\
&= \int_s^t e^{\lambda(r-s)}\left(\phi(x_r) - \phi(x_t)\right)\mathrm{d}r \\
&\overset{(2.3)}{\leq} -\int_s^t \left(e^{\lambda(r-s)}c(x_r, x)\right)'\mathrm{d}r + \int_s^t \left(e^{\lambda(r-s)}(\phi(x) - \phi(x_t)\right)\mathrm{d}r \\
&= -e^{\lambda(t-s)}c(x, x_t) + c(x, x_s) + E_\lambda(t-s)\left(\phi(x) - \phi(x_t)\right).
\end{aligned}$$

Asymptotic behaviour as $t \to +\infty$. Since $t \mapsto \phi(x_t)$ is nonincreasing and bounded from below, it converges to some $A \in \mathbb{R}$ as $t \to +\infty$ and $\{x_t\}_{t>0}$ is contained in a sublevel set of ϕ. In the case $\lambda > 0$, taking $0 < s \leq t$ (2.5) reads

$$0 \leq c(x, x_t) \leq e^{-\lambda(t-s)}E_\lambda(t-s)\big(\phi(x) - \phi(x_t)\big) + e^{-\lambda(t-s)}c(x, x_s) \quad \forall x \in \mathscr{D}(\phi). \tag{2.26}$$

Hence, for $x = x_s$ we obtain

$$0 \leq c(x_s, x_t) \leq \frac{e^{\lambda(t-s)} - 1}{\lambda e^{\lambda(t-s)}}\left(\phi(x_s) - \phi(x_t)\right) \leq \frac{|\phi(x_s) - \phi(x_t)|}{\lambda}. \tag{2.27}$$

Since c is symmetric, the inequality holds for any $t, s \geq 0$. If i) σ is Cauchy-compatible, then $x(\cdot)$ converges to some $\bar{x} \in X$ as $t \to +\infty$. Otherwise, let us assume that ii) the sublevel sets of ϕ are σ-compact. Fix $t_0 > 0$ and set $A_{t_0} := \{x \,|\, \phi(x) \leq \phi(x_{t_0})\}$. It is enough to show that, given two subnets $(x_{t_\eta})_\eta$ and $(x_{t_\gamma})_\gamma$ converging to points $\bar{x}, \bar{x}' \in A_{t_0}$, we necessarily have that $\bar{x} = \bar{x}'$. By (2.27), we have

$$0 \leq c(x_{t_\eta}, x_{t_\gamma}) \leq \frac{|\phi(x_{t_\eta}) - \phi(x_{t_\gamma})|}{\lambda} \quad \forall \eta, \gamma.$$

Using the joint σ- lower semicontinuity of c and the fact that $\lim_{t\to+\infty}\phi(x_t) = A$, we deduce that $c(\bar{x}, \bar{x}') = 0$ so that $\bar{x} = \bar{x}'$.

Let us show that $\bar{x}$ is the unique minimizer of ϕ. Taking the lim inf as $t \to +\infty$ in (2.26) and, using the σ- lower semicontinuity properties of c and ϕ, we get

$$0 \leq \lambda c(x, \bar{x}) \leq \phi(x) - \limsup_{t\to+\infty}\phi(x_t) \leq \phi(x) - \phi(\bar{x}) \quad \forall x \in \mathscr{D}(\phi), \tag{2.28}$$

which shows that $\bar{x}$ is a minimizer of ϕ and that it is unique by the nondegeneracy of c coming from (Diss) .

Similarly, for given $0 < t_0 < t$, considering (2.5) for $s = t_0$, and $x = \bar{x}$, we obtain

$$-c(\bar{x}, x_{t_0}) \leq e^{\lambda(t-t_0)}c(\bar{x}, x_t) - c(\bar{x}, x_{t_0}) \leq E_\lambda(t-t_0)\left(\phi(\bar{x}) - \phi(x_t)\right) \quad \text{for every } 0 < t_0 \leq t$$

which implies $\phi(x_t)-\phi(\bar{x}) \leq \frac{\lambda c(\bar{x},x_{t_0})}{e^{\lambda(t-t_0)}-1}$. In combination with (2.28) evaluated at $x = x_t$, the latter gives the first inequality in (2.19).

Note that (2.28) is nothing but relation (2.3) written for the constant trajectory $t \mapsto \bar{x}$, proving that it indeed belongs to $\mathrm{EVI}_\lambda(X,c,\phi)$. Therefore, the second part of (2.19) is simply the contractivity (2.16) for the curves $t \mapsto x_t$ and $t \mapsto \bar{x}$ with $s = t_0$.

In the case $\lambda = 0$, we use the fact that $t \mapsto \phi(x_t)$ is nonincreasing and the integral formulation (2.4) with $x = \bar{x}$ and $s = t_0$ to obtain that, for every $0 < t_0 \leq t$, it holds

$$(t-t_0)(\phi(x_t)-\phi(\bar{x})) \leq \int_{t_0}^t (\phi(x_r)-\phi(\bar{x}))\,\mathrm{d}r \overset{(2.4)}{\leq} c(\bar{x},x_{t_0}) - c(\bar{x},x_t) \overset{c\geq 0}{\leq} c(\bar{x},x_{t_0}). \tag{2.29}$$

This gives that $t \mapsto c(\bar{x},x_t)$ is nonincreasing and proves (2.20).

The second inequality in (2.21) follows by the already proven fact that $t \mapsto e^{2\lambda t}|\dot{x}_{t+}|_c^2$ is nonincreasing on $\mathbb{R}_+$. The first one follows from the same fact, energy identity (2.17), and has to be proven only for $\lambda > 0$:

$$\begin{aligned}\phi(x_t)-\phi(\bar{x}) &= -\int_t^{+\infty} \frac{\mathrm{d}}{\mathrm{d}r}\phi(x_r)\,\mathrm{d}r = \int_t^{+\infty} e^{-2\lambda r}e^{2\lambda r}|\dot{x}_{r+}|_c^2\,\mathrm{d}r \\ &\leq e^{2\lambda t}|\dot{x}_{t+}|_c^2 \int_t^{+\infty} e^{-2\lambda r}\,\mathrm{d}r = \frac{|\dot{x}_{t+}|_c^2}{2\lambda}.\end{aligned}$$

Using (2.18) with $x = \bar{x}$ and $s = t_0$ we obtain that, for every $0 < t_0 \leq t$, it holds

$$e^{\lambda(t-t_0)}c(\bar{x},x_t) + E_\lambda(t-t_0)\,(\phi(x_t)-\phi(\bar{x})) + \frac{(E_\lambda(t-t_0))^2}{2}|\dot{x}_{t+}|_c^2 \leq c(\bar{x},x_{t_0}). \tag{2.30}$$

As the first two terms on the l.h.s. are nonnegative, the third inequality in (2.21) eventually ensues.

Finally, let us assume that $\lambda = 0$ and ϕ has σ-compact sublevel sets, so that it has at least one minimizer $\bar{x}$. Since $\{x_t\}_{t>0}$ is contained in a sublevel set of ϕ, there exists a subnet $(x_{t_\eta})_\eta$ converging to a point $\bar{x}'$. By (2.20), using the lower semicontinuity of ϕ, we deduce that

$$\phi(\bar{x}') \leq \liminf_\eta \phi(x_{t_\eta}) \leq \liminf_\eta \left[\phi(\bar{x}) + \frac{c(\bar{x},x_{t_0})}{t_\eta - t_0}\right] = \phi(\bar{x}),$$

where $t_0 > 0$ is any value such that $t_\eta > t_0$ eventually in η. Therefore also $\bar{x}'$ is a minimum point of ϕ. Since $t \mapsto c(\bar{x}',x_t)$ is nonincreasing, there exists the limit $\ell := \lim_{t\to+\infty} c(\bar{x}',x_t) \in [0,+\infty)$. Since c is jointly σ-continuous we deduce that $\lim_\eta c(x_{t_\eta},\bar{x}') = 0$. Hence, we necessarily have that $\lim_{t\to+\infty} c(x_t,\bar{x}') = 0$. By compatibility, this implies that x_t σ-converges to $\bar{x}'$ as $t \to +\infty$. □

2.3. **On slopes and local characterizations.**

In this section, we temporarily depart from the study of EVI solutions and comment on some alternative formulations of local nature. Indeed, in the metric case $c = d^2/2$ one could consider a hierarchy of weaker notions, which go under the name of *Energy-Dissipation Equality/Inequality* (EDE and EDI, for short), as well as of *subdifferential formulation*, [AG13]. In the metric case, one has the chain of implications [AG13, Section 3.2]

$$EVI \Rightarrow EDE \Rightarrow EDI \Rightarrow \text{subdifferential formulation.}$$

Except for the EVI, which is global, the other formulations are of local nature, see Definition 2.12. Reversing these implications requires a global property, such as some notion of convexity. This will be considered in Proposition 2.19.

In addition, EDE and EDI formulations require the notion of *slope*, which still needs to be introduced in the present case of general costs, see Lemma 2.13 below.

2.3.1. *On general costs and slopes.*

Definition 2.12 (Local c-variational inequality)**.** *Let $\phi : X \to (-\infty,+\infty]$, and let c be a symmetric cost satisfying* (Diss). *We say that a function $x : (0,+\infty) \to \mathscr{D}(\phi)$ satisfies the* local c-variational inequality *if, for all $t > 0$ and all functions $z : [t,+\infty) \to \mathscr{D}(\phi)$ with $z_t = x_t$, we have*

$$0 \geq \limsup_{h\downarrow 0} \frac{c(z_{t+h},x_{t+h}) - c(z_{t+h},z_t) - c(x_t,x_{t+h})}{h^2} + \frac{\phi(z_t)-\phi(z_{t+h})}{h}. \tag{2.31}$$

The first term in (2.31) is remindful of a triangle-inequality term. Note however that in general we cannot separate the dependence in $x(\cdot)$ and $z(\cdot)$ in (2.31), in order to introduce a notion of slope. Nonetheless, some of the corresponding information is contained in (2.31), as the following lemma, specializing indeed to the metric case, shows.

Lemma 2.13 (Slope in the metric case). *Let $\lambda \geq 0$, $\phi : X \to (-\infty, +\infty]$, c be a symmetric cost satisfying (Diss), σ be compatible with the pair (c, ϕ), and assume, in addition, that c is jointly σ-continuous. If $x(\cdot)$ belongs to $\mathrm{EVI}_\lambda(X, c, \phi)$, then $x(\cdot)$ solves the local c-variational inequality in the sense of Definition 2.12.*

Moreover , if (X, d) is a geodesic metric space, $p \in (0, 2]$, and $c = \frac{d^p}{2}$, then, if $x(\cdot)$ solves the local c-variational inequality in the sense of Definition 2.12 we have that

$$|\partial\phi(x_t)|_{d,p} := \limsup_{x \xrightarrow{d} x_t} \left[\frac{(\phi(x_t) - \phi(x))_+}{d^{\frac{p}{2}}(x_t, x)} \right] \leq \limsup_{h \downarrow 0} \frac{d^{\frac{p}{2}}(x_{t+h}, x_t)}{h}, \tag{2.32}$$

with equality if $x(\cdot)$ belongs to $\mathrm{EVI}_\lambda(X, \frac{d^p}{2}, \phi)$.

Remark 2.14. Since for $p \in (0, 2]$, we have that $\delta := d^{\frac{p}{2}}$ is a distance. Lemma 2.13 shows that, for $c = \frac{\delta^2}{2}$, trajectories in $\mathrm{EVI}_\lambda(X, \frac{d^p}{2}, \phi)$ satisfy the energy identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(z_{t+}) = -|\dot{x}_{t+}|^2_{\delta^2/2} = -\frac{1}{2}|\partial\phi(x_t)|^2_{\delta,2}.$$

In this case, trajectories in $\mathrm{EVI}_\lambda(X, \frac{d^p}{2}, \phi)$ may be seen as *curves of maximal slope* for ϕ with respect to the *local slope* $|\partial\phi(x_t)|_{\delta,2}$ and the c-cost derivative $|\dot{x}_{t+}|^2_{\delta^2/2}$. This leads to the classical EDE or EDI formulations when integrated on time. The restriction to $p \in (0, 2]$ in (2.32) comes from the subadditivity of power functions, and the presence of $\frac{p}{2}$ from a simple estimate, see Lemma 2.15. Note that the slope $|\partial\phi(x_t)|_{d,p}$ is obtained constructively from (2.31). In particular, no chain-rule nor upper-gradient notions [AGS08] have been used.

Proof of Lemma 2.13. Let $t > 0$ and let $[t, +\infty) \ni s \mapsto z_s \in \mathscr{D}(\phi)$ be such that $z_t = x_t$. Given $h > 0$, consider (2.18) and apply it on $[t, t+h]$ by choosing $x = z_{t+h}$. We get

$$\begin{aligned} e^{\lambda h} c(z_{t+h}, x_{t+h}) - c(z_{t+h}, x_t) + \frac{(E_\lambda(h))^2}{2}|\dot{x}_{(t+h)+}|^2_c &\leq E_\lambda(h)\left(\phi(z_{t+h}) - \phi(x_{t+h})\right) \\ &= \frac{E_\lambda(h)}{h} h\left(\phi(z_{t+h}) - \phi(z_t) + \phi(x_t) - \phi(x_{t+h})\right). \end{aligned} \tag{2.33}$$

Divide by h^2 and take the lim sup as $h \downarrow 0$. Using that $\lim_{h\downarrow 0} \frac{E_\lambda(h)}{h} = 1$, $c(z_{t+h}, x_t) = c(z_{t+h}, z_t)$ since $x_t = z_t$ and the energy identity (2.17), we obtain

$$\begin{aligned} &\limsup_{h\downarrow 0} \left[\frac{c(z_{t+h}, x_{t+h}) - c(z_{t+h}, z_t)}{h^2} + \frac{\phi(z_t) - \phi(z_{t+h})}{h} \right] + \liminf_{h\downarrow 0} \frac{1}{2}|\dot{x}_{(t+h)+}|^2_c \\ &\leq -\frac{\mathrm{d}}{\mathrm{d}t}\phi(x_{t+}) \overset{(2.17)}{=} |\dot{x}_{t+}|^2_c. \end{aligned}$$

The latter entails (2.31) by moving all the terms to the l.h.s. and by using that

$$\begin{aligned} \liminf_{h\downarrow 0} \frac{1}{2}|\dot{x}_{(t+h)+}|^2_c &= \liminf_{h\downarrow 0} \lim_{r\downarrow 0} \frac{c(x_{t+h}, x_{t+h+r})}{r^2} \\ &\geq \limsup_{r\downarrow 0} \liminf_{h\downarrow 0} \frac{c(x_{t+h}, x_{t+h+r})}{r^2} \geq \limsup_{r\downarrow 0} \frac{c(x_t, x_{t+r})}{r^2} = \frac{1}{2}|\dot{x}_{t+}|^2_c, \end{aligned}$$

where we used that c is jointly σ-lower semicontinuous and that $x(\cdot)$ is σ-continuous.

We consider now the case where (X, d) is a geodesic metric space, $p \in [1, 2]$, and $c = d^p/2$. Fix $t > 0$ and $x \in \mathscr{D}(\phi)$ with $x \neq x_t$. Take a curve $z(\cdot)$ such that $d^p(z_t, z_{t+h}) = h^2 d^p(z_t, x)$.

We use the triangle inequality to get that the l.h.s. of (2.31) has a lower bound of the form

$$\begin{aligned} &|d(z_{t+h}, x_t) - d(x_{t+h}, x_t)|^p - d^p(z_{t+h}, z_t) - d^p(x_{t+h}, x_t) \\ &\leq d^p(z_{t+h}, x_{t+h}) - d^p(z_{t+h}, z_t) - d^p(x_{t+h}, x_t). \end{aligned}$$

We are going to separate the dependence in z and x through the following fact.

Lemma 2.15 (Bound via powers)**.** *Let $p > 0$ and $F : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ be given by $F(a,b) = |a-b|^p - a^p - b^p$. Then*

$$\begin{aligned} &F(a,b) \geq -2a^{p/2}b^{p/2} && \textit{if } p \in (0,2], \\ &F(a,b) \leq -2a^{p/2}b^{p/2} && \textit{if } p \in [2,+\infty). \end{aligned}$$

Proof. Let $p \in (0,2]$ and set $t = p/2 \in (0,1]$. The assertion follows by recalling that the map $r \in \mathbb{R}_+ \mapsto r^t$ is subadditive, as this implies that $|a-b|^{2t} \geq |a^t - b^t|^2$, which immediately gives $F(a,b) \geq -2a^{p/2}b^{p/2}$.

If $p \in (2,+\infty]$ we argue similarily, by recalling that in this case $t = p/2 > 1$ and the map $r \in \mathbb{R}_+ \mapsto r^t$ is superadditive. □

As a consequence of the Lemma we have

$$\begin{aligned} d^p(z_{t+h}, x_{t+h}) - d^p(z_{t+h}, z_t) - d^p(x_{t+h}, x_t) &\geq F(d(z_{t+h}, x_t), d(x_{t+h}, x_t)) \\ &\geq -2d^{p/2}(z_{t+h}, x_t)d^{p/2}(x_{t+h}, x_t). \end{aligned}$$

Inserting this bound in (2.31) we obtain

$$0 \geq \limsup_{h \downarrow 0} -\frac{hd^{p/2}(z_{t+h}, z_t)d^{p/2}(x_{t+h}, x_t)}{h^2} + \frac{\phi(z_t) - \phi(z_{t+h})}{h}. \tag{2.34}$$

Divide by $d^{p/2}(z_{t+h}, z_t)$ to get

$$\limsup_{h \downarrow 0} \frac{d^{p/2}(x_{t+h}, x_t)}{h} \geq \limsup_{h \downarrow 0} \frac{\phi(z_t) - \phi(z_{t+h})}{d^{p/2}(z_t, z_{t+h})}. \tag{2.35}$$

Since the l.h.s. is nonnegative and that the inequality holds for arbitrary $x = z_{t+h}$, we obtain

$$\limsup_{h \downarrow 0} \frac{d^{\frac{p}{2}}(x_{t+h}, x_t)}{h} \geq \limsup_{x \to x_t} \frac{(\phi(x_t) - \phi(x))_+}{d^{\frac{p}{2}}(x_t, x)} \tag{2.36}$$

which is relation (2.32). □

2.3.2. *A local characterization of* $\mathrm{EVI}_\lambda(X, c, \phi)$. We now turn to a local characterization of EVI solutions.

Proposition 2.16 (Local characterization)**.** *Let $\lambda \geq 0$, $\phi : X \to (-\infty, +\infty]$, c be a cost satisfying* (Diss)*, σ be compatible with the pair (c, ϕ), and let $x : \mathbb{R}_+ \to X$ belong to $\mathrm{EVI}_\lambda(X, c, \phi)$. For any $t_0 > 0$ and any curve $\gamma : [0, +\infty) \to X$ with $\gamma_0 = x_{t_0}$, we have that*

$$[\dot{x}, \gamma]_{c,t_0} := \liminf_{s \downarrow 0} \frac{1}{s} \frac{\mathrm{d}^+}{\mathrm{d}t} c(\gamma_s, x_t)_{|t=t_0} \leq \phi'(x_{t_0}; \gamma) := \liminf_{s \downarrow 0} \frac{\phi(\gamma_s) - \phi(x_{t_0})}{s} \tag{2.37}$$

where the inequality is intended in $[-\infty, +\infty]$.

Proof. Notice first that we can restrict ourselves to the case $\lambda = 0$ since c is nonnegative and to curves such that $\gamma_s \in \mathscr{D}(\phi)$ for all $s \in (0,1]$. Hence setting $x = \gamma_s$ in (2.3) and dividing by s, we get

$$\frac{1}{s} \frac{\mathrm{d}^+}{\mathrm{d}t} c(\gamma_s, x_t)_{|t=t_0} \leq \frac{\phi(\gamma_s) - \phi(x_{t_0})}{s}$$

taking the $\liminf$ on both sides, we obtain (2.37). □

Remark 2.17. Equation (2.37) corresponds to the generalization to the case of a general cost c of the classical subdifferential formulation. Indeed, consider first for simplicity that X is a Hilbert space with norm $\|\cdot\|$, let $c = \|\cdot\|^2/2$, and assume that ϕ is differentiable. Then, taking $v \in X$ and $\gamma_s = x_{t_0} + sv$, we obtain $[\dot{x}, \gamma]_{c,t_0} = \langle -\dot{x}_{t_0}, v \rangle \leq \langle \nabla\phi(x_{t_0}), v \rangle$. Since this holds for all $v \in X$, we get that $\dot{x}_{t_0} = -\nabla\phi(x_{t_0})$. In fact, one could obtain the same conclusion by arguing directly at the level of (2.31), taking $z_{t+h} = x_t + hv$.

More generally, assume that (X,d) is a locally Hilbertian manifold, that $c \in C^2$ with $c=0$ on the diagonal, and that $\phi \in C^1$. Setting $\dot\gamma_0 = v$ for arbitrary $v \in T_{x_{t_0}}M$ we get

$$\begin{aligned}\liminf_{s\downarrow 0} \frac{\phi(\gamma_s)-\phi(x_{t_0})}{s} &\geq [\dot x,\gamma]_{c,t_0} = \liminf_{s\downarrow 0} \frac{1}{s}\langle \nabla_2 c(\gamma_s,x_{t_0}),\dot x_{t_0}\rangle \\ &= \left\langle \lim_{s\downarrow 0} \frac{\nabla_2 c(\gamma_s,x_{t_0})-\nabla_2 c(\gamma_0,x_{t_0})}{s}, \dot x_{t_0}\right\rangle = \langle \nabla_{1,2} c(x_{t_0},x_{t_0})v,\dot x_{t_0}\rangle \qquad (2.38)\end{aligned}$$

which is precisely the definition of the elements of the Fréchet subdifferential $\partial\phi$, hence

$$\nabla_{2,1} c(x_{t_0},x_{t_0})\dot x_{t_0} \in \partial\phi(x_{t_0}).$$

We thus recover the seminal observation of Kim and McCann [KM10, Section 2] that the "geometric information of c" is contained in its mixed-Hessian $\nabla_{2,1}c$. Moreover two costs with the same mixed Hessian will induce the same gradient flows, as shown for $\mathcal{W}_c$ and $\mathcal{W}_2$ by [RW24].

We will now consider the converse to Proposition 2.16, that is whether the local equation (2.37) implies that $x(\cdot) \in \mathrm{EVI}_\lambda(X,c,\phi)$. To this aim, some global property will be needed. Concerning ϕ, this will be some suitable form of λ-convexity. Nonetheless, as the cost c is integrated along trajectories, some form of convexity for c can be useful, as well. The latter is naturally connected to curvature conditions of c. Note, however, that the (Alexandrov) curvature does not affect the validity of (2.39) in geodesic metric spaces, see [MS20, Lemma 3.13].

More precisely, our goal is to prove that, under some suitable assumption on c, for some class of curves $\gamma(\cdot)$ such that $\gamma_0 = x_{t_0}$ and $\gamma_1 = x$ is left free, we have

$$[\dot x,\gamma]_{c,t_0} := \liminf_{s\downarrow 0} \frac{1}{s}\frac{\mathrm{d}^+}{\mathrm{d}t} c(\gamma_s,x_t)_{|t=t_0} \geq \frac{\mathrm{d}^+}{\mathrm{d}t} c(\gamma_1,x_t)_{|t=t_0}. \qquad (2.39)$$

This would indeed correspond to [MS20, eq. (3.56), Lemma 3.13], which holds for a complete geodesic space (X,d), $c = d^2/2$, and with γ being any geodesic. We hence anticipate that one is asked to restrict the class of admissible curves $\gamma(\cdot)$ in the case of general costs c, as well. In preparation for Proposition 2.19, we recall the following Definition from [LTV25, Definitions 2.6, 2.7].

Definition 2.18 (NNCC space and variational c-segment)**.** *Let $c\colon X\times X \to [-\infty,+\infty]$ be given. We say that $(X\times X,c)$ is a* Nonnegatively Cross-Curved (NNCC) space *if for every $(x_0,x_1,\bar y) \in X\times X\times X$ such that $c(x_0,\bar y)$ and $c(x_1,\bar y)$ are finite, there exists a function $\gamma:[0,1]\to X$ with $\gamma_0 = x$ and $\gamma_1 = x_1$ such that*

$$\begin{aligned}c(\gamma_s,\bar y) - c(\gamma_s,y) \leq (1-s)&[c(\gamma(0),\bar y) - c(\gamma(0),y)] \\ &+ s\,[c(\gamma(1),\bar y) - c(\gamma(1),y)] \quad \forall y\in X, s\in(0,1), \qquad (2.40)\end{aligned}$$

with the rule $(+\infty)+(-\infty) = +\infty$ for the r.h.s. In this case, we say that the function $\gamma(\cdot)$ is a variational c-segment*.*

Relation (2.40) is often referred to as the *NNCC inequality*. Being required for all $y\in X$, it is a quite strong property. In particular, variational c-segments may not exist. Correspondingly, situations where variational c-segments exist are of special interest. In the metric setting $c = d^2/2$, NNCC spaces are a specific subclass of complete PC metric spaces ([LTV25, Proposition 2.26]). In particular, the Wasserstein space is NNCC [LTV25, Theorem 3.11].

Proposition 2.19 ((2.37) implies EVI for NNCC spaces)**.** *Let $\lambda \geq 0$, $\phi : X \to (-\infty,+\infty]$, c be a cost satisfying* (Diss)*, and σ be compatible with the pair (c,ϕ). Let $x:\mathbb{R}_+ \to \mathscr{D}(\phi)$ be a σ-continuous function. Assume that*

i) *$(X\times X,c)$ is a NNCC space;*
ii) *for all $x_0, x\in X$ there exists a variational c-segment $\gamma:[0,1]\to X$ with $\gamma_0 = x_0$ and $\gamma_1 = x$, and $M:[0,1]\to\mathbb{R}\cup\{+\infty\}$ with*

$$\phi(\gamma_t) \leq (1-t)\phi(x_0) + t\phi(x) - \lambda t c(x,x_0) + M(t) \qquad (2.41)$$

such that $\liminf_{t\downarrow 0}\frac{M(t)}{t} = 0$;

iii) (2.37) *holds along these variational c-segments.*

Then, $x(\cdot) \in \mathrm{EVI}_\lambda(X, c, \phi)$.

Note that relation (2.41) holds if $t \mapsto \phi(\gamma_t) - \lambda t^2 c(x, x_0)$ is convex. Moreover, it follows also if $t \mapsto \phi(\gamma_t) - \lambda c(\gamma_t, x_0)$ is convex and $\liminf_{t\downarrow 0} \frac{\lambda c(\gamma_t, x_0)}{t} = 0$. These two cases coincide when (X, d) is a geodesic metric space, $c = d^2/2$ and γ is a geodesic. The former is closer to the condition on generalized geodesics of [AGS08], while the latter is more inline with the theory of NNCC spaces and (2.40) (see also Remark 3.11).

Proof of Proposition 2.19. Let $t_0 > 0$, $h > 0$, $x \in X$ and take $x_0 = \bar{y} = x_{t_0}$, $x_1 = x$, $y = x_{t_0+h}$ in (2.40), noticing that, since $\gamma(\cdot)$ does not depend on y,

$$c(\gamma_s, x_{t_0}) - c(\gamma_s, x_{t_0+h}) \leq (1-s)[c(x_{t_0}, x_{t_0}) - c(x_{t_0}, x_{t_0+h})] + s\,[c(x, x_{t_0}) - c(x, x_{t_0+h})].$$

As $c(x_{t_0}, x_{t_0}) = 0$ and $c(x_{t_0}, x_{t_0+h}) \geq 0$, we can drop the (1-s)-term of the r.h.s., and obtain

$$c(\gamma_s, x_{t_0+h}) - c(\gamma_s, x_{t_0}) \geq s\,[c(x, x_{t_0+h}) - c(x, x_{t_0})].$$

Divide by h and take the $\limsup$ as $h \downarrow 0$ to obtain

$$\frac{\mathrm{d}^+ c(\gamma_s, x_t)}{\mathrm{d}t}_{|t=t_0} \geq s \frac{\mathrm{d}^+ c(x, x_t)}{\mathrm{d}t}_{|t=t_0}. \tag{2.42}$$

Dividing by s and taking the $\liminf$ for $s \downarrow 0$, we obtain (2.39).

Consider now (2.41), divide by t, and take the $\liminf$ as $t \downarrow 0$ to obtain

$$\frac{\mathrm{d}^+}{\mathrm{d}t} c(x, x_t)_{|t=t_0} \overset{(2.39)}{\leq} [\dot{x}, \gamma]_{c,t_0} \overset{(2.37)}{\leq} \phi'(x_{t_0}; \gamma) \overset{(2.41)}{\leq} \phi(x) - \phi(x_{t_0}) - \lambda c(x, x_{t_0}) \tag{2.43}$$

so $x(\cdot)$ satisfies (2.3), hence $x(\cdot) \in \mathrm{EVI}_\lambda(X, c, \phi)$. □

Corollary 2.20. *Let* $0 \leq \lambda_1 \leq \lambda_2$, $\phi : X \to (-\infty, +\infty]$, c *be a cost satisfying* (Diss), *and* σ *be compatible with the pair* (c, ϕ). *Then,* $EVI_{\lambda_2}(X, c, \phi) \subset EVI_{\lambda_1}(X, c, \phi)$. *Conversely, for* (X, c) *being a NNCC space, if* $x(\cdot) \in EVI_{\lambda_1}(X, c, \phi)$ *and* ϕ *is* λ_2*-convex along variational* c*-segments as in* (2.41)*, then* $x(\cdot) \in EVI_{\lambda_2}(X, c, \phi)$.

Proof. The first part follows immediately from (2.3). The converse stems from (2.43) applied to $\lambda = \lambda_2$, since (2.37) is independent of λ in the case $\lambda \geq 0$. □

3. Existence of EVI solutions as limit of splitting schemes

In this section, we are interested in describing a natural minimizing movements scheme that converges to the EVI. Unlike previous works, we will not focus solely on the implicit (Euler) iteration, but also on a more general splitting scheme. The key idea, following [LAF23], is to perform a *majorization-minimization* of the functional $\phi := f + g$ where $f : X \to \mathbb{R}$ and $g : X \to (-\infty, +\infty]$, noticing that for any set X, $\tau > 0$, and any cost function $c : X \times X \to \mathbb{R}$ we have

$$\phi(x) = f(x) + g(x) \leq g(x) + \frac{c(x, y)}{\tau} + f^{c/\tau}(y) =: \Phi(x, y) \tag{3.1}$$

using the c-transform $f^{c/\tau}(y) := \sup_{x' \in X} f(x') - \frac{c(x', y)}{\tau}$. Given $x_0 \in X$, we then perform an alternating minimization of Φ, assuming that the iterates exist,

$$y^\tau_{n+1} \in \operatorname*{argmin}_{y \in X} g(x_n) + \frac{c(x_n, y)}{\tau} + f^{c/\tau}(y), \tag{3.2}$$

$$x^\tau_{n+1} \in \operatorname*{argmin}_{x \in X} g(x) + \frac{c(x, y_{n+1})}{\tau} + f^{c/\tau}(y_{n+1}). \tag{3.3}$$

The alternating minimization of Φ does not provide a minimizer of ϕ in general. This is however the case if ϕ is c/τ-concave (Definition 3.1), which entails that $\phi(x) = \inf_{y \in X} \Phi(x, y)$.

In the relevant subcase $f = 0$ we only have implicit iterations. If $c(\cdot, y)$ is lower bounded for each $y \in X$, then, introducing $\tilde{c}(x, y) := c(x, y) - \inf_{x' \in X} c(x', y)$, we have that $\tilde{c} \geq 0$ and

$\Phi(x,y) = g(x) + \frac{\tilde{c}(x,y)}{\tau}$ in (3.1), which clearly defines a minimizing movement scheme. To ensure that no information about g is lost, we require that $0 = \inf_{y\in X} \tilde{c}(x,y)$.

Léger and the first-named author in [LAF23] extensively studied the discrete iterations with $\tau > 0$ fixed and proved convergence of $(\phi(x_n))_{n\in\mathbb{N}}$ to the infimum of Φ under the so-called *five-point* property. We are instead interested in taking the limit $\tau \to 0$ and recovering a trajectory in $\mathrm{EVI}_\lambda(X, c, \phi)$. In Section 3.1, we prove that this is possible if f and g satisfy that f is c/τ-concave for all small τ and some cross-convexity property from [LAF23] holds. The latter reduces to a form of discrete EVI. Similarly to the discrete EVI in the metric case [AG13], this cross-convexity is implied by a more readable compatibility condition between the functional and the cost, described by some convexity of two functionals along the same curve, see Section 3.2.

Since we consider separate properties on f and g due to the splitting structure, we consider also two auxiliary variables $\xi, z \in X$ corresponding to (3.2) with $f = 0$ and (3.3) with $g = 0$. Their role is made clear in Section 3.1.

3.1. **Definition of the scheme through alternating minimization.** We start with the definitions of the fundamental objects we are going to use. Throughout the section, X will be a nonempty set and $c : X \times X \to \mathbb{R}$ will be a given cost. Note that in the first part of this section we do not need to assume the cost to be nonnegative, as we did in the previous Section 2.

Definition 3.1 (c-transform, c-concavity)**.** *The* c-transform *of a function* $U: X \to \mathbb{R}$ *is the function* $U^c: X \to (-\infty, +\infty]$ *defined by*

$$U^c(y) := \sup_{x\in X} U(x) - c(x,y), \quad y \in X. \tag{3.4}$$

We say that a function $U: X \to \mathbb{R}$ *is c-concave if there exists a function* $h: X \to (-\infty, +\infty]$ *such that*

$$U(x) = \inf_{y\in X} c(x,y) + h(y) \quad \textit{for every } x \in X. \tag{3.5}$$

Note that, if U *is c-concave, then* U^c *is the smallest* h *such that* (3.5) *holds.*

Since c has finite values, c-concave functions cannot take the value $+\infty$. The next lemma shows that a c-concave function is also c/τ-concave for $\tau \in (0,1]$, provided the cost satisfies a suitable condition.

Lemma 3.2. *Let* $c' : X \times X \to \mathbb{R}$ *be given. Every c-concave function is* c'*-concave if and only if* $x \mapsto c(x,y)$ *is* c'*-concave for every* $y \in X$*. In particular, given* $\tau \in (0,1]$*, any c-concave function is* c/τ*-concave if and only if, for all* $y \in X$*, there exists* $h_y : X \to (-\infty, +\infty]$ *such that* $\tau c(\cdot, y) = \inf_{y'\in X} c(\cdot, y') + h_y(y')$*.*

If (X, d) *is an intrinsic metric space, namely, there exist* ϵ*-midpoints in the following sense*

$$\forall \epsilon > 0,\, \forall\, x \in X,\, x' \in X,\, \exists y \in X,\, \max(d(x,y), d(x',y)) \le \frac{1}{2} d(x,x') + \epsilon, \tag{3.6}$$

then $\frac{d^2(\cdot,x)}{2^n}$ *is* d^2*-concave for all* $x \in X$ *and* $n \in \mathbb{N}$*.*

Proof. Suppose every c-concave function is c'-concave. For every $y \in X$, the function $x \mapsto c(x,y)$ is c-concave (it is enough to take h as the indicator function of $\{y\}$ in the definition of c-concavity) whence it is c'-concave.

Suppose now that $x \mapsto c(x,y)$ is c'-concave for every $y \in X$. This is equivalent to say that, for every $y \in X$, there exists a function $h_y : X \to (-\infty, +\infty]$ such that $c(x,y) = \inf_{y'\in X} c'(x,y') + h_y(y')$. If $U : X \to \mathbb{R}$ is c-concave, there exists $h : X \to \mathbb{R} \cup \{+\infty\}$ such that

$$U(x) = \inf_{y\in X} c(x,y) + h(y) \quad \text{for every } x \in X.$$

We deduce

$$U(x) = \inf_{y\in X} \left[\inf_{y'\in X} c'(x,y') + h_y(y')\right] + h(y) = \inf_{y'\in X} \left\{c'(x,y') + \inf_{y\in X} \left[h_y(y') + h(y)\right]\right\}, \quad x \in X.$$

Notice that the quantity $h'(y') := \inf_{y\in X} [h_y(y') + h(y)]$ cannot attain the value $-\infty$ since otherwise we would obtain $U(x) = -\infty$, in contrast with $U(x) \in \mathbb{R}$.

Let (X,d) be an intrinsic metric space. By the triangle inequality, we have for any $x, x', y \in X$

$$\frac{1}{2}d^2(x,x') \leq d^2(x,y) + d^2(x',y).$$

Taking the infimum over y we deduce that $\frac{1}{2}d^2(x,x') \leq \inf_{y\in X} d^2(x,y) + d^2(x',y)$. Take y to be an ϵ-midpoint. We get $d^2(x,y_\epsilon) + d^2(x',y_\epsilon) \leq 2(\frac{1}{2}d(x,x')+\epsilon)^2$. Hence taking the limit as $\epsilon \downarrow 0$, we obtain $\frac{1}{2}d^2(x,x') = \inf_{y\in X} d^2(x,y) + d^2(x',y)$.

Replacing d by $\frac{d}{\sqrt{2}}$, the same proof shows that $\frac{1}{4}d^2(x,\cdot)$ is $\frac{d^2}{2}$ concave, and an immediate induction gives the result for all 2^n with $n \in \mathbb{N}$. $\square$

While c-concavity is a classical notion, here used for explicit schemes, we now introduce the key notion of [LAF23]: cross-convexity. It is based on iterates and should be seen as requiring the analogue of the discrete EVI in gradient flows on metric spaces [AGS08, Corollary 4.1.3].

Definition 3.3. *Let $U : X \to \mathbb{R}$ be a function and let $x_0 \in X$ and $y_0 \in X$. We define the (possibly empty) sets*

$$\begin{aligned}
\mathcal{P}_c(U,y_0) &:= \operatorname*{argmin}_{x\in X}\{c(x,y_0) - U(x)\},\\
\mathcal{Q}_c(U,x_0) &:= \operatorname*{argmin}_{y\in X}\{c(x_0,y) + U^c(y)\},\\
\mathcal{R}_c(x_0) &:= \operatorname*{argmin}_{y\in X} c(x_0,y),\\
\mathcal{S}_c(x_0) &:= \{\xi_0 \in X \mid x_0 \in \operatorname*{argmin}_{x\in X} c(x,\xi_0)\}.
\end{aligned}$$

Definition 3.4 (c-cross-concavity/convexity)**.** *Let $\mu \geq 0$ be given. A function $U : X \to \mathbb{R}$ is said to be*

- μ-strongly c-cross-concave *if*
$$U(x) - U(x_1) \leq [c(x,y_0) - c(x_1,y_0)] - [c(x,y_1) - c(x_1,y_1)] - \mu\,[c(x,y_1) - c(x_1,y_1)]$$
for every $(x,y_0) \in X\times X$ and every $x_1 \in \mathcal{P}_c(U,y_0)$, $y_1 \in \mathcal{R}_c(x_1)$;
- *μ-strongly c-cross-convex if*
$$U(x) - U(x_0) \geq -[c(x,\xi_0) - c(x_0,\xi_0)] + [c(x,y_0) - c(x_0,y_0)] + \mu\,[c(x,\xi_0) - c(x_0,\xi_0)]$$
for every $(x,x_0) \in X\times X$ and every $\xi_0 \in \mathcal{S}_c(x_0)$, $y_0 \in \mathcal{Q}_c(U,x_0)$.

Because of the definition of the iterates and of the focus on y_0 and $\mathcal{P}_c$ for cross-concavity and on x_0 and $\mathcal{Q}_c$ for cross-convexity, the two notions are not symmetric: having $-U$ c-cross-concave does not imply that U is c-cross-convex. In Section 3.2, we provide some simpler curve-based sufficient assumptions for these two notions to hold.

Example 1.1 (continued). For $\tau > 0$ and $c = \frac{d^2}{2\tau}$, c-cross-concavity is precisely the discrete EVI considered in [AGS08, Corollary 4.1.3]. Cross-convexity is similar concept, but applies to an explicit scheme. The $\frac{d^2}{2\tau}$-concavity can be understood as a local $\frac{d^2}{2\tau}$-bound on the growth.

Example 1.2 (continued). Let X be a subset of a Banach space, X^* the dual of the closure of its span, and assume there exists a strictly convex $u : X \to \mathbb{R}$ such that for all $y \in X$ there exists $u'(y) \in X^*$ such that for all x

$$c(x,y) = u(x) - u(y) - \langle u'(y), x-y\rangle. \tag{3.7}$$

In particular if $u'(y)$ is the Gateaux derivative of u at y, then (3.7) states that c is the Bregman divergence of u. The KL is not Gateaux differentiable for nondiscrete $\mathbb{X} \subset \mathbb{R}^d$, but (3.7) still holds, see [AFKL22, Example 2] with $u(\mu) = \mathrm{KL}(\mu|\rho)$ and $u'(\mu) = 1 + \ln(\frac{\mathrm{d}\mu}{\mathrm{d}\rho})$.

Cross-concavity is the three-point inequality in mirror descent, see e.g [CT93, Lemma 3.2]. It reads similarly to cross-convexity which takes the form: for all $x_0, x \in X$, since $\mathcal{S}_c(x_0) = \{x_0\}$ by strict convexity of u,

$$[U(x) - \mu u(x)] - [U(x_0) - \mu u(x_0)] \geq \langle u'(y_0) - u'(x) - \mu u'(x_0), x - x_0\rangle \tag{3.8}$$

with $y_0 \in \mathcal{Q}_c(U, x_0)$. On the other hand, c-concavity reads $U(x) = \inf_{y\in X} c(x, y) + h(y)$ which implies that $U - u$ is concave upper semicontinuous as an infimum of affine functions.

We work under the following fundamental hypotheses, which will be progressively reinforced by additional assumptions when needed.

Assumption 3.5. *Let $c : X \times X \to \mathbb{R}$, $f : X \to \mathbb{R}$, $g : X \to (-\infty, +\infty]$ be given and assume that there exist $\bar\tau > 0$ and $\lambda_f, \lambda_g \in \mathbb{R}$ such that:*

(1) f is $\lambda_f\tau$-strongly c/τ-cross-convex and c/τ-concave for every $0 < \tau < \bar\tau$.
(2) $-g$ is $\lambda_g\tau$-strongly c/τ-cross-concave for every $0 < \tau < \bar\tau$.
(3) for every $x_0 \in X$ there exists $y_0 \in \mathcal{Q}_{c/\tau}(f, x_0)$ and $x_1 \in \mathcal{P}_{c/\tau}(-g, y_0)$ such that $\mathcal{S}_c(x_1) \neq \emptyset$ and $\mathcal{R}_c(x_1) \neq \emptyset$.

Under the above Assumption 3.5, starting from $x_0 \in X$ we can implement the following scheme: for $i \in \mathbb{N}$ and $0 < \tau < \bar\tau$ we define

$$x_0^\tau := x_0, \quad y_i^\tau \in \mathcal{Q}_{c/\tau}(f, x_i^\tau), \quad x_{i+1}^\tau \in \mathcal{P}_{c/\tau}(-g, y_i^\tau), \quad \xi_i^\tau \in \mathcal{S}_c(x_i^\tau), \quad z_i^\tau \in \mathcal{R}_c(x_i^\tau). \tag{3.9}$$

$$\begin{tikzcd} y_{i-1} \arrow[dr, "g"] & & y_i \\ & x_i \arrow[ur, "f"] \arrow[dr, "0"] & \\ \xi_i \arrow[ur, "0"] & & z_i \end{tikzcd}$$

TABLE 1. Diagram of the iterates of the splitting scheme (3.9), omitting the role of c and τ.

While $x_{i+1}^\tau \in \mathcal{P}_{c/\tau}(-g, y_i^\tau)$ clearly corresponds to an implicit iteration on g, we refer to [LAF23, Section 3] for the interpretation of $y_i^\tau \in \mathcal{Q}_{c/\tau}(f, x_i^\tau)$ as an explicit iteration on f. In particular, it is shown in [LAF23] that, if X is an Euclidean space, for c/τ-concave f, this iteration corresponds to gradient, mirror or Riemmanian descent for c taken respectively to be the squared Euclidean norm, a Bregman divergence, or a squared Riemannian distance.

Definition 3.6. *Under Assumption 3.5, for $0 < \tau < \bar\tau$, given points $(x_i^\tau)_i \subset X$ and $(y_i^\tau)_i, (z_i^\tau)_i \subset X$ as in (3.9), we define the following interpolating curves $\bar x^\tau, \bar y^\tau, \bar z^\tau : [0, +\infty) \to X$ as*

$$\bar x_t^\tau := x_{\lfloor t/\tau\rfloor}^\tau, \quad \bar y_t^\tau := y_{\lfloor t/\tau\rfloor}^\tau, \quad \bar z_t^\tau := z_{\lfloor t/\tau\rfloor}^\tau \quad t \in [0, +\infty).$$

Our aim is to prove the following theorem.

Theorem 3.7 (Existence of EVI solutions via splitting scheme)**.** *Let c be a cost satisfying* (Diss) *with Assumption 3.5 in place; suppose in addition that*

(1) *$\phi = f + g$ is lower bounded;*
(2) *there exists a Hausdorff topology σ on X such that σ is compatible with (c, ϕ) and c is σ-continuous;*
(3) *Either*
   (a) σ is Cauchy-compatible with (c, ϕ) and c is symmetric, or
   (b) the sublevel sets of ϕ are σ-sequentially compact.

Then, for every $x_0 \in X$ and $\tau \in (0, \bar\tau)$ the family of curves $(\bar x^{\tau/2^n})$ (constructed according to Definition 3.6 starting from x_0) converges (up to a subsequence in case (b)) pointwise (w.r.t. σ) as $\tau \downarrow 0$ to a σ-continuous curve $x : [0, +\infty) \to X$ which satisfies

$$c(x, x_t) - c(x, x_s) + (\lambda_f + \lambda_g)\int_s^t c(x, x_r)\,\mathrm{d}r \leq (t - s)\phi(x) - \int_s^t \phi(x_r)\,\mathrm{d}r \quad \forall\, 0 \leq s \leq t,\ x \in X. \tag{3.10}$$

If $\mathrm{EVI}_\lambda(X, c, \phi)$ *contains a unique trajectory* $x(\cdot)$ *starting from* x_0*, then,* $\bar{x}^\tau(\cdot)$ σ*-converges pointwise to this* $x(\cdot)$*, which does not depend on* τ*. In case (a), we also have the following error estimate*

$$c\,(\bar{x}^\tau_t, x_t) \leq 2\tau(\phi(x_0) - \inf \phi) \quad \forall t \geq 0,\ \tau \in (0, \bar{\tau}). \tag{3.11}$$

The proof of Theorem 3.7 is deferred to Section 3.4. In terms of assumptions, the key differences with Section 2 are that we assume ϕ to be lower bounded in order to have a lower bound on $(\phi(x^\tau_n))_n$ independent of τ and n. We also require c to be σ-continuous so as to take limits in $-c(x, \bar{z}^\tau_t)$, where lower semicontinuity of c would not be sufficient to conclude.

Before moving on, let us mention that the assumptions of Theorem 3.7 may be weakened in some specific situations. A specific setting where this seems to be possible is the case in which c dominates a (power of a) distance d^2. In this case, the a-priori bounds obtained below (see (3.25), for instance) would guarantee that the approximating trajectories are equicontinuous w.r.t to d^2. In combination with a compactness assumption on the subslevels of ϕ, this would allow to use an Ascoli–Arzelà argument to obtain convergence. In the linear-space setting of X being a Hilbert or a Banach space, one may even avoid asking for the compactness of the sublevels of ϕ and alternatively handle the limit passage by lower-semicontinuity arguments. We give below an application of Theorem 3.7 to Example 1.2.

Corollary 3.8. *Let* $\mathbb{X} \subset \mathbb{R}^d$ *be a closed subset,* $0 < a < b < +\infty$*,* $\rho \in \mathcal{P}(\mathbb{X})$*, and let* $X = \{\mu \in \mathcal{P}(\mathbb{X}) \,|\, \mu = h\rho$ *with* $h(\cdot) \in [a, b]$ ρ*-a.s.*$\}$*. Let* σ_ρ *be the strong* $L^1(\rho)$ *topology on* X*, and set* $c = \mathrm{KL}$*. Let* $f, g : X \to \mathbb{R}$ *be convex,* σ_ρ*-lower semicontinuous, and such that* g *has* σ_ρ *compact sublevel sets and is lower-bounded and* f *is lower bounded. Assume furthermore that there exists* $\bar{\tau} > 0$ *such that for every* $0 < \tau < \bar{\tau}$*, there exists* $h_\tau : X \to (-\infty, +\infty)$ *with*

$$\frac{\mathrm{KL}(\mu|\rho)}{\tau} - f(\mu) = \max_{\mu' \in X} \left\langle \mu, \frac{1}{\tau} \ln\left(\frac{\mathrm{d}\mu'}{\mathrm{d}\rho}\right)\right\rangle - h_\tau(\mu'). \tag{3.12}$$

Then, for every $x_0 \in X$ *and* $\tau \in (0, \bar{\tau})$ *the family of curves* $(\bar{x}^{\tau/2^n})$ *(constructed according to Definition 3.6 starting from* x_0*) converges up to a subsequence pointwise (w.r.t.* σ*) as* $\tau \downarrow 0$ *to a* σ*-continuous curve* $x : [0, +\infty) \to X$ *satisfying* (3.10)*.*

Corollary 3.8 proves that the limit curve belongs $\mathrm{EVI}_\lambda(X, c, \phi)$. Based on the discussion around (2.38), taking derivatives $\nabla_{2,1}$ in (3.7), this flow formally corresponds to the mirror flow

$$-[\nabla^2(\mathrm{KL}(\cdot|\rho))(x_{t_0})]\dot{x}_{t_0} \in \partial\phi(x_{t_0})$$

restricted to the set X.

Proof. We just have to check the assumptions of Theorem 3.7. Since c clearly satisfies (Diss) and (1), (2) and (3)b hold, it only remains to check Assumption 3.5. Since the sublevels of g are compact for σ_ρ and $g(\cdot) + \frac{\mathrm{KL}(\cdot|\mu_0)}{\tau}$ is σ_ρ-lower semicontinuous for all $\mu_0 \in X$, the set $\mathcal{P}_{c/\tau}(-g, y_0)$ is nonempty. We use Lemmas 3.12 and 3.13 below, along with the discussion in Example 3.2 below on $c = \mathrm{KL}$. Equation (3.12) is just a rewriting of the c/τ-concavity of (3.5). □

3.2. **Sufficient conditions: compatibility of energy and cost.** We now provide some more readable sufficient assumptions, inspired by [AG13, Section 3.2.4], also known as C2G2 (Compatible Convexity along Generalized Geodesics) in [San17].

Assumption 3.9 (Compatibility to obtain cross-concavity)**.** *The function* $g : X \to (-\infty, +\infty]$ *is such that there exists* $\lambda \in \mathbb{R}$ *such that, for any* $y_0 \in Y$ *and* $x_1 \in X$ *and* $y_1 \in \mathcal{R}_c(x_1)$*, and for all* $x \in X$*, there exists functions* $\gamma : [0, 1] \to X$ *and* $M : [0, 1] \to (-\infty, +\infty]$ *such that, for every* $t \in [0, 1]$

$$g(\gamma(t)) \leq (1 - t)g(x_1) + tg(x) - \lambda t[c(x, y_1) - c(x_1, y_1)] + M_t, \tag{3.13}$$

$$c(\gamma(t), y_0) \leq (1 - t)c(x_1, y_0) + tc(x, y_0) - t[c(x, y_1) - c(x_1, y_1)] + M_t \tag{3.14}$$

with $\liminf_{t\downarrow 0}\frac{M_t}{t}=0$. *Specifically,* $\gamma(\cdot)$ *fulfills*

$$\liminf_{t\downarrow 0}\frac{g(\gamma(t))-g(x_1)}{t}\le g(x)-g(x_1)-\lambda[c(x,y_1)-c(x_1,y_1)], \tag{3.15}$$

$$\liminf_{t\downarrow 0}\frac{c(\gamma(t),y_0)-c(x_1,y_0)}{t}\le c(x,y_0)-c(x_1,y_0)-[c(x,y_1)-c(x_1,y_1)]. \tag{3.16}$$

Assumption 3.10 (Compatibility to obtain cross-convexity)**.** *There exists* $\bar\tau>0$ *such that* f *is* c/τ*-concave for all* $\tau\in(0,\bar\tau)$. *There exists* $\lambda\in\mathbb{R}$ *such that, for any* $x_0\in X$, *and all* $\xi_0\in\mathcal{S}_c(x_0)$, $y_0\in\mathcal{Q}_c(f,x_0)$, *for all* $x\in X$ *and* $\tau\in(0,\bar\tau)$, *there exists functions* $\gamma:[0,1]\to X$, $z:[0,1]\to Y$ *and* $M:[0,1]\to(-\infty,+\infty]$ *such that* $\gamma(0)=x_0$, $z(0)=y_0$ *and, for every* $t\in[0,1]$, $f(\gamma(t))-f^{c/\tau}(z(t))=\frac{1}{\tau}c(\gamma(t),z(t))$ *and*

$$f(\gamma(t))\le(1-t)f(x_0)+tf(x)-\lambda t[c(x,\xi_0)-c(x_0,\xi_0)]+M_t, \tag{3.17}$$

$$c(\gamma(t),z(t))\ge(1-t)c(x_0,z(t))+tc(x,z(t))-t[c(x,\xi_0)-c(x_0,\xi_0)]-M_t, \tag{3.18}$$

with $\liminf_{t\to 0}[c(x,z(t))-c(x_0,z(t))]\le c(x,z(0))-c(x_0,z(0))$ *and* $\liminf_{t\downarrow 0}\frac{M_t}{t}=0$.

Remark 3.11. In the analysis, it is paramount to choose which functions should be convex along which curves. Nonetheless, we mostly use assumptions for $t\to 0$ only. By introducing the function $t\mapsto M_t$ we intend to cover both convexity of $t\mapsto\phi(\gamma_t)-\lambda t^2c(x,x_0)$ and of $t\mapsto\phi(\gamma_t)-\lambda c(\gamma_t,x_0)$ as already discussed around (2.41) after Proposition 2.19. For instance, for (3.13)–(3.14) and $p>1$, taking $M_t=\max(1,\lambda)t^p[c(x,y_1)-c(x_1,y_1)]$, we recover the (p,λ)-convexity of Ohta and Zhao in the case $c=d^p$ [OZ23, Definition 4.1].

Lemma 3.12. *Let* $\tau>0$ *and assume that for all* $y_0\in Y$ *the set* $\mathcal{P}_{c/\tau}(-g,y_0)$ *is nonempty. Then, Assumption 3.9 implies the* $\lambda\tau$*-strong* c/τ*-cross-concavity of* $-g$.

Proof. Fix $x_1\in\mathcal{P}_{c/\tau}(-g,y_0)$, multiply (3.14) by $1/\tau$ and sum with (3.13) to obtain

$$\begin{aligned}g(x_1)+\frac{1}{\tau}c(x_1,y_0)&\le g(\gamma(t))+\frac{1}{\tau}c(\gamma(t),y_0)\\&\le(1-t)g(x_1)+tg(x)-\lambda tc(x,y_1)\\&\quad+\frac{1-t}{\tau}c(x_1,y_0)+\frac{t}{\tau}c(x,y_0)-\frac{t}{\tau}[c(x,y_1)-c(x_1,y_1)]+(1+\frac{1}{\tau})M_t\end{aligned} \tag{3.19}$$

Observe that the terms not depending on t simplify, divide by t, and take the limit $t\to 0$. We obtain that, for every $x\in X$ and every $x_1\in\mathcal{P}_c(-g,y_0)$ and $y_1\in\mathcal{R}_c(x_1)$,

$$g(x_1)+\frac{1}{\tau}c(x_1,y_0)\le g(x)+\frac{1}{\tau}c(x,y_0)-\frac{\lambda\tau+1}{\tau}[c(x,y_1)-c(x_1,y_1)] \tag{3.20}$$

which is precisely the cross-concavity of $-g$. □

Lemma 3.13. *Take* $\tau\in(0,\bar\tau)$. *Then, Assumption 3.10 implies the* c/τ*-cross-convexity of* f.

Proof. Use that $f(x_0)-\frac{c(x_0,z(t))}{\tau}\le f^{c/\tau}(z(t))=f(\gamma(t))-\frac{c(\gamma(t),z(t))}{\tau}$, multiply (3.18) by $1/\tau$ and sum with (3.17) to obtain

$$\begin{aligned}f(x_0)&\le f(\gamma(t))-\frac{c(\gamma(t),z(t))}{\tau}+\frac{c(x_0,z(t))}{\tau}\\&\le(1-t)f(x_0)+tf(x)+\frac{1-\lambda\tau}{\tau}t[c(x,\xi_0)-c(x_0,\xi_0)]\\&\quad+\frac{t}{\tau}c(x_0,z(t))-\frac{t}{\tau}c(x,z(t))+(1+\frac{1}{\tau})M_t.\end{aligned} \tag{3.21}$$

Take the limit $t\to 0$. We obtain that, for every $x\in X$, every $\xi_0\in\mathcal{S}_c(x_0)$, and $y_0=z(0)\in\mathcal{Q}_{c/\tau}(f,x_0)$,

$$f(x_0)\le f(x)+\frac{1-\lambda\tau}{\tau}[c(x,\xi_0)-c(x_0,\xi_0)]+\frac{1}{\tau}\liminf_{t\to 0}[c(x,z(t))-c(x_0,z(t))]. \tag{3.22}$$

As the last term is bounded by $(c(x,z(0))-c(x_0,z(0)))$, this proves the cross-convexity of f. □

Remark 3.14 (Explicit vs. implicit). Notice that the existence of $z(t)$ in Assumption 3.10 implies that f is c/τ-concave over $\gamma([0,1])$, but since $x_0 = \gamma(0)$ is left free, we should nevertheless request c/τ-concavity over the whole X. Moreover, the c/τ-concavity of f does not imply the existence of $z(t)$ which we hence have to assume. Informally, this second trajectory $z(t)$, the smoothness required by the lim inf, and the role of τ do not appear in Assumption 3.9 as implicit schemes for g do not rely on smoothness and require only a lower bound on the Hessian of g. On the contrary, explicit schemes require smoothness and the Hessian of f is assumed to fulfill a lower and an upper bound. We refer to [LAF23] for more details on this aspect.

Example 1.1 (continued). For geodesic metric spaces (X,d) and $c = d^2$, (3.14) and (3.18) are implied by Alexandrov curvature conditions, for $\gamma(\cdot)$ being a geodesic between x_0 and x.

In particular, if X is a Hilbert space and $c(x,y) = \frac{\|x-y\|^2}{2}$, this implies that (3.18), with $c(x_0, z(0)) = 0$, and (3.14), with $c(x_1, z_1) = 0$, are both satisfied and turn into equalities. We then see that (3.17) and (3.13) are nothing but the usual λ-convexity of f or g for $M_t = t^2[c(x,y_1) - c(x_1,y_1)]$. If X is a smooth reflexive Banach space, then Assumption 3.9 for $g = \|\cdot\|^2/2$, setting $f = 0$ (resp. Assumption 3.10 for $f = \|\cdot\|^2, 2$, setting $g = 0$), implies that X is Hilbert. Indeed, using Theorem 3.7 one can find a trajectory in $\mathrm{EVI}_\lambda(X,c,g)$ starting from any $x_0 \in X$, and we can apply [vRT12, Corollary 4.5].

When considering a geodesic space X with nonzero curvature, an interesting phenomenon arises. If X is nonpositively curved (NPC) in the sense of Alexandrov, we can consider a geodesic $\gamma(t)$ for Assumption 3.9, but not for Assumption 3.10; the opposite holds if X is positively curved (PC). In other words, when X is NPC, every λ-convex function g satisfies Assumption 3.9, and one finds a EVI solution by Theorem 3.7, as also stated in [MS20, Theorem 3.14]. Conversely, in PC spaces, Assumption 3.9 implies that $-g$ is c/τ-cross-concave for $\tau \in (0, 1/\lambda^-)$, which in turn implies that g is λ-convex, as shown in [LAF23, Proposition 4.14].

If X is the Wasserstein space equipped with the distance W_2, which is positively curved, [AGS08, Remark 9.2.8] discusses how Assumption 3.9 with $M_t = t^2[c(x,y_1) - c(x_1,y_1)]$ corresponds to convexity along generalized geodesics. This assumption implies [AGS08, Assumption 4.0.1], which leads to the discrete EVI and consequently to the continuous EVI. However, they also emphasize that convexity along generalized geodesics is a stronger requirement, as it implies convexity along standard geodesics. Note that (3.14) holds in both NPC spaces (along geodesics) and in NNCC spaces (along variational d^2-segments), with NNCC being a subclass of PC spaces.

In summary, explicit schemes are particularly well-suited for positively curved spaces and c-concave, λ-convex functions (which are therefore c-cross-convex) while in NPC spaces, λ-convex functions are more naturally handled via implicit schemes. The strong focus on implicit schemes in the Wasserstein space is likely motivated by the fact that some relevant functionals, such as the entropy, show some convexity but are not W_2^2-concave. Furthermore, as our discussion suggests, implicit schemes are conceptually simpler: they require only a single bound and demand less regularity. Finally, as shown in [LTV25, Theorem 3.11], the Wasserstein space is NNCC (as defined in Definition 2.18), so its structure is shaped not only by its positive curvature but also by its NNCC nature.

Example 1.2 (continued). For c as in (3.7) with X convex, in particular $c = \mathrm{KL}$, taking $\gamma(t) = (1-t)x_1 + tx$, one can argue similarly to [LTV25, Example 2.9] to show that $(X \times X, c)$ satisfies Definition 2.18 with (2.40) being an equality. So (3.14) and (3.18) hold with equality. Taking f and g convex will then give (3.13) and (3.17). The c/τ-concavity on f implies that $\frac{u}{\tau} - f$ is convex l.s.c., the existence of $z(t)$ corresponds to having a nonempty subdifferential at all points.

Example 1.3 (continued). For S_ϵ, preliminary work following [LLM+24] seems to indicate that potential functionals of the form $\mathcal{V}(\mu) = \int V(x)d\mu(x)$ with V convex C^1 are candidates for satisfying Assumption 3.9 or 3.10. A rigorous verification of this fact remains an open problem at this stage.

3.3. **Existence of the flow for implicit schemes.** In this section, we focus on implicit iterations and may hence simplify Assumption 3.5 as follows.

Assumption 3.15. *Let $c : X \times Y \to \mathbb{R}$ and $g : X \to \mathbb{R}$ be given and assume that there exist $\bar{\tau} > 0$ and $\lambda \geq 0$ such that*

(1) $-g$ *is* $\lambda\tau$*-strongly* c/τ*-cross-concave for every* $0 < \tau < \bar{\tau}$*;*
(2) for every $x_0 \in X$ *and every* $0 < \tau < \bar{\tau}$ *there exists some* $y_0 \in \mathcal{R}_c(x_0)$ *such that* $\mathcal{P}_{c/\tau}(-g, y_0) \neq \emptyset$.

Under the above Assumption 3.15, given an arbitrary starting point $x_0 \in X$ and $\tau \in (0, \bar{\tau})$ we consider the following scheme, which corresponds to (3.9) with $z_i^\tau = y_i^\tau$ since $f = 0$.

Definition 3.16 (Scheme for $\phi = g$). *For all* $i \in \mathbb{N}$ *we iteratively define*

$$x_0^\tau := x_0, \quad y_i^\tau \in \mathcal{R}_c(x_i^\tau), \quad x_{i+1}^\tau \in \mathcal{P}_{c/\tau}(-g, y_i^\tau). \tag{3.23}$$

Note that Assumption 3.15 guarantees the existence of two sequences $(x_i^\tau)_i \subset X$ and $(y_i^\tau)_i \subset X$ solving the scheme (3.23). We define the interpolating curves $\bar{x}^\tau : [0, +\infty) \to X$ and $\bar{y}^\tau : [0, +\infty) \to X$ as

$$\bar{x}_t^\tau := x_{\lfloor t/\tau \rfloor}^\tau, \quad \bar{y}_t^\tau := y_{\lfloor t/\tau \rfloor}^\tau, \quad t \in [0, +\infty). \tag{3.24}$$

Lemma 3.17. *Under Assumption 3.15, let* $\bar{x}^\tau$ *and* $\bar{y}^\tau$ *be defined as in (3.24). For any* $0 \le s = m\tau < n\tau = t$*, we have that*

$$\begin{aligned}\frac{c(x, \bar{y}_t^\tau) - c(x, \bar{y}_s^\tau)}{t - s} + \frac{1}{t-s} \sum_{i=m}^{n-1} \left[c(x_{i+1}^\tau, y_i^\tau) - c(x_{i+1}^\tau, y_{i+1}^\tau)\right] + \frac{\lambda\tau}{t-s} \sum_{i=m}^{n-1} c(x, y_{i+1}^\tau) \\ \le g(x) - \frac{\tau}{t-s} \sum_{i=m}^{n-1} g(x_{i+1}^\tau) \quad \forall x \in X.\end{aligned} \tag{3.25}$$

Relation (3.25) is a discrete version of the EVI in integral form (2.4), our goal is to show that we can indeed take the limit in τ.

Proof. Since $-g$ is $\lambda\tau$-strongly c/τ-cross-concave, by the definition of the scheme, we have

$$\begin{aligned}\frac{c(x, y_{i+1}^\tau) - c(x, y_i^\tau)}{\tau} + \frac{c(x_{i+1}^\tau, y_i^\tau) - c(x_{i+1}^\tau, y_{i+1}^\tau)}{\tau} + \lambda c(x, y_{i+1}^\tau) \\ \le g(x) - g(x_{i+1}^\tau) \quad \forall\, x \in X,\, i \in \mathbb{N}.\end{aligned} \tag{3.26}$$

We sum (3.26) from $i = m$ to $i = n - 1$ to get

$$\begin{aligned}\frac{c(x, \bar{y}_t^\tau) - c(x, \bar{y}_s^\tau)}{\tau} + \frac{1}{\tau} \sum_{i=m}^{n-1} c(x_{i+1}^\tau, y_i^\tau) - \frac{1}{\tau} \sum_{i=m}^{n-1} c(x_{i+1}^\tau, y_{i+1}^\tau) + \lambda \sum_{i=m}^{n-1} c(x, y_{i+1}^\tau) \\ \le (n - m) g(x) - \sum_{i=m}^{n-1} g(x_{i+1}^\tau)\end{aligned} \tag{3.27}$$

and dividing by $(n - m)$ we obtain (3.25). ☐

If c satisfies (Diss), then (3.23) defines a descent scheme over g, as well as a first step toward a Cauchy estimate involving the symmetrization of c.

Lemma 3.18 (g nonincreasing and first Cauchy-like estimate). *Let* c *be a cost satisfying* (Diss) *and let Assumption 3.15 hold. Then, the map* $t \mapsto g(\bar{x}_t^\tau)$ *is nonincreasing. Moreover, for any* $0 \le n\tau = t$ *and* $p \in \mathbb{N}$*, we have*

$$\begin{aligned}&[c(\bar{x}_t^{\tau/2^p}, \bar{y}_t^\tau) + c(\bar{x}_t^\tau, \bar{y}_t^{\tau/2^p})] \\ &\quad + \sum_{i=0}^{n-1} [c(\bar{x}_{i\tau}^{\tau/2^p}, \bar{y}_{(i+1)\tau}^\tau) - c(\bar{x}_{(i+1)\tau}^{\tau/2^p}, \bar{y}_{i\tau}^\tau) + c(\bar{x}_{i\tau}^\tau, \bar{y}_{(i+1)\tau}^{\tau/2^p}) - c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{i\tau}^{\tau/2^p})] \\ &\quad + 2 \sum_{i=0}^{n-1} c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{i\tau}^\tau) + 2 \sum_{j=0}^{n2^p - 1} c(\bar{x}_{(j+1)\tau/2^p}^{\tau/2^p}, \bar{y}_{j\tau/2^p}^{\tau/2^p}) \\ &\le \tau(2g(x_0) - g(\bar{x}_t^\tau) - g(\bar{x}_t^{\tau/2^p})).\end{aligned} \tag{3.28}$$

Proof. Assumption (Diss) and the fact that $y_i^\tau \in \mathcal{R}_c(x_i^\tau)$ imply that $c(x_i^\tau, y_i^\tau) = 0$. As $x_{i+1}^\tau \in \mathcal{P}_{c/\tau}(-g, y_i^\tau)$ and $c \geq 0$ we have that

$$g(x_{i+1}^\tau) \overset{c\geq 0}{\leq} \frac{c(x_{i+1}^\tau, y_i^\tau)}{\tau} + g(x_{i+1}^\tau) \overset{x_{i+1}^\tau \in \mathcal{P}_{c/\tau}(-g,y_i^\tau)}{\leq} \frac{c(x_i^\tau, y_i^\tau)}{\tau} + g(x_i^\tau) \overset{c(x_i^\tau,y_i^\tau)=0}{=} g(x_i^\tau).$$

This proves that $(g(x_i^\tau))_i$ is nonincreasing.

Since $c \geq 0$ from (Diss), by using (3.27) with $t = (i+1)\tau$, $s = i\tau$, $n = i+1$, and $m = i$, and recalling that $c(x_{i+1}^\tau, y_{i+1}^\tau) = 0$ as $y_{i+1}^\tau \in \mathcal{R}_c(x_{i+1}^\tau)$, for all $i \in \mathbb{N}$ we get

$$\frac{c(x, \bar{y}_{(i+1)\tau}^\tau) - c(x, \bar{y}_{i\tau}^\tau)}{\tau} + \frac{c(x_{i+1}^\tau, y_i^\tau)}{\tau} \leq g(x) - g(\bar{x}_{(i+1)\tau}^\tau) \quad \forall\, x \in X. \tag{3.29}$$

Write now (3.25) for the time step $\tau/2^p$, and for the choices $n = (i+1)2^p$ and $m = i2^p$. Still using that $c \geq 0$ and the fact that $c(x_{i+1}^\tau, y_{i+1}^\tau) = 0$ we get

$$\begin{aligned} &\frac{c(x, \bar{y}_{(i+1)\tau}^{\tau/2^p}) - c(x, \bar{y}_{i\tau}^{\tau/2^p})}{\tau} + \frac{1}{\tau} \sum_{j=i2^p}^{(i+1)2^p-1} c(\bar{x}_{(j+1)\tau/2^p}^{\tau/2^p}, \bar{y}_{j\tau/2^p}^{\tau/2^p}) \\ &\qquad \leq g(x) - g(\bar{x}_{(i+1)\tau}^{\tau/2^p}) \quad \forall\, x \in X,\ i \in \mathbb{N}. \end{aligned} \tag{3.30}$$

By summing (3.29) with $x = \bar{x}_{i\tau}^{\tau/2^p}$ and (3.30) with $x = \bar{x}_{(i+1)\tau}^\tau$, and then (3.29) with $x = \bar{x}_{(i+1)\tau}^{\tau/2^p}$ and (3.30) with $x = \bar{x}_{i\tau}^\tau$ we obtain

$$\begin{aligned} &c(\bar{x}_{i\tau}^{\tau/2^p}, \bar{y}_{(i+1)\tau}^\tau) - c(\bar{x}_{i\tau}^{\tau/2^p}, \bar{y}_{i\tau}^\tau) + c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{(i+1)\tau}^{\tau/2^p}) - c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{i\tau}^{\tau/2^p}) \\ &\quad + c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_i^\tau) + \sum_{j=i2^p}^{(i+1)2^p-1} c(\bar{x}_{(j+1)\tau/2^p}^{\tau/2^p}, \bar{y}_{j\tau/2^p}^{\tau/2^p}) \leq \tau(g(\bar{x}_{i\tau}^{\tau/2^p}) - g(\bar{x}_{(i+1)\tau}^{\tau/2^p})), \\ &c(\bar{x}_{(i+1)\tau}^{\tau/2^p}, \bar{y}_{(i+1)\tau}^\tau) - c(\bar{x}_{(i+1)\tau}^{\tau/2^p}, \bar{y}_{i\tau}^\tau) + c(\bar{x}_{i\tau}^\tau, \bar{y}_{(i+1)\tau}^{\tau/2^p}) - c(\bar{x}_{i\tau}^\tau, \bar{y}_{i\tau}^{\tau/2^p}) \\ &\quad + c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_i^\tau) + \sum_{j=i2^p}^{(i+1)2^p-1} c(\bar{x}_{(j+1)\tau/2^p}^{\tau/2^p}, \bar{y}_{j\tau/2^p}^{\tau/2^p}) \leq \tau(g(\bar{x}_{i\tau}^\tau) - g(\bar{x}_{(i+1)\tau}^\tau)). \end{aligned}$$

Taking the sum of the above inequalities and summing from $i = 0$ to $i = n-1$ we obtain (3.28). □

Proposition 3.19 (Cauchy estimate). *Let c be a cost satisfying (Diss) and let Assumption 3.15 hold. Assume additionally that c decomposes as $c = c_1 + c_2$ with*

(1) $c_1 \geq 0$ symmetric: $c_1(x_1, x_2) = c_1(x_2, x_1)$ for all $x_1, x_2 \in X$;
(2) c_2 fulfilling the triangle inequality:

$$c_2(x_1, x_2) \leq c_2(x_1, x_3) + c_2(x_3, x_2) \quad \forall x_1,\ x_2,\ x_3 \in X.$$

Then, for every $0 \leq n\tau = t$ and every $p \in \mathbb{N}$, we have

$$c(\bar{x}_t^{\tau/2^p}, \bar{x}_t^\tau) + c(\bar{x}_t^\tau, \bar{x}_t^{\tau/2^p}) \leq \tau(2g(x_0) - g(\bar{x}_t^\tau) - g(\bar{x}_t^{\tau/2^p})).$$

Proof. From the nondegeneracy of c coming from (Diss) we get that $x_i^\tau = y_i^\tau$ for all $\tau \in (0, \bar{\tau})$ and all $i \in \mathbb{N}$. The assertion of the proposition follows from (3.28), as soon as we check that the quantity A_c defined by

$$\begin{aligned} A_c := \sum_{i=0}^{n-1} [&c(\bar{x}_{i\tau}^{\tau/2^p}, \bar{y}_{(i+1)\tau}^\tau) - c(\bar{x}_{(i+1)\tau}^{\tau/2^p}, \bar{y}_{i\tau}^\tau) + c(\bar{x}_{i\tau}^\tau, \bar{y}_{(i+1)\tau}^{\tau/2^p}) - c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{i\tau}^{\tau/2^p})] \\ &+ 2\sum_{i=0}^{n-1} c(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{i\tau}^\tau) + 2\sum_{j=0}^{n2^p-1} c(\bar{x}_{(j+1)\tau/2^p}^{\tau/2^p}, \bar{y}_{j\tau/2^p}^{\tau/2^p}) \end{aligned} \tag{3.31}$$

is nonnegative. Note that $A_c = A_{c_1} + A_{c_2}$. From the symmetry of $c_1 \geq 0$ we obtain

$$A_{c_1} = 2\sum_{i=0}^{n-1} c_1(\bar{x}_{(i+1)\tau}^\tau, \bar{y}_{i\tau}^\tau) + 2\sum_{j=0}^{n2^p-1} c_1(\bar{x}_{(j+1)\tau/2^p}^{\tau/2^p}, \bar{y}_{j\tau/2^p}^{\tau/2^p}) \geq 0. \tag{3.32}$$

Using the triangle inequality and the fact that $x_i^\tau = y_i^\tau$ and $x_i^{\tau/2^p} = y_i^{\tau/2^p}$, for every $i = 0, \dots, n-1$ one has that

$$c_2(\bar{x}^\tau_{(i+1)\tau}, \bar{y}^\tau_{i\tau}) + c_2(\bar{x}^\tau_{i\tau}, \bar{y}^{\tau/2^p}_{(i+1)\tau}) + c_2(\bar{x}^{\tau/2^p}_{(i+1)\tau}, \bar{y}^{\tau/2^p}_{i\tau}) \geq c_2(\bar{x}^\tau_{(i+1)\tau}, \bar{y}^{\tau/2^p}_{i\tau}),$$
$$c_2(\bar{x}^{\tau/2^p}_{(i+1)\tau}, \bar{y}^{\tau/2^p}_{i\tau}) + c_2(\bar{x}^{\tau/2^p}_{i\tau}, \bar{y}^\tau_{(i+1)\tau}) + c_2(\bar{x}^\tau_{(i+1)\tau}, \bar{y}^\tau_{i\tau}) \geq c_2(\bar{x}^{\tau/2^p}_{(i+1)\tau}, \bar{y}^\tau_{i\tau}).$$

At the same time, again the triangle inequality gives

$$2\sum_{j=0}^{n2^p-1} c_2(\bar{x}^{\tau/2^p}_{(j+1)\tau/2^p}, \bar{y}^{\tau/2^p}_{j\tau/2^p}) \geq 2\sum_{i=0}^{n-1} c_2(\bar{x}^{\tau/2^p}_{(i+1)\tau}, \bar{y}^{\tau/2^p}_{i\tau}).$$

By combining these inequalities one gets that

$$\begin{aligned} A_{c_2} \geq \sum_{i=0}^{n-1}&[c_2(\bar{x}^{\tau/2^p}_{i\tau}, \bar{y}^\tau_{(i+1)\tau}) - c_2(\bar{x}^{\tau/2^p}_{(i+1)\tau}, \bar{y}^\tau_{i\tau}) + c_2(\bar{x}^\tau_{i\tau}, \bar{y}^{\tau/2^p}_{(i+1)\tau}) - c_2(\bar{x}^\tau_{(i+1)\tau}, \bar{y}^{\tau/2^p}_{i\tau})] \\ &+ 2\sum_{i=0}^{n-1} c_2(\bar{x}^\tau_{(i+1)\tau}, \bar{y}^\tau_{i\tau}) + 2\sum_{i=0}^{n-1} c_2(\bar{x}^{\tau/2^p}_{(i+1)\tau}, \bar{y}^{\tau/2^p}_{i\tau}) \geq 0. \end{aligned} \tag{3.33}$$

Inequalities (3.32)–(3.33) imply that $A_c = A_{c_1} + A_{c_2} \geq 0$, whence the thesis. □

Theorem 3.20 (Existence of EVI solutions via implicit scheme)**.** *Let c be a cost satisfying* (Diss) *and let Assumption 3.15 hold. Suppose in addition that*

(1) g is lower bounded;
(2) there exist a topology σ on X such that σ is compatible with (c, g) and c is additionally σ-continuous;
(3) Either
  (a) σ is Cauchy-compatible with (c, g) and c decomposes as in Proposition 3.19, or
  (b) the sublevel sets g are sequentially compact for σ.

Then, for every $x_0 \in X$ and $\tau \in (0, \bar\tau)$ the family of curves $(\bar{x}^{\tau/2^n})$ (constructed according to (3.23) *starting from x_0) converges (up to a subsequence in case (b)) pointwise (w.r.t. σ) as $\tau \downarrow 0$ to a σ-continuous curve $x : [0, +\infty) \to X$ which may depend on τ and satisfies*

$$c(x, x_t) - c(x, x_s) + \lambda \int_s^t c(x, x_r)\,\mathrm{d}r \leq (t-s)g(x) - \int_s^t g(x_r)\,\mathrm{d}r \quad \forall\, 0 \leq s \leq t,\ x \in X. \tag{3.34}$$

If $\mathrm{EVI}_\lambda(X, c, g)$ contains a unique trajectory $x(\cdot)$ starting from x_0, then, $\bar{x}^\tau(\cdot)$ σ-converges pointwise to this $x(\cdot)$, which does not depend on τ. In case (a), we also have the following error estimate

$$c\,(\bar{x}^\tau_t, x_t) \leq 2\tau(g(x_0) - \inf g) \quad \forall t \geq 0,\ \tau \in (0, \bar\tau). \tag{3.35}$$

Proof. We apply Lemma 3.17, written for the time step $\tau/2^n$. Discarding the second nonnegative term in the l.h.s., testing against any $x \in X$, we obtain, whenever $0 \leq s \leq t$ and $j, k \in \mathbb{N}$ are such that $j\tau \leq s2^n < (j+1)\tau$ and $k\tau \leq t2^n < (k+1)\tau$, that

$$\begin{aligned} &c(x, \bar{x}^{\tau/2^n}_t) - c(x, \bar{x}^{\tau/2^n}_s) + \lambda\tau\sum_{i=j}^{k-1} c(x, x^{\tau/2^n}_{i+1}) + \tau\sum_{i=j}^{k-1} g(x^{\tau/2^n}_{i+1}) \\ &\quad = c(x, \bar{x}^{\tau/2^n}_t) - c(x, \bar{x}^{\tau/2^n}_s) + \int_{j\tau 2^{-n}}^{(k-1)\tau 2^{-n}} \Big(\lambda c(x, \bar{x}^{\tau/2^n}_r) + g(\bar{x}^{\tau/2^n}_r)\Big)\,\mathrm{d}r \\ &\quad \leq (t-s)g(x). \end{aligned} \tag{3.36}$$

Notice that, choosing $x = \bar{x}^{\tau/2^n}_s$ and dropping the term $\lambda c \geq 0$, we get

$$c(\bar{x}^{\tau/2^n}_t, \bar{x}^{\tau/2^n}_s) \leq (t-s)(g(\bar{x}^{\tau/2^n}_s) - A) \leq (t-s)(g(x_0) - A) \quad \forall\, 0 \leq s \leq t, \tag{3.37}$$

where A is a lower bound for g. Now we split the proof in the two cases *(a)* and *(b)* as above.

In case *(a)*, by Proposition 3.19, we get that

$$c\left(\bar{x}_t^{\tau/2^n}, \bar{x}_t^{\tau/2^m}\right) \le \frac{\tau(g(x_0)-A)}{2^{m\wedge n-1}} \quad \forall t \ge 0,\ \tau \in (0,\bar{\tau}),\ m,n \in \mathbb{N}. \tag{3.38}$$

This shows that the sequence $(\bar{x}_t^{\tau/2^n})_{n\in\mathbb{N}}$ is c-Cauchy for every $t \ge 0$ so that, by Cauchy-compatibility, it converges (w.r.t. σ) as $n \to +\infty$ to a point $x_t \in X$.

In case *(b)*, let us set

$$\varphi_n(t) := \sup\left\{\sum_{i=1}^N c(\bar{x}_{t_{i-1}}^{\tau/2^n}, \bar{x}_{t_i}^{\tau/2^n}) : 0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = t,\ N \in \mathbb{N}\right\}, \quad t \ge 0.$$

Clearly, φ_n is a nondecreasing function such that $\varphi_n(t) \le t(g(x_0)-A)$ by (3.37) for every $t \ge 0$, and

$$c(\bar{x}_t^{\tau/2^n}, \bar{x}_s^{\tau/2^n}) \le \varphi_n(t) - \varphi_n(s) \quad \text{for every } 0 \le s \le t. \tag{3.39}$$

By Helly's theorem, up to passing to a unrelabeled subsequence, we can assume that there exists a nondecreasing function $\varphi : [0,+\infty) \to \mathbb{R}$ such that $\varphi_n(t) \to \varphi(t)$ for every $t \ge 0$. Take a countable dense subset A of $[0,+\infty)$ such that φ is continuous on A^c. By the diagonal principle and the sequential compactness of $\{x \,|\, g(x) \le g(x_0)\}$, we can find a subsequence $n_k \uparrow +\infty$ and points $(x_t)_{t\in A} \subset X$ such that $\bar{x}_t^{\tau/2^{n_k}}$ σ-converges to x_t as $k \to +\infty$. We now show that, for every $t \in A^c$, we have that there exists a point $x_t \in X$ such that $\bar{x}_t^{\tau/2^{n_k}}$ σ-converges to x_t. Let $t \in A^c$ and let $n_{k_h} \uparrow +\infty$ and $u_t \in X$ be such that $\bar{x}_t^{\tau/2^{n_{k_h}}}$ σ-converges to u_t as $h \to +\infty$. Let $(t_j)_j \subset A$ be such that $t_j \downarrow t$ as $j \to +\infty$ and let us write (3.39) for $t < t_j$, namely,

$$c(\bar{x}_{t_j}^{\tau/2^{n_{k_h}}}, \bar{x}_t^{\tau/2^{n_{k_h}}}) \le \varphi_{n_{k_h}}(t_j) - \varphi_{n_{k_h}}(t) \quad \text{for every } j,h \in \mathbb{N}.$$

Using the σ-continuity of c, we can pass to the limit as $h \to +\infty$ and we get

$$c(x_{t_j}, u_t) \le \varphi(t_j) - \varphi(t).$$

By the continuity of φ at t, we can pass to the limit as $j \to +\infty$ and we obtain that

$$\lim_{j\to+\infty} c(x_{t_j}, u_t) = 0.$$

Hence, by compatibility of c with σ, we have that u_t is the σ-limit of x_{t_j} and does not depends on the chosen subsequence, thus identifying it as the limit of $\bar{x}_t^{\tau/2^{n_k}}$ as $k \to +\infty$. This in particular proves the existence of $x : [0,+\infty) \to X$ such that $\bar{x}_t^{\tau/2^{n_k}}$ σ-converges to x_t for every $t \ge 0$.

In both cases *(a)* and *(b)*, we can pass to the limit as $n_k \to +\infty$ in (3.36) and in (3.37) and get that $x(\cdot)$ satisfies (3.34) and, by virtue of the compatibility of c with σ and of (3.37), it is σ-continuous. Assume that $\mathrm{EVI}_\lambda(X,c,g)$ contains a unique trajectory $x(\cdot)$ starting from x_0. In case *(a)*, we just take $n = 0$ and take $m \to +\infty$ in (3.38) to derive (3.35). Equation (3.35) also shows the pointwise convergence when taking the limit $\tau \to 0$. In case *(b)*, every converging subsequence w.r.t. n of $(\bar{x}_t^{\tau/2^n})_{t\in A}$ converges to a curve in $\mathrm{EVI}_\lambda(X,c,g)$, hence to $x(\cdot)$. This also shows that, for every $t > 0$ and every neighborhood V of x_t, there exists $N \in \mathbb{N}$ such that, for every $\tau \in (0, \frac{\bar{\tau}}{2^N})$, $\bar{x}_t^\tau \in V$. This concludes the proof. □

3.4. **Existence of the flow for splitting schemes.** In this section, we prove Theorem 3.7. The main steps are nearly identical to the implicit case $f = 0$. We refer to Section 3.1 for the relations between the iterates.

Lemma 3.21. *Under Assumption 3.5 the scheme in (3.9) satisfies for every $i \in \mathbb{N}$ and every $0 < \tau < \bar{\tau}$*

$$\frac{1}{\tau}[c(x,y_i^\tau) - c(x,\xi_i^\tau)] + \frac{1}{\tau}[c(x_i^\tau,\xi_i^\tau) - c(x_{i+1}^\tau, y_i^\tau)] + \lambda_f(c(x,\xi_i^\tau) - c(x_i^\tau,\xi_i^\tau))$$
$$\le f(x) - f(x_{i+1}^\tau) \quad \forall\, x \in X,\ \xi_i^\tau \in \mathcal{S}_c(x_i^\tau) \tag{3.40}$$

with $y_i^\tau, x_i^\tau, \xi_i^\tau$ as per (3.9).

If we additionally assume that c satisfies (Diss)*, i.e., $z_i^\tau \in \mathcal{R}_c(x_i^\tau)$ exists with $c(x_i^\tau, z_i^\tau) = 0$, then we have $z_i^\tau \in \mathcal{S}_c(x_i^\tau)$ and*

$$\frac{1}{\tau}[c(x, z_{i+1}^\tau) - c(x, z_i^\tau)] + \lambda_f c(x, z_i^\tau) + \lambda_g c(x, z_{i+1}^\tau) \leq \phi(x) - \phi(x_{i+1}^\tau) \quad \forall\, x \in X, \tag{3.41}$$

with $\phi = f + g$.

Proof. Fix $\xi_i^\tau \in \mathcal{S}_c(x_i^\tau)$ and y_i^τ defined as per (3.9). By definition of the c-transform we can bound $f(x_{i+1}^\tau) \leq \frac{c(x_{i+1}^\tau, y_i^\tau)}{\tau} + f^{c/\tau}(y_i^\tau)$. Since f is c/τ-concave, we get that $f(x_i^\tau) = \frac{c(x_i^\tau, y_i^\tau)}{\tau} + f^{c/\tau}(y_i^\tau)$. Thus

$$f(x_{i+1}^\tau) \leq f(x_i^\tau) - \frac{c(x_i^\tau, y_i^\tau)}{\tau} + \frac{c(x_{i+1}^\tau, y_i^\tau)}{\tau}. \tag{3.42}$$

The $\lambda_f\tau$-strongly c/τ-cross-convexity of f at x_i^τ gives

$$f(x_i^\tau) \leq f(x) + \frac{1}{\tau}[c(x, \xi_i^\tau) - c(x_i^\tau, \xi_i^\tau)] - \frac{1}{\tau}[c(x, y_i^\tau) - c(x_i^\tau, y_i^\tau)] - \lambda_f(c(x, \xi_i^\tau) - c(x_i^\tau, \xi_i^\tau)). \tag{3.43}$$

Summing (3.42) and (3.43), we get (3.40).

On the other hand, by the x-update (cf. (3.9)) we have

$$g(x_{i+1}^\tau) + \frac{c(x_{i+1}^\tau, y_i^\tau)}{\tau} \leq g(x_i^\tau) + \frac{c(x_i^\tau, y_i^\tau)}{\tau}. \tag{3.44}$$

Summing (3.42) and (3.44) yields

$$f(x_{i+1}^\tau) + g(x_{i+1}^\tau) \leq f(x_i^\tau) + g(x_i^\tau), \tag{3.45}$$

Using strong cross-concavity of $-g$ at y_i^τ gives us

$$\begin{aligned} g(x_{i+1}^\tau) \leq g(x) - \frac{1}{\tau}[c(x, z_{i+1}^\tau) - c(x_{i+1}^\tau, z_{i+1}^\tau)] + \frac{1}{\tau}[c(x, y_i^\tau) - c(x_{i+1}^\tau, y_i^\tau)] \\ - \lambda_g(c(x, z_{i+1}^\tau) - c(x_{i+1}^\tau, z_{i+1}^\tau)). \end{aligned} \tag{3.46}$$

Since $c(x_i^\tau, z_i^\tau) = 0$ and $c \geq 0$ by (Diss), we have that for all $x' \in X$, $c(x_i^\tau, z_i^\tau) = 0 \leq c(x', z_i^\tau)$. So $z_i^\tau \in \mathcal{S}_c(x_i^\tau)$ and we can choose $\xi_i^\tau = z_i^\tau$. In particular, y_i^τ depends on z_i^τ, and we can use z_i^τ in (3.40).

After summing (3.40) with $\xi_i^\tau = z_i^\tau$ and (3.46) and using that $c(x_i^\tau, z_i^\tau) = 0$, we obtain (3.41). □

Remark 3.22 (EVI for explicit/splitting scheme). If $g = 0$ we can take $\xi_i^\tau = y_{i-1}^\tau$ in (3.40) and prove that $(f(x_i^\tau))_i$ is a decreasing sequence, see (3.42). The term $[c(x_i^\tau, y_{i-1}^\tau) - c(x_{i+1}^\tau, y_i^\tau)]$ is then telescopic but has unknown sign. Nevertheless, if (Diss) holds, then this term vanishes since $c(x_{i+1}^\tau, y_i^\tau) = 0$, and we obtain

$$\frac{1}{\tau}[c(x, y_i^\tau) - c(x, y_{i-1}^\tau)] + \lambda_f(c(x, \xi_i^\tau) - c(x_i^\tau, \xi_i^\tau)) \leq f(x) - f(x_{i+1}^\tau) \tag{3.47}$$

which is a our candidate discrete EVI for the explicit scheme.

Our EVI (3.41) for the splitting scheme differs from the EVI (3.26) of the implicit scheme as the term $c(x_{i+1}^\tau, y_i^\tau)$ is missing. This is because (3.41) is obtained adding up inequalities, among which the c-concavity of f, which for $f = 0$ boils down to the uninformative $c(x_{i+1}^\tau, y_i^\tau) \geq 0$.

Lemma 3.23. *Let c be a cost satisfying* (Diss) *and let Assumption 3.5 hold. Let $\bar{x}^\tau$ and $\bar{z}^\tau$ be as in Definition 3.6 for an arbitrary starting point $x_0 \in X$. We have, for any $0 \leq s = m\tau < n\tau = t$, that*

$$\frac{c(x, \bar{z}_t^\tau) - c(x, \bar{z}_s^\tau)}{t - s} + \frac{\lambda_g \tau}{t - s} \sum_{i=m}^{n-1} c(x, z_{i+1}^\tau) + \frac{\lambda_f \tau}{t - s} \sum_{i=m}^{n-1} c(x, z_i^\tau) \leq \phi(x) - \phi(x_n^\tau) \quad \forall\, x \in X. \tag{3.48}$$

Proof. Using Lemma 3.21, we get

$$\frac{c(x, z_{i+1}^\tau) - c(x, z_i^\tau)}{\tau} + \lambda_g c(x, z_{i+1}^\tau) + \lambda_f c(x, z_i^\tau) \leq \phi(x) - \phi(x_{i+1}^\tau) \quad \forall\, x \in X\; i \in \mathbb{N}. \tag{3.49}$$

We sum (3.49) from $i = m$ to $i = n - 1$ to get

$$\frac{c(x, \bar{z}_t^\tau) - c(x, \bar{z}_s^\tau)}{\tau} + \sum_{i=m}^{n-1} [\lambda_g c(x, z_{i+1}^\tau) + \lambda_f c(x, z_i^\tau)] \le (n - m)\phi(x) - \sum_{i=m}^{n-1} \phi(x_{i+1}^\tau)$$

and dividing by $(n - m)$ we have

$$\frac{c(x, \bar{z}_t^\tau) - c(x, \bar{z}_s^\tau)}{t - s} + \frac{\tau}{t - s} \sum_{i=m}^{n-1} [\lambda_g c(x, z_{i+1}^\tau) + \lambda_f c(x, z_i^\tau)]$$
$$\le \phi(x) - \frac{\tau}{t - s} \sum_{i=m}^{n-1} \phi(x_{i+1}^\tau) \le \phi(x) - \phi(x_n^\tau), \quad (3.50)$$

where we used that $\phi(x_{i+1}^\tau) \le \phi(x_i^\tau)$ for every $i \in \mathbb{N}$. □

Proposition 3.24. *Let c be a cost satisfying (Diss) and let Assumption 3.5 hold. Assume additionally that $\lambda_f \ge 0$, $\lambda_g \ge 0$, and that c is symmetric. Then, for any $0 \le s = m\tau < n\tau = t$ and $p \in \mathbb{N}$, we have*

$$c(x_t^{\tau/2^p}, z_t^\tau) \le \tau(\phi(x_0) - \phi(x_{n\tau}^\tau)). \quad (3.51)$$

Proof. Since $c \ge 0$, we can drop the λ-term in the l.h.s. of (3.49) (3.50). We write these inequalities for two sequences, based on steps τ and $\tau/2^p$, respectively, getting

$$\frac{c(x, \bar{z}_{(i+1)\tau}^\tau) - c(x, \bar{z}_{i\tau}^\tau)}{\tau} \le \phi(x) - \phi(x_{(i+1)\tau}^\tau) \quad \forall\, x \in X, \quad (3.52)$$

$$\frac{c(u, z_{(i+1)\tau}^{\tau/2^p}) - c(u, z_{i\tau}^{\tau/2^p})}{\tau} \le \phi(u) - \phi(x_{(i+1)\tau}^{\tau/2^p}) \quad \forall\, u \in X. \quad (3.53)$$

Summing (3.52) with $x = x_{(i+1)\tau}^{\tau/2^p}$ and (3.53) with $u = x_{i\tau}^\tau$ we obtain

$$c(x_{(i+1)\tau}^{\tau/2^p}, z_{(i+1)\tau}^\tau) - c(x_{(i+1)\tau}^{\tau/2^p}, z_{i\tau}^\tau) + c(x_{i\tau}^\tau, z_{(i+1)\tau}^{\tau/2^p}) - c(x_{i\tau}^\tau, z_{i\tau}^{\tau/2^p}) \le \tau(\phi(x_{i\tau}^\tau) - \phi(x_{(i+1)\tau}^\tau)).$$

By symmetry, we have $c(x_{(i+1)\tau}^{\tau/2^p}, z_{i\tau}^\tau) = c(x_{i\tau}^\tau, z_{(i+1)\tau}^{\tau/2^p})$, so that summing from $i = 0 \ldots n - 1$, we obtain

$$c(x_t^{\tau/2^p}, z_t^\tau) \le \tau(\phi(x_0) - \phi(x_{n\tau}^\tau)).$$ □

Starting from the discrete EVI (3.48) and using the Cauchy property in (3.51), the proof of Theorem 3.7 is then identical to that of Theorem 3.20.

Acknowledgments

This research was funded in whole or in part by the Austrian Science Fund (FWF) projects 10.55776/F65, 10.55776/I5149, 10.55776/P32788, as well as by the OeAD-WTZ project CZ 09/2023. For open-access purposes, the authors have applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

References

[AFG24] P.-C. Aubin-Frankowski and S. Gaubert. Tropical reproducing kernels and optimization. *Integral Equations Operator Theory*, 96(2):Paper No. 19, 31, 2024.

[AFKL22] P.-C. Aubin-Frankowski, A. Korba, and F. Léger. Mirror descent with relative smoothness in measure spaces, with application to Sinkhorn and EM. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, Red Hook, NY, USA, 2022. Curran Associates Inc.

[AG13] L. Ambrosio and N. Gigli. A user's guide to optimal transport. In *Modelling and optimisation of flows on networks*, volume 2062 of Lecture Notes in Math., pages 1–155. Springer, Heidelberg, 2013.

[AGS08] 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, second edition, 2008.

[BP17] M. Bessenyei and Z. Páles. A contraction principle in semimetric spaces. *J. Nonlinear Convex Anal.*, 18(3):515–524, 2017.

[BP22] M. Bessenyei and E. Pénzes. Generalized fractals in semimetric spaces. *Nonlinear Anal.*, 220:Paper No. 112853, 15, 2022.

[Br67] L. M. Brègman. A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming. *Ž. Vyčisl. Mat i Mat. Fiz.*, 7:620–631, 1967.

[CDPS17] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.*, 49(2):1385–1418, 2017.

[CJT18] K. Chrzaszcz, J. Jachymski, and F. Turoboś. On characterizations and topology of regular semimetric spaces. *Publ. Math. Debrecen*, 93(1-2):87–105, 2018.

[CM11] P. Clément and J. Maas. A Trotter product formula for gradient flows in metric spaces. *J. Evol. Equ.*, 11(2):405–427, 2011.

[Cra17] K. Craig. Nonconvex gradient flow in the Wasserstein metric and applications to constrained nonlocal interactions. *Proc. Lond. Math. Soc. (3)*, 114(1):60–102, 2017.

[CRZ09] I. V. Chenchiah, M. O. Rieger, and J. Zimmer. Gradient flows in asymmetric metric spaces. *Nonlinear Anal.*, 71(11):5820–5834, 2009.

[CSS23] G. Cavagnari, G. Savaré, and G. E. Sodini. Dissipative probability vector fields and generation of evolution semigroups in Wasserstein spaces. *Probab. Theory Related Fields*, 185(3-4):1087–1182, 2023.

[CT93] G. Chen and M. Teboulle. Convergence analysis of a proximal-like minimization algorithm using Bregman functions. *SIAM J. Optim.*, 3(3):538–543, 1993.

[DS14] S. Daneri and G. Savaré. Lecture notes on gradient flows and optimal transport. In *Optimal transportation*, volume 413 of London Math. Soc. Lecture Note Ser., pages 100–144. Cambridge Univ. Press, Cambridge, 2014.

[FSV+18] J. Feydy, T. Séjourné, F.-X. Vialard, S. Amari, A. Trouvé, and G. Peyré. Interpolating between optimal transport and MMD using Sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, 2018.

[GPC18] A. Genevay, G. Peyre, and M. Cuturi. Learning generative models with Sinkhorn divergences. In A. Storkey and F. Perez-Cruz, editors, *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of Proceedings of Machine Learning Research, pages 1608–1617. PMLR, 09–11 Apr 2018.

[KM10] Y.-H. Kim and R. J. McCann. Continuity, curvature, and the general covariance of optimal transportation. *J. Eur. Math. Soc. (JEMS)*, 12(4):1009–1040, 2010.

[Le14] C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[LAF23] F. Léger and P.-C. Aubin-Frankowski. Gradient descent with a general cost, 2023. https://arxiv.org/abs/2305.04917.

[LLM+24] H. Lavenant, J. Luckhardt, G. Mordant, B. Schmitzer, and L. Tamanini. The Riemannian geometry of Sinkhorn divergences, 2024. https://arxiv.org/abs/2405.04987.

[LTV25] F. Léger, G. Todeschi, and F.-X. Vialard. Nonnegative cross-curvature in infinite dimensions: synthetic definition and spaces of measures, 2025. https://arxiv.org/abs/2409.18112.

[MS20] M. Muratori and G. Savaré. Gradient flows and evolution variational inequalities in metric spaces. I: Structural properties. *J. Funct. Anal.*, 278(4):108347, 67, 2020.

[NS23] S. Neumayer and G. Steidl. *From Optimal Transport to Discrepancy*, pages 1791–1826. Springer International Publishing, 2023.

[NY83] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, Inc., New York, 1983.

[OZ23] S.-I. Otha and W. Zhao. Gradient flows in asymmetric metric spaces and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 45, 2023.

[PC19] G. Peyré and M. Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

[RMS08] R. Rossi, A. Mielke, and G. Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 7(1):97–169, 2008.

[RS24] R. Rossi and G. Savaré. Abstract action spaces and their topological and dynamic properties. *Discrete Contin. Dyn. Syst. Ser. S*, 17(1):395–420, 2024.

[RW24] C. Rankin and T.-K. L. Wong. JKO schemes with general transport costs, 2024. https://arxiv.org/abs/2402.17681.

[San17] F. Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bull. Math. Sci.*, 7(1):87–154, 2017.

[SS05] E. M. Stein and R. Shakarchi. *Real analysis*, volume 3 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[vEH14] T. van Erven and P. Harremoës. Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. Inform. Theory*, 60(7):3797–3820, 2014.

[vRT12] M.-K. von Renesse and J. M. Tölle. On an EVI curve characterization of Hilbert spaces. *J. Math. Anal. Appl.*, 385(1):589–598, 2012.

Pierre-Cyril Aubin-Frankowski: CERMICS, ENPC, Institut Polytechnique de Paris, France
Email address: pierre-cyril.aubin@enpc.fr

Giacomo Enrico Sodini: Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria
Email address: giacomo.sodini@univie.ac.at

Ulisse Stefanelli: Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, Vienna Research Platform on Accelerating Photoreaction Discovery, University of Vienna, Währingerstrasse 17, 1090 Wien, Austria.
Email address: ulisse.stefanelli@univie.ac.at