Roberto Colombo (ORCID)

Partial regularity for minimizers of a class of discontinuous Lagrangians

Abstract. We study a one-dimensional Lagrangian problem including the variational reformulation, derived in a recent work of Ambrosio–Baradat–Brenier, of the discrete Monge–Ampère gravitational model, which describes the motion of interacting particles whose dynamics is ruled by the optimal transport problem. The more general action-type functional we consider contains a discontinuous potential term related to the descending slope of the opposite squared distance function from a generic discrete set in $\mathbb{R}^d$. We exploit the underlying geometrical structure provided by the associated Voronoi decomposition of the space to obtain $C^{1,1}$-regularity for local minimizers out of a finite number of shock times.

Contents

1. Introduction

In recent years, action functionals of the form

$$I_f(\gamma) = \int_0^1 |\dot{\gamma}|^2 + |\nabla f(\gamma)|^2 \tag{1.1}$$

have received the attention of many authors, due to their appearence in several areas of Mathematics. In the theory of gradient flows, for instance, they correspond to the integral form of the energy dissipation (see [5]), and they are also related to the so-called entropic regularization of the Wasserstein distance, when f is a multiple

R. Colombo (✉): Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56126 Pisa, Italy
e-mail: roberto.colombo@sns.it

of the logarithmic entropy, defined on the space of probability measures with finite quadratic moment (see [9,12,13]). In all these cases, the main obstruction to the application of standard results of Calculus of Variations stems from the lack of differentiability, even continuity, of the Lagrangian with respect to γ.

When $f : X \to (-\infty, +\infty]$ is a λ-convex function defined on a metric space (X, d), the term $|\nabla f(x)|$ has to be interpreted as the descending slope of f at x, namely

$$|\nabla^- f|(x) := \limsup_{y \to x} \frac{[f(x) - f(y)]^+}{d(x, y)}, \tag{1.2}$$

and, if X is a Hilbert space, it also coincides with the norm of the minimal selection in the subdifferential $\partial f(x)$, known as the extended gradient of f at x (see [5]). In this general framework, stability of the functionals I_f with respect to Γ-convergence of the functions f was investigated in [2–4]. In particular, [2] addressed a rigorous derivation, along the lines of [8], of a dynamical system of interacting particles strictly related to the optimal transport problem, known as the discrete Monge–Ampère gravitational (MAG) model (see Sect. 2.2). What emerged from this work is that the dynamics of MAG can be conveniently studied as the Euler–Lagrange equation associated to an action functional of type (1.1), where $f = f_K$ is the (-1)-convex function given by the (halved) opposite squared distance from a specific discrete set $K \subset \mathbb{R}^d$, namely

$$f_K(x) := -\frac{\mathrm{dist}_K^2(x)}{2} = -\min_{y \in K} \frac{|x - y|^2}{2}. \tag{1.3}$$

Clearly, f_K is not everywhere differentiable in general, so that ∇f_K, denoting the extended gradient of f_K, is not even continuous, and standard results of Calculus of Variations are not directly applicable in this case.

The present work aims at a systematic analysis of the properties of local minimizers for the functional I_{f_K}, where K is a generic discrete set K in $\mathbb{R}^d$. Our results apply in particular to solutions of the discrete MAG model, thereby addressing the general n-dimensional case, left open in [2], where the most involved part of the analysis was carried out only in dimension 1.

The plan of the paper is the following. In Sect. 2 we present the general framework and the motivations of our work. More precisely, in Sect. 2.1 we provide a contextualization of the problem in the general Hilbertian setting, with particular emphasis on the variational properties of functionals of type (1.1), when f is a λ-convex function. The main references for this part are [3,5]. Then, in Sect. 2.2, we introduce the Monge–Ampère gravitational model in the flat torus $\mathbb{T}^n$ as a modification of the classical Newtonian gravitation in which the linear Poisson equation is replaced by the fully non-linear Monge–Ampère equation (see (2.7)). Following the ideas of [8], we underline the intriguing link of this dynamical system with the optimal transport problem, whose powerful tools can be used in order to derive a Lagrangian reformulation of MAG particularly meaningful in the discrete setting, where additional foundation to the model is given by the results in [2]. By means of a least action principle, we then interpret solutions of the discrete MAG model as local minimizers of the functional I_{f_K}, where f_K is the opposite squared distance

function from a discrete set $K \subset \mathbb{R}^d$, as defined in (1.3). Section 3, being the core of the paper, is then devoted to the analysis of local minimizers for the Lagrangian problem associated to I_{f_K}, when K is a finite collection of points in $\mathbb{R}^d$:

$$K = \{p_1, \ldots, p_N\}.$$

We crucially consider the Voronoi partition of the space carried by K, which encodes the underlying geometrical structure of the problem, and exploit it in order to obtain, in Proposition 3.8, the existence of some specific directions along which momentum is locally conserved by the dynamics. As a byproduct, we show in Corollary 3.9 that a local minimizer γ is regular as long as it stays in a single Voronoi cell, possibly developing singularities only at those times in which the optimality class changes. As it is shown later, the set K also carries a partition of $\mathbb{R}^d$ into "potential zones" (see Proposition 3.6). This second partition is in general less fine than the Voronoi one, and coincides with it when the set K is "balanced", like for instance a cubic lattice. We then define $S(\gamma)$ to be the set of "shock times", at which the curve γ jumps from a Voronoi cell to another, and $NDS(\gamma) \subseteq S(\gamma)$ the set of "non-degenerate shock times", at which γ not only changes the Voronoi cell, but also the potential zone. With this in mind, our main regularity results Theorem 3.15 and Corollary 3.16 can be collected in a single statement as follows:

Theorem. (Partial regularity) *Let γ be a local minimizer of I_{f_K} with endpoints constraints. Then*

(i) *γ has a finite number of non-degenerate shock times out of which it is $C^{1,1}$.*
(ii) *Under the additional assumption that K is balanced, γ has a finite number of shock times out of which it is C^∞.*

This result in particular provides an extention to any space dimension of [2, Theorem 13], where regularity out of a finite number of shock times was proved for minimizers of a one-dimensional version of the MAG model.

2. General framework and motivations

2.1. Action functionals depending on the gradient of convex functions

λ-*convex functions.* Given a Hilbert space H, we consider a function $f : H \to (-\infty, +\infty]$, and denote by $\mathrm{dom}(f)$ its finiteness domain. We say that f is λ-*convex* if $x \mapsto f(x) - \frac{\lambda}{2}|x|^2$ is convex. It is easily seen that λ-convex functions are precisely those functions that satisfy the perturbed convexity inequality

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) - \frac{\lambda}{2}t(1-t)|x-y|^2, \quad t \in [0, 1].$$

By $\partial f(x)$ we denote the *Gateaux subdifferential* of f at $x \in \mathrm{dom}(f)$, namely the (possibly empty) closed convex set

$$\partial f(x) := \left\{ \xi \in H : \liminf_{t \to 0^+} \frac{f(x + tv) - f(x)}{t} \geq \xi \cdot v, \ \forall v \in H \right\}.$$

We denote by $\mathrm{dom}(\partial f)$ the domain of the subdifferential. For a λ-convex function, we can exploit the monotonicity of difference quotients to derive the equivalent non-asymptotic definition of the subdifferential

$$\partial f(x) := \left\{ \xi \in H : f(y) \geq f(x) + \langle \xi, y - x \rangle + \frac{\lambda}{2}|y - x|^2, \ \forall y \in H \right\}.$$

Whenever $x \in \mathrm{dom}(\partial f)$, there exists a unique element ξ with minimal norm in $\partial f(x)$, obtained by projecting 0 on $\partial f(x)$. This element is called the *extended gradient* of f at x, and is denoted by $\nabla f(x)$. The concept of extended gradient is strictly related to the one of *descending slope* of f at $x \in \mathrm{dom}(f)$, namely

$$|\nabla^- f|(x) := \limsup_{y \to x} \frac{[f(x) - f(y)]^+}{|x - y|}.$$

In fact, for λ-convex functions, it can be proved that $\partial f(x)$ is not empty if and only if $|\nabla^- f|(x) < +\infty$, and that, in this case, the following equalities hold (see [5]):

$$|\nabla f(x)| = |\nabla^- f|(x) = \sup_{y \neq x} \frac{[f(x) - f(y) + \frac{\lambda}{2}|x - y|^2]^+}{|x - y|}. \tag{2.1}$$

By setting $|\nabla f|$ equal to $+\infty$ out of $\mathrm{dom}(\partial f)$, we easily deduce from (2.1) that the function $H \ni x \mapsto |\nabla f(x)| \in [0, +\infty]$ is lower semicontinuous, being the supremum of a collection of continuous functions.

In this paper we deal with the following specialization of the above setting. Given a closed set $K \subseteq H$, we consider the (halved) *opposite squared distance function* from K, namely

$$f_K(x) := -\frac{\mathrm{dist}_K^2(x)}{2} = -\inf_{y \in K} \frac{|x - y|^2}{2}. \tag{2.2}$$

The infimum is not attained in general, unless K is either convex or compact, or H is finite-dimensional. By defining the convex function

$$g_K(x) := \sup_{y \in K} \left(\langle x, y \rangle - \frac{|y|^2}{2} \right),$$

we derive from the equality $g_K(x) = f_K(x) + \frac{|x|^2}{2}$ that f_K is (-1)-convex.

A class of action functionals. We now introduce, in the general Hilbertian setting, the class of action functionals that we are going to study throughout the paper. We fix a function $h : [0, +\infty] \to [0, +\infty]$ representing a "potential shape". Then, for $\delta > 0$ and $f : H \to (-\infty, +\infty]$ proper, λ-convex and lower semicontinuous, we consider the functional $I_f^\delta : C([0, \delta], H) \to [0, +\infty]$ defined by

$$I_f^\delta(\gamma) := \begin{cases} \displaystyle\int_0^\delta |\dot{\gamma}|^2 + h(|\nabla f|^2(\gamma)) & \text{if } \gamma \in AC([0, \delta], H), \\ +\infty & \text{otherwise.} \end{cases} \tag{2.3}$$

Compared to the one of type (1.1) studied so far in the literature, we consider here the enriched class of functionals in which the potential shape h is allowed to be different from the identity. In the sequel we assume h to be continuous, and C^1 when restricted to $[0, +\infty)$.

As we have in mind to study this type of functionals from the variational point of view, it is crucial to realize that (2.3) is lower semicontinuous with respect to the $C([0, \delta], H)$ topology. This in fact easily follows from the lower semicontinuity of the classical action and the above characterization of the extended gradient (2.1). Then, for $x_0, x_\delta \in H$, the infimum

$$\Gamma_f^\delta(x_0, x_\delta) := \inf \left\{ I_f^\delta(\gamma) : \gamma(0) = x_0, \gamma(\delta) = x_\delta \right\}$$

is attained under suitable coercivity conditions. Note in particular that this is the case if H is finite-dimensional.

Due to the lack of continuity of the potential term, however, very little is known about the regularity for minimizers of this type of functionals, even in the finite-dimensional case. We could ask for instance whether some higher regularity or at least a sort of *Euler–Lagrange equation* like formally

$$\ddot{\gamma} = h'(|\nabla f(\gamma)|^2)\nabla^2 f(\gamma)\nabla f(\gamma) \tag{2.4}$$

could be derived for local minimizers. It is worth mentioning here that a rigorous Euler–Lagrange equation like the one in (2.4) has been derived in [10] in the context of the entropic regularization of the Wasserstein distance, where f is the Boltzmann–Shannon relative entropy. In the very specific case in which f is the opposite squared distance function from a discrete set in $\mathbb{R}^d$, we will prove in the sequel that local minimizers are piecewise $C^{1,1}$, and that, out of a finite number of singularities, (2.4) holds taking the modulus on both sides and replacing the equality with a $\leq$ sign (see Theorem 3.15). Nevertheless, in the general setting, one can exploit the fact that the functional I_f^δ is autonomous in order to perform "horizontal" variations of the independent variable, and eventually derive the *Du Bois–Reymond equation* for a local minimizer γ (see [1]):

$$\frac{d}{dt}\left\{ |\dot{\gamma}|^2 - h(|\nabla f|^2(\gamma)) \right\} = 0 \tag{2.5}$$

in the sense of distributions in $(0, \delta)$. Equivalently, there exists a constant $c \in \mathbb{R}$ such that

$$|\dot{\gamma}|^2 = h(|\nabla f|^2(\gamma)) + c$$

a.e. in $(0, \delta)$. This implies in particular that every local minimizer of I_f^δ is Lipschitz continuous, provided that $|\nabla f|$ is bounded on bounded sets.

We end this part by quoting a result from [3] addressing the matter of stability for the class of functionals considered so far. By adding endpoints constraints $x_0, x_\delta \in H$, we define the functional $I_{f,x_0,x_\delta}^\delta : C([0, \delta], H) \to [0, +\infty]$ such that

$$I_{f,x_0,x_\delta}^\delta(\gamma) := \begin{cases} \int_0^\delta |\dot{\gamma}|^2 + h(|\nabla f|^2(\gamma)) & \text{if } \gamma \in AC([0, \delta], H) \text{ and } \gamma(0) = x_0, \gamma(\delta) = x_\delta, \\ +\infty & \text{otherwise.} \end{cases}$$

Theorem 2.1. (Stability, [3]) *Let f_j, f be uniformly λ-convex functions, and let $x_{j,0}, x_{j,\delta}, x_0, x_\delta \in H$. Suppose that*

(i) $f_j \to f$ *w.r.t. Mosco convergence.*
(ii) $\lim_{j \to \infty} x_{j,i} = x_i$, *for $i = 0, \delta$.*
(iii) $\sup_j |\nabla f_j|(x_{j,i}) < \infty$, *for $i = 0, \delta$.*

Then $I^\delta_{f_j, x_{j,0}, x_{j,\delta}}$ Γ-converge to $I^\delta_{f, x_0, x_\delta}$ in the $C([0, \delta], H)$ topology.

As a byproduct, under an additional equi-coercivity assumption, this theorem grants convergence of minimal values to minimal values and of minimizers to minimizers. Notice that Theorem 2.1 is stated in [3] for $h = \text{id}$, but the same proof is seen to work in the general case with only minor modifications.

2.2. The Monge–Ampère gravitational model

In a periodic spatial domain like the flat torus $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$, we can describe classical Newtonian gravitation of a unity of mass in a "parametric" way as follows. We first choose a reference probability space $(\mathcal{A}, \lambda)$ of labels for the gravitating particles. Then we assign to each particle $a \in \mathcal{A}$ its position $X_t(a) \in \mathbb{T}^n$ at time t. Tipical choices for the reference space are the unit cube $[0, 1]^n$ with the n-dimensional Lebesgue measure in the continuous case, and a finite set of points with the renormalized counting measure in the discrete case. Denoting by $\mu_t := (X_t)_\# \lambda$ the image measure of λ by X_t, the Newtonian model can be written as

$$\begin{cases} \frac{d^2}{dt^2} X_t(a) = -\nabla \phi_t(X_t(a)), \\ \Delta \phi_t = \mu_t - 1. \end{cases} \tag{2.6}$$

Here ϕ_t is the gravitational potential generated by μ_t, defined on $\mathbb{T}^n$. Notice that due to the periodicity of the space, the average density 1 has been subtracted from the right-hand side of the Poisson equation, in order to let the uniform measure $\mathscr{L}^n$ be a stationary solution of the system. This is a perfectly meaningful assumption, because by symmetry, the attractive force of the uniform density has to be zero everywhere on $\mathbb{T}^n$.

In this section we are interested in the related *Monge–Ampère gravitational model* (MAG in short), which is simply obtained from (2.6) by replacing the Poisson equation with the fully non-linear Monge–Ampère equation:

$$\begin{cases} \frac{d^2}{dt^2} X_t(a) = -\nabla \phi_t(X_t(a)), \\ \det(\mathbb{I} + \nabla^2 \phi_t) = \mu_t. \end{cases} \tag{2.7}$$

Notice that (2.6) can be recovered from (2.7) by expanding the determinant in the Monge–Ampère equation and keeping only the linear term:

$$\det(\mathbb{I} + \nabla^2 \phi_t) \approx 1 + \text{Tr}(\nabla^2 \phi_t) = 1 + \Delta \phi_t.$$

We refer to [8] and the references therein for a broader introduction to this dynamical system, as well as for a comparison with the classical Newtonian model.

The MAG model in optimal transportation terms. System (2.7) appears to have an intriguing geometrical interpretation if we look at it from the optimal transportation point of view. In order to better illustrate this link, we first quote the following specialization to the flat torus $\mathbb{T}^n$ of the classical Brenier-McCann theorem on the existence and uniqueness of optimal transport maps on Riemannian manifolds (see [7,11,15]). Let us begin with some notation. We denote by $\pi : \mathbb{R}^n \to \mathbb{T}^n$ the projection to the quotient. We say that a vector field $F : \mathbb{R}^n \to \mathbb{R}^n$ is $\mathbb{Z}^n$-translation invariant if $F(\cdot + z) = F(\cdot) + z$, for every $z \in \mathbb{Z}^n$. If this is the case we make a little abuse of notation by considering F also as a vector field from $\mathbb{T}^n$ to itself. Given a Borel probability measure λ on $\mathbb{T}^n$, we consider the Hilbert space

$$H_\lambda := L^2(\mathbb{T}^n, \lambda; \mathbb{R}^n)$$

and its closed subset K_λ given by all the λ-preserving vector fields

$$K_\lambda := \{Y \in H_\lambda : (\pi \circ Y)_\# \lambda = \lambda\}.$$

Finally, we recall that f_{K_λ} denotes the opposite squared distance function from K_λ, as defined in (2.2).

Theorem 2.2. (Existence and uniqueness of optimal transport maps in $\mathbb{T}^n$) *Let μ and λ be Borel probability measures on $\mathbb{T}^n$, and suppose that $\mu \ll \mathscr{L}^n$. Then*

(i) *There exists a locally Lipschitz convex function $\psi : \mathbb{R}^n \to \mathbb{R}$, such that $\phi(x) := \psi(x) - \frac{|x|^2}{2}$ is $\mathbb{Z}^n$-periodic (therefore $\nabla\psi(x) = x + \nabla\phi(x)$ is $\mathbb{Z}^n$-translation invariant), and $T := \nabla\psi : \mathbb{T}^n \to \mathbb{T}^n$ is the unique optimal transport map from μ to λ.*

(ii) *If $\mu = \rho\mathscr{L}^n$ and $\lambda = \eta\mathscr{L}^n$ are both absolutely continuous w.r.t. the Lebesgue measure, then ϕ solves the Monge–Ampère equation*

$$\det(\mathbb{I} + \nabla^2\phi)\eta(T(x)) = \rho(x) \tag{2.8}$$

in the almost everywhere sense. Furthermore, if ρ and η are of class $C^{0,\alpha}$ and $\rho, \eta > 0$, then ϕ is of class $C^{2,\beta}$, for $0 < \beta < \alpha$, and solves (2.8) in the classical sense.

(iii) *Let $Y \in H_\lambda$ be such that $(\pi \circ Y)_\# \lambda = \mu$. Then $T \circ Y$ is the unique projection of Y on K_λ, and*

$$\|T \circ Y - Y\|_{H_\lambda} = W_2(\mu, \lambda), \tag{2.9}$$

where W_2 is the Wasserstein distance in the space $\mathscr{P}_2(\mathbb{T}^n)$. Moreover, the map f_{K_λ} is Gateaux differentiable at Y and it holds

$$\nabla f_{K_\lambda}(Y) = T \circ Y - Y = \nabla\phi \circ Y. \tag{2.10}$$

In order to reformulate the Monge–Ampère gravitational model in optimal transportation terms, we look to the continuous case, in which the reference space is given by $(\mathbb{T}^n, \mathscr{L}^n)$. Fix then $\lambda = \mathscr{L}^n$ in the Theorem above, and consider a parametrization $X_t : \mathbb{T}^n \to \mathbb{T}^n$ such that $(X_t)_\# \lambda = \mu_t$ and $\mu_t = \rho_t\mathscr{L}^n$ is absolutely continuous w.r.t. the Lebesgue measure. If $Y_t \in H_\lambda$ is any lifting of X_t, that is to say a map

that satisfies $\pi \circ Y_t = X_t$, then Theorem 2.2 grants that the Kantorovich potential ϕ_t solves the Monge–Ampère equation

$$\det(\mathbb{I} + \nabla^2 \phi_t) = \mu_t,$$

and $-\nabla \phi_t = Y_t - T_t \circ Y_t$, where T_t is the unique optimal transport map from μ_t to λ. So we see that (2.7) reduces to

$$\frac{d^2}{dt^2} Y_t(a) = Y_t(a) - T_t(Y_t(a)). \tag{2.11}$$

Moreover, from (2.10) we obtain

$$|\nabla f_{K_\lambda}(Y)|^2 = |T \circ Y - Y|^2 = -2 f_{K_\lambda}(Y),$$

suggesting an interpretation of (2.11) as the Euler–Lagrange equation associated to the functional

$$\int_0^\delta |\dot{\gamma}|^2 + |\nabla f_{K_\lambda}(\gamma)|^2, \quad \gamma : [0, \delta] \to H_\lambda. \tag{2.12}$$

This variational reformulation appears natural in the attempt to give a meaning to system (2.7) also in the discrete setting, where, as it is well-known, Theorem 2.2 fails.

The discrete MAG model. As already mentioned before, one of the aims of this work is to go deeper in the analysis of the discrete version of the Monge–Ampère gravitational model, first introduced in [8] and then formalized in [2]. Here we choose as reference measure

$$\lambda = \frac{1}{m} \sum_{i=1}^m \delta_{a_i},$$

where the a_i's are distinct points on $\mathbb{T}^n$ (think for instance to a regular lattice approximating the uniform measure). In this case, the space H_λ is easily seen to be finite-dimensional, and isomorphic to $\mathbb{R}^{nm}$, through the identification of a map $Y \in H_\lambda$ with the m-uple $(Y(a_1), \ldots, Y(a_m)) \in (\mathbb{R}^n)^m$. Under this correspondence, K_λ is represented by the discrete set of all points $(b_1, \ldots, b_m)$ in $\mathbb{R}^{nm}$ such that $\pi(\{b_1, \ldots, b_m\}) = \{a_1, \ldots, a_m\}$. By regarding, a bit improperly, the a_i's as elements of $[0, 1)^n$, the set K_λ can be written as the union of $m!$ cubic lattices in $\mathbb{R}^{nm}$:

$$K_\lambda = \bigcup_{\sigma \in \mathfrak{S}_m} (a_{\sigma(1)}, \ldots, a_{\sigma(m)}) + \mathbb{Z}^{nm}.$$

In this discrete scenario, the MAG model describes the motion of m particles of equal mass $1/m$ in the torus $\mathbb{T}^n$, whose dynamics is ruled by the optimal transport problem as follows. The position of the ith particle at time t is denoted by $x_i(t) =$

$X_t(a_i)$, and a lifting of $x_i(t)$ to $\mathbb{R}^n$ by $y_i(t) = Y_t(a_i)$. The equivalent of (2.11) in this setting is, at least formally,

$$\ddot{y}_i(t) = y_i(t) - b_i^{\mathrm{opt}}(t), \quad i \in \{1, \ldots, m\}, \tag{2.13}$$

where $(b_1^{\mathrm{opt}}(t), \ldots, b_m^{\mathrm{opt}}(t))$ is the closest point to $(y_1(t), \ldots, y_m(t))$ in K_λ. The system (2.13) is easily seen to be ill posed, because of the general non-uniqueness of the projection on K_λ, ultimately due to the non-uniqueness of the optimal transport map in the discrete setting, in contrast with the absolutely continuous one. As already pointed out in [2], in order to fix this problem, it is convenient to switch to a variational reformulation of the dynamical system, by considering an action functional of type (2.12). Therefore, relying on a least action principle, we say that $y \in AC([0, \delta], \mathbb{R}^{nm})$ is a solution of the discrete MAG model if it is a local minimizer of the functional

$$\int_0^\delta |\dot{\gamma}|^2 + |\nabla f_{K_\lambda}(\gamma)|^2$$

subject to endpoints constraints. In the next section, we are going to study a more general functional in which K_λ is replaced by a generic discrete set K in $\mathbb{R}^d$.

Before concluding this part, we would like to briefly turn the attention of the reader to an analogous Lagrangian problem in the space of probability measures $(\mathscr{P}(\mathbb{T}^n), W_2)$. This can be obtained from MAG by dropping the parametric description of the gravitating matter, required by the Hilbertian setting of Sect. 2.1, and directly considering the evolution of a probability measure μ_t in $\mathbb{T}^n$.

A related Lagrangian problem in $(\mathscr{P}(\mathbb{T}^n), W_2)$. Far from being limited to the Hilbertian context, functionals of type (1.1) can be considered in a much more general metric setting, provided that we interpret $|\nabla f(x)|$ as the descending slope $|\nabla^- f|(x)$ defined in (1.2), and $|\dot{\gamma}|$ as the metric derivative of an absolutely continuous curve $\gamma : [0, 1] \to X$. We avoid to repeat all the constructions in this new scenario (see [4] for a systematic introduction), and prefer to immediately specialize to our case of interest. We take $(X, d) = (\mathscr{P}(\mathbb{T}^n), W_2)$ the space of probability measures on $\mathbb{T}^n$ endowed with the Wasserstein distance induced by the optimal transport problem with quadratic cost. As it is well-known, (X, d) is compact, geodesic and positively curved (see [5]). Given a "reference" probability measure $\lambda \in \mathscr{P}(\mathbb{T}^n)$, we consider the (halved) opposite squared distance function from λ, namely

$$f_\lambda(\mu) = -\frac{W_2^2(\mu, \lambda)}{2}.$$

Since (X, d) is positively curved, we easily deduce that f_λ is (-1)-convex (in the metric setting, convexity has to be intended along geodesics). Moreover, we can bound the descending slope of f_λ at μ as follows:

$$\begin{aligned} |\nabla^- f_\lambda|(\mu) &= \limsup_{\nu \to \mu} \frac{[W_2^2(\nu, \lambda) - W_2^2(\mu, \lambda)]^+}{2W_2(\mu, \nu)} \\ &\leq \limsup_{\nu \to \mu} \frac{W_2(\nu, \lambda) + W_2(\mu, \lambda)}{2} = W_2(\mu, \lambda). \end{aligned} \tag{2.14}$$

A precise characterization of $|\nabla^- f_\lambda|(\mu)$ is given in [5, Theorem 10.4.12], and involves the minimal L^2 norm of the barycentric projection of optimal transport plans. Inspired by the MAG model, and in particular by formulas (2.9) and (2.10), one could study the Lagrangian problem associated to the lower semicontinuous functional $I^\delta_{f_\lambda,\mu_0,\mu_\delta} : C([0,\delta], X) \to [0, +\infty]$ defined by

$$I^\delta_{f_\lambda,\mu_0,\mu_\delta}(\gamma) = \begin{cases} \int_0^\delta |\dot\gamma|^2 + |\nabla^- f_\lambda|^2(\gamma) & \text{if } \gamma \in AC([0,\delta], X), \text{ and } \gamma(0) = \mu_0, \gamma(\delta) = \mu_\delta, \\ +\infty & \text{otherwise.} \end{cases}$$

From the compactness of X, we immediately obtain the existence of minimizers of $I^\delta_{f_\lambda,\mu_0,\mu_\delta}$. In addition, by exploiting a generalization of Theorem 2.1 to the general metric setting provided by [4, Theorem 17], as well as the bound on the descending slope (2.14), we obtain the following stability result:

Proposition 2.3. *Let* $\lambda_j, \lambda \in \mathscr{P}(\mathbb{T}^d)$ *be reference measures, and* $\mu_{j,0}, \mu_{j,\delta}, \mu_0,$ $\mu_\delta \in \mathscr{P}(\mathbb{T}^d)$ *be endpoints. Suppose that*

(i) $\lambda_j \to \lambda$ *in* W_2.
(ii) $\mu_{j,i} \to \mu_i$ *in* W_2, *for* $i = 0, \delta$.

Then $I^\delta_{f_{\lambda_j},\mu_{j,0},\mu_{j,\delta}}$ *Γ-converge to* $I^\delta_{f_\lambda,\mu_0,\mu_\delta}$ *in the* $C([0,\delta], X)$ *topology. Moreover, we have convergence of minimal values to minimal values and of minimizers to minimizers.*

3. The case of the opposite squared distance function in $\mathbb{R}^d$

In this section the main results of the paper will be derived. We study functionals of type (2.1) in the special case in which $H = \mathbb{R}^d$ and $f = f_K$ is the opposite squared distance function from a closed subset $K \subseteq \mathbb{R}^d$. Motivated by the variational reformulation of the discrete MAG model, derived in the previous section, we will in particular focus on the case in which K is a discrete collection of points in $\mathbb{R}^d$. We highlight here that the functional associated to a general closed set can be approximated, in the sense of Γ-convergence, by functionals associated to discrete sets (see Corollary 3.4 below). In the discrete setting, we will exploit the geometrical structure given by the associated Voronoi decomposition of the space in order to get regularity for local minimizers out of a finite number of "shock times".

Given a closed set $K \subseteq \mathbb{R}^d$, we consider the opposite squared distance function from K, defined by

$$f_K(x) := -\frac{\text{dist}_K^2(x)}{2} = -\min_{y \in K} \frac{|x - y|^2}{2}. \tag{3.1}$$

Notice that the infimum in (2.2) is always attained here, due to the local compactness of the ambient space. The convex function

$$g_K(x) := \max_{y \in K} \left(x \cdot y - \frac{|y|^2}{2} \right) \tag{3.2}$$

satisfies $g_K(x) = f_K(x) + \frac{|x|^2}{2}$, thus implying the (-1)-convexity of f_K. We fix a potential shape $h : [0, +\infty) \to [0, +\infty)$ of class C^1 and consider the action functional $I^\delta_{f_K, x_0, x_\delta} : C([0, \delta], \mathbb{R}^d) \to [0, +\infty]$ defined by

$$I^\delta_{f_K, x_0, x_\delta}(\gamma) := \begin{cases} \int_0^\delta |\dot\gamma|^2 + h(|\nabla f_K|^2(\gamma)) & \text{if } \gamma \in AC([0, \delta], \mathbb{R}^d) \text{ and } \gamma(0) = x_0, \gamma(\delta) = x_\delta, \\ +\infty & \text{otherwise.} \end{cases}$$

We stress that ∇f_K has to be intended as an extended gradient, because f_K is differentiable only at those points in which the projection on K is unique.

In order to get a useful characterization of ∇f_K, we need a well-known Lemma of convex analysis providing an explicit formula for the subdifferential at x of the maximum of a family of convex functions, under suitable assumptions (see [14]).

Lemma 3.1. (Subdifferential of the sup function) *Let* $\{g_\alpha : \mathbb{R}^d \to \mathbb{R}\}_{\alpha \in \mathcal{A}}$ *be a collection of convex functions indexed on a compact metric space* $\mathcal{A}$*, and suppose that* $\alpha \mapsto g_\alpha(x)$ *is upper semicontinuous for every* $x \in \mathbb{R}^d$*. We consider the supremum function*

$$g := \sup_{\alpha \in \mathcal{A}} g_\alpha.$$

Then, if the supremum in the definition of $g(x)$ *is attained, the following formula holds for the subdifferential of* g *at* x*:*

$$\partial g(x) = \text{conv}\left(\bigcup \{\partial g_\alpha(x) : g_\alpha(x) = g(x)\}\right).$$

We call $\text{opt}_K(x)$ the compact subset of K containing all the points that minimize the distance from x:

$$\text{opt}_K(x) := \{y \in K : \text{dist}_K(x) = |x - y|\} = \left\{y \in K : g_K(x) = x \cdot y - \frac{|y|^2}{2}\right\}.$$

In the sequel we will refer to $\text{opt}_K(x)$ as the *optimality class* of x. Applying Lemma 3.1 we get:

Proposition 3.2. (Subdifferential of the opposite squared distance function) *Let* $K \subseteq \mathbb{R}^d$ *be a closed set, and let* f_K, g_K *be defined as in* (3.1) *and* (3.2)*. Then*

(i) *The subdifferential of* g_K *at* x *is given by*

$$\partial g_K(x) = \text{conv}(\text{opt}_K(x)).$$

(ii) *The subdifferential of* f_K *at* x *is given by*

$$\partial f_K(x) = \text{conv}(\text{opt}_K(x)) - x.$$

Moreover, denoting by $\eta_K(x)$ *the unique projection of* x *on the closed convex set* $\text{conv}(\text{opt}_K(x))$*, the following formula holds for the extended gradient of* f_K *at* x*:*

$$\nabla f_K(x) = \eta_K(x) - x. \tag{3.3}$$

(iii) *The point $\eta_K(x)$ depends only on the optimality class of x. That is to say, $\eta_K(x) = \eta_K(y)$, whenever $\mathrm{opt}_K(x) = \mathrm{opt}_K(y)$.*

Proof. Point (i) easily follows from Lemma 3.1 if K is compact. To deal with the general case it is enough to notice that for every $x \in \mathbb{R}^d$ and every radius $R > \mathrm{dist}_K(x)$, we have $g_K = g_{K \cap \overline{B_R(x)}}$ in a neighborhood of x. The formula for $\partial f_K(x)$ is a consequence of the rule for the subdifferential of the sum of two functions, one of which smooth. Then, by definition, $\nabla f_K(x)$ is the projection of 0 on $\partial f_K(x) = \mathrm{conv}(\mathrm{opt}_K(x)) - x$, and formula (3.3) follows after a translation of x. Let us now address point (iii). Suppose that x and y share the same optimality class, $\mathrm{opt}_K(x) = \mathrm{opt}_K(y)$. Consider the affine space A spanned by $\mathrm{opt}_K(x)$ and its orthogonal space B passing through x. From the hypothesis on x and y we deduce that also y belongs to B. Then, denoting by p the point of intersection of A and B, by orthogonality we have:

$$|x - z|^2 = |x - p|^2 + |p - z|^2,$$
$$|y - z|^2 = |y - p|^2 + |p - z|^2 \quad \text{for every} \quad z \in \mathrm{conv}(\mathrm{opt}_K(x)).$$

Hence, both distances are minimized by the point z obtained by projecting p on $\mathrm{conv}(\mathrm{opt}_K(x))$, thus $\eta_K(x) = \eta_K(y)$. □

Remark 3.3. From (3.3) we deduce that the potential term $|\nabla f_K(x)|^2$ is always less than or equal to $-2 f_K(x) = \mathrm{dist}_K^2(x)$, and equality holds if and only if x has a unique projection on K. It is interesting to see what this means for the MAG model. Using the notation of Sect. 2.2, the potential of a configuration $(y_1, \ldots, y_m) \in \mathbb{R}^{nm}$ is always smaller than $W_2(\mu, \lambda)$, where, setting $x_i = \pi(y_i)$,

$$\lambda = \frac{1}{m} \sum_{i=1}^{m} \delta_{a_i}, \quad \mu = \frac{1}{m} \sum_{i=1}^{m} \delta_{x_i}$$

and W_2 is the Wasserstein distance in $\mathscr{P}_2(\mathbb{T}^n)$. Moreover, equality holds if and only if there exists a unique optimal transport map from μ to λ. So we see that in the context of the MAG model, the potential term should be interpreted as a "measure of the ambiguity in the optimal transport problem". Tipical manifestations of ambiguity in the discrete scenario appear when two or more particles collapse, thus sharing the same position in $\mathbb{T}^n$. Compare also this phenomenon with the continuous framework of Theorem 2.2, where this ambiguity does not occur.

To end this part, we briefly come back to the matter of stability in this specialized context, stating the following Corollary of Theorem 2.1:

Corollary 3.4. *Let K_j, K be closed subsets of $\mathbb{R}^d$, and let $x_{j,0}, x_{j,\delta}, x_0, x_\delta \in \mathbb{R}^d$. Suppose that*

(i) *$K_j \to K$ in the sense of Hausdorff in every compact set.*
(ii) *$x_{j,i} \to x_i$ for $i = 0, \delta$.*

Then $I^\delta_{f_{K_j}, x_{j,0}, x_{j,\delta}}$ Γ-converge to $I^\delta_{f_K, x_0, x_\delta}$ in the $C([0, \delta], \mathbb{R}^d)$ topology. Moreover, we have convergence of minimal values to minimal values and of minimizers to minimizers.

As a consequence, the functional associated to a closed set K can be approximated by functionals associated to K_j, where each K_j is a finite collection of points in $\mathbb{R}^d$. It is the scope of the following subsection to focus on this simpler situation.

3.1. The discrete case

From now on, we restrict our analysis to the case in which K is given by a collection of N distinct points in $\mathbb{R}^d$:

$$K = \{p_1, \ldots, p_N\}.$$

We are particularly interested in studying properties of local minimizers for $I^\delta_{f_K, x_0, x_\delta}$ because of the link with the variational reformulation of the discrete MAG model (see the discussion above). There, K was an infinite discrete set, but we can clearly restrict our analysis, which is essentially local, to the case in which K is finite, due to the compactness of the range of every continuous curve $\gamma : [0, \delta] \to \mathbb{R}^d$. Let us fix K, so as to be allowed to omit all the pedices involving it. Then, for instance, we will write f, g, η, opt in the place of $f_K, g_K, \eta_K, \text{opt}_K$.

Polyhedra, Voronoi cells and potential zones. We say that $P \subseteq \mathbb{R}^d$ is a *polyhedron* if it is a non-empty closed convex set admitting a representation of the form

$$P = \bigcap_{j=1}^{\ell} \left\{x \in \mathbb{R}^d : T_j(x) \leq 0\right\}, \tag{3.4}$$

where $\ell \in \mathbb{N}$ and $T_j : \mathbb{R}^d \to \mathbb{R}$ are affine functions. A bounded polyhedron is called a *polytope*.

The *Voronoi partition* associated to a finite collection of points $K = \{p_1, \ldots, p_N\}$ is the finite decomposition $\{V_H\}_{H \in \mathcal{P}(K)}$ of $\mathbb{R}^d$, indexed by the set $\mathcal{P}(K)$ of the parts of K, and such that

$$V_H = \left\{x \in \mathbb{R}^d : \text{opt}(x) = H\right\}.$$

We call V_H the *Voronoi cell* corresponding to the *optimality class H*. The following are well-known facts about this remarkable cellular decomposition of the space (see [6]).

Proposition 3.5. (Properties of the Voronoi partition) *Let $H \in \mathcal{P}(K)$ be such that the Voronoi cell V_H is non-empty. Then*

(i) *V_H is a convex set. Moreover, denoting by A_H the affine space spanned by H, and by B_H the affine space*

$$B_H := \left\{x \in \mathbb{R}^d : |x - p_i| = |x - p_j|, \ \forall p_i, p_j \in H\right\},$$

we have that A_H is orthogonal to B_H, they have complementary dimensions in $\mathbb{R}^d$, and V_H is relatively open in B_H. We call p_H the unique intersection point of A_H and B_H.

(ii) *The closure $\overline{V_H}$ is a polyhedron, whose relative boundary in B_H is precisely given by the disjoint union of all the Voronoi cells V_L with $H \subsetneq L$.*

In order to introduce the second fundamental decomposition associated to K, we also need to define, for every $\eta \in \mathbb{R}^d$, the sets:

$$Q_\eta := \left\{ x \in \mathbb{R}^d : \eta(x) = \eta \right\},$$

$$P_\eta := \left\{ x \in \mathbb{R}^d : \eta \in \partial g(x) \right\}.$$

The following proposition encodes the underlying geometrical structure conferred to our variational problem by the particular choice we made for the potential. It will be of fundamental importance in deriving regularity results for local minimizers of the functional $I^\delta_{f_K, x_0, x_\delta}$.

Proposition 3.6. (Voronoi cells and potential zones) *The following facts hold:*

(i) *The map η is constant in each Voronoi cell, and hence has a finite range, that we denote by $\mathcal{E}$.*

(ii) *$\left\{ Q_\eta \right\}_{\eta \in \mathcal{E}}$ is a partition of $\mathbb{R}^d$, and $x \in Q_\eta$ if and only if $\nabla f(x) = \eta - x$. In the sequel we will refer to the Q_η's as potential zones.*

(iii) *For every $\eta \in \mathcal{E}$, both Q_η and P_η are union of Voronoi cells and it holds $Q_\eta \subseteq P_\eta$.*

(iv) *For every $\eta \in \mathcal{E}$, P_η is a polyhedron.*

(v) *Let β be the positive constant defined by*

$$\beta := \min_{\substack{\eta \neq \bar\eta \\ \eta, \bar\eta \in \mathcal{E}}} |\eta - \bar\eta|^2. \tag{3.5}$$

Then we have

$$|\bar\eta - x|^2 \geq |\eta - x|^2 + \beta \quad \text{for every distinct } \eta, \bar\eta \in \mathcal{E} \text{ and } x \in Q_\eta \cap P_{\bar\eta}. \tag{3.6}$$

Proof. Point (i), (ii) and (iii) are direct consequences of Proposition 3.2. To prove point (iv) it is enough to notice that P_η is a closed convex set that can be written as the union of a finite number of polyhedra (the closures of the Voronoi cells contained in P_η). Finally, point (v) easily follows from the fact that η is the projection of x on the closed convex set $\partial g(x)$ containing $\bar\eta$. □

So we see that K carries two partitions of $\mathbb{R}^d$, one finer than the other: the first into Voronoi cells and the second into potential zones. Simple examples show that for a general K they do not coincide (see for instance Example 3.14 hereafter). If they coincide, we say that K is *balanced*. In such a case, the map η defines a bijection between Voronoi cells and potential zones, that is to say,

$$\eta(x) = \eta(y) \iff \text{opt}(x) = \text{opt}(y).$$

Clearly, a sufficient condition for K to be balanced is given by

$$\text{opt}(\eta(x)) = \text{opt}(x) \quad \text{for every } x \in \mathbb{R}^d. \tag{3.7}$$

It is worth noting that in dimension $d = 1$ every K is balanced, and that the same is true in any dimension for cubic lattices.

Conserved quantities. In this paragraph we underline the presence of some conserved quantities for local minimizers of our variational problem. They naturally arise by testing the local minimality against variations along some specific directions. The following Lemma collects two crucial observations in order to suitably perform such variations:

Lemma 3.7. (Local properties of the Voronoi diagram) *The following facts hold:*

(i) *If $x \in V_H$, for some $H \subseteq K$, then there exists a neighborhood U of x such that $\mathrm{opt}(y) \subseteq H$, for every $y \in U$.*

(ii) *If $x \in V_H$, then $x + \epsilon v \in V_H$, provided that the vector v is parallel to B_H, and ϵ is sufficiently small.*

Proof. Point (i) follows from the fact that a point of K which is not optimal for x is neither optimal for y, provided that y is chosen close enough to x. Point (ii) is instead a direct consequence of the fact that V_H is relatively open in the affine space B_H. □

By using very classical variational arguments as well as Lemma 3.7 we derive the following

Proposition 3.8. (Conservation laws) *Let γ be a local minimizer of $I^{\delta}_{f_K, x_0, x_\delta}$. Then*

(i) *(Conservation of the energy). There exists a constant $c \in \mathbb{R}$ such that*

$$|\dot{\gamma}|^2 - h(|\nabla f(\gamma)|^2) = c \quad a.e. \text{ in } (0, \delta).$$

In particular, γ is Lipschitz continuous.

(ii) *(Local conservation of momentum). Let $(t_1, t_2) \subseteq [0, \delta]$ be a time interval. Suppose that there exists an optimality class $H \subseteq K$, with $V_H \neq \emptyset$, such that for every $s \in (t_1, t_2)$ the inclusion $\mathrm{opt}(\gamma(s)) \subseteq H$ holds. Let γ_H be the curve obtained by projecting γ on the affine space B_H. Then γ_H is a $C^{1,1}$ curve in (t_1, t_2) and, in this interval, it satisfies*

$$\ddot{\gamma}_H = h'(|\gamma - \eta(\gamma)|^2)(\gamma_H - p_H).$$

In particular, for each time $t \in (0, \delta)$, denoting $H = \mathrm{opt}(\gamma(t))$, there exists a neighborhood of t in which the component $\dot{\gamma}_H$ of the momentum parallel to B_H is continuous.

Proof. Point (i) states that γ solves the Du Bois–Reymond equation (2.5). This can be shown by testing the local minimality through "horizontal" variations of the independent variable of the form $\gamma_\epsilon = \gamma \circ \rho_\epsilon^{-1}$, where $\rho_\epsilon = \mathrm{id} + \epsilon \varphi$, $\varphi \in C_c^\infty((0, \delta))$, and ϵ is small enough so that ρ_ϵ is a diffeomorphism. The Lipschitz continuity of γ then follows from (3.3) via $|\nabla f_K(x)| \leq \mathrm{dist}_K(x)$ and the continuity of γ. To get point (ii), instead, we need to perform "vertical" variations of the form $\gamma_\epsilon = \gamma + \epsilon \varphi v$, where $\varphi \in C_c^\infty((t_1, t_2))$ and v is any vector parallel to the affine space B_H. We then use point (ii) of Lemma 3.7 to get $\eta(\gamma_\epsilon) = \eta(\gamma)$, for ϵ sufficiently small. □

Point (ii) of the previous proposition implies in particular that γ is regular as long as it stays in a single Voronoi cell. More precisely:

Corollary 3.9. (Regularity inside a Voronoi cell) *Let γ be a local minimizer of $I^\delta_{f_K, x_0, x_\delta}$. Suppose that $\mathrm{opt}(\gamma(s)) = H$ is constant in (t_1, t_2) and let $\eta_H = \eta(V_H)$. Then, in this interval, $\gamma = \gamma_H$, $\eta(\gamma) = \eta_H$ is constant, and γ is a C^2 solution of*

$$\ddot{\gamma} = h'(|\gamma - \eta_H|^2)(\gamma - p_H).$$

Moreover, if h is C^∞, then $\gamma \in C^\infty((t_1, t_2))$.

As a natural consequence, any singularity for a local minimizer of the functional $I^\delta_{f_K, x_0, x_\delta}$ appears only when the optimality class "changes". In the next paragraph we will try to give a more precise meaning to this statement.

Shock times and minimal deviation. Given a curve $\gamma : [0, \delta] \to \mathbb{R}^d$, and a time $t \in [0, \delta]$, we say that

- t is a *shock* for γ if $\mathrm{opt}(\gamma)$ is not constant on I for every neighborhood I of t in $[0, \delta]$.
- t is a *non-degenerate shock* for γ if $\eta(\gamma)$ is not constant on I for every neighborhood I of t in $[0, \delta]$.
- t is an *effective shock* for γ if there are two distinct potential zones Q_η, $Q_{\bar{\eta}}$, and two distinct Voronoi cells $V_H \subseteq Q_\eta$, $V_{\bar{H}} \subseteq Q_{\bar{\eta}}$ such that, for some $\epsilon > 0$, one of the following holds:
  - $\gamma((t - \epsilon, t)) \subset V_H$ and $\gamma([t, t + \epsilon)) \subset V_{\bar{H}}$. In this case, $H \subsetneq \bar{H}$ and we say that t is a *left effective shock*.
  - $\gamma((t - \epsilon, t]) \subset V_H$ and $\gamma((t, t + \epsilon)) \subset V_{\bar{H}}$. In this case, $\bar{H} \subsetneq H$ and we say that t is a *right effective shock*.

We denote by $S(\gamma)$, $NDS(\gamma)$ and $ES(\gamma)$ respectively the sets of shocks, non-degenerate shocks and effective shocks for γ. Notice that $ES(\gamma) \subseteq NDS(\gamma) \subseteq S(\gamma)$, and that $S(\gamma)$ and $NDS(\gamma)$ are compact. According to the definitions above, during a shock there must be a change of Voronoi cell, while, during a non-degenerate shock there is also a change of potential zone. Clearly we have $S(\gamma) = NDS(\gamma)$ provided that K is balanced. Finally, we have an effective shock when a neat passage occurs from a Voronoi cell to an adjacent one with different potential. By conservation of the energy, we expect the dynamics to develop a singularity in the kinetic term here. This is the content of the following proposition, which is a direct consequence of the conservation laws stated in Proposition 3.8.

Proposition 3.10. (Minimal deviation during an effective shock) *Suppose that the potential shape h is strictly increasing. Let γ be a local minimizer of $I^\delta_{f_K, x_0, x_\delta}$. Suppose that $t \in (0, \delta)$ is a left effective shock for γ, in which γ jumps from V_H to $V_{\bar{H}}$. Let $\eta_H = \eta(V_H)$ and $\eta_{\bar{H}} = \eta(V_{\bar{H}})$. Defined $\dot{\gamma}_-(t)$ and $\dot{\gamma}_+(t)$ respectively by*

$$\dot{\gamma}_-(t) := \lim_{s \to t^-} \dot{\gamma}(s) \qquad \dot{\gamma}_+(t) := \lim_{s \to t^+} \dot{\gamma}(s),$$

then $\dot{\gamma}_+(t)$ is the component of $\dot{\gamma}_-(t)$ parallel to $B_{\bar{H}}$. Moreover

$$|\dot{\gamma}_-(t) - \dot{\gamma}_+(t)|^2 = |\dot{\gamma}_-(t)|^2 - |\dot{\gamma}_+(t)|^2 = h(|\gamma(t) - \eta_H|^2) - h(|\gamma(t) - \eta_{\bar{H}}|^2) > 0$$

Clearly an analogous result holds for right effective shocks.

Remark 3.11. The fact that $\dot\gamma_+(t)$ is the component of $\dot\gamma_-(t)$ parallel to $B_{\bar H}$ has the following interpretation: in the dynamics, when there is a neat passage from a given Voronoi cell to an adjacent one (hence lower-dimensional), the component of the momentum $\dot\gamma$ parallel to the second cell is continuous, while the orthogonal one has a jump. In the specific case of a superadditive shape h, (for instance when $h = $ id, as in the MAG model), we can also derive a uniform lower bound for such a jump:

$$|\dot\gamma_-(t) - \dot\gamma_+(t)|^2 = h(|\gamma(t) - \eta_H|^2) - h(|\gamma(t) - \eta_{\bar H}|^2) \geq h(|\eta_H - \eta_{\bar H}|^2) \geq h(\beta),$$

where β was defined in (3.5).

Remark 3.12. In the MAG dynamics, shocks tipically happen when two or more particles collide or separate, generating an instant change in the optimality class. For instance an effective shock occurs when two particles collide and remain sticked together. Notice that Proposition 3.8 tells us that energy and momentum are conserved in a collision.

To end this part, we show, through a couple of simple examples, that all of the three types of shock times defined in this paragraph may occur for a minimizer of $I^\delta_{f_K, x_0, x_\delta}$.

Example 3.13. (Effective and non-effective shocks) Take $h = $ id, $\delta = 1$, $d = 1$, $K = \{-1, 1\}$, $x_0 = -c$, $x_1 = c$, with $c \geq 0$. We see that $\mathbb{R}$ is partitioned into three Voronoi cells, namely: $V_{\{-1\}} = (-\infty, 0)$, $V_{\{1\}} = (0, +\infty)$ and $V_{\{-1,1\}} = \{0\}$. The potential term takes the form

$$|\nabla f(x)|^2 = \begin{cases} |x + 1|^2 & \text{if} \quad x \in (-\infty, 0), \\ 0 & \text{if} \quad x = 0, \\ |x - 1|^2 & \text{if} \quad x \in (0, +\infty). \end{cases}$$

It is easily seen that a minimizer γ of $I^\delta_{f_K, x_0, x_\delta}$ has to be non-decreasing, and thus $\gamma^{-1}(0) = [t_0, t_1]$ is a closed interval (possibly degenerate if $t_0 = t_1$), and $S(\gamma) = NDS(\gamma) = \{t_0, t_1\}$. Then, there are two possible qualitatively different behaviors of γ, according to whether $t_0 = t_1$ or not. If $t_0 = t_1$, then γ has a single non-degenerate non-effective shock time. If instead $t_0 \neq t_1$, then t_0 and t_1 are respectively left and right effective shocks. Now, direct computations show that the first case occurs if for instance $c = 1$. On the other hand, we can prove that the second case necessarily occurs if c is chosen sufficiently small. As a matter of fact, for $c < 1$, by the monotonicity of γ, the minimum value $\Gamma(-c, c)$ of the functional can be bounded below as follows:

$$\Gamma(-c, c) \geq (1 - t_1 + t_0)(1 - c)^2.$$

At the same time, by Corollary 3.4, we have

$$\Gamma(-c, c) \to \Gamma(0, 0) = 0, \quad \text{as } c \to 0^+.$$

Therefore, $t_1 > t_0$, provided that we choose c small enough.

Example 3.14. (Degenerate shocks) Take $h = \mathrm{id}$, $\delta = 1$, $d = 2$, $K = \{(1,0), (0,1), (-1,0)\}$, $x_0 = (0,-1)$, $x_1 = (0,0)$. Here we notice that the two distinct Voronoi cells $V_{\{(1,0),(0,1),(-1,0)\}} = \{(0,0)\}$ and $V_{\{(1,0),(-1,0)\}} = \{0\} \times (-\infty, 0)$ share the same potential zone $Q_\eta = \{0\} \times (-\infty, 0]$ with $\eta = (0,0)$. We claim that every minimizer γ of $I^\delta_{f_K, x_0, x_\delta}$ must live in Q_η. As a consequence, we would deduce that the first time τ at which $\gamma(\tau) = (0,0)$ has to be a degenerate shock for γ. To prove the claim we test the minimality of γ against the competitor $\delta(t) = \pi(\gamma(t))$, where π denotes the projection on Q_η. It can be easily checked via $\pi(K) = \{\eta\}$ and the contractivity of π that, for every t,

$$|\nabla f(\delta(t))| = |\eta(\delta(t)) - \delta(t)| = |\eta - \delta(t)|$$
$$= |\pi(\eta(\gamma(t))) - \pi(\gamma(t))| \le |\eta(\gamma(t)) - \gamma(t)| = |\nabla f(\gamma(t))|.$$

Moreover, again by the contractivity of γ: $|\dot\delta| \le |\dot\gamma|$ a.e., with equality if and only if $\gamma = \delta$. Comparing the actions of γ and δ and imposing the minimality of γ, we see that necessarily $\gamma = \delta$, that is to say γ lives in Q_η.

3.2. Regularity results

Here we state and prove our main regularity results. Recall that $K = \{p_1, \ldots, p_N\}$ is a finite collection of points in $\mathbb{R}^d$, and $h : [0, +\infty) \to [0, +\infty)$ is a C^1 potential shape.

Theorem 3.15. ($C^{1,1}$-regularity out of a finite number of non-degenerate shock times) *Let γ be a local minimizer of $I^\delta_{f_K, x_0, x_\delta}$. Suppose that h is strictly increasing. Then:*

(i) *The set $NDS(\gamma)$ of non-degenerate shock times of γ is finite. That is to say, there is a finite number of times $0 \le t_1 < \cdots < t_\ell \le \delta$ such that $\eta(\gamma)$ is constant in each connected component of $[0, \delta] \setminus \{t_1, \ldots, t_\ell\}$.*

(ii) *Setting $t_0 = 0$ and $t_{\ell+1} = \delta$, then γ is $C^{1,1}$-regular in the interval $[t_i, t_{i+1}]$ for every $i \in \{0, \ldots, \ell\}$. Moreover, if we let $\eta_i \in \mathcal{E}$ be such that $\gamma((t_i, t_{i+1})) \subseteq Q_{\eta_i}$, then we can estimate*

$$|\ddot\gamma(r)| \le h'(|\gamma(r) - \eta_i|^2)|\gamma(r) - \eta_i| \quad \text{for a.e.} \quad r \in (t_i, t_{i+1}).$$

Actually, if K is balanced and h is smooth, Theorem 3.15 can be improved to reach piecewise smooth regularity for any local minimizer γ. In fact, since K is balanced, the equality $NDS(\gamma) = S(\gamma)$ holds and point i) of Theorem 3.15 implies that γ has a finite number of shock times. On the other hand, Corollary 3.9 together with the smoothness of h ensures that γ is smooth in each connected component of $[0, \delta] \setminus S(\gamma)$, where clearly $\mathrm{opt}(\gamma(t))$ is constant.

Corollary 3.16. (C^∞-regularity out of a finite number of shock times) *Let γ be a local minimizer of $I^\delta_{f_K, x_0, x_\delta}$. Suppose that K is balanced and h is C^∞ and strictly increasing. Then:*

(i) *The set $S(\gamma)$ of shock times of γ is finite. That is to say, there is a finite number of times $0 \leq t_1 < \cdots < t_\ell \leq \delta$ such that the optimality class $\mathrm{opt}(\gamma)$ is constant in each connected component of $[0, \delta] \setminus \{t_1, \ldots, t_\ell\}$.*

(ii) *Setting $t_0 = 0$ and $t_{\ell+1} = \delta$, then γ is C^∞ in the interval $[t_i, t_{i+1}]$ for every $i \in \{0, \ldots, \ell\}$. Moreover, if $t_{i+1} > t_i$, denoting by H_i the optimality class of γ in the interval (t_i, t_{i+1}), and defining $\eta_{H_i} = \eta(V_{H_i})$, then γ solves*

$$\ddot{\gamma} = h'(|\gamma - \eta_{H_i}|^2)(\gamma - p_{H_i})$$

in (t_i, t_{i+1}).

Remark 3.17. Corollary 3.16 offers a generalization of [2, Theorem 13], in which smooth regularity out of a finite number of shock times was derived for solutions of a one-dimensional version of the discrete MAG model. In their framework, h was simply the identity, while K consisted of all the $m!$ points of $\mathbb{R}^d$ obtainable by permuting the components of a fixed vector $A = (a_1, \ldots, a_d) \in \mathbb{R}^d$, with $a_1 < \cdots < a_d$. Exploiting the rearrangement inequality provided by the order structure of $\mathbb{R}$, it is not difficult to show that hypothesis (3.7) holds in this case, therefore implying that K is balanced.

The rest of the paper is devoted to the proof of Theorem 3.15, which is organized as follows. We first prove the two Lemmas 3.18 and 3.19. As a byproduct of Lemma 3.19 we deduce that a local minimizer is $C^{1,1}$-regular as long as it stays in a single potential zone (see Corollary 3.20). From this the implication point i) $\Longrightarrow$ point ii) in Theorem 3.15 easily descends. Finally, we prove point i) in the form of the equivalent local statement given by Claim (3.14).

Two Lemmas. The first Lemma has a strong geometric flavour, highlighting a metric property of two given intersecting polytopes. It will be a crucial ingredient in the proof of point i) of Theorem 3.15.

Lemma 3.18. (Reciprocal distance of intersecting polytopes) *Let $A, B \subset \mathbb{R}^d$ be two polytopes with $A \cap B \neq \emptyset$. Then there exists a sufficiently large constant $M > 0$ such that*

$$\mathrm{dist}_{A \cap B}(x) \leq M \, \mathrm{dist}_B(x) \quad \text{for every } x \in A.$$

Proof. Let $P \subset \mathbb{R}^d$ be a polytope endowed with a representation of the form (3.4), where we assume without loss of generality that each of the affine functions T_j has Lipschitz constant equal to 1. Let us consider the associated vector-valued function $z_P : \mathbb{R}^d \to [0, +\infty)^\ell$ defined by

$$z_P(x) = \left(T_1(x)^+, \ldots, T_\ell(x)^+\right). \tag{3.8}$$

As a first step we show that there exists a constant $c_P > 0$ such that

$$T_j(x)^+ \leq \mathrm{dist}_P(x) \leq c_P |z_P(x)|_1 \quad \text{for every} \quad j \in \{1, \ldots, \ell\} \quad \text{and every} \quad x \in \mathbb{R}^d. \tag{3.9}$$

(recall that the 1-norm of a vector $z = (z_1, \ldots, z_\ell) \in \mathbb{R}^\ell$ is defined as $|z|_1 := |z_1| + \cdots + |z_\ell|$). The left inequality follows by the 1-Lipschitz assumption on

T_j, while the right one can be obtained using a compactness argument as follows. Suppose by contradiction that there exists a sequence of points $x_n \in \mathbb{R}^d \setminus P$ such that

$$\text{dist}_P(x_n) \geq n|z_P(x_n)|_1.$$

Up to replacing x_n with

$$\pi_P(x_n) + \frac{x_n - \pi_P(x_n)}{\text{dist}_P(x_n)},$$

we can assume that $\text{dist}_P(x_n) = 1$. Now, if x is any cluster point for x_n, we obtain that $\text{dist}_P(x) = 1$ and $z_P(x) = 0$, which is clearly a contradiction. We are now in the position to prove the Lemma. Let $\ell \in \mathbb{N}$ be the number of affine functions in a representation of B of the form (3.4), with 1-Lipschitz affine functions T_j. By exploiting the estimates in (3.9), for every $x \in A$ we can bound

$$\text{dist}_{A\cap B}(x) \leq c_{A\cap B}|z_{A\cap B}(x)|_1 \leq \ell c_{A\cap B}\,\text{dist}_B(x),$$

where $c_{A\cap B}$ is a positive constant. The thesis follows by choosing $M = \ell c_{A\cap B}$. $\square$

The second Lemma, of independent interest, concerns the regularity of local minimizers of action functionals restricted to curves living in a given closed convex set. Because of boundary effects, such constrained minimizers are in general not C^2, even if the Lagrangian is smooth. Nevertheless, they must be at least $C^{1,1}$, whenever the Lagrangian is C^1.

Lemma 3.19. (Regularity for a problem with a convex constraint) *Let $P \subseteq \mathbb{R}^d$ be a closed convex set, and let $\Psi : \mathbb{R}^d \to [0, +\infty)$ be of class C^1 in a neighborhood of P. Given two points $x_0, x_\delta \in P$, we consider the functional $\mathcal{G} : C([0, \delta], \mathbb{R}^d) \to [0, +\infty]$ defined by*

$$\mathcal{G}(\gamma) = \begin{cases} \int_0^\delta |\dot\gamma|^2 + \Psi(\gamma) & \text{if } \gamma \in AC([0, \delta], \mathbb{R}^d), \gamma(0) = x_0, \gamma(\delta) = x_\delta, \gamma([0, \delta]) \subseteq P, \\ +\infty & \text{otherwise.} \end{cases}$$

Then every local minimizer γ of $\mathcal{G}$ is of class $C^{1,1}$ and we can estimate

$$|\ddot\gamma(r)| \leq \frac{1}{2}|\nabla\Psi(\gamma(r))| \quad \text{for a.e. } r \in (0, \delta).$$

Proof. We call π_P the projection on P. We start by defining, for each point $x \in P$, the "blow up" of P at x, namely the closed convex cone P_x defined by the formula

$$P_x + x = \begin{cases} \mathbb{R}^d & \text{if } x \in \mathring{P}, \\ \bigcap\{\mathcal{H} \text{ half space} : P \subset \mathcal{H}, x \in \partial\mathcal{H}\} & \text{if } x \in \partial P. \end{cases}$$

We then call S_x the projection on P_x, which turns out to be positive homogeneous and 1-Lipschitz. It is not difficult to realize that

$$\frac{P - x}{\epsilon} \xrightarrow{\epsilon \to 0^+} P_x \tag{3.10}$$

in the sense of Hausdorff on every compact set. Since the inclusion $(P - x)/\epsilon \subseteq P_x$ always holds, in order to show the convergence (3.10), we only need to check, using a separation argument, that, for every $\epsilon_j \to 0^+$, and every point $z \in P_x$, there exists a sequence $y_j \in (P - x)/\epsilon_j$ converging to z. Hence, the projection on $(P - x)/\epsilon$ pointwise converges to S_x as $\epsilon \to 0^+$. By exploiting also the homogeneity of S_x we eventually obtain that, for every $v \in \mathbb{R}^d$,

$$\pi_P(x + \epsilon v) = x + \epsilon S_x(v) + o(\epsilon), \quad \text{as } \epsilon \to 0^+. \tag{3.11}$$

Now, let γ be a local minimizer for the functional $\mathcal{G}$. Given a test function $\varphi \in C_c^\infty((0, \delta); \mathbb{R}^d)$, we consider the following competitors:

$$\gamma_\epsilon := \pi_P(\gamma + \epsilon \varphi), \quad \text{for } \epsilon > 0.$$

Since π_P is 1-Lipschitz, we have $|\gamma_\epsilon'| \leq |\gamma' + \epsilon \varphi'|$. Then

$$\mathcal{G}(\gamma_\epsilon) \leq \mathcal{G}(\gamma) + 2\epsilon \int_0^\delta \gamma' \cdot \varphi' + \frac{\Psi(\gamma_\epsilon) - \Psi(\gamma)}{2\epsilon} + \epsilon^2 \int_0^\delta |\varphi'|^2.$$

Using the previous pointwise expansion (3.11), as well as the dominated convergence Theorem, we obtain

$$\limsup_{\epsilon \to 0^+} \frac{\mathcal{G}(\gamma_\epsilon) - \mathcal{G}(\gamma)}{2\epsilon} \leq \int_0^\delta \gamma' \cdot \varphi' + \frac{1}{2} \nabla \Psi(\gamma) \cdot S_\gamma(\varphi). \tag{3.12}$$

From the local minimality of γ, it follows immediately that the right-hand side in (3.12) is non-negative. Finally, we can use the contractivity of the projections S_x to get the inequality

$$\langle \gamma'', \varphi \rangle \leq \int_0^\delta \frac{1}{2} |\nabla \Psi(\gamma)| |\varphi| \quad \text{for every} \quad \varphi \in C_c^\infty((0, \delta); \mathbb{R}^d),$$

and the thesis follows. $\qquad\square$

To conclude this paragraph we propose the following easy consequence of Lemma 3.19.

Corollary 3.20. (Regularity in a potential zone) *Let h be a non-decreasing potential shape, and let γ be a local minimizer of $I_{f_K, x_0, x_\delta}^\delta$. Suppose that there exists an $\eta \in \mathcal{E}$ such that $\gamma((s, t)) \subseteq Q_\eta$, where $(s, t) \subset [0, \delta]$. Then γ is $C^{1,1}$-regular on $[s, t]$ and the following estimate holds*

$$|\ddot{\gamma}(r)| \leq h'(|\gamma(r) - \eta|^2)|\gamma(r) - \eta| \quad \text{for a.e. } r \in (s, t).$$

Proof. We clearly have the inclusion $\gamma([s, t]) \subset P_\eta$. Then we observe that, for every absolutely continuous curve $\rho : [s, t] \to P_\eta$, due to the monotonicity of h and inequality (3.6), the following inequality holds:

$$\int_s^t |\dot{\rho}|^2 + h(|\nabla f|^2(\rho)) \leq \int_s^t |\dot{\rho}|^2 + h(|\eta - \rho|^2). \tag{3.13}$$

By a comparison argument, knowing that γ is a local minimizer of the functional on the left-hand side of (3.13), and that $\rho = \gamma$ saturates the inequality, we deduce that γ is a local minimizer for the functional on the right-hand side of (3.13), if one restricts to curves living in the closed convex set P_η. At this point the proof is easily concluded by invoking Lemma 3.19. $\qquad\square$

Proof of the regularity result. In this paragraph we prove our main regularity result, Theorem 3.15. Point (ii) is immediately derived from point (i) via Corollary 3.20. Thus, only point (i) remains to be proved. To do so, we can restrict to prove the following equivalent local statement:

Claim : for each time $t \in (0, \delta]$,

there exists $\epsilon > 0$ such that $\eta(\gamma)$ is constant in $(t - \epsilon, t)$. (3.14)

In fact, from Claim (3.14) and the analogous one for right intervals (obtainable by exploiting the autonomicity of the functional), it clearly descends that $NDS(\gamma)$ is discrete. Then, by compactness, we derive that $NDS(\gamma)$ is finite. The proof of Claim (3.14) will be elementary accomplished by a quite intricated series of "cut and paste" constructions of competitors. Let us first outline the general heuristic idea before entering in the details of the rigorous proof in the forthcoming pages. We will divide the proof in a few steps:

Step 1. Suppose by contradiction that t is a cluster point for the "jumps" in the potential. Then, approaching t from the left, γ would infinitely often visit high and low potential zones. If only one low potential zone were visited asymptotically, then it would be convenient to stay in it definitely. Therefore we can assume that γ infinitely often visits at least two different low potential zones, approaching t from the left. Moreover, in alternating between different low potential zones, γ necessarily spends a non-negligible amount of time in high potential ones.

Step 2. Approaching t from the left, the percentage of time spent by γ in high potential zones tends to zero, thus enforcing at least two different low potential zones to be very near to each other in order to make it possible for the Lipschitz curve γ to jump from one to the other in a short time.

Step 3. By Lemma 3.18 we will be eventually allowed to slightly deviate γ to reach an even lower potential zone (the interface between the two), hence contradicting the local optimality of γ and reaching the desired absurdum.

Proof of Claim (3.14). Throughout the proof, γ will be a fixed local minimizer of $I^{\delta}_{f_K, x_0, x_\delta}$ and L the Lipschitz constant of γ. Moreover, for the sake of simplicity, we indicate $I^{\delta}_{f_K, x_0, x_\delta}$ as $\mathcal{F}$. As a preliminary observation, notice that we can choose an $R > 0$ large enough such that $\gamma([0, \delta]) \subset B_R$, and then, by point (v) in Proposition 3.6, we can find an $\alpha > 0$ small enough such that

$$|\bar{\eta} - x| \geq |\eta - x| + \alpha \quad \text{for every } x \in B_{2R} \cap Q_\eta \cap P_{\bar{\eta}} \text{ and every distinct } \eta, \bar{\eta} \in \mathcal{E}. \tag{3.15}$$

By translation invariance, we can assume without loss of generality that $\gamma(t) = 0$, thus simplifying some further computations. Then we consider the asymptotic lowest potential threshold

$$a := \liminf_{s \to t^-} |\eta(\gamma(s))|.$$

We call $\tilde{\mathcal{E}}$ the subset of $\mathcal{E}$ indexing the potential zones that are visited infinitely often before t. Namely,

$$\tilde{\mathcal{E}} := \left\{ \eta \in \mathcal{E} : \gamma^{-1}(Q_\eta) \cap (0, t) \text{ accumulates in } t \right\}.$$

Notice that the thesis is equivalent to $\#\tilde{\mathcal{E}} = 1$. The set $\tilde{\mathcal{E}}$ can be partitioned into two subsets $\tilde{\mathcal{E}}_1$ and $\tilde{\mathcal{E}}_2$ defined by

$$\tilde{\mathcal{E}}_1 := \left\{ \eta \in \tilde{\mathcal{E}} : |\eta| = a \right\},$$
$$\tilde{\mathcal{E}}_2 := \left\{ \eta \in \tilde{\mathcal{E}} : |\eta| > a \right\}.$$

The set $\tilde{\mathcal{E}}_1$, which is clearly non-empty by the definition of a, corresponds to those potential zones which are infinitely often visited by γ before t and that share the asymptotic lowest potential threshold a. We expect the curve γ to spend most of the time there.

We make the following technical choices of constants in order to simplify later arguments. We fix $r, \mu, \epsilon > 0$ small enough so that these requirements are satisfied:

R1) For every $x \in B_r$ we have $\text{opt}(x) \subseteq \text{opt}(0)$.
R2) $\gamma((t - \epsilon, t)) \subset B_r$.
R3) For every $x \in B_r$, and for every $\eta \in \mathcal{E}$, we have $|\eta| - \mu < |x - \eta| < |\eta| + \mu$.
R4) $3\mu \le \alpha$.
R5) For every choice of $\eta, \bar{\eta} \in \mathcal{E}$, exactly one of the following holds:

$$|\eta| = |\bar{\eta}|, \quad \min\{|\eta|, |\bar{\eta}|\} \le \max\{|\eta|, |\bar{\eta}|\} - 3\mu.$$

R6) $(t - \epsilon, t) \subseteq \bigcup_{\eta \in \tilde{\mathcal{E}}} \gamma^{-1}(Q_\eta)$.

Notice that thanks to R1), for every $x \in B_r$ the segment $[x, 0)$ will be entirely contained in the Voronoi cell $V_{\text{opt}(x)}$, while R2) assures us that the curve belongs to this good area. Conditions R3), R4) and R5) will be useful to effectively distinguish between different potential zones. Finally, condition R6) implies that the potential zones touched by γ in the interval $(t - \epsilon, t)$ are exactly those touched asymptotically. Thus, in particular, we have

$$a = \min_{s \in (t - \epsilon, t)} |\eta(\gamma(s))|.$$

By condition R6), the interval $(t - \epsilon, t)$ can be partitioned into the following two sets:

$$C_1 := (t - \epsilon, t) \cap \bigcup_{\eta \in \tilde{\mathcal{E}}_1} \gamma^{-1}(Q_\eta),$$
$$C_2 := (t - \epsilon, t) \cap \bigcup_{\eta \in \tilde{\mathcal{E}}_2} \gamma^{-1}(Q_\eta).$$

We observe that C_1 is closed in $(t - \epsilon, t)$. In fact, if $s_j \in C_1$ and $s_j \to s_\infty \in (t - \epsilon, t)$, then by lower semicontinuity of the modulus of the extended gradient, we have $|\eta(\gamma(s_\infty))| \le a + \mu$, whence $|\eta(\gamma(s_\infty))| \le a + 2\mu$, which implies $|\eta(\gamma(s_\infty))| = a$, that is $s_\infty \in C_1$. Then C_2 is open and can be written as a finite or countable (possibly empty) union of open intervals.

Remember that the thesis is equivalent to $\#\tilde{\mathcal{E}} = 1$. We now assume that $\#\tilde{\mathcal{E}} \geq 2$ and try to find a contradiction.

Step 1. (Reduction to the case in which $\#\tilde{\mathcal{E}}_2 \geq 1$ and $\#\tilde{\mathcal{E}}_1 \geq 2$).
We first show that $\#\tilde{\mathcal{E}}_2 \geq 1$. If this were not the case, then we would have $|\eta(\gamma(s))| = a$ for every $s \in (t - \epsilon, t)$ and $\#\tilde{\mathcal{E}}_1 \geq 2$. Then we could find two distinct $\eta, \bar{\eta}$ both belonging to $\tilde{\mathcal{E}}_1$ and a time $s \in (t - \epsilon, t)$ such that $x := \gamma(s) \in Q_\eta \cap P_{\bar{\eta}}$. Now, by (3.15), this would imply that

$$|x - \eta| \leq |x - \bar{\eta}| - \alpha,$$

and we reach a contradiction through the following chain of inequalities:

$$a = |\eta| \qquad \leq |x - \eta| + \mu \leq |x - \bar{\eta}| - \alpha + \mu$$
$$\leq |\bar{\eta}| + 2\mu - \alpha = a + 2\mu - \alpha \leq a - \mu < a.$$

We now show that $\#\tilde{\mathcal{E}}_1 \geq 2$. Suppose by contradiction that $\tilde{\mathcal{E}}_1 = \{\eta\}$. Then we can build a better competitor $\tilde{\gamma}$ by performing arbitrarily small perturbations of γ in the following way. We choose $s \in (t - \epsilon, t)$ as close to t as we want, such that $\gamma(s) \in Q_\eta$. Then we modify γ only in the interval $[s, t]$, by replacing it with its projection on the closed convex set P_η. Namely, denoting by π_η the projection on P_η, we define

$$\tilde{\gamma}(u) = \begin{cases} \gamma(u) & \text{for} \quad u < s \quad \text{or} \quad u > t, \\ \pi_\eta(\gamma(u)) & \text{for} \quad u \in [s, t]. \end{cases}$$

Now, we clearly have $|\dot{\tilde{\gamma}}| \leq |\dot{\gamma}|$, by the 1-Lipschitz property of π_η. On the other hand, for what concerns the potential, in the interval $[s, t]$ it holds

$$|\nabla f(\gamma(u))| = |\nabla f(\tilde{\gamma}(u))| \quad \text{if} \quad u \in C_1,$$
$$|\nabla f(\gamma(u))| > |\nabla f(\tilde{\gamma}(u))| \quad \text{if} \quad u \in C_2.$$

The second expression easily follows from the fact that if $s \in C_2$, then $|\eta(\gamma(s))| \geq a + 3\mu$. From these estimates, the strict monotonicity of the potential shape h, and the local minimality of γ, we obtain that $C_2 \cap (s, t)$ must be negligible for the one-dimensional Lebesgue measure. But this contradicts the fact that $C_2 \cap (s, t)$ is open and accumulates in t. Thus Step 1 is completed.

To make the point, after Step 1 the situation is the following. There must exist

- Two elements $\eta, \bar{\eta} \in \tilde{\mathcal{E}}_1$, with $\eta \neq \bar{\eta}$;
- Two distinct Voronoi cells V_H and $V_{\bar{H}}$, with $V_H \subseteq Q_\eta$ and $V_{\bar{H}} \subseteq Q_{\bar{\eta}}$;
- A sequence of open intervals $(s_\ell, r_\ell) \subset (t - \epsilon, t)$ accumulating in t and such that, for every ℓ, the following conditions hold:
  - $r_\ell < s_{\ell+1}$;
  - $\gamma(s_\ell) \in V_H$;
  - $\gamma(r_\ell) \in V_{\bar{H}}$;
  - $(s_\ell, r_\ell) \subset C_2$.

Step 2. (We have $(r_\ell - s_\ell) = O((t - s_\ell)^2))$.
This will be handled once again by comparison with a suitably chosen competitor.
We define, for every ℓ, the competitor γ_ℓ, obtained by modifying γ only in the interval $[s_\ell, t]$, where γ is replaced with its projection on the segment $[\gamma(s_\ell), 0]$. That is to say, calling π_ℓ the projection on $[\gamma(s_\ell), 0]$, we set

$$\gamma_\ell(u) = \begin{cases} \gamma(u) & \text{for} \quad u < s_\ell \text{ or } u > t, \\ \pi_\ell(\gamma(u)) & \text{for} \quad u \in [s_\ell, t]. \end{cases}$$

Then $|\dot{\gamma}_\ell| \leq |\dot{\gamma}|$ as before. Regarding the potential part, we have the following estimates, for $u \in (s_\ell, t)$:

$$|\nabla f(\gamma(u))|^2 \geq |\eta(\gamma(u))|^2 - 2|\eta(\gamma(u))||\gamma(u)| - |\gamma(u)|^2$$
$$\geq |\eta(\gamma(u))|^2 - L(2S + r)(t - u),$$
$$|\nabla f(\gamma_\ell(u))|^2 \leq |\eta(\gamma_\ell(u))|^2 + 2|\eta(\gamma_\ell(u))||\gamma_\ell(u)| + |\gamma_\ell(u)|^2$$
$$\leq |\eta(\gamma_\ell(u))|^2 + L(2S + r)(t - u),$$

where we set $S := \max\{|\eta| : \eta \in \tilde{\mathcal{E}}\}$. We now show that $|\eta(\gamma_\ell(u))| \leq a$ for every $u \in (s_\ell, t)$. By the definition of γ_ℓ it is enough to show that $|\eta(x)| \leq a$ for every $x \in [\gamma(s_\ell), 0]$. Remember that, by condition R1), $\text{opt}(\gamma(s_\ell)) \subseteq \text{opt}(0)$ so that $0 \in \overline{V_H} \subseteq P_\eta$. In particular, $[\gamma(s_\ell), 0) \subseteq V_H$, hence $\eta(x) = \eta$ and $|\eta(x)| = a$ for every $x \in [\gamma(s_\ell), 0)$. Moreover, since $0 \in P_\eta$, $|\eta(0)| \leq |\eta| = a$. This justifies the following estimates:

$$|\eta(\gamma(u))|^2 \geq a^2 \geq |\eta(\gamma_\ell(u))|^2 \quad \text{for } u \in (s_\ell, t),$$
$$|\eta(\gamma(u))|^2 \geq a^2 + 9\mu^2 \geq |\eta(\gamma_\ell(u))|^2 + 9\mu^2 \quad \text{for } u \in (s_\ell, r_\ell).$$

After renaming the constant $c_0 := L(2S + r)$, we obtain, for sufficiently large ℓ:

$$0 \geq \mathcal{F}(\gamma) - \mathcal{F}(\gamma_\ell)$$
$$\geq \int_{s_\ell}^{r_\ell} \left\{ h(a^2 + 9\mu^2 - c_0(t - u)) - h(a^2 + c_0(t - u)) \right\}$$
$$+ \int_{r_\ell}^{t} \left\{ h(a^2 - c_0(t - u)) - h(a^2 + c_0(t - u)) \right\}$$
$$\geq c_1(r_\ell - s_\ell) - c_2(t - s_\ell)^2.$$

Here $c_1 := h(a^2 + 8\mu^2) - h(a^2 + \mu^2) > 0$ and ℓ is large enough so that $c_0(t - s_\ell) \leq \mu^2$. The constant c_2 is instead defined as

$$c_2 := c_0 \, \text{Lip}\left(h; [a^2 - \mu^2, a^2 + \mu^2]\right)$$

By defining $c_3 := c_2/c_1$, we eventually get

$$(r_\ell - s_\ell) \leq c_3(t - s_\ell)^2 \quad \text{for } \ell \text{ large enough}.$$

Step 3. (A slight deviation of γ through $\overline{V_H} \cap \overline{V_{\bar{H}}}$ reduces the action, thus contradicting its local minimality).

Notice that $\overline{V_H}$ and $\overline{V_{\bar{H}}}$ are two polyhedra whose intersection contains 0. Possibly replacing them with their intersection with a large d-dimensional cube, we can assume that they are polytopes. Then we can apply Lemma 3.18 to deduce the existence of a constant $M > 0$ and a sequence of points $x_\ell \in \overline{V_H} \cap \overline{V_{\bar{H}}}$ such that

$$|x_\ell - \gamma(s_\ell)| \leq M|\gamma(r_\ell) - \gamma(s_\ell)| \leq MLc_3(t - s_\ell)^2 = c_4(t - s_\ell)^2.$$

We call $\epsilon_\ell := c_4(t - s_\ell)$ and assume that ℓ is large enough so that $\epsilon_\ell \in (0, 1)$. We crucially consider the following competitor:

$$\delta_\ell(u) = \begin{cases} \gamma(u) & \text{for} \quad u < s_\ell \text{ or } u > t, \\ \gamma(s_\ell) + \frac{x_\ell - \gamma(s_\ell)}{\epsilon_\ell(t - s_\ell)}(u - s_\ell) & \text{for} \quad u \in [s_\ell, s_\ell + \epsilon_\ell(t - s_\ell)), \\ x_\ell + \frac{-x_\ell}{(1 - \epsilon_\ell)(t - s_\ell)}(u - s_\ell - \epsilon_\ell(t - s_\ell)) & \text{for} \quad u \in [s_\ell + \epsilon_\ell(t - s_\ell), t]. \end{cases}$$

Notice that in the interval $[s_\ell, t]$, the curve δ_ℓ is simply a piecewise linear modification of γ, going from $\gamma(s_\ell)$ to x_ℓ in time $\epsilon_\ell(t - s_\ell)$, and then from x_ℓ to 0 in the remaining time. We will see that δ_ℓ has strictly less action than γ for ℓ large enough, thus reaching the desired contradiction. We first want to be sure that $\overline{V_H} \cap \overline{V_{\bar{H}}}$ is a very low potential zone, so that we can lower the action of γ by a slight deviation through it. This can be seen as follows. For sure η and $\bar{\eta}$ both belong to $\partial g(x_\ell)$, thus also $\frac{\eta + \bar{\eta}}{2} \in \partial g(x_\ell)$. But then

$$|\eta(x_\ell)| \leq |\eta(x_\ell) - x_\ell| + |x_\ell| \leq \left|\frac{\eta + \bar{\eta}}{2} - x_\ell\right| + |x_\ell| \leq \left|\frac{\eta + \bar{\eta}}{2}\right| + 2|x_\ell|$$

$$= \left(a^2 - \left|\frac{\eta - \bar{\eta}}{2}\right|^2\right)^{\frac{1}{2}} + 2|x_\ell|,$$

where we used in the very last equality that $|\eta| = |\bar{\eta}| = a$. Therefore, for ℓ large enough, we can assume that $x_\ell \in B_r$, and $|\eta(x_\ell)| < a$. The last one in particular implies that

$$|\eta(x)| \leq a - 3\mu \quad \text{for every } x \in [x_\ell, 0] \text{ and } \ell \text{ large enough.}$$

We can also estimate

$$|x_\ell|^2 \leq (|\gamma(s_\ell)| + |x_\ell - \gamma(s_\ell)|)^2 \leq |\gamma(s_\ell)|^2 + 2L\epsilon_\ell(t - s_\ell)^2 + \epsilon_\ell^2(t - s_\ell)^2.$$

Let us then compare the action of δ_ℓ with the one of γ. We start from the kinetic part:

$$|\dot{\delta}_\ell(u)|^2 = \left(\frac{|x_\ell - \gamma(s_\ell)|}{\epsilon_\ell(t - s_\ell)}\right)^2 \leq 1 \quad \text{for } u \in [s_\ell, s_\ell + \epsilon_\ell(t - s_\ell)),$$

$$|\dot{\delta}_\ell(u)|^2 = \left(\frac{|x_\ell|}{(1 - \epsilon_\ell)(t - s_\ell)}\right)^2$$

$$\leq \frac{|\gamma(s_\ell)|^2}{(1 - \epsilon_\ell)^2(t - s_\ell)^2} + \frac{2L\epsilon_\ell}{(1 - \epsilon_\ell)^2} + \frac{\epsilon_\ell^2}{(1 - \epsilon_\ell)^2} \quad \text{for } u \in [s_\ell + \epsilon_\ell(t - s_\ell), t).$$

Whence, integrating:

$$\int_{s_\ell}^t |\dot\gamma|^2 \geq \frac{|\gamma(s_\ell)|^2}{t - s_\ell},$$

$$\int_{s_\ell}^t |\dot\delta_\ell|^2 \leq \epsilon_\ell(t - s_\ell) + \frac{|\gamma(s_\ell)|^2}{(1 - \epsilon_\ell)(t - s_\ell)} + \frac{2L\epsilon_\ell(t - s_\ell)}{(1 - \epsilon_\ell)} + \frac{\epsilon_\ell^2(t - s_\ell)}{(1 - \epsilon_\ell)}$$

$$= \frac{|\gamma(s_\ell)|^2}{(1 - \epsilon_\ell)(t - s_l)} + O(\epsilon_\ell^2).$$

On the other hand, for the potential part, computations very similar to the ones of Step 2 give

$$|\nabla f(\gamma(u))|^2 \geq a^2 - c_0(t - u) \quad \text{for } u \in [s_\ell, t),$$
$$|\nabla f(\delta_\ell(u))|^2 \leq a^2 + c_0(t - u) \quad \text{for } u \in [s_\ell, s_\ell + \epsilon_\ell(t - s_\ell)),$$
$$|\nabla f(\delta_\ell(u))|^2 \leq a^2 - 9\mu^2 + c_0(t - u) \quad \text{for } u \in [s_\ell + \epsilon_\ell(t - s_\ell), t).$$

Integration yields

$$\int_{s_\ell}^t h(|\nabla f(\gamma)|^2) \geq h(a^2)(t - s_\ell) - c_5(t - s_\ell)^2,$$

$$\int_{s_\ell}^t h(|\nabla f(\delta_\ell)|^2) \leq h(a^2)(t - s_\ell) + c_5(t - s_\ell)^2 - c_6(1 - \epsilon_\ell)(t - s_\ell).$$

Here ℓ is chosen large enough so that $c_0(t - s_\ell) \leq \mu^2$. Moreover, we set

$$c_5 := \frac{c_0}{2} \operatorname{Lip}\left(h; [a^2 - \mu, a^2 + \mu]\right), \quad c_6 := h(a^2) - h(a^2 - 8\mu^2) > 0.$$

Finally, collecting all the estimates together, we obtain

$$\mathcal{F}(\gamma) - \mathcal{F}(\delta_\ell) \geq \frac{|\gamma(s_\ell)|^2}{t - s_\ell}\left(\frac{-\epsilon_\ell}{(1 - \epsilon_\ell)}\right) + c_6(1 - \epsilon_\ell)(t - s_\ell) + O(\epsilon_\ell^2)$$

$$\geq \left(-\frac{L^2}{c_4(1 - \epsilon_\ell)}\epsilon_\ell + \frac{c_6}{c_4}(1 - \epsilon_\ell) + O(\epsilon_\ell)\right)\epsilon_\ell.$$

Now the contradiction comes from the fact that the right term is strictly positive for ℓ large enough. This concludes the proof.

□

Declarations

References

[1] Ambrosio, L., Ascenzi, O., Buttazzo, G.: Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands. J. Math. Anal. Appl. **142**, 301–316 (1989)

[2] Ambrosio, L., Baradat, A., Brenier, Y.: Monge–Ampère gravitation as a Γ-limit of good rate functions. Anal. PDE **9**, 2005–2040 (2023)

[3] Ambrosio, L., Baradat, A., Brenier, Y.: Γ-convergence for a class of action functionals induced by gradients of convex functions. Rend. Lincei. Mat. Appl. **32**, 97–108 (2021)

[4] Ambrosio, L., Brena, C.: Stability of a class of action functionals depending on convex functions. Discrete Contin. Dyn. Syst. **43**, 993–1005 (2023)

[5] Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures, 2nd edn. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2008)

[6] Aurenhammer, F.: Voronoi diagrams, a survey of a fundamental geometric data structure. ACM Comput. Surv. **23**, 345–405 (1991)

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

[8] Brenier, Y.: A double large deviation principle for Monge–Ampére gravitation. Bull. Inst. Math. Acad. Sin. **11**, 23–41 (2016)

[9] Chen, Y., Georgiou, T.T., Pavon, M.: On the relation between optimal transport and Schrödinger bridges: a stochastic control viewpoint. J. Optim. Theory Appl. **169**, 671–691 (2016)

[10] Conforti, G.: A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost. Prob. Theory Related Fields **174**, 1–47 (2019)

[11] Cordero-Erausquin, D.: Sur le transport de mesures périodiques. C. R. Acad. Sci. Paris Sér. I Math. **329**, 199–202 (1999)

[12] Gentil, I., Léonard, C., Ripani, L.: About the analogy between optimal transport and minimal entropy. Ann. Fac. Sci. Toulouse **3**, 569–600 (2017)

[13] Gigli, N., Tamanini, L.: Benamou–Brenier and duality formulas for the entropic cost on $\mathsf{RCD}^*(K, N)$ spaces. Prob. Theory Related Fields **176**, 1–34 (2020)

[14] Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis. Grundlehren Text Editions, 1st edn. Springer, Berlin (2001)

[15] McCann, R.: Polar factorization of maps on Riemannian manifolds. Geom. Funct. Anal. **11**, 589–608 (2001)