
Regularity of a kind of marginal functions in Hilbert spaces

Fátima F. Pereira and Vladimir V. Goncharov [†]

CIMA-UE, Rua Romão Ramalho 59, 7000-671, Évora, Portugal
e-mail: fmf@uevora.pt and goncha@uevora.pt

Dedicated to Panos M. Pardalos in honor of his 60th birthday

Summary. We study well-posedness of some *mathematical programming problem* depending on a parameter that generalizes in a certain sense the metric projection onto a closed nonconvex set. We are interested in regularity of the set of minimizers as well as of the value function, which can be seen, on one hand, as the viscosity solution to a Hamilton-Jacobi equation, while, on the other, as the minimal time in some related optimal time control problem. The regularity includes both the *Fréchet differentiability* of the value function and the *Hölder continuity of its (Fréchet) gradient*.

Key words: marginal function, metric projection, optimal time control problem, Hamilton-Jacobi equation, viscosity solution, uniform rotundity, duality mapping, proximal normals, Fréchet differentiability, Hölder continuity

Mathematical Subject Classification (2000): 49L25, 49J52, 49N15

1 Introduction

Given a real-valued function $f : X \times Y \rightarrow \mathbb{R}$ and a multivalued mapping $C : X \rightrightarrows Y$ (X and Y are Banach spaces) the general *marginal function* is defined as

$$T(x) := \inf_{y \in C(x)} f(x, y), \quad (1)$$

where inf can be certainly substituted by sup. The *marginal mapping* instead associates to each $x \in X$ the set of minimizers (or maximizers):

[†] Corresponding author. Tel.: +351 936514934; Fax: +351 266745393. He belongs also to the Institute of Systems Dynamics and Control Theory of Siberian Branch of RAS, ul. Lermontov 134, 664033 Irkutsk, Russia

$$\Pi(x) := \{y \in C(x) : T(x) = f(x, y)\}. \quad (2)$$

Regularity properties of marginal functions and mappings are important due to numerous applications in control theory, theory of games, mathematical economics, stochastic analysis, etc. We refer to [1, 2] for their general topological properties regarding the continuity and the lipschitzeanity.

A lot of works (see, e.g., [28, 10, 38, 29, 24, 31, 32] and the bibliography therein) was devoted to representation of various kind of subdifferentials of the marginal function through the respective subdifferentials of the function $f(\cdot, \cdot)$. The authors studied also subdifferential regularity of (1) in the sense of coincidence of different subdifferentials, and other properties such as the *approximate convexity* [33] or *generic differentiability* [19, 41, 24].

As about the *differentiability* of $T(\cdot)$ at a given point, notice that it can be treated in different ways. Namely, the *Fréchet differentiability* means the possibility to reduce the (*Fréchet*) *subdifferential* to the (eventually continuous) singleton $\nabla T(x)$ called the *Fréchet derivative*, or *gradient*. On the other hand, one may take an interest in both existence and uniqueness of the *proximal subgradient* that is a stronger property.

The Fréchet differentiability of the marginal function was particularly well studied when $X = Y = H$ is a Hilbert space with the norm $\|\cdot\|$; $C(x) = C \subset H$ is a closed set and $f(x, y) = \|x - y\|$. In this case $T(\cdot)$ is nothing else than the distance of the point to C , denoted further by $d_C(\cdot)$, whereas $\Pi(x) = \pi_C(x)$ is the (multivalued) *metric projection* of x onto C . In general, the set $\pi_C(x)$ can be certainly empty that does not occur when C is convex. Moreover, in the convex case $\pi_C(x)$ is a singleton, which is Lipschitz continuous w.r.t. x (with the Lipschitz constant 1) on the whole space H , and the gradient $\nabla d_C(x)$ is continuous (and locally lipschitzean) out of C .

If, instead, C is no longer convex then the latter property, in general, fails. However, there is a class of so named φ -convex sets (called also *prox-regular*, *proximally smooth*, sets *with positive reach*, etc.), for which the projection $\pi_C(x)$ is well defined and continuous (in fact, locally lipschitzean) on some (open) neighbourhood \mathcal{U} of C (equivalently, the distance $d_C(\cdot)$ is of class $C_{loc}^{1,1}$ on $\mathcal{U} \setminus C$). For the first time such sets were considered in the pioneer work [23] by H. Federer in finite dimensions, while afterwards various characterizations of them were given in Hilbert and Banach setting (see, e.g., [8, 37, 11, 34, 13, 4, 5] and the bibliography therein). This class of sets is well studied up to now, and we refer to the nice survey [15] for their basic properties.

The next step is to minimize the function $f(x, y) = \rho_F(x - y)$ in the place of the norm, where $F \subset H$ is a closed convex bounded set such that the origin belongs to the interior $\text{int } F$, and $\rho_F(\cdot)$ is the *Minkowski functional* (*gauge function*),

$$\rho_F(\xi) = \inf \{\lambda > 0 : \xi \in \lambda F\}. \quad (3)$$

In the latter case the value function (denoted by $\mathfrak{T}_C^F(x)$) of the respective minimization problem can be seen as the minimal time, which is necessary to achieve the *target set* $C \subset H$ from a point x by trajectories of the differential inclusion (with a constant convex right-hand side)

$$\dot{x}(t) \in -F. \quad (4)$$

Observe that another interpretation via viscosity theory for Hamilton-Jacobi equations can be given. Namely, $\mathfrak{T}_C^F(\cdot)$ is nothing else than the (unique) *viscosity solution* to the equation

$$\rho_{F^0}(\nabla u(x)) - 1 = 0, \quad x \in H \setminus C, \quad (5)$$

with $u(x) = 0$ on C (here F^0 is the *polar set*). This is a natural generalization of the so called *eikonal equation* arising from the geometric optics. For instance, if $H = \mathbb{R}^3$ and

$$F^0 = \left\{ \xi \in H : \sum_{i=1}^3 c_i^2 \xi_i^2 \leq 1 \right\}$$

then (5) describes the propagation of a light wave from a point source placed at the origin in (anisotropic) medium with the constant coefficients of refraction of light rays parallel to coordinate axes (denoted by c_i).

Notice that in the past many authors studied such *best approximation problem*

$$\min \{ \rho_F(x - y) : y \in C \}. \quad (6)$$

For example, in [22, 17] the generic properties of (6) were established, while in the works [20, 16, 39, 40] the directional derivatives as well as various subdifferentials of the value (time-minimum) function $\mathfrak{T}_C^F(\cdot)$ were computed. Furthermore, G. Colombo and P.R. Wolenski gave in [16] the first sufficient condition guaranteeing the local well-posedness of the problem (6) as well as the regularity of $\mathfrak{T}_C^F(\cdot)$. Afterwards, in [25, 26] the authors essentially sharpened this condition and represented it as a certain ballance between curvatures of the dynamics F and the target set C . Toward this end some quantitative results on convex duality in a Hilbert space were obtained in [25]. Besides that another (independent) "first order" hypothesis ensuring the well-posedness was proposed. It is written in terms of the ballance between (external) normals to the sets F and C .

It turned out that the latter hypothesis can be easily adapted to a more general problem where a supplementary additive term appears. Namely, given a sufficiently regular function $\theta : C \rightarrow \mathbb{R}$ we are led to consider the mathematical programming problem

$$\min \{ \rho_F(x - y) + \theta(y) : y \in C \}, \quad (7)$$

whose value function under an additional "slope assumption" is nothing else than the *viscosity solution* to the same *Hamilton-Jacobi equation* (5) but with the (general) boundary condition $u(x) = \theta(x)$, $x \in C$. Although the latter fact is well-known (see, e.g., [30, 9]), for the sake of completeness we give in Section 3 its detailed proof emphasizing thereby one of the crucial interpretations of the problem (7). Section 4 instead is devoted to another interpretation in terms of an optimal time control problem, which somehow extends the problem with the constant dynamics (4) mentioned above.

In the next Section 5 we introduce basic assumptions, under which the well-posedness and regularity results are obtained. Notice that the main hypothesis is, roughly speaking, a sort of (Lipschitz) compatibility of the normal vectors to F , on one hand, and of the (proximal) subdifferential to the restriction $\theta|_C$, on the other. Moreover, an auxiliary statement similar to Lemma 5.1 [25] is placed here. The (local) well-posedness of the problem (7) is proved in Section 6, while Section 7 is devoted to the regularity of the value function that includes its Fréchet differentiability and the (Hölder) continuity of the gradient near the target. We conclude

in Section 8 with two examples, which illustrate the applicability and the novelty of obtained results even in finite dimensions.

The main results of the paper (without detailed proofs) were announced earlier in [27].

2 Preliminaries

Let us emphasize first the setting of the problem. Everywhere in our considerations we assume that H is a Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the norm $\|\cdot\|$, that $F \subset H$ is a convex closed bounded set containing the origin in the interior and that $C \subset H$ is an arbitrary nonempty closed subset. Given a sufficiently regular (e.g., Lipschitz continuous) real-valued function $\theta : C \rightarrow \mathbb{R}$ we are interested in the well-posedness of the problem (7), i.e., in the existence, uniqueness and stability of its minimizers as well as in some kind of regularity of the value function (denoted further by $\hat{u}(\cdot)$).

We define the *support function* $\sigma_F : H \rightarrow \mathbb{R}^+$,

$$\sigma_F(\xi^*) := \sup \{ \langle \xi, \xi^* \rangle : \xi \in F \},$$

and recall the well known identity

$$\rho_F(\xi) = \sigma_{F^0}(\xi), \quad \xi \in H, \quad (8)$$

where F^0 is the *polar set* of F . Hence

$$\frac{1}{\|F\|} \|\xi\| \leq \rho_F(\xi) \leq \|F^0\| \|\xi\|, \quad \xi \in H, \quad (9)$$

where $\|F\| := \sup \{ \|\xi\| : \xi \in F \}$, and $\rho_F(\cdot)$ is lipschitzean with the Lipschitz constant $\|F^0\|$.

In what follows we use also the so called *duality mapping* $\mathfrak{J}_F : \partial F^0 \rightarrow \partial F$ that associates to each $\xi^* \in \partial F^0$ the set of all linear functionals $\langle \xi, \cdot \rangle$ with $\xi \in \partial F$ supporting F^0 in ξ^* . In other words,

$$\mathfrak{J}_F(\xi^*) := \{ \xi \in \partial F : \langle \xi, \xi^* \rangle = 1 \}.$$

Denoting by $\mathbf{N}_F(\xi)$ the *normal cone* to F at the point $\xi \in \partial F$ and by $\partial \rho_F(\xi)$ the *subdifferential* of the function $\rho_F(\cdot)$ in the sense of Convex Analysis we have other characterizations of the duality mapping:

$$\mathfrak{J}_F(\xi^*) = \partial \rho_{F^0}(\xi^*); \quad (10)$$

$$\mathfrak{J}_{F^0}(\xi) = \mathfrak{J}_F^{-1}(\xi) = \mathbf{N}_F(\xi) \cap \partial F^0. \quad (11)$$

In particular, using positive homogeneity of the gauge function one easily deduce from (10) and (11) that

$$\partial \rho_F(\xi) = \mathbf{N}_F\left(\frac{\xi}{\rho_F(\xi)}\right) \cap \partial F^0, \quad \xi \neq 0. \quad (12)$$

Following [25, Definition 3.2], for each dual pair (ξ, ξ^*) (i.e., $\xi^* \in \partial F^0$ and $\xi \in \mathfrak{J}_F(\xi^*)$) let us define the *modulus of rotundity*

$$\widehat{\mathfrak{C}}_F(r, \xi, \xi^*) := \inf \{ \langle \xi - \eta, \xi^* \rangle : \eta \in F, \|\xi - \eta\| \geq r \}, \quad r > 0.$$

The set F is said to be *strictly convex* (or *rotund*) at ξ w.r.t. ξ^* if

$$\widehat{\mathfrak{C}}_F(r, \xi, \xi^*) > 0 \quad \text{for all } r > 0. \quad (13)$$

If (13) is fulfilled then ξ is an *exposed point* of F and, in particular, ξ is the unique element of $\mathfrak{J}_F(\xi^*)$. So, in this case ξ is well defined whenever ξ^* is fixed. Observe that there is a strong connection between the rotundity of F and the smoothness of F^0 . Namely (see [25, Proposition 3.3 (iii)]), F is strictly convex at ξ w.r.t. ξ^* iff $\rho_{F^0}(\cdot)$ is Fréchet differentiable at ξ^* with $\nabla \rho_{F^0}(\xi^*) = \xi$. In this case we say also that F^0 is *smooth* at ξ^* (w.r.t. ξ).

Given a set $U \subset \partial F^0$ we say that F is *uniformly rotund* w.r.t. U if

$$\beta_U(r) := \inf \left\{ \widehat{\mathfrak{C}}_F(r, \xi, \xi^*) : \xi^* \in U \right\} > 0 \quad \text{for all } r > 0.$$

In [26, Proposition 2.1] the dual version of the latter property was given: the gauge F is uniformly rotund w.r.t. U if and only if the duality mapping $\mathfrak{J}_F(\cdot)$ is single-valued on U and uniformly continuous in the following sense

$$\sup_{\eta \in \mathfrak{J}_F(\eta^*)} \|\mathfrak{J}_F(\xi^*) - \eta\| \rightarrow 0 \quad \text{as } \|\xi^* - \eta^*\| \rightarrow 0, \quad \xi^* \in U, \quad \eta^* \in \partial F^0 \quad (14)$$

(we clearly identify $\mathfrak{J}_F(\xi^*)$ with its element whenever it is a singleton). Recalling the characterization of the duality mapping through the subdifferential of the Minkowski functional (see (10)) we derive from (14) that the uniform rotundity of F w.r.t. U implies, in particular, the uniform continuity of the *Fréchet gradient* $\nabla \rho_{F^0}(\cdot)$ on the set U .

In Sections 4 and 7 we will use also the distance between sets $A, B \subset H$. So, let us remind the *Pompeiu-Hausdorff metric*:

$$\begin{aligned} \mathcal{D}(A, B) &:= \max \left\{ \sup_{x \in A} d_B(x), \sup_{y \in B} d_A(y) \right\} = \\ &= \inf \{ r > 0 : A \subset B + r\overline{\mathbf{B}} \text{ and } B \subset A + r\overline{\mathbf{B}} \}. \end{aligned} \quad (15)$$

Here and in what follows $\overline{\mathbf{B}}$ means the closed unit ball in H . It is well known that the family $\text{conv } H$ of all nonempty convex closed bounded subsets of H supplied with the above distance is *isometrically embedded* into the space of real continuous functions defined on H as a complete cone, and the respective isometry is given by the formula:

$$\mathcal{D}(A, B) = \sup_{\|v\|=1} |\sigma_A(v) - \sigma_B(v)|. \quad (16)$$

Given now $F \in \text{conv } H$ with nonempty interior and $v \in \text{int } F$ let us denote by

$$r_F(v) := \sup \{ r > 0 : v + r\overline{\mathbf{B}} \subset F \}. \quad (17)$$

Being the set $(F - v)^0$ convex closed and bounded we have the following (local) Lipschitz inequality for the mapping $v \mapsto (F - v)^0$:

$$\mathcal{D}((F - v_1)^0, (F - v_2)^0) \leq \frac{1}{r_F(v_1)r_F(v_2)} |v_1 - v_2|, \quad (18)$$

$v_1, v_2 \in \text{int } F$. It was obtained in [21, Lemma 2] for $H = \mathbb{R}^n$ but readily can be adapted to the Hilbert case.

In the rest of this section let us give some concepts and notations of Nonsmooth Analysis, which will be used in the sequel.

Given a proper lower semicontinuous function $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$ we denote by $\partial^p f(x)$, $\partial^- f(x)$, $\partial^l f(x)$ and $\partial^c f(x)$ the *proximal*, *Fréchet*, *limiting (Mordukhovich)* and *Clarke-Rockafellar subdifferential* of $f(\cdot)$ at a point x , $f(x) < +\infty$, respectively. Let us recall their definitions (see, e.g., [12, 31, 10]):

- $\partial^p f(x) := \{\zeta \in H : \exists \eta > 0, \sigma \geq 0 \text{ such that } f(y) \geq f(x) + \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \quad \forall y, \|y - x\| \leq \eta\};$
- $\partial^- f(x) := \left\{ \zeta \in H : \liminf_{y \rightarrow x} \frac{f(y) - f(x) - \langle \zeta, y - x \rangle}{\|y - x\|} \geq 0 \right\};$
- $\partial^l f(x) := w\text{-}\limsup_{y \xrightarrow{f} x} \partial^- f(y) = \left\{ w\text{-}\lim_{i \rightarrow \infty} \zeta_i : \zeta_i \in \partial^- f(x_i), x_i \rightarrow x, f(x_i) \rightarrow f(x) \right\};$
- $\partial^c f(x) := \{\zeta \in H : \langle \zeta, v \rangle \leq f^\dagger(x; v) \quad \forall v \in H\}.$

Here "w-lim" stands for the weak limit, and

$$f^\dagger(x; v) := \lim_{\varepsilon \rightarrow 0^+} \limsup_{y \xrightarrow{f} x, h \rightarrow 0^+} \inf_{\|w - v\| \leq \varepsilon} \frac{f(y + hw) - f(y)}{h}$$

is the *Rockafellar's generalized directional derivative*; $y \xrightarrow{f} x$ means the convergence $y \rightarrow x$ together with $f(y) \rightarrow f(x)$.

Moreover, in order to treat viscosity solutions in the next section we define the *Fréchet superdifferential* $\partial^+ f(x)$ as the counterpart to $\partial^- f(x)$ assuming that the function $f(\cdot)$ is upper semicontinuous at x :

$$\partial^+ f(x) := \left\{ \zeta \in H : \limsup_{y \rightarrow x} \frac{f(y) - f(x) - \langle \zeta, y - x \rangle}{\|y - x\|} \leq 0 \right\}. \quad (19)$$

It is well known (see, e.g., [31, p. 90]) that for a continuous function $f(\cdot)$ both $\partial^- f(x)$ and $\partial^+ f(x)$ are nonempty simultaneously if and only if $f(\cdot)$ is *Fréchet differentiable* at the point x . In this case $\partial^- f(x) = \partial^+ f(x) = \{\nabla f(x)\}$.

Notice that the inclusions

$$\partial^p f(x) \subset \partial^- f(x) \subset \partial^l f(x) \subset \partial^c f(x) \quad (20)$$

are always valid, and that $f^\dagger(x; v)$ is reduced to the *Clarke's directional derivative*

$$f^\circ(x; v) := \limsup_{y \rightarrow x, h \rightarrow 0^+} \frac{f(y + hv) - f(y)}{h} \quad (21)$$

whenever $f(\cdot)$ is lipschitzean around x . In the latter case $\partial^c f(x)$ is bounded and can be represented as the convex closed hull of the limiting subdifferential. Taking

into account that in turn $\partial^l f(x)$ can be expressed through proximal subgradients in the place of Fréchet ones (see [31, p. 240]), we have

$$\partial^c f(x) = \overline{\text{co}} \left\{ w\text{-}\lim_{i \rightarrow \infty} \zeta_i : \zeta_i \in \partial^p f(x_i), x_i \rightarrow x \right\} \quad (22)$$

(see also [12, p. 88]). A function $f(\cdot)$ is said to be *proximally (lower, Clarke) regular* at x if $\partial^p f(x) = \partial^l f(x)$ (respectively, $\partial^- f(x) = \partial^l f(x)$ or $\partial^- f(x) = \partial^c f(x)$).

If $f(\cdot)$ is convex then all the subdifferentials above coincide with the subdifferential in the sense of Convex Analysis. If instead $f(\cdot)$ is (Fréchet) continuously differentiable at x then we can only affirm that $\partial^- f(x) = \partial^l f(x) = \partial^c f(x) = \{\nabla f(x)\}$ whereas the proximal subdifferential may be empty. However, this does not occur if the gradient $\nabla f(\cdot)$ is lipschitzean near x . So, in the latter case also $\partial^p f(x) = \{\nabla f(x)\}$. Let us observe that even Hölder continuity of $\nabla f(\cdot)$ with an exponent $0 < \alpha < 1$ does not guarantee the proximal regularity.

Given an open set $U \subset H$ in what follows we denote by $\mathcal{C}_{loc}^{1,\alpha}(U)$, $0 < \alpha \leq 1$, the class of all continuously differentiable functions $f(\cdot)$ whose gradient $\nabla f(\cdot)$ is locally Hölderean on U with the exponent α . In this case we say that $f(\cdot)$ is of class $\mathcal{C}_{loc}^{1,\alpha}$ on U .

The various notions of normal cones to a closed set C (all of them coincide with the cone $\mathbf{N}_C(x)$ if C is convex) can be given through the respective subdifferentials of the indicator function $\mathbf{I}_C(\cdot)$ equal to 0 on C and to $+\infty$ elsewhere. Thus, the *proximal, Fréchet (ou weak Bouligand), limiting (or Mordukhovich) and Clarke normal cones* are defined and denoted by $\mathbf{N}_C^p(x)$, $\mathbf{N}_C^c(x)$, $\mathbf{N}_C^l(x)$, $\mathbf{N}_C^e(x)$, respectively. Various properties of the normal cones (as well as of the subdifferentials of lower semicontinuous functions) can be found, e.g., in [10, 2, 36, 12, 31, 6]). Similarly as for the subdifferentials, a closed set C is said to be *proximally (normally, Clarke) regular* at $x \in \partial C$ if $\mathbf{N}_C^p(x) = \mathbf{N}_C^l(x)$ (respectively, $\mathbf{N}_C^c(x) = \mathbf{N}_C^l(x)$ or $\mathbf{N}_C^e(x) = \mathbf{N}_C^c(x)$).

We say that a closed set $C \subset H$ has *smooth (or \mathcal{C}^1) boundary* at $x_0 \in \partial C$ if for each $x \in \partial C$ enough close to x_0 the limiting cone $\mathbf{N}_C^l(x)$ is reduced to $\mathbf{n}_C(x) \mathbb{R}^+$ with some continuous function $\mathbf{n}_C(\cdot)$, $\|\mathbf{n}_C(x)\| = 1$. If, moreover, $\mathbf{n}_C(\cdot)$ is Hölder continuous with an exponent $0 < \alpha \leq 1$ then we say that C has *$\mathcal{C}^{1,\alpha}$ -boundary* at x_0 .

In what follows by the *restriction* $\theta|_C$ we mean the function equal to $\theta(x)$ on C and to $+\infty$ elsewhere. If $\theta(\cdot)$ is defined also out of C then clearly $\theta|_C = \theta + \mathbf{I}_C$. Due to this representation and to the proximal "sum rule" $\partial^p f + \partial^p g \subset \partial^p (f + g)$ we have that

- the subdifferential $\partial^p (\theta|_C)(x)$ is unbounded whenever $\mathbf{N}_C^p(x) \neq \{0\}$;
- $\partial^p (\theta|_C)(x) = \partial^p \theta(x)$ whenever $x \in \text{int } C$.

3 Marginal function as the viscosity solution

Let $\Omega \subset H$ be a nonempty open set and $\Gamma : \Omega \times \mathbb{R} \times H \rightarrow \mathbb{R}$ be a continuous mapping. Let us remind that a continuous function $u : \overline{\Omega} \rightarrow \mathbb{R}$ is said to be *viscosity solution* of the (stationary) *Hamilton-Jacobi equation*

$$\Gamma(x, u(x), \nabla u(x)) = 0 \quad (23)$$

if the following two conditions are fulfilled:

- (I) $\Gamma(x, u(x), p) \leq 0$ for each $x \in \Omega$ and each $p \in \partial^+ u(x)$;
- (II) $\Gamma(x, u(x), p) \geq 0$ for each $x \in \Omega$ and each $p \in \partial^- u(x)$.

For the main results of the viscosity theory for Hamilton-Jacobi equations we refer to [3] and to the bibliography therein (for a concise survey of viscosity solutions in finite dimensions see also the tutorial lessons by A. Bressan [7]). In particular, it is known that for each suitable boundary function $\theta(\cdot)$ there exists a unique viscosity solution $u(\cdot)$ of the equation (23) such that $u|_{\partial\Omega} = \theta$. Moreover, this solution is stable w.r.t. θ . Notice also that the class of viscosity solutions is consistent with other types of solutions. For instance, any continuously differentiable (by Fréchet) function $u(\cdot)$ satisfying (23) everywhere in Ω (so named *classical solution*) is also a viscosity solution.

Now we consider a hamiltonian Γ depending only on the gradient. Then, as shown in [9], under some geometric conditions the Hamilton-Jacobi equation (23) can be reduced to a particular case, where the hamiltonian takes the form $\rho_{F^0}(\xi) - 1$ with an appropriate gauge F (see (5) with $\Omega = H \setminus C$).

So, in what follows we deal with the boundary value problem for the equation (5), assuming that the boundary function $\theta : C \rightarrow \mathbb{R}$ satisfies the *slope condition* with respect to F , namely,

$$\theta(x) - \theta(y) \leq \rho_F(x - y) \quad \forall x, y \in C. \quad (24)$$

Remark 1. By (9) the inequality (24) implies the Lipschitz continuity of the function $\theta(\cdot)$ on C with the Lipschitz constant $\|\rho_{F^0}\|$. Moreover, the function

$$\hat{u}(x) := \inf_{y \in C} \{\rho_F(x - y) + \theta(y)\} \quad (25)$$

is a sort of extension of $\theta(\cdot)$ to the whole H with keeping the property (24) (a generalization of McShane lemma, see also [9, Lemma 4.1]).

Indeed, on one hand, it follows directly from (24) that $\hat{u}(x) = \theta(x)$ for all $x \in C$. On the other hand, given $y \in H$ and $\varepsilon > 0$ we find $z_y \in C$ such that

$$\hat{u}(y) \geq \rho_F(y - z_y) + \theta(z_y) - \varepsilon.$$

Then for each $x \in H$

$$\begin{aligned} \hat{u}(x) - \hat{u}(y) &\leq \rho_F(x - z_y) - \rho_F(y - z_y) + \varepsilon \\ &\leq \rho_F(x - y) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we arrive at the same slope condition as (24):

$$\hat{u}(x) - \hat{u}(y) \leq \rho_F(x - y) \quad \forall x, y \in H. \quad (26)$$

It implies, in particular, that $\hat{u}(x)$ admits finite value for each $x \in H$ (in fact, $\hat{u}(x) \geq \theta(x_0) - \rho_F(x_0 - x)$, $x \in H$, where $x_0 \in C$ is an arbitrary fixed point).

Theorem 1. *If the inequality (24) is fulfilled then the convolution (25) is the (unique) viscosity solution of the equation (5) such that $\hat{u}(x) = \theta(x)$, $x \in C$.*

Proof. In conformity with the definition above the proof splits into two parts as follows.

- (I) Let us fix $x \notin C$ and $p \in \partial^+ \hat{u}(x)$. Then, given $\varepsilon > 0$ by using the formula (19) we find $\delta > 0$ such that

$$\hat{u}(y) - \hat{u}(x) - \langle p, y - x \rangle \leq \varepsilon \|y - x\|$$

whenever $\|y - x\| \leq \delta$. In the case $y \neq x$ dividing the latter inequality by $\rho_F(x - y)$ and taking into account that

$$\frac{\hat{u}(y) - \hat{u}(x)}{\rho_F(x - y)} \geq -1$$

(see (26)) we obtain

$$-1 + \left\langle p, \frac{x - y}{\rho_F(x - y)} \right\rangle \leq \varepsilon \frac{\|x - y\|}{\rho_F(x - y)}.$$

Hence,

$$-1 + \sup_{0 < \|y - x\| < \delta} \left\langle p, \frac{x - y}{\rho_F(x - y)} \right\rangle \leq \varepsilon \sup_{0 < \|y - x\| < \delta} \frac{\|y - x\|}{\rho_F(y - x)}. \quad (27)$$

Since, by the positive homogeneity of the gauge function,

$$\sup_{0 < \|y - x\| < \delta} \frac{\|y - x\|}{\rho_F(y - x)} = \sup_{z \neq 0} \frac{\|z\|}{\rho_F(z)} = \|F\|$$

and

$$\sup_{0 < \|y - x\| < \delta} \left\langle p, \frac{x - y}{\rho_F(x - y)} \right\rangle = \sup_{z \in \partial F} \langle p, z \rangle = \sigma_F(p),$$

it follows from (27) and (8) that

$$-1 + \rho_{F^0}(p) \leq \varepsilon \|F\|.$$

Letting $\varepsilon \rightarrow 0^+$ we obtain $\rho_{F^0}(p) \leq 1$.

- (II) Let us fix now $p \in \partial^- \hat{u}(x)$ ($x \notin C$ is given). We should prove that $\rho_{F^0}(p) \geq 1$. Assuming the contrary, we choose $\varepsilon > 0$ so small that $\rho_{F^0}(p) < 1 - \varepsilon$. Then, by the definition of the Fréchet subdifferential and by (9) there exists δ , $0 < \delta < d_C(x)$, with

$$\hat{u}(x) - \hat{u}(y) + \langle p, y - x \rangle \leq \frac{\varepsilon}{4} \rho_F(x - y) \quad (28)$$

whenever $\|y - x\| \leq \delta$. On the other hand, let us take $z_x \in C$ such that

$$\hat{u}(x) \geq \rho_F(x - z_x) + \theta(z_x) - \frac{\varepsilon \delta}{8 \|F\|},$$

and, consequently,

$$\hat{u}(x) - \hat{u}(y) \geq \rho_F(x - z_x) - \rho_F(y - z_x) - \frac{\varepsilon \delta}{8 \|F\|} \quad \forall y \notin C. \quad (29)$$

Since $\|z_x - x\| > \delta$ (by the choice of $\delta > 0$), there exists $y_x \in [x, z_x]$ with $\|y_x - x\| = \delta/2$. Representing this point as $\lambda x + (1 - \lambda) z_x$ for some $0 < \lambda < 1$, we clearly have

$$\begin{aligned}\rho_F(x - y_x) &= (1 - \lambda) \rho_F(x - z_x); \\ \rho_F(y_x - z_x) &= \lambda \rho_F(x - z_x).\end{aligned}$$

Consequently (see also (9)),

$$\rho_F(x - z_x) - \rho_F(y_x - z_x) = \rho_F(x - y_x) \geq \frac{\delta}{2\|F\|}. \quad (30)$$

Therefore, applying successively (30), (29) and (28), we obtain

$$\begin{aligned}\rho_F(x - y_x) - \frac{\varepsilon\delta}{8\|F\|} + \langle p, y_x - x \rangle \\ = \rho_F(x - z_x) - \rho_F(y_x - z_x) - \frac{\varepsilon\delta}{8\|F\|} + \langle p, y_x - x \rangle \\ \leq \hat{u}(x) - \hat{u}(y_x) + \langle p, y_x - x \rangle \leq \frac{\varepsilon}{4} \rho_F(x - y_x).\end{aligned} \quad (31)$$

It follows from the inequality $\rho_{F^0}(p) \geq \left\langle p, \frac{x - y_x}{\rho_F(x - y_x)} \right\rangle$ (see (8)) and from the choice of $\varepsilon > 0$ that

$$\langle p, x - y_x \rangle < \rho_F(x - y_x)(1 - \varepsilon).$$

Hence, recalling (31) and (30) we obtain

$$\begin{aligned}\rho_F(x - y_x) &\leq \frac{\varepsilon\delta}{8\|F\|} + \langle p, x - y_x \rangle + \frac{\varepsilon}{4} \rho_F(x - y_x) \\ &< \frac{\varepsilon}{4} \rho_F(x - y_x) + \frac{\varepsilon}{4} \rho_F(x - y_x) + \rho_F(x - y_x)(1 - \varepsilon) \\ &= \rho_F(x - y_x) \left(1 - \frac{\varepsilon}{2}\right),\end{aligned}$$

which is a contradiction.

Combining the parts (I) and (II) proves the Theorem. \square

Remark 2. Observe that the slope condition (24) is always fulfilled if $\theta(\cdot)$ is defined and Lipschitz continuous on an open convex neighbourhood U of C and either $\nabla\theta(x) \in F^0$ for a.e. $x \in U$ (in finite dimensions), or $\partial^c\theta(x) \subset F^0 \quad \forall x \in U$ (in general case). This immediately follows from Lebourg's theorem (see [10, p. 41]). Vice versa (we use this property in the sequel), if $\theta(\cdot)$ is defined and satisfies (24) on a neighbourhood $U(\hat{x})$ of $\hat{x} \in C$ then $\partial^c\theta(\hat{x}) \subset F^0$. Indeed, it follows directly from (24) that given arbitrary $v \in H$ for each $x \in U(\hat{x})$ and sufficiently small $h > 0$ we have

$$\frac{\theta(x + hv) - \theta(x)}{h} \leq \rho_F(v).$$

Then, passing to \limsup as $h \rightarrow 0+$ and $x \rightarrow \hat{x}$ we conclude from both (8) and (21) that $f^o(\hat{x}; v) \leq \sigma_{F^0}(v)$. So, the definition of the Clarke subdifferential gives:

$$\partial^c\theta(\hat{x}) \subset \{\zeta \in H : \langle v, \zeta \rangle \leq \sigma_{F^0}(v) \quad \forall v \in H\} = F^0.$$

4 Marginal function as the minimal time

In this section we relate the function (25) with an optimal time control problem having, in general, non constant (autonomous) dynamics. However, in order to do this we should impose much stronger hypothesis on the function $\theta(\cdot)$.

Theorem 2. *Let $U \subset H$ be an open convex set with $U \supset C$, and $\theta : U \rightarrow \mathbb{R}$ be a (Fréchet) continuously differentiable function such that*

$$\nabla\theta(x) \in \text{int } F^0 \quad \forall x \in U. \quad (32)$$

Then for each $x \in U$ the equality

$$\hat{u}(x) = \mathfrak{T}_C^{F,\theta}(x) + \theta(x) \quad (33)$$

holds, where $\mathfrak{T}_C^{F,\theta}(x)$ is the minimal time necessary to achieve (the boundary of) the set C from the point $x \in U$ by trajectories of the differential inclusion

$$\dot{x}(t) \in (-F^0 + \nabla\theta(x(t)))^0 \quad (34)$$

remaining inside U .

Proof. Observe first that under the condition (32) the right-hand side of the inclusion (34) is bounded for each $x \in U$ that is essential for proving (33). Moreover, it is easy to see that in this case the inequality (24) is strict.

Let us prove first that

$$\hat{u}(x) \leq \mathfrak{T}_C^{F,\theta}(x) + \theta(x). \quad (35)$$

To this end fix $x \in U$ and assume that $\mathfrak{T}_C^{F,\theta}(x) < +\infty$ (otherwise (35) is trivial). So, there exists an instant, in which one can achieve the target C from x by some trajectory of the inclusion (34) do not leaving the neighbourhood U . Let us take an arbitrary such instant $T > 0$ and a trajectory $x(\cdot)$ of (34) with $x(0) = x$, $x(T) \in C$ and $x(t) \in U$, $t \in [0, T]$.

On the other hand, since the function $\hat{u}(\cdot)$ satisfies on H the inequality (26), it admits nonempty, convex and closed Clarke subdifferential $\partial^c \hat{u}(z) \subset F^0$, $z \in H$ (see Remark 2). In other words, setting $g(z) := \hat{u}(z) - \theta(z)$ we have

$$-\partial^c g(z) \subset -F^0 + \nabla\theta(z) \quad (36)$$

(see [10, Propositions 2.3.1 and 2.3.2]). Now, the relations (34) and (36) imply that $\langle -p, \dot{x}(t) \rangle \leq 1$ for all $p \in \partial^c g(x(t))$ and a.e. $t \in [0, T]$. It follows then from [10, p. 27] that

$$g^o(x(t), -\dot{x}(t)) = \max_{p \in \partial^c g(x(t))} \langle p, -\dot{x}(t) \rangle \leq 1. \quad (37)$$

Let us consider the superposition $t \mapsto g(x(t))$, which is Lipschitz continuous because $x(t) \in U$, $t \in [0, T]$, and the right-hand side of (34) is bounded. Therefore, $t \mapsto g(x(t))$ admits derivative at a.e. $t \in [0, T]$. By (21) and (37) we successively obtain

$$\begin{aligned}
\frac{d}{dt}g(x(t)) &= -\limsup_{h \rightarrow 0^+} \frac{g(x(t-h)) - g(x(t))}{h} \\
&= -\limsup_{h \rightarrow 0^+} \frac{g(x(t) - h\dot{x}(t)) - g(x(t))}{h} \\
&\geq -g^o(x(t), -\dot{x}(t)) \geq -1
\end{aligned}$$

a.e. on $[0, T]$. Hence, by integrating the latter inequality on the interval $[0, T]$ and taking into account that $g(x(T)) = 0$ (due to the boundary condition on C) we have

$$-g(x) = g(x(T)) - g(x(0)) = \int_0^T \frac{d}{dt}g(x(t)) dt \geq -T.$$

Since the instant $T > 0$ was chosen arbitrarily, we arrive at (35).

In order to prove the opposite inequality we fix $x \in U$, $\varepsilon > 0$ and choose $z_x \in C$ such that

$$\hat{u}(x) \geq \rho_F(x - z_x) + \theta(z_x) - \varepsilon. \quad (38)$$

We should find a trajectory $x(\cdot)$ of the inclusion (34) remaining in U such that $x(0) = x$ and $x(T) = z_x \in C$ where

$$T := \rho_F(x - z_x) + \theta(z_x) - \theta(x) > 0 \quad (39)$$

(see (32)). To this end we define first an approximative sequence $\{x_n(\cdot)\}$ as follows.

Given $n = 1, 2, \dots$ let us divide the segment $[x, z_x]$ in small parts by the points $x_i^n := x + \frac{i}{n}(z_x - x) \in U$, $i = 1, 2, \dots, n$. Denote by

$$T_i^n := \rho_F(x - x_i^n) + \theta(x_i^n) - \theta(x)$$

and observe that $T_0^n = 0$, $T_n^n = T$, $T_i^n > 0$ and

$$h_i^n := T_i^n - T_{i-1}^n = \rho_F(x_{i-1}^n - x_i^n) + \theta(x_i^n) - \theta(x_{i-1}^n) > 0, \quad (40)$$

$i = 1, 2, \dots, n$, due to the strict slope condition. Defining on $[0, T]$ the continuous piecewise affine function

$$x_n(t) := x_{i-1}^n + \frac{t - T_{i-1}^n}{h_i^n}(x_i^n - x_{i-1}^n), \quad t \in [T_{i-1}^n, T_i^n], \quad (41)$$

we clearly have $x_n(0) = x$, $x_n(T) = z_x \in C$, $x_n(T_i^n) = x_i^n$, and

$$x_n(t) \in [x, z_x] \subset U, \quad t \in [0, T]. \quad (42)$$

The composed function $t \mapsto \theta(x_n(t))$ is continuously differentiable on each interval $]T_{i-1}^n, T_i^n[$, $i = 1, 2, \dots, n$, and by the mean value theorem there exists $\tau_i^n \in]T_{i-1}^n, T_i^n[$ such that

$$\begin{aligned}
\theta(x_i^n) &= \theta(x_{i-1}^n) + \frac{d}{dt}\theta(x_n(t)) \Big|_{t=\tau_i^n} h_i^n \\
&= \theta(x_{i-1}^n) + \langle \nabla\theta(x_n(\tau_i^n)), x_i^n - x_{i-1}^n \rangle
\end{aligned} \quad (43)$$

(see (41)). Combining (40), (43) and (8) we have that

$$\begin{aligned}
h_i^n &= \theta(x_i^n) - \theta(x_{i-1}^n) + \rho_F(x_{i-1}^n - x_i^n) \\
&= \langle -\nabla\theta(x_n(\tau_i^n)), x_{i-1}^n - x_i^n \rangle + \sigma_{F^0}(x_{i-1}^n - x_i^n) \\
&\geq \langle -\xi^* + \nabla\theta(x_n(\tau_i^n)), x_i^n - x_{i-1}^n \rangle
\end{aligned}$$

whenever $\xi^* \in F^0$. Consequently,

$$\dot{x}_n(t) = \frac{x_i^n - x_{i-1}^n}{h_i^n} \in (-F^0 + \nabla\theta(x_n(\tau_i^n)))^0 \quad (44)$$

for all $t \in]T_{i-1}^n, T_i^n[$, $i = 1, 2, \dots, n$.

Thus, it remains to prove the convergence of $\{x_n(\cdot)\}$ (up to a subsequence) to a desired trajectory. To this end we observe, first, that the functions $x_n(\cdot)$ admit values in the same compact set $[x, z_x]$ (see (42)). Furthermore, since there exists $r > 0$ with

$$\nabla\theta(z) + r\bar{\mathbf{B}} \subset F^0 \quad (45)$$

whenever $z \in [x, z_x]$ (see (32)), passing to the polar sets and recalling (44) we have that

$$\|\dot{x}_n(t)\| \leq \left\| (-F^0 + \nabla\theta(x_n(\tau_i^n)))^0 \right\| \leq \frac{1}{r}, \quad t \in]T_{i-1}^n, T_i^n[, \quad i = 1, 2, \dots, n,$$

$n = 1, 2, \dots$. Therefore, applying successively Ascoli and Banach-Alaoglu theorems without loss of generality we can assume that both the sequence $\{x_n(\cdot)\}$ converges uniformly on $[0, T]$ to some Lipschitz continuous function $x(\cdot)$ and $\{\dot{x}_n(\cdot)\}$ converges weakly in the space $L^\infty([0, T], H)$ to the derivative $\dot{x}(\cdot)$, which exists almost everywhere on $[0, T]$. Then, by Mazur's Lemma there exists a sequence $\{v_n(\cdot)\}$ of convex combinations of the functions $\dot{x}_n(\cdot)$, converging to $\dot{x}(\cdot)$ strongly in $L^\infty([0, T], H)$ and, consequently, almost everywhere on $[0, T]$. Let us denote by \mathcal{T} the set of full measure in $[0, T]$, does not containing the node points T_i^n , and such that for $t \in \mathcal{T}$ the derivative $\dot{x}(t)$ exists and $v_n(t) \rightarrow \dot{x}(t)$, $n \rightarrow \infty$.

Fix now $t \in \mathcal{T}$ and choose i_n , $n = 1, 2, \dots$, with $t \in]T_{i_n-1}^n, T_{i_n}^n[$. Since by (40), (9) and (24)

$$h_{i_n}^n = T_{i_n}^n - T_{i_n-1}^n \leq 2 \|F^0\| \|x_{i_n-1}^n - x_{i_n}^n\| = \frac{2}{n} \|F^0\| \|z_x - x\| \rightarrow 0,$$

we have that $\tau_{i_n}^n \rightarrow t$ as $n \rightarrow \infty$. Then also $x_n(\tau_{i_n}^n) \rightarrow x(t)$ and $\nabla\theta(x_n(\tau_{i_n}^n)) \rightarrow \nabla\theta(x(t))$, $n \rightarrow \infty$ (we use here the continuity of the gradient $\nabla\theta(\cdot)$). Hence, recalling (18), (17) and (45) we obtain

$$\begin{aligned} & \mathcal{D} \left((-F^0 + \nabla\theta(x_n(\tau_{i_n}^n)))^0, (-F^0 + \nabla\theta(x(t)))^0 \right) \\ & \leq \frac{1}{r^2} \|\nabla\theta(x_n(\tau_{i_n}^n)) - \nabla\theta(x(t))\| \rightarrow 0. \end{aligned}$$

In particular, given $\delta > 0$ one can choose $N = N(\delta)$ such that

$$(-F^0 + \nabla\theta(x_n(\tau_{i_n}^n)))^0 \subset (-F^0 + \nabla\theta(x(t)))^0 + \delta\bar{\mathbf{B}} \quad (46)$$

whenever $n \geq N$. Combining (46) with (44) by the convexity of the right-hand side of (46) it follows that

$$v_n(t) \in (-F^0 + \nabla\theta(x(t)))^0 + \delta\bar{\mathbf{B}}$$

for all $n \geq N$. Passing now to the limit and taking into account the arbitrariness of $\delta > 0$, we conclude that $x(\cdot)$ is indeed a solution of the differential inclusion (34) with $x(0) = x$ and $x(T) = z_x \in C$. Moreover, $x(\cdot)$ admits values in U because all

the approximate solutions $x_n(\cdot)$ map $[0, T]$ into the compact segment $[x, z_x] \subset U$. Thus (see (38) and (39)),

$$\mathfrak{T}_C^{F, \theta}(x) \leq T \leq \hat{u}(x) - \theta(x) + \varepsilon,$$

and getting $\varepsilon \rightarrow 0+$ we prove the second part of the Theorem. \square

Remark 3. Notice that although the right-hand side in (34) is nonconstant and the trajectories realizing the optimal time are, in general, nonaffine (as one can see from the second part of the proof above), this time optimal control problem satisfies an essential property that the target set is achieved for the shortest time along one fixed direction (or close to that).

5 Auxiliary statement and standing assumptions

Our goal is to study regularity properties of $\hat{u}(\cdot)$ (see (25)), which (see the Sections 3 and 4) can be seen either as viscosity solution of a stationary Hamilton-Jacobi equation or as translated value function in an associated optimal time control problem. Such regularity is strictly related to the existence, uniqueness and stability of minimizers of $y \mapsto \rho_F(x - y) + \theta(y)$ on C . In the particular case $\theta \equiv 0$ this relation was well studied in [16, 25, 26], while for a general marginal function $T(x)$ and a compact set $C = C(x)$ (see (1)) we find a justification of this property, for instance, in the result by F. Clarke on representation of the generalized gradient $\partial^c T(x)$ as a family of integrals of $f(x, \cdot)$ with respect to all Radon measures supported on the set of minimizers $\Pi(x)$ (see [10, p. 86]). In the sequel the set $\Pi(x)$ for the problem (7) will be denoted by $\pi_C^{F, \theta}(x)$, and we keep the same notation for its element if $\pi_C^{F, \theta}(x)$ is a singleton.

Our standing hypothesis in what follows is a slightly strengthened slope condition (compare with (24)):

(H) there exists $0 < \gamma < \frac{1}{\|F\| \|F^0\|}$ such that

$$\theta(x) - \theta(y) \leq \gamma \rho_F(x - y) \quad (47)$$

for all $x, y \in C$.

Extending if necessary $\theta(\cdot)$ in a suitable way (see Remark 1), without loss of generality we can assume that the function $\theta(\cdot)$ is defined and satisfies (47) on the whole space H . Due to Remark 2 the inequality (47) can be equivalently written as the inclusion

$$\partial^c \theta(x) \subset \gamma F^0, \quad x \in H. \quad (48)$$

By (9) and the convexity of F^0 it follows from (48) that

$$\partial^c \theta(x) + \frac{1-\gamma}{\|F\|} \overline{\mathbf{B}} \subset \partial^c \theta(x) + (1-\gamma) F^0 \subset F^0, \quad (49)$$

or, recalling the definition (17),

$$r_{F^0}(\zeta) \geq \frac{1-\gamma}{\|F\|} \quad \forall \zeta \in \partial^c \theta(x), \quad x \in H. \quad (50)$$

On the other hand, passing in (49) to polar sets we have

$$\|(F^0 - \zeta)^0\| \leq \frac{\|F\|}{1-\gamma} \quad \forall \zeta \in \partial^c \theta(x), \quad x \in H. \quad (51)$$

In particular, if $\theta(\cdot)$ is (Fréchet) differentiable at x then it follows from (50) and (51) that

$$r_{F^0}(\nabla \theta(x)) \geq \frac{1-\gamma}{\|F\|} \quad (52)$$

and

$$\|(F^0 - \nabla \theta(x))^0\| \leq \frac{\|F\|}{1-\gamma}, \quad (53)$$

respectively (these estimates will be used further in Section 7).

Besides **(H)** in what follows we need a certain "slope-preserving" compatibility of the set C and the function $\theta(\cdot)$. Namely, set the following hypothesis:

($\hat{\mathbf{H}}$) for each $x \in \partial C$ there exist a (possibly empty) convex set $\Gamma(x) \subset \gamma F^0$ and a (possibly trivial) convex cone $\mathbf{N}_C^\theta(x)$ such that

$$\partial^p(\theta|_C)(x) = \Gamma(x) + \mathbf{N}_C^\theta(x). \quad (54)$$

So, the subdifferential $\partial^p(\theta|_C)(x)$ is empty if and only if $\Gamma(x) = \emptyset$, while it is bounded iff $\mathbf{N}_C^\theta(x) = \{0\}$. Let us denote by $\partial^\theta C$ the part of ∂C consisting of the points x where the cone $\mathbf{N}_C^\theta(x)$ is nontrivial. Notice that in the case $\theta \equiv 0$ we have $\Gamma(x) = \{0\}$, $\mathbf{N}_C^\theta(x) = \mathbf{N}_C^p(x)$ for each $x \in \partial C$, and $\partial^\theta C = \partial^* C$ is the *reduced boundary* in the sense of [25, 26].

Observe that the equality (54) holds with $\Gamma(x) = \partial^p \theta(x)$ and $\mathbf{N}_C^\theta(x) = \mathbf{N}_C^p(x)$, in particular, whenever either both $\theta(\cdot)$ and C are proximally regular (because for the proximal subdifferentials the inclusion $\partial^p(\theta|_C)(x) \supset \partial^p \theta(x) + \mathbf{N}_C^p(x)$ always holds, while for the limiting ones we have $\partial^l(\theta|_C)(x) \subset \partial^l \theta(x) + \mathbf{N}_C^l(x)$ (see, e.g., [12, p. 62])), or $\theta(\cdot)$ is of class $C^{1,1}$ near a given point (that can be proved easily by the same line as Proposition 2.11 [12, p. 38]). In the latter case, moreover, $\Gamma(x) = \{\nabla \theta(x)\}$. However, the decomposition property (**$\hat{\mathbf{H}}$**) can be valid in an essentially nonsmooth case as well (see Example 2 in Section 8).

Let us prove now an auxiliary assertion giving a property of minimizing sequences in (25), which generalizes the similar result [25, Lemma 5.1] obtained for the case $\theta \equiv 0$. We use here some tools of Variational and Proximal Analysis.

Lemma 1. *Let us suppose the standing assumptions **(H)** and **($\hat{\mathbf{H}}$)**. Then given a point $z \in H \setminus C$ and a minimizing sequence $\{x_n\} \subset C$ for the function $x \mapsto \rho_F(z-x) + \theta(x)$ on C one can find another minimizing sequence $\{x'_n\} \subset \partial^\theta C$ and sequences $\{x''_n\}$, $\{\mathbf{v}_n\}$, $\{\xi_n^*\}$ such that*

$$\mathbf{v}_n \in \partial^p (\theta|_C) (x'_n) \cap \partial F^0, \quad (55)$$

$$\xi_n^* \in \partial \rho_F (z - x''_n) \quad (56)$$

and

$$\|x'_n - x_n\| + \|x''_n - x_n\| \rightarrow 0, \quad (57)$$

$$\|\mathbf{v}_n - \xi_n^*\| \rightarrow 0 \quad (58)$$

as $n \rightarrow \infty$.

Proof. Let us take an arbitrary sequence $\varepsilon_n \rightarrow 0^+$ with

$$\rho_F (z - x_n) + \theta (x_n) \leq \hat{u} (z) + \varepsilon_n$$

and applying the *Ekeland Variational Principle* [18, Corollary 11] choose a sequence $\{y_n\} \subset C$ such that

$$\begin{aligned} \rho_F (z - y_n) + \theta (y_n) &\leq \hat{u} (z) + \varepsilon_n; \\ \|x_n - y_n\| &\leq \sqrt{\varepsilon_n} \end{aligned} \quad (59)$$

and

$$\rho_F (z - y_n) + \theta (y_n) \leq \rho_F (z - y) + \theta (y) + \sqrt{\varepsilon_n} \|y - y_n\| \quad (60)$$

for all $y \in C$, $n = 1, 2, \dots$.

The inequality (60) means that y_n minimizes the functional

$$F_n (y) := \rho_F (z - y) + \theta|_C (y) + \sqrt{\varepsilon_n} \|y - y_n\|$$

on H . Then the *necessary condition of optimality* in proximal form yields $0 \in \partial^p F_n (y_n)$. Decomposing the proximal subdifferential in accordance with the *fuzzy sum rule* (see Theorem 8.3 [12, p. 56]) we find sequences $\{x'_n\} \subset C$ and $\{x''_n\} \subset H$, $\|x'_n - y_n\| \leq \sqrt{\varepsilon_n}$, $\|x''_n - y_n\| \leq \sqrt{\varepsilon_n}$, such that

$$\begin{aligned} 0 &\in -\partial \rho_F (z - x''_n) + \sqrt{\varepsilon_n} \frac{x''_n - y_n}{\|x''_n - y_n\|} + \partial^p (\theta|_C) (x'_n) + \sqrt{\varepsilon_n} \bar{\mathbf{B}} \\ &\subset -\partial \rho_F (z - x''_n) + \partial^p (\theta|_C) (x'_n) + 2\sqrt{\varepsilon_n} \bar{\mathbf{B}}. \end{aligned}$$

Hence, there exist vectors $\mathbf{v}'_n \in \partial^p (\theta|_C) (x'_n)$ and $\xi_n^* \in \partial \rho_F (z - x''_n)$ with

$$\|\mathbf{v}'_n - \xi_n^*\| \leq 2\sqrt{\varepsilon_n}. \quad (61)$$

It follows from (59), (47) and (9) that $\{x'_n\}$ is a minimizing sequence of $x \mapsto \rho_F (z - x) + \theta (x)$ on C . Indeed,

$$\begin{aligned} \rho_F (z - x'_n) + \theta (x'_n) &\leq \rho_F (z - y_n) + \theta (y_n) \\ &\quad + \rho_F (y_n - x'_n) + \gamma \|F^0\| \|y_n - x'_n\| \\ &\leq \hat{u} (z) + (\gamma + 1) \|F^0\| \sqrt{\varepsilon_n} + \varepsilon_n. \end{aligned}$$

By using the hypothesis $(\hat{\mathbf{H}})$ we deduce that $x'_n \in \partial^\theta C$ since otherwise $\mathbf{v}'_n \in \Gamma (x'_n) \subset \gamma F^0$ contradicting the choice of ξ_n^* because $\|\mathbf{v}'_n - \xi_n^*\| \rightarrow 0$ (see (61)) and $\xi_n^* \in \partial F^0$ (see (12)). So, by (54) \mathbf{v}'_n can be decomposed in a sum $\mathbf{w}_n + \mathbf{u}_n$ where $\mathbf{w}_n \in \Gamma (x'_n)$ and $\mathbf{u}_n \in \mathbf{N}_C^\theta (x'_n)$ with $\mathbf{u}_n \neq 0$. Finally, let us define the vectors

$$\mathbf{v}_n := \frac{\mathbf{u}_n}{\rho_{F^0 - \mathbf{w}_n}(\mathbf{u}_n)} + \mathbf{w}_n \in \mathbf{N}_C^\theta(x'_n) + \Gamma(x'_n) = \partial^p(\theta|_C)(x'_n). \quad (62)$$

Obviously, $\mathbf{v}_n \in \partial F^0$ implying together with (62) the property (55). Furthermore, applying (9), the hypothesis $(\hat{\mathbf{H}})$ and the relations (48), (51) with $\Gamma(x'_n)$ in the place of $\partial^c \theta(x'_n)$ we successively obtain

$$\begin{aligned} \|\mathbf{v}'_n - \mathbf{v}_n\| &= \frac{\|\mathbf{u}_n\|}{\rho_{F^0 - \mathbf{w}_n}(\mathbf{u}_n)} \left| \rho_{F^0 - \mathbf{w}_n}(\mathbf{v}'_n - \mathbf{w}_n) - \rho_{F^0 - \mathbf{w}_n}(\xi_n^* - \mathbf{w}_n) \right| \\ &\leq \|F^0 - \mathbf{w}_n\| \left\| (F^0 - \mathbf{w}_n)^0 \right\| \|\mathbf{v}'_n - \xi_n^*\| \\ &\leq \frac{1 + \gamma}{1 - \gamma} \|F\| \|F^0\| \|\mathbf{v}'_n - \xi_n^*\|, \end{aligned} \quad (63)$$

$n = 1, 2, \dots$. Combining now (63) and (61) we arrive at (58), and the Lemma is proved. \square

6 Existence, uniqueness and stability of minimizers

Given $x_0 \in \partial C$ let us set now the local assumptions, under which the results on well-posedness and regularity near x_0 hold:

$(\mathbf{H}_1(x_0))$ the mapping $x \mapsto \mathfrak{J}_F(\partial^p(\theta|_C)(x) \cap \partial F^0)$ is **single-valued** and **lipschitzean** (with Lipschitz constant $L = L(x_0) > 0$) on the set

$$C_\delta(x_0) := \left\{ x \in \partial^\theta C : \|x - x_0\| \leq \delta \right\}, \quad \delta > 0;$$

$(\mathbf{H}_2(x_0))$ F is **uniformly rotund** w.r.t. the set

$$\mathfrak{U}_\delta(x_0) := \bigcup_{x \in C_\delta(x_0)} \partial^p(\theta|_C)(x) \cap \partial F^0. \quad (64)$$

Observe that (like the case $\theta \equiv 0$) in finite dimensions the hypothesis $(\mathbf{H}_2(x_0))$ holds automatically if one requires just the strict convexity of F w.r.t. each vector $\xi^* \in \mathfrak{U}_\delta(x_0)$ that trivially follows from $(\mathbf{H}_1(x_0))$. So, the assumption $(\mathbf{H}_2(x_0))$ can be required only in an infinite dimensional space H , while in \mathbb{R}^n it is superfluous.

Theorem 3. *Under the standing assumptions (\mathbf{H}) and $(\hat{\mathbf{H}})$ let us fix $x_0 \in \partial C$ and assume that the local hypotheses $(\mathbf{H}_1(x_0))$ and $(\mathbf{H}_2(x_0))$ are fulfilled. Then there exists a neighbourhood $\mathcal{U}(x_0)$ of x_0 where the mapping $\pi_C^{F, \theta}(\cdot)$ is single-valued and locally lipschitzean.*

Proof. Due to the choice of $\gamma > 0$ (see (\mathbf{H})) we can assume $\delta > 0$ from the hypotheses $(\mathbf{H}_1(x_0))$ – $(\mathbf{H}_2(x_0))$ so small that

$$\delta \gamma \|F^0\| < \frac{1 - \gamma \|F\| \|F^0\|}{L}.$$

Now, using the upper semicontinuity of $\hat{u}(\cdot)$ and the equality $\hat{u}(x_0) = \theta(x_0)$ (see Theorem 1), we can define the (open) neighbourhood

$$\mathcal{U}(x_0) := \left\{ x \in H : \|x - x_0\| < \frac{(1 - \gamma \|F\| \|F^0\|) \delta}{2 \|F\| \|F^0\|}, \right. \\ \left. \hat{u}(x) < \theta(x_0) + \frac{1 - \gamma \|F\| \|F^0\|}{L} - \delta \gamma \|F^0\| \right\}. \quad (65)$$

Fix $z \in \mathcal{U}(x_0) \setminus C$ and a minimizing sequence $\{x_n\} \subset C$ for the function $x \mapsto \rho_F(z - x) + \theta(x)$ on C . By Lemma 1 let us choose another minimizing sequence $\{x'_n\} \subset \partial^\theta C$ and sequences $\{x''_n\} \subset H$, $\mathbf{v}_n \in \partial^\rho(\theta|_C)(x'_n) \cap \partial F^0$, $\xi_n^* \in \partial \rho_F(z - x''_n)$ satisfying (57) and (58).

Let us show first that $x'_n \in C_\delta(x_0)$ for $n \geq 1$ large enough. To this end we denote by

$$0 < \varepsilon_n := \rho_F(z - x'_n) + \theta(x'_n) - \hat{u}(z) \rightarrow 0+ \quad (66)$$

and using the inequalities (9) and (47) successively write

$$\begin{aligned} \rho_F(x_0 - x'_n) &\leq \rho_F(x_0 - z) + \rho_F(z - x'_n) \\ &= \rho_F(x_0 - z) + \hat{u}(z) - \theta(x'_n) + \varepsilon_n \\ &\leq \rho_F(x_0 - z) + \rho_F(z - x_0) + \theta(x_0) - \theta(x'_n) + \varepsilon_n \\ &\leq 2 \|F^0\| \|z - x_0\| + \gamma \|F^0\| \|x'_n - x_0\| + \varepsilon_n. \end{aligned} \quad (67)$$

Hence, again by (9) we have

$$\left(\frac{1}{\|F\|} - \gamma \|F^0\| \right) \|x'_n - x_0\| \leq \varepsilon_n + 2 \|F^0\| \|z - x_0\|,$$

and by the choice of z (see (65)) conclude that $\|x'_n - x_0\| < \delta$.

Then, due to one of the characterizations of the convex subdifferential (see (12)) $\xi_n^* \in \mathbf{N}_F(\xi_n) \cap \partial F^0$, where

$$\xi_n := \frac{z - x''_n}{\rho_F(z - x''_n)},$$

and, consequently (see (10)), $\xi_n \in \mathfrak{J}_F(\xi_n^*)$, $n = 1, 2, \dots$.

Set now

$$\beta_n := \max \{ \varepsilon_n, \|x'_n - x_n\| + \|x''_n - x_n\|, \|\mathfrak{J}_F(\mathbf{v}_n) - \xi_n\| \} \quad (68)$$

and deduce from (57), (58), the hypothesis $(\mathbf{H}_2(x_0))$ and [26, Proposition 2.1] that $\beta_n \rightarrow 0+$ as $n \rightarrow \infty$.

Taking into account the representation $x''_n = z - \xi_n \rho_F(z - x''_n)$, for given $m, n \geq 1$ we write

$$\|x''_n - x''_m\| \leq \rho_F(z - x''_n) \|\xi_n - \xi_m\| + \|\xi_m\| |\rho_F(z - x''_n) - \rho_F(z - x''_m)|. \quad (69)$$

Let us estimate each term of the latter inequality. First, by the definition of ε_n (see (66)), (9), (47) and (68) we obtain that

$$\begin{aligned} \rho_F(z - x''_n) &\leq \|F^0\| \|x''_n - x'_n\| + \hat{u}(z) - \theta(x'_n) + \varepsilon_n \\ &\leq (\|F^0\| + 1) \beta_n + \hat{u}(z) - \theta(x_0) + \gamma \delta \|F^0\| \end{aligned} \quad (70)$$

and that

$$\begin{aligned}
 & |\rho_F(z - x''_n) - \rho_F(z - x''_m)| \\
 & \leq |\rho_F(z - x'_n) - \rho_F(z - x'_m)| + \|F^0\|(\beta_n + \beta_m) \\
 & \leq |\rho_F(z - x'_n) + \theta(x'_n) - \hat{u}(z)| + \|F^0\|(\beta_n + \beta_m) \\
 & \quad + |\rho_F(z - x'_m) + \theta(x'_m) - \hat{u}(z)| + |\theta(x'_n) - \theta(x'_m)| \\
 & \leq (\|F^0\| + 1)(\beta_n + \beta_m) + \gamma \|F^0\| \|x'_n - x'_m\|. \tag{71}
 \end{aligned}$$

Furthermore, since $\mathbf{v}_n \in \partial^p(\theta|_C)(x'_n) \cap \partial F^0$, applying the main hypothesis $(\mathbf{H}_1(x_0))$ we have

$$\begin{aligned}
 \|\xi_n - \xi_m\| & \leq \|\mathfrak{J}_F(\mathbf{v}_n) - \mathfrak{J}_F(\mathbf{v}_m)\| + \beta_n + \beta_m \\
 & \leq L \|x'_n - x'_m\| + \beta_n + \beta_m. \tag{72}
 \end{aligned}$$

After substituting (70)-(72) into (69) and joining all the infinitesimal constants we finally arrive at:

$$\begin{aligned}
 \|x'_n - x'_m\| & \leq \|x''_n - x''_m\| + \beta_n + \beta_m \leq [L(\hat{u}(z) - \theta(x_0) + \gamma\delta \|F^0\|) \\
 & \quad + \gamma \|F\| \|F^0\|] \|x'_n - x'_m\| + \mu_{n,m} \tag{73}
 \end{aligned}$$

where $\mu_{n,m} \rightarrow 0+$ as $n, m \rightarrow \infty$. Since

$$\gamma \|F\| \|F^0\| + L(\hat{u}(z) - \theta(x_0) + \gamma\delta \|F^0\|) < 1$$

by the choice of z (see (65)), we conclude from (73) that $\{x'_n\}$ (and, consequently, $\{x_n\}$) is a Cauchy sequence in H .

In fact, we have proved that each minimizing sequence of the function $x \mapsto \rho_F(z - x) + \theta(x)$ on C is a Cauchy sequence. Therefore, its limit \bar{x} is the (unique) element of the set of minimizers $\pi_C^{F,\theta}(z)$. Moreover, by using the same argument we can prove the continuity of the mapping $z \mapsto \pi_C^{F,\theta}(z)$ on $\mathcal{U}(x_0)$. Indeed, taking a sequence $\{z_n\} \subset \mathcal{U}(x_0)$, $z_n \rightarrow z \in \mathcal{U}(x_0)$, and denoting by \bar{z}_n the unique element of $\pi_C^{F,\theta}(z_n)$, we observe that $\{\bar{z}_n\}$ is a minimizing sequence of $x \mapsto \rho_F(z - x) + \theta(x)$ on C . Indeed, we have

$$\begin{aligned}
 \hat{u}(z) & \leq \rho_F(z - \bar{z}_n) + \theta(\bar{z}_n) \leq \rho_F(z - z_n) + \rho_F(z_n - \bar{z}_n) + \theta(\bar{z}_n) \\
 & \leq \hat{u}(z) + 2 \|F^0\| \|z - z_n\|,
 \end{aligned}$$

where the latter inequality follows from the Lipschitz continuity of the function $\hat{u}(\cdot)$. So, $\{\bar{z}_n\}$ converges to the (unique) element of $\pi_C^{F,\theta}(z)$.

In the second part of the proof we show that the single-valued function $\pi_C^{F,\theta}(\cdot)$ is actually Lipschitz continuous on $\mathcal{U}(x_0)$. To do this fix an arbitrary point $x \in \mathcal{U}(x_0)$ and choose $\tau > 0$ and $0 < \varepsilon \leq \frac{\tau}{2\|F^0\|}$ so small that

$$\hat{u}(x) - \theta(x_0) + \gamma\delta \|F^0\| + \tau < \frac{1 - \gamma \|F\| \|F^0\|}{L} \tag{74}$$

and

$$\frac{2 \|F\| \|F^0\|}{1 - \gamma \|F\| \|F^0\|} (\|x - x_0\| + \varepsilon) < \delta. \tag{75}$$

Let us take $z_1, z_2 \in \mathcal{U}(x_0)$, $\|z_i - x\| < \varepsilon$, $i = 1, 2$, and assume first that both (different) points z_1 and z_2 are out of C . Setting $\beta := \|z_1 - z_2\|/2 > 0$, in virtue of the hypothesis $(\mathbf{H}_2(x_0))$ and [26, Proposition 2.1 (ii)] we find $0 < \nu \leq \varepsilon \wedge \beta$ such that

$$\|\mathfrak{J}_F(\eta^*) - \xi\| \leq \beta$$

whenever $\xi \in \mathfrak{J}_F(\xi^*)$, $\xi^* \in \partial F^0$ and $\eta^* \in \mathfrak{U}_\delta(x_0)$ with $\|\xi^* - \eta^*\| \leq \nu$. Without loss of generality one may suppose that

$$\nu + \frac{2\|F\|\|F^0\|}{1 - \gamma\|F\|\|F^0\|}(\|x - x_0\| + \varepsilon) < \delta \quad (76)$$

and that $(z_i + \nu\overline{B}) \cap C = \emptyset$, $i = 1, 2$. Set also $\bar{z}_i := \pi_C^{F, \theta}(z_i)$. Now we apply the tools used for proving Lemma 1 but without recurrence to the Ekeland Principle (because the exact minimizer is already known). Namely, \bar{z}_i minimizes the function

$$F_i(x) := \rho_F(z_i - x) + \theta|_C(x)$$

on H . Therefore $0 \in \partial^p F_i(\bar{z}_i)$, $i = 1, 2$. By the fuzzy sum rule similarly as in the proof of Lemma 1 we find points $z'_i \in \partial^\theta C$ and $z''_i \in H$ both close to \bar{z}_i (say $\|z'_i - \bar{z}_i\| + \|z''_i - \bar{z}_i\| \leq \nu$) and vectors $\mathbf{v}_i \in \partial^p(\theta|_C)(z'_i) \cap \partial F^0$, $\xi_i^* \in \partial \rho_F(z_i - z''_i)$ such that

$$\|\mathbf{v}_i - \xi_i^*\| \leq \nu. \quad (77)$$

Let us show now that $z'_i \in C_\delta(x_0)$. Similarly as in (67) we have:

$$\begin{aligned} \rho_F(x_0 - \bar{z}_i) &\leq \rho_F(z_i - \bar{z}_i) + \rho_F(x_0 - z_i) \\ &= \hat{u}(z_i) - \theta(\bar{z}_i) + \rho_F(x_0 - z_i) \\ &\leq \rho_F(z_i - x_0) + \theta(x_0) - \theta(\bar{z}_i) + \|F^0\| \|z_i - x_0\| \\ &\leq 2\|F^0\| \|z_i - x_0\| + \gamma\|F^0\| \|\bar{z}_i - x_0\|, \end{aligned}$$

and, hence,

$$\frac{1 - \gamma\|F^0\|\|F\|}{\|F\|} \|\bar{z}_i - x_0\| \leq 2\|F^0\| \|z_i - x_0\|.$$

Recalling (76), from the latter inequality we obtain

$$\begin{aligned} \|z'_i - x_0\| &\leq \|z'_i - \bar{z}_i\| + \|\bar{z}_i - x_0\| \\ &\leq \nu + \frac{2\|F^0\|\|F\|}{1 - \gamma\|F^0\|\|F\|}(\varepsilon + \|x - x_0\|) < \delta. \end{aligned}$$

Thus $z'_i \in C_\delta(x_0)$ and $\mathbf{v}_i \in \mathfrak{U}_\delta(x_0)$ (see (64)).

Setting now $\xi_i := \frac{z_i - z''_i}{\rho_F(z_i - z''_i)}$ we see that $\xi_i \in \mathfrak{J}_F(\xi_i^*)$, and it follows from (77) and from the choice of $\nu > 0$ that

$$\|\mathfrak{J}_F(\mathbf{v}_i) - \xi_i\| \leq \beta. \quad (78)$$

Joining together the inequalities (78) for $i = 1, 2$ and using the hypothesis $(\mathbf{H}_1(x_0))$, we have

$$\begin{aligned} \|\xi_1 - \xi_2\| &\leq 2\beta + L\|z'_1 - z'_2\| \\ &\leq 2\beta + L(2\nu + \|\bar{z}_1 - \bar{z}_2\|). \end{aligned} \quad (79)$$

In order to estimate the distance $\|\bar{z}_1 - \bar{z}_2\|$ we use first the proximity of each minimizer \bar{z}_i to $z_i'' = z_i - \xi_i \rho_F(z_i - z_i'')$. Namely,

$$\begin{aligned} \|\bar{z}_1 - \bar{z}_2\| &\leq 2\nu + \|z_1'' - z_2''\| \\ &\leq 2\nu + \|z_1 - z_2\| + \|\xi_1 \rho_F(z_1 - z_1'') - \xi_2 \rho_F(z_2 - z_2'')\| \\ &\leq 2\nu + \|z_1 - z_2\| + \rho_F(z_1 - z_1'') \|\xi_1 - \xi_2\| \\ &\quad + \|F\| |\rho_F(z_1 - z_1'') - \rho_F(z_2 - z_2'')|. \end{aligned} \quad (80)$$

On the other hand, similarly to (70) and (71) we successively have

$$\begin{aligned} \rho_F(z_1 - z_1'') &\leq \rho_F(z_1 - \bar{z}_1) + \|F^0\| \|\bar{z}_1 - z_1''\| \\ &\leq \|F^0\| \nu + \hat{u}(z_1) - \theta(\bar{z}_1) \\ &\leq \|F^0\| (\nu + \|z_1 - x\|) + \hat{u}(x) - \theta(x_0) + \gamma \|F^0\| \|\bar{z}_1 - x_0\| \\ &\leq \hat{u}(x) - \theta(x_0) + \gamma \|F^0\| \delta + \tau \end{aligned} \quad (81)$$

(recall that $\nu \leq \varepsilon \leq \frac{\tau}{2\|F^0\|}$), and

$$\begin{aligned} &|\rho_F(z_1 - z_1'') - \rho_F(z_2 - z_2'')| \\ &\leq |\rho_F(z_1 - \bar{z}_1) - \rho_F(z_2 - \bar{z}_2)| + 2\|F^0\| \nu \\ &\leq |\hat{u}(z_1) - \hat{u}(z_2)| + |\theta(\bar{z}_1) - \theta(\bar{z}_2)| + 2\|F^0\| \nu \\ &\leq \|F^0\| (2\nu + \|z_1 - z_2\|) + \gamma \|F^0\| \|\bar{z}_1 - \bar{z}_2\|. \end{aligned} \quad (82)$$

Taking into account the inequalities (81), (79), (82) and recalling that $\nu \leq \beta = \|z_1 - z_2\|/2$ we deduce from (80):

$$[1 - L(\hat{u}(x) - \theta(x_0) + \gamma \|F^0\| \delta + \tau) - \gamma \|F\| \|F^0\|] \|\bar{z}_1 - \bar{z}_2\| \leq K \|z_1 - z_2\|$$

where

$$K = K(x) := 2(1 + \|F\| \|F^0\|) + (L + 1)(\hat{u}(x) - \theta(x_0) + \gamma \|F^0\| \delta + \tau) > 0.$$

Finally,

$$\mu = \mu(x) := 1 - L(\hat{u}(x) - \theta(x_0) + \gamma \|F^0\| \delta + \tau) - \gamma \|F\| \|F^0\| > 0$$

by (74), and we arrive at the (local) Lipschitz inequality

$$\left\| \pi_C^{F,\theta}(z_1) - \pi_C^{F,\theta}(z_2) \right\| \leq \frac{K(x)}{\mu(x)} \|z_1 - z_2\|. \quad (83)$$

In the case when one of the points z_i (say z_2) belongs to C , we obviously have $\pi_C^{F,\theta}(z_2) = z_2$ and

$$\begin{aligned} \|\bar{z}_1 - \bar{z}_2\| &= \|\bar{z}_1 - z_2\| \leq \|F\| \rho_F(z_1 - \bar{z}_1) + \|z_1 - z_2\| \\ &= \|F\| (\hat{u}(z_1) - \theta(\bar{z}_1)) + \|z_1 - z_2\| \\ &= \|F\| (\hat{u}(z_1) - \hat{u}(z_2)) + \|F\| (\theta(z_2) - \theta(\bar{z}_1)) + \|z_1 - z_2\| \\ &\leq (\|F\| \|F^0\| + 1) \|z_1 - z_2\| + \gamma \|F\| \|F^0\| \|\bar{z}_1 - \bar{z}_2\|. \end{aligned}$$

Hence, (83) holds as well with

$$K(x) := \|F\| \|F^0\| + 1$$

and

$$\mu(x) := 1 - \gamma \|F\| \|F^0\| > 0.$$

Theorem is completely proved. \square

Remark 4. Notice that the Lipschitz constant of the mapping $\pi_C^{F,\theta}(\cdot)$ depends essentially on the distance from the boundary of the neighbourhood $\mathcal{U}(x_0)$ controlled by the parameter τ . In fact, $\mathcal{U}(x_0)$ is defined by means of two inequalities: the first one gives direct proximity to the boundary point x_0 , while the second derives from the upper semicontinuity of the marginal function $\hat{u}(\cdot)$ at x_0 . Thus, the Lipschitz constant of $\pi_C^{F,\theta}(\cdot)$ depending on $x \in \mathcal{U}(x_0)$ tends to $+\infty$ ($\mu(x) \rightarrow 0$) whenever the strict upper semicontinuity inequality

$$\hat{u}(x) < \hat{u}(x_0) + \frac{1 - \gamma \|F\| \|F^0\|}{L} - \delta \gamma \|F^0\|$$

tends to become an equality, i.e., the value of the function $\hat{u}(\cdot)$ at x is most distant from its value at x_0 . This generalizes the well-known property of the metric projections onto prox-regular sets (see, e.g., [8]).

Remark 5. If the conditions $(\mathbf{H}_1(x_0))$ and $(\mathbf{H}_2(x_0))$ are fulfilled at each point $x_0 \in \partial C$ then the marginal mapping $x \mapsto \pi_C^{F,\theta}(x)$ is single-valued and locally Lipschitzian on the open neighbourhood

$$\mathfrak{A}(C) := \text{int } C \cup \bigcup_{x_0 \in \partial C} \mathcal{U}(x_0)$$

of the target set.

7 Regularity of the value function

At the beginning of this section we study the Clarke (and lower) regularity of the function $\hat{u}(\cdot)$ at a given point \hat{x} out of the target set under an *a priori* assumption that for each x near \hat{x} the infimum in (7) is attained at a unique point, and a kind of stability of the minimizer takes place. Furthermore, we give a representation formula for the Clarke (Fréchet or Mordukhovich) subdifferential of $\hat{u}(\cdot)$ at \hat{x} in terms of the respective constructions for F , $\theta(\cdot)$ and C . A similar result was obtained in [26] in the case $\theta \equiv 0$ (see also [14]).

For one step of the proof we need the following simple observation.

Lemma 2. Fix $x \notin C$ such that $\pi_C^{F,\theta}(x)$ is a singleton (say \bar{x}) and denote by $\bar{\xi} := \frac{x-\bar{x}}{\rho_F(x-\bar{x})}$. Then for all $0 \leq t \leq \rho_F(x-\bar{x})$ the inequality

$$\hat{u}(x - t\bar{\xi}) \leq \hat{u}(x) - t \quad (84)$$

holds.

Proof. Setting $y_t := x - t\bar{\xi}$ for $0 \leq t \leq \rho_F(x-\bar{x})$, we have

$$\begin{aligned} \hat{u}(y_t) &\leq \rho_F(y_t - \bar{x}) + \theta(\bar{x}) = \rho_F\left((x - \bar{x})\left(1 - \frac{t}{\rho_F(x-\bar{x})}\right)\right) + \theta(\bar{x}) \\ &= \rho_F(x - \bar{x}) - t + \theta(\bar{x}) = \hat{u}(x) - t, \end{aligned}$$

and (84) is proved. \square

Theorem 4. Let us fix $\hat{x} \notin C$ and assume that the mapping $x \mapsto \pi_C^{F,\theta}(x)$ is single-valued in a neighbourhood $U(\hat{x})$ of \hat{x} and such that

$$\lim_{r \rightarrow 0^+} \frac{\omega(\hat{x}; r)}{\sqrt{r}} = 0, \quad (85)$$

where $\omega(\hat{x}; r)$ is the modulus of continuity of $\pi_C^{F,\theta}(\cdot)$ at the point \hat{x} , namely,

$$\omega(\hat{x}; r) := \sup \left\{ \left\| \pi_C^{F,\theta}(x) - \pi_C^{F,\theta}(\hat{x}) \right\| : \|x - \hat{x}\| \leq r \right\}.$$

Suppose also that the restriction $\theta|_C$ is proximally regular at $\bar{x} := \pi_C^{F,\theta}(\hat{x})$. Then the function $\hat{u}(\cdot)$ is Clarke (and, hence, lower) regular at \hat{x} . Furthermore, the following formula takes place:

$$\partial^c \hat{u}(\hat{x}) = \partial^l \hat{u}(\hat{x}) = \partial^- \hat{u}(\hat{x}) = \partial \rho_F(\hat{x} - \bar{x}) \cap \partial^- (\theta|_C)(\bar{x}) \neq \emptyset. \quad (86)$$

Proof. Our proof is divided into several steps.

Step 1. Let us show first that $\partial^- \hat{u}(x) \subset \partial \rho_F(x - \bar{x})$ for each $x \in U(\hat{x})$ where $\bar{x} := \pi_C^{F,\theta}(x)$. To this end we use the representation of the subdifferential $\partial \rho_F(x - \bar{x})$ via the normal cone to F (see (12)). Since the function $\hat{u}(\cdot)$ satisfies the slope condition (26), by (20) and Remark 2 we have $\partial^- \hat{u}(x) \subset \partial^c \hat{u}(x) \subset F^0$. On the other hand, by Theorem 1 $\hat{u}(\cdot)$ is the viscosity solution of (5). So, in particular,

$$\rho_{F^0}(p) \geq 1 \quad (87)$$

for each $p \in \partial^- \hat{u}(x)$. Thus $\partial^- \hat{u}(x) \subset \partial F^0$.

Besides that, given $p \in \partial^- \hat{u}(x)$ let us choose $\varepsilon > 0$ and $\delta > 0$ such that

$$\hat{u}(y) - \hat{u}(x) - \langle p, y - x \rangle \geq -\varepsilon \|x - y\|$$

for all y , $\|y - x\| \leq \delta$. In particular, setting $y = x - t\bar{\xi}$, where

$$\bar{\xi} := \frac{x - \bar{x}}{\rho_F(x - \bar{x})},$$

and applying Lemma 2 we have that

$$\begin{aligned}\hat{u}(x) - t &\geq \hat{u}(x - t\bar{\xi}) \\ &\geq \hat{u}(x) - t\langle p, \bar{\xi} \rangle - \varepsilon t \|\bar{\xi}\|\end{aligned}$$

for sufficiently small $t > 0$. Hence, letting $\varepsilon \rightarrow 0+$ we arrive at $1 \leq \langle p, \bar{\xi} \rangle$ and, consequently,

$$\langle p, \xi - \bar{\xi} \rangle \leq 0$$

whenever $\xi \in F$, i.e., $p \in \mathbf{N}_F(\bar{\xi})$. So, due to (12) we conclude that $\partial^- \hat{u}(x) \subset \partial \rho_F(x - \bar{x})$.

Step 2. We prove the inclusion $\partial^p \hat{u}(x) \subset \partial^p(\theta|_C)(\bar{x})$, $x \in U(\hat{x})$. Given $p \in \partial^p \hat{u}(x)$ let us choose $\eta > 0$ and $\sigma > 0$ such that

$$\hat{u}(y) - \hat{u}(x) - \langle p, y - x \rangle \geq -\sigma \|y - x\|^2$$

for all y , $\|y - x\| \leq \eta$. In particular, for $y = z - \bar{x} + x$, where $z \in C$, $\|z - \bar{x}\| \leq \eta$, we have:

$$\begin{aligned}-\sigma \|z - \bar{x}\|^2 &\leq \hat{u}(z - \bar{x} + x) - \hat{u}(x) - \langle p, z - \bar{x} \rangle \\ &\leq \rho_F(x - \bar{x}) + \theta(z) - \rho_F(x - \bar{x}) - \theta(\bar{x}) - \langle p, z - \bar{x} \rangle,\end{aligned}$$

or, in other words,

$$-\sigma \|z - \bar{x}\|^2 \leq \theta|_C(z) - \theta|_C(\bar{x}) - \langle p, z - \bar{x} \rangle$$

for all $z \in H$ with $\|z - \bar{x}\| \leq \eta$ that means $p \in \partial^p(\theta|_C)(\bar{x})$.

Thus, joining Steps 1 and 2 we conclude that

$$\partial^p \hat{u}(x) \subset \partial \rho_F(x - \bar{x}) \cap \partial^p(\theta|_C)(\bar{x}) \quad (88)$$

for each x close to \hat{x} .

Step 3. Let us prove now a kind of opposite inclusion

$$\partial \rho_F(x - \bar{x}) \cap \partial^p(\theta|_C)(\bar{x}) \subset \partial^- \hat{u}(x) \quad (89)$$

but only at the point $x = \hat{x}$. To this end fix p from the left-hand side of (89). Then, in particular,

$$\rho_F(\xi) \geq \rho_F(\hat{x} - \bar{x}) + \langle p, \xi - \hat{x} + \bar{x} \rangle \quad (90)$$

for all $\xi \in H$. Given $x \notin C$ sufficiently close to \hat{x} let us set $\xi = x - \bar{x}$ in (90) and rewrite the latter inequality as

$$\rho_F(x - \bar{x}) - \rho_F(\hat{x} - \bar{x}) - \langle p, x - \hat{x} \rangle \geq -\langle p, \bar{x} - \hat{x} \rangle. \quad (91)$$

On the other hand, one can choose $\eta > 0$ and $\sigma > 0$ such that

$$\theta(z) \geq \theta(\bar{x}) + \langle p, z - \bar{x} \rangle - \sigma \|z - \bar{x}\|^2 \quad (92)$$

whenever $z \in C$ with $\|z - \bar{x}\| \leq \eta$. Due to the continuity of the mapping $\pi_C^{F, \theta}(\cdot)$ at \hat{x} the inequality (92) holds for $z = \bar{x}$ with x enough close to \hat{x} . Combining this with both (91) and the condition (85), we successively obtain:

$$\begin{aligned}
 & \liminf_{x \rightarrow \hat{x}} \frac{\hat{u}(x) - \hat{u}(\hat{x}) - \langle p, x - \hat{x} \rangle}{\|x - \hat{x}\|} \\
 &= \liminf_{x \rightarrow \hat{x}} \frac{\rho_F(x - \bar{x}) + \theta(\bar{x}) - \rho_F(\hat{x} - \bar{x}) - \theta(\bar{x}) - \langle p, x - \hat{x} \rangle}{\|x - \hat{x}\|} \\
 &\geq \liminf_{x \rightarrow \hat{x}} \frac{-\langle p, \bar{x} - \hat{x} \rangle + \theta(\bar{x}) - \theta(\hat{x})}{\|x - \hat{x}\|} \\
 &\geq \liminf_{x \rightarrow \hat{x}} \frac{-\sigma \|\bar{x} - \hat{x}\|^2}{\|x - \hat{x}\|} \geq -\sigma \left(\lim_{r \rightarrow 0^+} \frac{\omega(\hat{x}, r)}{\sqrt{r}} \right)^2 = 0,
 \end{aligned}$$

i.e., $p \in \partial^- \hat{u}(\hat{x})$.

Taking into account the proximal regularity of $\theta|_C$ at the point \bar{x} and the inclusions (20) we deduce from (89) that

$$\partial \rho_F(\hat{x} - \bar{x}) \cap \partial^- (\theta|_C)(\bar{x}) \subset \partial^- \hat{u}(\hat{x}) \subset \partial^c \hat{u}(\hat{x}). \quad (93)$$

Step 4. To complete the proof recall the representation formula (22) for the Clarke subdifferential through the proximal subgradients in neighbour points and apply the inclusion (88). So, $\partial^c \hat{u}(\hat{x})$ is contained in the closed convex hull of the set of all weak limits of sequences $\zeta_i \in \partial \rho_F(x_i - \bar{x}_i) \cap \partial^p (\theta|_C)(\bar{x}_i)$ such that $x_i \rightarrow \hat{x}$ as $i \rightarrow \infty$. Furthermore, since $\bar{x}_i \rightarrow \bar{x}$, the subdifferential of a convex function has *s×w sequentially closed graph* and

$$w\text{-} \limsup_{x \rightarrow \hat{x}, x \in C} \partial^p (\theta|_C)(x) = \partial^l (\theta|_C)(\bar{x})$$

(see [31, p. 240]), we have

$$\begin{aligned}
 \partial^c \hat{u}(\hat{x}) &\subset \overline{\text{co}} \left(\partial \rho_F(\hat{x} - \bar{x}) \cap \partial^l (\theta|_C)(\bar{x}) \right) \\
 &= \partial \rho_F(\hat{x} - \bar{x}) \cap \partial^- (\theta|_C)(\bar{x}), \quad (94)
 \end{aligned}$$

where the latter equality follows from the lower regularity of the function $\theta|_C$ (it is a consequence of the proximal regularity) and from the convexity of the Fréchet subdifferential $\partial^- (\theta|_C)(\bar{x})$. Hence, in particular, the right-hand side of (94) is nonempty because $\hat{u}(\cdot)$ is a lipschitzean function.

Combining now (94) with (93) proves the Theorem. \square

Corollary 1. *If the condition (85) is fulfilled not only at the point \hat{x} itself but at each $x \in U(\hat{x})$ (in particular, if $\pi_C^{F, \theta}(\cdot)$ is Hölder continuous with an exponent $\beta > 1/2$ on this neighbourhood) then the equality*

$$\partial^c \hat{u}(\hat{x}) = \partial \rho_F(\hat{x} - \bar{x}) \cap \partial^- (\theta|_C)(\bar{x}) \quad (95)$$

takes place whenever $\theta|_C$ is just lower regular at \bar{x} . If, moreover, $\theta|_C$ is lower regular at each point of ∂C close to \bar{x} then the same equality as (95) holds also at each $x \in U(\hat{x})$.

Proof. Indeed, under the assumptions of the Corollary the inclusion (89) is valid at all points x close to \hat{x} , and passing to the *weak Kuratowski-Painlevé upper limits* we have

$$\begin{aligned} \partial\rho_F(\hat{x} - \bar{x}) \cap \partial^-(\theta|_C)(\hat{x}) &= \partial\rho_F(\hat{x} - \bar{x}) \cap \partial^l(\theta|_C)(\hat{x}) \\ &\subset \partial^l\hat{u}(\hat{x}) \subset \partial^c\hat{u}(\hat{x}), \end{aligned}$$

while the opposite inclusion is already proved (see (94)). The second assertion is obvious. \square

Observe that in the framework of Corollary 1 by strengthening the condition (85) we can achieve the Clarke regularity of $\hat{u}(\cdot)$ as well.

Proposition 1. *Given $\hat{x} \notin C$ let us assume that the mapping $x \mapsto \pi_C^{F,\theta}(x)$ is single-valued in a neighbourhood $U(\hat{x})$ of \hat{x} and satisfies the Lipschitz type inequality*

$$\left\| \pi_C^{F,\theta}(x) - \pi_C^{F,\theta}(\hat{x}) \right\| \leq L \|x - \hat{x}\| \quad (96)$$

for all $x \in U(\hat{x})$ with some constant $L > 0$. If, moreover, the restriction $\theta|_C$ is lower regular at $\bar{x} := \pi_C^{F,\theta}(\hat{x})$ then the statement of Theorem 4 holds.

Proof. The equalities in (86) split essentially into two inclusions. The first one is (94), which is obtained by using only the lower regularity of $\theta|_C$, and the second is

$$\partial\rho_F(\hat{x} - \bar{x}) \cap \partial^-(\theta|_C)(\bar{x}) \subset \partial^-\hat{u}(\hat{x}) \quad (97)$$

(compare with (89)).

Taking $p \in \partial\rho_F(\hat{x} - \bar{x}) \cap \partial^-(\theta|_C)(\bar{x})$ similarly as in the proof of Theorem 4 (see Step 3) we write the inequality (91) for all $x \notin C$ sufficiently close to \hat{x} . Furthermore, given $\varepsilon > 0$ we choose $\eta > 0$ such that

$$\theta(z) \geq \theta(\bar{x}) + \langle p, z - \bar{x} \rangle - \frac{\varepsilon}{L} \|z - \bar{x}\| \quad (98)$$

whenever $z \in C$ with $\|z - \bar{x}\| \leq \eta$ (compare with (92)). By the continuity of $\pi_C^{F,\theta}(\cdot)$ let us choose $\delta > 0$ such that $\|\bar{x} - \hat{x}\| \leq \eta$ whenever $\|x - \hat{x}\| \leq \delta$. Setting then $z = \bar{x}$ in (98) and taking into account the inequality (96) we have:

$$\theta(\bar{x}) \geq \theta(\hat{x}) + \langle p, \bar{x} - \hat{x} \rangle - \varepsilon \|x - \hat{x}\|. \quad (99)$$

Joining together (91) and (99) we obtain that

$$\begin{aligned} \hat{u}(x) - \hat{u}(\hat{x}) - \langle p, x - \hat{x} \rangle &= \rho_F(x - \bar{x}) + \theta(\bar{x}) \\ &\quad - \rho_F(\hat{x} - \bar{x}) - \theta(\bar{x}) - \langle p, x - \hat{x} \rangle \\ &\geq -\langle p, \bar{x} - \hat{x} \rangle + \theta(\bar{x}) - \theta(\hat{x}) \geq -\varepsilon \|x - \hat{x}\| \end{aligned}$$

for all x , $\|x - \hat{x}\| \leq \delta$. So, $p \in \partial^-\hat{u}(\hat{x})$, and the inclusion (97) is proved. \square

We see that under the assumption of Corollary 1 the question of the Fréchet continuous differentiability of the value function $\hat{u}(\cdot)$ is reduced to the single-valuedness and the continuity of the mapping

$$\Phi(x) := \partial^-(\theta|_C)(\bar{x}) \cap \partial\rho_F(x - \bar{x}) \quad (100)$$

near a given point \hat{x} . Indeed, if $\Phi(x) = \partial^c \hat{u}(x)$ is a singleton then by [10, Proposition 2.2.4] $\hat{u}(\cdot)$ is *strictly differentiable* and $\partial^c \hat{u}(x) = \{\nabla_H \hat{u}(x)\}$, where $\nabla_H \hat{u}(x)$ stands for the (strict) *Hadamard derivative* coinciding with the Fréchet one by the continuity. Observe that in finite dimensions the mapping (100) is continuous as soon as it is single-valued. This follows from the lower regularity of $\theta|_C$ and from the properties of the subdifferentials $\partial^l(\theta|_C)$ and $\partial\rho_F$.

Thus, recalling (12) we have a representation formula for the gradient $\nabla \hat{u}(\cdot)$ in a neighbourhood of \hat{x} through the (unique) minimizer \bar{x} of the function $y \mapsto \rho_F(x - y) + \theta(y)$ on C :

$$\nabla \hat{u}(x) = \partial^-(\theta|_C)(\bar{x}) \cap \mathbf{N}_F\left(\frac{x - \bar{x}}{\rho_F(x - \bar{x})}\right) \cap \partial F^0. \quad (101)$$

Although the condition guaranteeing the continuous Fréchet differentiability in such a form (single-valuedness and continuity of the mapping (100)) and the formula (101) have a certain theoretical interest, their practical applicability is very restrictive because they are given in terms of an *a priori* unknown minimizer. To overcome this difficulty we propose first an alternate hypothesis regarding the regularity properties of either the function $\theta|_C$ or the gauge F . Namely, observing that the right-hand side in (101) is reduced to a singleton whenever either the Fréchet subdifferential $\partial^-(\theta|_C)(\bar{x})$ or the normal cone $\mathbf{N}_F\left(\frac{x - \bar{x}}{\rho_F(x - \bar{x})}\right)$ (both unbounded) becomes a semiline, we arrive at the following result.

Theorem 5. *Given $\hat{x} \in H \setminus C$ assume that in some neighbourhood $U(\hat{x})$ of \hat{x} the mapping $x \mapsto \pi_C^{F,\theta}(x)$ is single-valued and Hölder continuous with an exponent $\beta > 1/2$, and that the restriction $\theta|_C$ is lower regular at $\bar{x} = \pi_C^{F,\theta}(x)$ for each x close to \hat{x} . Then the function $\hat{u}(\cdot)$ is (Fréchet) continuously differentiable on $U(\hat{x})$ if at least one of the conditions below holds:*

- (i) F is smooth at $\xi := \frac{x - \bar{x}}{\rho_F(x - \bar{x})}$ for each $x \in U(\hat{x})$;
- (ii) C has smooth boundary at \bar{x} , and the function $\theta(\cdot)$ is of class C^1 near this point.

Furthermore, in the first case

$$\nabla \hat{u}(x) = \nabla \rho_F(x - \bar{x}) \quad (102)$$

(it coincides with the unique normal vector to F at the point ξ , belonging to the boundary ∂F^0), while in the second

$$\nabla \hat{u}(x) = \nabla \theta(\bar{x}) + \lambda(\bar{x}) \mathbf{n}_C(\bar{x}), \quad (103)$$

where $\lambda = \lambda(\bar{x}) > 0$ is the unique positive root of the equation

$$\rho_{F^0}(\nabla \theta(\bar{x}) + \lambda \mathbf{n}_C(\bar{x})) = 1, \quad (104)$$

and $\mathbf{n}_C(\bar{x})$ is the (unique) unit normal vector to C at \bar{x} .

Proof. The first assertion follows directly from (101) because the smoothness of F at ξ means exactly that the convex subdifferential $\partial\rho_F(\xi)$ is reduced to the unique point $\xi^* \in \mathfrak{J}_{F^0}(\xi) = \mathbf{N}_F(\xi) \cap \partial F^0$ (see (10)), which is nothing else than the Fréchet gradient $\nabla \rho_F(x - \bar{x})$. The continuity of $\nabla \hat{u}(\cdot)$ instead follows from both [25, Proposition 3.3 (ii)] and the continuity of $\pi_C^{F,\theta}(\cdot)$.

In the second case let us decompose the Fréchet subdifferential of $\theta|_C$ into the sum of the gradient $\nabla\theta(\bar{x})$ and the (Fréchet) normal cone to C (see [31, p. 112]):

$$\partial^-(\theta|_C)(\bar{x}) = \nabla\theta(\bar{x}) + \mathbf{N}_C^\sigma(\bar{x}).$$

Since the boundary ∂C is assumed to be smooth at \bar{x} , and \bar{x} is close to \hat{x} whenever $x \in U(\hat{x})$, we have that

$$\mathbf{N}_C^\sigma(\bar{x}) = \mathbf{N}_C^l(\bar{x}) = \{\lambda \mathbf{n}_C(\bar{x}) : \lambda \geq 0\},$$

where $\mathbf{n}_C(\cdot)$ is a continuous function defined on ∂C near \bar{x} with $\|\mathbf{n}_C(\bar{x})\| = 1$. So, the intersection $\partial^-(\theta|_C)(\bar{x}) \cap \partial F^0$ is the singleton $\nabla\theta(\bar{x}) + \lambda(\bar{x}) \mathbf{n}_C(\bar{x})$ where $\lambda(\bar{x}) > 0$ can be uniquely determined from the equation (104), and the gradient $\nabla\hat{u}(x)$ takes the form (103) (see (101)). In order to show continuity let us fix a sequence $\{x_n\}$ converging to $x \in U(\hat{x})$. Then $\{\bar{x}_n\}$ converges to $\bar{x} \in \partial C$. Let us denote by $\lambda_n = \lambda(\bar{x}_n)$ the respective positive root of (104) and observe that the sequence $\{\lambda_n\}$ is bounded. Consequently, some of its subsequences (assume that $\{\lambda_n\}$ itself) converges to $\bar{\lambda} \geq 0$. Passing to limit in the equality

$$\rho_{F^0}(\nabla\theta(\bar{x}_n) + \lambda_n \mathbf{n}_C(\bar{x}_n)) = 1$$

and using continuity of the involved functions we arrive at

$$\rho_{F^0}(\nabla\theta(\bar{x}) + \bar{\lambda} \mathbf{n}_C(\bar{x})) = 1.$$

Hence $\bar{\lambda} > 0$ (see (48)), and by the uniqueness $\bar{\lambda} = \lambda(\bar{x})$. Therefore $\nabla\hat{u}(x_n) \rightarrow \nabla\hat{u}(x)$, and the continuity is proved. \square

Recall now that under the standing assumptions (\mathbf{H}) , $(\hat{\mathbf{H}})$ and the local hypotheses $(\mathbf{H}_1(x_0))$, $(\mathbf{H}_2(x_0))$ the (single-valued) mapping $x \mapsto \pi_C^{F,\theta}(x)$ is locally lipschitzean, so satisfies the hypotheses of both Corollary 1 and Proposition 1 near a fixed point $x_0 \in C$ (see Theorem 3). Taking into account that $\pi_C^{F,\theta}(x)$ is close to x_0 whenever x approaches x_0 , we give a version of the regularity theorem, which do not use explicitly the minimizers.

Theorem 6. *Let us fix $x_0 \in \partial C$ and suppose all the hypotheses of Theorem 3 to be valid. Assume, in addition, that $\theta|_C$ is lower regular near x_0 , and that for some $\delta > 0$ at least one of the conditions below is fulfilled:*

- (i) F is smooth at each $\xi \in \mathfrak{J}_F(\partial^-(\theta|_C)(x) \cap \partial F^0)$, $x \in \partial C$, $\|x - x_0\| \leq \delta$;
- (ii) C has smooth boundary, and $\theta(\cdot)$ is of class \mathcal{C}^1 on $\partial C \cap (x_0 + \delta \bar{\mathbf{B}})$.

Then the marginal function $\hat{u}(\cdot)$ is (Fréchet) continuously differentiable on a neighbourhood of x_0 (outside of C), and the gradient $\nabla\hat{u}(x)$ can be computed by the formula (102) or (103), respectively.

Proof. By Theorem 3 there exists a neighbourhood $\mathcal{U}(x_0)$ of x_0 such that for each $x \in \mathcal{U}(x_0)$ the set $\pi_C^{F,\theta}(x)$ is a singleton, say $\{\bar{x}\}$, and $\|\bar{x} - x_0\| \leq \delta$. Now, in the case (i) we apply Corollary 1 at the point $x \in \mathcal{U}(x_0) \setminus C$ and consider the unique vector

$$\xi^* \in \partial^-(\theta|_C)(\bar{x}) \cap \mathbf{N}_F(\xi) \cap \partial F^0$$

where $\xi := \frac{x - \bar{x}}{\rho_F(x - \bar{x})}$ (see (101)). In particular, $\xi^* \in \mathfrak{J}_{F^0}(\xi)$, or dually $\xi \in \mathfrak{J}_F(\xi^*)$. So, we are led to the hypothesis (i) of Theorem 5. The case (ii) instead is directly reduced to Theorem 5 (ii). \square

By duality, in the place of the *smoothness* of F we may require here the *rotundity* of F^0 with respect to each $\xi \in \mathfrak{J}_F(\partial^-(\theta|_C)(x) \cap \partial F^0)$ (compare with the condition $(\mathbf{H}_2(x_0))$).

If the hypotheses $(\mathbf{H}_1(x_0))$ and $(\mathbf{H}_2(x_0))$ hold at each point $x_0 \in \partial C$, and the restriction $\theta|_C$ is lower regular everywhere on the boundary ∂C , then in order to have the continuous differentiability of $\hat{u}(\cdot)$ in some neighbourhood $\mathfrak{A} \supset C$ (outside of C) one can alternate the conditions **(i)** and **(ii)** from one point $x_0 \in \partial C$ to other.

In conclusion let us strengthen the hypotheses on F , C and $\theta(\cdot)$ in order to have more regularity for the value function $\hat{u}(\cdot)$. Remind that in the case $\theta \equiv 0$ and $F = \bar{\mathbf{B}}$ under well-posedness assumptions, which are reduced to the φ -convexity of C , the function $\hat{u}(\cdot) = d_C(\cdot)$ is of class $\mathcal{C}^{1,1}$ near C , while in the case of an arbitrary gauge the lipschitzeanity (hölderianity, in general) of $\nabla \hat{u}(\cdot)$ depends on the order of smoothness of the input data (see [26, Theorems 5.6 and 5.7]). The same happens in the case $\theta \neq 0$.

Theorem 7. *Given $x_0 \in \partial C$ let us assume that all the hypotheses of Theorem 6 hold. Moreover, suppose that in the case **(i)** the gradient $\nabla \rho_F(\cdot)$ is Hölder continuous with an exponent $0 < \alpha \leq 1$ on the set*

$$\mathfrak{M}_\delta(x_0) := \bigcup_{x \in \partial C, \|x - x_0\| \leq \delta} \mathfrak{J}_F(\partial^-(\theta|_C)(x) \cap \partial F^0)$$

(equivalently, the unit normal vector to F moves in a hölderean way along the part $\mathfrak{M}_\delta(x_0)$ of the boundary ∂F), while in the case **(ii)** both the gradient $\nabla \theta(\cdot)$ and the normal $\mathbf{n}_C(\cdot)$ are Hölder continuous (with the exponent $0 < \alpha \leq 1$) near x_0 . Then the value function $\hat{u}(\cdot)$ is of class $\mathcal{C}_{loc}^{1,\alpha}$ in a neighbourhood of x_0 (outside of C).

Proof. After application of Theorem 6 the proving consists in the verification (locally) the Hölder inequality for the gradient $\nabla \hat{u}(x)$ in both cases. Let $\mathcal{U}(x_0)$ be the neighbourhood of x_0 constructed in Theorem 3. Without loss of generality assume that $\delta > 0$ from the formula (65) is the same as in the conditions **(i)** and **(ii)** of Theorem 6. Fix $x \in \mathcal{U}(x_0) \setminus C$ and choose $\bar{\delta} > 0$ such that $x + \bar{\delta}\bar{\mathbf{B}} \subset \mathcal{U}(x_0) \setminus C$. We have $\mathcal{U}(x_0) \subset x_0 + \delta\bar{\mathbf{B}}$. Moreover, for each $z \in x + (\bar{\delta}/2)\bar{\mathbf{B}}$ the (unique) minimizer $\bar{z} := \pi_C^{F,\theta}(z)$ also belongs to $x_0 + \delta\bar{\mathbf{B}}$ as shown in the first part of the proof of Theorem 3. Now, let us consider the respective estimates in each case separately.

(i) Given $z_1, z_2 \in x + (\bar{\delta}/2)\bar{\mathbf{B}}$ we denote by

$$\xi_i := \frac{z_i - \bar{z}_i}{\rho_F(z_i - \bar{z}_i)}, \quad \bar{z}_i := \pi_C^{F,\theta}(z_i), \quad i = 1, 2,$$

and by the positive homogeneity of the Minkowski functional deduce from (102) that

$$\|\nabla \hat{u}(z_1) - \nabla \hat{u}(z_2)\| \leq \mathfrak{h} \|\xi_1 - \xi_2\|^\alpha, \quad (105)$$

where $\mathfrak{h} > 0$ is the Hölder constant of $\nabla \rho_F(\cdot)$ on $\mathfrak{M}_\delta(x_0)$. Setting for the sake of brevity $\rho_i := \rho_F(z_i - \bar{z}_i)$, we further have

$$\begin{aligned} \|\xi_1 - \xi_2\| &\leq \frac{1}{\rho_1 \rho_2} (\|\bar{z}_1 - z_1\| |\rho_2 - \rho_1| + \rho_1 \|(\bar{z}_1 - \bar{z}_2) + (z_2 - z_1)\|) \\ &\leq \frac{1}{\rho_2} (\|F\| \|F^0\| + 1) \|(\bar{z}_1 - \bar{z}_2) + (z_2 - z_1)\| \\ &\leq \frac{1}{\rho_2} (\|F\| \|F^0\| + 1) (\|\bar{z}_1 - \bar{z}_2\| + \|z_1 - z_2\|). \end{aligned} \quad (106)$$

Notice that $\|z_2 - \bar{z}_2\| > \bar{\delta}/2$ because otherwise $\|\bar{z}_2 - x\| \leq \bar{\delta}$, contradicting the choice of $\bar{\delta} > 0$. Consequently, $\rho_2 \geq \frac{\bar{\delta}}{2\|F\|}$ (see (9)). Hence, using the Lipschitz continuity of $\pi_C^{F,\theta}(\cdot)$ (with the Lipschitz constant $K > 0$) we obtain from (106) that

$$\|\xi_1 - \xi_2\| \leq \frac{2\|F\|(\|F\|\|F^0\| + 1)(K+1)}{\bar{\delta}} \|z_1 - z_2\|.$$

Joining the latter inequality with (105) we arrive at

$$\|\nabla\hat{u}(z_1) - \nabla\hat{u}(z_2)\| \leq \mathfrak{H} \|z_1 - z_2\|^\alpha, \quad (107)$$

where the constant $\mathfrak{H} > 0$ essentially depends on x (through $\bar{\delta}$ and K) and tends to $+\infty$ as the point x approaches the target C .

- (ii) In this case we prove a Hölder inequality like (107) in the neighbourhood $x + \bar{\delta}\bar{\mathbf{B}}$. To this end we apply the Lipschitz continuity of $\pi_C^{F,\theta}(\cdot)$ on $\mathcal{U}(x_0) \supset x + \bar{\delta}\bar{\mathbf{B}}$ and the Hölder continuity of both $\nabla\theta(\cdot)$ and $\mathbf{n}_C(\cdot)$ on $\partial C \cap (x_0 + \bar{\delta}\bar{\mathbf{B}})$. Let us take $z_1, z_2 \in x + \bar{\delta}\bar{\mathbf{B}}$ and set as usual $\bar{z}_i := \pi_C^{F,\theta}(z_i)$, $i = 1, 2$. Then it follows from (103) that

$$\begin{aligned} \|\nabla\hat{u}(z_1) - \nabla\hat{u}(z_2)\| &\leq \|\nabla\theta(\bar{z}_1) - \nabla\theta(\bar{z}_2)\| + |\lambda(\bar{z}_1) - \lambda(\bar{z}_2)| \\ &\quad + \lambda(\bar{z}_1) \|\mathbf{n}_C(\bar{z}_1) - \mathbf{n}_C(\bar{z}_2)\| \end{aligned} \quad (108)$$

where $\lambda(\bar{z}_i) > 0$, $i = 1, 2$, satisfy the equality

$$\rho_{F^0}(\nabla\theta(\bar{z}_i) + \lambda(\bar{z}_i)\mathbf{n}_C(\bar{z}_i)) = 1. \quad (109)$$

Notice that (109) is equivalent to

$$\frac{1}{\lambda(\bar{z}_i)} = \rho_{F^0 - \nabla\theta(\bar{z}_i)}(\mathbf{n}_C(\bar{z}_i)).$$

Then, due to (9) and to the hypothesis **(H)** (see (48))

$$\lambda(\bar{z}_i) \leq \|F^0 - \nabla\theta(\bar{z}_i)\| \leq (1 + \gamma) \|F^0\|. \quad (110)$$

On the other hand, by the Lipschitz continuity of the gauge function, (8), (53), (16), (18) and (52) we successively have:

$$\begin{aligned} \left| \frac{1}{\lambda(\bar{z}_1)} - \frac{1}{\lambda(\bar{z}_2)} \right| &\leq \left| \rho_{F^0 - \nabla\theta(\bar{z}_1)}(\mathbf{n}_C(\bar{z}_1)) - \rho_{F^0 - \nabla\theta(\bar{z}_1)}(\mathbf{n}_C(\bar{z}_2)) \right| \\ &\quad + \left| \rho_{F^0 - \nabla\theta(\bar{z}_1)}(\mathbf{n}_C(\bar{z}_2)) - \rho_{F^0 - \nabla\theta(\bar{z}_2)}(\mathbf{n}_C(\bar{z}_2)) \right| \\ &\leq \left\| (F^0 - \nabla\theta(\bar{z}_1))^0 \right\| \|\mathbf{n}_C(\bar{z}_1) - \mathbf{n}_C(\bar{z}_2)\| + \\ &\quad \left| \sigma_{(F^0 - \nabla\theta(\bar{z}_1))^0}(\mathbf{n}_C(\bar{z}_2)) - \sigma_{(F^0 - \nabla\theta(\bar{z}_2))^0}(\mathbf{n}_C(\bar{z}_2)) \right| \\ &\leq \frac{\|F\|}{1 - \gamma} \|\mathbf{n}_C(\bar{z}_1) - \mathbf{n}_C(\bar{z}_2)\| + \\ &\quad + \mathcal{D}\left((F^0 - \nabla\theta(\bar{z}_1))^0, (F^0 - \nabla\theta(\bar{z}_2))^0\right) \\ &\leq \frac{\|F\|}{1 - \gamma} \|\mathbf{n}_C(\bar{z}_1) - \mathbf{n}_C(\bar{z}_2)\| \\ &\quad + \left(\frac{\|F\|}{1 - \gamma} \right)^2 \|\nabla\theta(\bar{z}_1) - \nabla\theta(\bar{z}_2)\|. \end{aligned} \quad (111)$$

Applying the Hölder inequality for both $\mathbf{n}_C(\cdot)$, $\nabla\theta(\cdot)$ with the exponent $0 < \alpha \leq 1$ and a Hölder constant $\mathfrak{h} > 0$ we obtain from (110) and (111) that

$$\begin{aligned} |\lambda(\bar{z}_1) - \lambda(\bar{z}_2)| &= \lambda(\bar{z}_1)\lambda(\bar{z}_2) \left| \frac{1}{\lambda(\bar{z}_1)} - \frac{1}{\lambda(\bar{z}_2)} \right| \\ &\leq \frac{(1+\gamma)^2}{1-\gamma} \left(1 + \frac{\|F\|}{1-\gamma} \right) \|F\| \|F^0\|^2 \mathfrak{h} \|\bar{z}_1 - \bar{z}_2\|^\alpha. \end{aligned} \quad (112)$$

Again by the Hölder continuity of the functions $\nabla\theta(\cdot)$ and $\mathbf{n}_C(\cdot)$ and by the inequalities (108), (110) and (112) it follows that

$$\|\nabla\hat{u}(z_1) - \nabla\hat{u}(z_2)\| \leq \bar{\mathfrak{h}} \|\bar{z}_1 - \bar{z}_2\|^\alpha$$

with some constant $\bar{\mathfrak{h}} > 0$, which is proportional to \mathfrak{h} . Recalling also the Lipschitz inequality for minimizers (with a Lipschitz constant $K > 0$) we arrive, finally, at (107) where $\mathfrak{H} := \bar{\mathfrak{h}}K^\alpha$.

So the Theorem is completely proved. \square

Observe that unlike the item (i) the Hölder constant in the case (ii) depends possibly on how close to the boundary ∂C the point x is just through the Lipschitzianity of the mapping $\pi_C^{F,\theta}(\cdot)$ and the Hölderianity of both $\nabla\theta(\cdot)$ and $\mathbf{n}_C(\cdot)$ (remind that in the case of metric projections onto convex and prox-regular sets the gradient $\nabla\hat{u}(x)$ exactly coincides with the unit normal $\mathbf{n}_C(\bar{x})$). However, in both cases \mathfrak{H} tends to $+\infty$ as the point x approaches the boundary of the neighbourhood $\mathcal{U}(x_0)$, where the regularity hypotheses $(\mathbf{H}_1(x_0))$ and $(\mathbf{H}_2(x_0))$ fail (see Remark 4).

Remark 6. Notice that Theorems 5 – 7 have been proved under the lower regularity assumption for the function $\theta|_C$, while in [26] we supposed the target set C to be proximal regular that is a much stronger property.

8 Examples

In this section we illustrate the obtained results with two simple examples restricting ourselves just to the case $H = \mathbb{R}^2$. We recommend to compare them with the examples given earlier for the case $\theta \equiv 0$ (see [25, 26]).

Example 1.

$$\begin{aligned} F &:= \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \leq 1 - \xi_1^4, -1 \leq \xi_1 \leq 1\}; \\ C &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq x_2^2\}; \\ \theta(x) &:= \frac{1}{2} \operatorname{arctg}(x_1^2 + x_2^2), x = (x_1, x_2) \in C. \end{aligned}$$

In [25, Example 8.3] the well-posedness problem for the same target set C and the same dynamics F with $\theta \equiv 0$ was considered. Now we somewhat complicate the problem by introducing a (smooth) nonlinear boundary function.

Let us verify first the standing slope condition (24), guaranteeing that the value function $\hat{u}(\cdot)$ is the viscosity solution of the associated Hamilton-Jacobi equation (5) with the boundary datum $u|_C = \theta$ (see Theorem 1). To this end we calculate the gradient

$$\nabla\theta(x) = \left(\frac{x_1}{1 + (x_1^2 + x_2^2)^2}, \frac{x_2}{1 + (x_1^2 + x_2^2)^2} \right)$$

and the dual gauge function

$$\rho_{F^0}(\xi^*) = \sigma_F(\xi^*) = \begin{cases} 3 \frac{(|\xi_1^*|/4)^{4/3}}{|\xi_2^*|^{1/3}} + |\xi_2^*| & \text{if } |\xi_2^*| \geq |\xi_1^*|/4; \\ |\xi_1^*| & \text{if } |\xi_2^*| < |\xi_1^*|/4, \end{cases} \quad (113)$$

$\xi^* = (\xi_1^*, \xi_2^*) \in \mathbb{R}^2$. Comparing the radii of two circles centred at the origin, one inscribed into F and other circumscribed around it, we find that $\|F\| \leq 7/6$ and $\|F^0\| \leq 9/8$. Substituting $\nabla\theta(x)$ in the place of ξ^* in (113) we obviously have

$$\rho_{F^0}(\nabla\theta(x)) \leq \begin{cases} \frac{3}{4} \frac{|x_1|}{1 + (x_1^2 + x_2^2)^2} + \frac{|x_2|}{1 + (x_1^2 + x_2^2)^2} & \text{if } |x_2| \geq \frac{|x_1|}{4}; \\ \frac{|x_1|}{1 + (x_1^2 + x_2^2)^2} & \text{if } |x_2| < \frac{|x_1|}{4}. \end{cases} \quad (114)$$

The function in the right-hand side of (114) attains its maximum at the point $\left(\frac{3^{3/4}}{5}, \frac{4}{3^{1/4 \cdot 5}}\right)$, and the maximum is $\frac{5 \cdot 3^{3/4}}{16} < \frac{35}{48}$. Thus, $\nabla\theta(x) \in \gamma F^0$ with

$$\gamma := \frac{35}{48} < \frac{1}{\|F\| \|F^0\|}, \quad (115)$$

and we have not merely the slope condition (24) but also the (stronger) standing hypothesis **(H)** required for the well-posedness and regularity results. Moreover, by [12, p. 38] the second standing assumption **(H̃)** holds as well with $\Gamma(x) = \{\nabla\theta(x)\}$ and

$$\begin{aligned} \mathbf{N}_C^\theta(x) &= \mathbf{N}_C^p(x) = \mathbf{N}_C^l(x) \\ &= \{\lambda \mathbf{n}_C(x) : \lambda \geq 0\} = \{(\lambda, -2\lambda x_2) : \lambda \geq 0\}, \end{aligned}$$

where $\mathbf{n}_C(x)$ is the unit normal vector to C ,

$$\mathbf{n}_C(x) = \frac{1}{\sqrt{1 + 4x_2^2}} (1, -2x_2),$$

$x = (x_1, x_2) \in \partial C$. Thus,

$$\partial^p(\theta|_C)(x) = \left\{ \left(\frac{x_1}{1 + (x_1^2 + x_2^2)^2} + \lambda, \frac{x_2}{1 + (x_1^2 + x_2^2)^2} - 2\lambda x_2 \right) : \lambda \geq 0 \right\}.$$

Let us fix now $x_0 = (x_1^0, x_2^0) \in \partial C$ and verify the local hypotheses of Theorem 3, or, rather, just the hypothesis **(H₁)(x₀)** according to the observation before Theorem 3. To this end we compute first the value $\mathfrak{J}_F(\xi^*)$, $\xi^* = (\xi_1^*, \xi_2^*) \in \partial F^0$, restricting

ourselves just to the case $\xi_1^* > 0$, since the first coordinate of the (unique) element of $\partial^p(\theta|_C)(x) \cap \partial F^0$, $x \in C$, is positive. By the formula (10) due to the continuous differentiability of $\rho_{F^0}(\cdot)$ we have that $\mathfrak{J}_F(\xi^*) = \{\nabla \rho_{F^0}(\xi^*)\}$, while the direct derivation of (113) gives

$$\nabla \rho_{F^0}(\xi^*) = \left(f\left(\frac{\xi_1^*}{4|\xi_2^*|}\right), \operatorname{sgn}(\xi_2^*) g\left(\frac{\xi_1^*}{4|\xi_2^*|}\right) \right), \quad (116)$$

where $f(\cdot)$ and $g(\cdot)$ are real functions defined on $]0, +\infty[$ by

$$f(t) := \begin{cases} t^{1/3} & \text{if } 0 < t \leq 1, \\ 1 & \text{if } t > 1, \end{cases} \quad (117)$$

$$g(t) := \begin{cases} 1 - t^{4/3} & \text{if } 0 < t \leq 1, \\ 0 & \text{if } t > 1. \end{cases} \quad (118)$$

Then we substitute in the place of ξ^* in (116) the unique subgradient of $\theta|_C$ belonging to ∂F^0 , i.e.,

$$\xi_1^* = \frac{x_1}{1 + (x_1^2 + x_2^2)^2} + \lambda; \quad (119)$$

$$\xi_2^* = \frac{x_2}{1 + (x_1^2 + x_2^2)^2} - 2\lambda x_2, \quad (120)$$

where $\lambda = \lambda(x)$ is the (unique) positive root of the equation $\rho_{F^0}(\xi_1^*, \xi_2^*) = 1$, $x \in \partial C$ (i.e., $x_1 = x_2^2$) with $\|x - x_0\| \leq \delta$. Similarly as in the proof of Theorem 7 (see (110) and (53)) we obtain the following estimates for the parameter λ (see also (115)):

$$\frac{1}{5} < \frac{13}{56} \leq \frac{1-\gamma}{\|F\|} \leq \lambda(x) \leq (1+\gamma)\|F^0\| \leq \frac{249}{128} < 2 \quad (121)$$

and establish the lipschitzeanity of the function $\lambda(\cdot)$ near x_0 (see (112)). It follows from (119)-(121) that

- $\frac{\xi_1^*}{4|\xi_2^*|} \geq 1$ whenever $|\xi_2^*| \leq \frac{1}{20}$;
- $\frac{\xi_1^*}{4|\xi_2^*|} \geq \frac{\lambda(x)}{4\|F^0\|} \geq s := \frac{2}{45}$ whenever $x \in \partial C$.

Taking this into account and observing that the functions (117) and (118) are lipschitzean on $[s, +\infty[$ with the Lipschitz constant $1/3 \max\{4, s^{-2/3}\}$, that they are constant for $t \geq 1$, and that the mapping $\xi^* \mapsto \frac{\xi_1^*}{4|\xi_2^*|}$ is lipschitzean on the set

$$\left\{ \xi^* \in \partial F^0 : |\xi_2^*| \geq \frac{1}{20} \right\},$$

we conclude that the gradient $\nabla \rho_{F^0}(\cdot)$ (see (116)) is lipschitzean on

$$\left\{ \xi^* \in \partial F^0 : \frac{\xi_1^*}{4|\xi_2^*|} \geq s \right\}.$$

Consequently, estimating further the second derivative of the function $\theta(\cdot)$ we obtain the lipschitzeanity of the composed mapping

$$x \mapsto \nabla \rho_{F^0} \left(\frac{x_1}{1 + (x_1^2 + x_2^2)^2} + \lambda(x), \frac{x_2}{1 + (x_1^2 + x_2^2)^2} - 2\lambda(x)x_2 \right)$$

on the set $C_\delta(x_0)$.

Thus, all the conditions of Theorem 3 are fulfilled, and we can affirm that the function

$$\frac{1}{2} \arctg(z_1^2 + z_2^2) + \rho_F(x_1 - z_1, x_2 - z_2)$$

admits an unique minimizer $\pi_C^{F,\theta}(x)$ on C , which is Lipschitz continuous w.r.t. x in a neighbourhood of each point $x_0 \in \partial C$ (out of C). This neighbourhood is given by the formula (65), where the Lipschitz constant $L > 0$ of the mapping

$$x \mapsto \mathfrak{J}_F(\partial^p(\theta|_C)(x) \cap \partial F^0), \quad x \in C_\delta(x_0),$$

can be computed by using the above arguments.

Furthermore, the restriction $\theta|_C$ is obviously lower (even proximally) regular on ∂C , and the condition (ii) of Theorems 6 and 7 holds (the condition (i) is violated in the "angle" point $(1, 0)$). Therefore, applying Theorem 7 we see that the value function $\hat{u}(x)$ in the above mathematical programming problem, which can be interpreted also as the viscosity solution to the Hamilton-Jacobi equation

$$\min \left\{ \left| \frac{\partial u}{\partial x_1} \right|, \frac{3}{4} \left| \frac{\partial u}{\partial x_1} \right| \sqrt{4 \left| \frac{\partial u}{\partial x_2} \right|} + \left| \frac{\partial u}{\partial x_2} \right| \right\} = 1,$$

$$u(x_1^2, x_2) = \frac{1}{2} \arctg(x_1^2 + x_2^2)$$

(see (113)), is of class $C_{loc}^{1,1}$ on an open set $\{(x_1, x_2) : x_2^2 < x_1 < x_2^2 + \eta(x_2)\}$, where $\eta(\cdot)$ is a positive real function.

By the next example we test the case of nonsmooth both a target C and a boundary function $\theta(\cdot)$. It shows, in particular, that the hypotheses of Theorem 3 can be fulfilled even if the target C has an "inward" angle point.

Example 2.

$$F := \{x \in \mathbb{R}^2 : \|x\| \leq 1 \text{ and } \langle \mathbf{v}, x \rangle + \mu \|x - \mathbf{v}\| \leq 1\}; \quad (122)$$

$$C := \{x \in \mathbb{R}^2 : \min\{x_1, x_2\} \leq 0\};$$

$$\theta(x) := \max\{\langle \mathbf{a}, x \rangle, \langle \mathbf{b}, x \rangle\}, \quad x = (x_1, x_2) \in C.$$

Here $0 < \mu < 1$ and $\mathbf{v}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ are such that $\|\mathbf{v}\| = 1$, $\mathbf{v}_i > 0$, $\|\mathbf{a}\| < 1$, $\|\mathbf{b}\| < 1$, $\mathbf{a}_i \geq 0$, $\mathbf{b}_i \geq 0$, $i = 1, 2$, and $\mathbf{a}_1 \neq \mathbf{b}_1$, $\mathbf{a}_2 \neq \mathbf{b}_2$. Our goal is to find conditions on the choice of the parameters $\mathbf{v}, \mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ and μ , under which the well-posedness and regularity results of the previous sections hold.

Since F can be represented as $\bar{\mathbf{B}} \cap (\mathbf{v} + K_{\mathbf{v},\mu})$ where

$$K_{\mathbf{v},\mu} := \{x : \langle -\mathbf{v}, x \rangle \geq \mu \|x\|\}$$

is a closed convex cone in \mathbb{R}^2 , from elementary geometric considerations we obtain that F^0 is the convexification of the unit circle $\bar{\mathbf{B}}$ and a symmetric segment of the tangent line at the point $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2)$. Namely,

$$\begin{aligned} F^0 &= \overline{\text{co}} \left(\bar{\mathbf{B}} \cup (\mathbf{v} + K_{\mathbf{v}, \mu})^0 \right) \\ &= \overline{\text{co}} \left(\bar{\mathbf{B}} \cup \left\{ (\mathbf{v}_1 + \lambda \mathbf{v}_2, \mathbf{v}_2 - \lambda \mathbf{v}_1) : |\lambda| \leq \frac{\mu}{\sqrt{1 - \mu^2}} \right\} \right). \end{aligned} \quad (123)$$

We obviously have $\|F\| = 1$, $\|F^0\| = \frac{1}{\sqrt{1 - \mu^2}}$, and $\mathfrak{J}_F(\xi^*) = \{\mathbf{v}\}$ for each $\xi^* = (\mathbf{v}_1 + \lambda \mathbf{v}_2, \mathbf{v}_2 - \lambda \mathbf{v}_1) \in \partial F^0$ with $|\lambda| < \frac{\mu}{\sqrt{1 - \mu^2}}$.

The target set C admits the unit normal vector

$$\mathbf{n}_C(x) = \begin{cases} (1, 0) & \text{if } x_1 = 0, x_2 > 0, \\ (0, 1) & \text{if } x_1 > 0, x_2 = 0 \end{cases} \quad (124)$$

at each point of the boundary ∂C except the origin, where the proximal and the Fréchet normal cones are trivial.

The function $\theta(\cdot)$ is convex and admits the piecewise constant gradient

$$\nabla \theta(x) = \begin{cases} \mathbf{a} & \text{if } \langle \mathbf{a} - \mathbf{b}, x \rangle > 0, \\ \mathbf{b} & \text{if } \langle \mathbf{a} - \mathbf{b}, x \rangle < 0, \end{cases} \quad (125)$$

while

$$\partial \theta(x) = \partial^c \theta(x) = \{\lambda \mathbf{a} + (1 - \lambda) \mathbf{b}, 0 \leq \lambda \leq 1\}$$

whenever $\langle \mathbf{a} - \mathbf{b}, x \rangle = 0$ (see [10, Theorem 2.5.1]). Hence, we deduce the first condition, under which the standing hypothesis **(H)** is fulfilled (see (48)):

$$\gamma := \max \{\|\mathbf{a}\|, \|\mathbf{b}\|\} < \frac{1}{\|F\| \|F^0\|} = \sqrt{1 - \mu^2}. \quad (126)$$

Since at each point $x \in \partial C$ with $x \neq 0$ both the function $\theta(\cdot)$ and the set C are *proximally regular* (moreover, $\theta(\cdot)$ is of class \mathcal{C}^2), we have that

$$\begin{aligned} \partial^p(\theta|_C)(x) &= \partial \theta(x) + \mathbf{N}_C^p(x) \\ &= \{\nabla \theta(x) + \lambda \mathbf{n}_C(x) : \lambda \geq 0\}. \end{aligned} \quad (127)$$

Taking into account (125) and (124) we may further represent (127) in an alternate form depending on the mutual location of the vectors \mathbf{a} and \mathbf{b} . Let us restrict ourselves just to the case when $\mathbf{a}_1 < \mathbf{b}_1$ and $\mathbf{a}_2 > \mathbf{b}_2$. Then

$$\partial^p(\theta|_C)(x) = \begin{cases} \{(\mathbf{a}_1 + \lambda, \mathbf{a}_2) : \lambda \geq 0\} & \text{if } x_1 = 0, x_2 > 0, \\ \{(\mathbf{b}_1, \mathbf{b}_2 + \lambda) : \lambda \geq 0\} & \text{if } x_1 > 0, x_2 = 0. \end{cases} \quad (128)$$

At the origin instead directly by the definition we compute

$$\begin{aligned} \partial^p(\theta|_C)(0) &= \partial^c(\theta|_C)(0) = \text{co} \{\mathbf{a}, \mathbf{b}\} + \{(v_1, v_2) : v_1 \geq 0, v_2 \geq 0\} \\ &= \partial \theta(0) + \mathbf{N}_C^c(0), \end{aligned}$$

so the hypothesis $(\hat{\mathbf{H}})$ also holds, and we have $\Gamma(x) = \partial\theta(x)$ and $\mathbf{N}_C^\theta(x) = \mathbf{N}_C^c(x)$ at each $x \in \partial C$. Notice that here $\partial^\theta C = \partial C$ (roughly speaking, the convex function $\theta(\cdot)$ "regularizes" the "inward" angle point of the target set C).

Given $\delta > 0$ for each $x \in C_\delta(0)$ the subdifferential (128) is a subset of $\partial^p(\theta|_C)(0)$ (a part of its boundary). By simple geometric considerations we find that the condition

$$\max \left\{ \frac{|\mathbf{a}_1 - \mathbf{v}_1|}{\mathbf{v}_2}, \frac{|\mathbf{b}_2 - \mathbf{v}_2|}{\mathbf{v}_1} \right\} \leq \frac{\mu}{\sqrt{1 - \mu^2}} \quad (129)$$

ensures that the intersection $\partial^p(\theta|_C)(0) \cap \partial F^0$ is contained in the line segment $\left\{ (\mathbf{v}_1 + \lambda \mathbf{v}_2, \mathbf{v}_2 - \lambda \mathbf{v}_1) : |\lambda| \leq \frac{\mu}{\sqrt{1 - \mu^2}} \right\}$ (see (123)). Namely, under (129)

$$\partial^p(\theta|_C)(0) \cap \partial F^0 = [A, B]$$

where

$$A = \left(\mathbf{a}_1, \frac{1 - \mathbf{a}_1 \mathbf{v}_1}{\mathbf{v}_2} \right) \quad \text{and} \quad B = \left(\frac{1 - \mathbf{b}_2 \mathbf{v}_2}{\mathbf{v}_1}, \mathbf{b}_2 \right).$$

Furthermore, if the inequality (129) is strict then (see above)

$$\mathfrak{J}_F(\partial^p(\theta|_C)(0) \cap \partial F^0) = \{\mathbf{v}\}.$$

So, the hypothesis $(\mathbf{H}_1(0))$ is trivially fulfilled. Notice that the verification of the respective hypothesis at each point $x_0 \in \partial C$, $x_0 \neq 0$, is reduced to the (more general) case $x_0 = 0$. So, everything said above is sufficient to be able to apply Theorem 3 and to conclude that under the assumptions

$$\max \{ \|\mathbf{a}\|, \|\mathbf{b}\| \} < \sqrt{1 - \mu^2} \quad (130)$$

$$\max \left\{ \frac{|\mathbf{a}_1 - \mathbf{v}_1|}{\mathbf{v}_2}, \frac{|\mathbf{b}_2 - \mathbf{v}_2|}{\mathbf{v}_1} \right\} < \frac{\mu}{\sqrt{1 - \mu^2}} \quad (131)$$

the minimization problem for the function

$$\max \{ \langle \mathbf{a}, z \rangle, \langle \mathbf{b}, z \rangle \} + \rho_F(x - z) \quad (132)$$

subject to $\min \{z_1, z_2\} \leq 0$, where F is defined by (122), admits an unique minimizer, which is Lipschitz continuous w.r.t. x from an open neighbourhood \mathfrak{A} of the constraint set.

For instance, setting $\mathbf{v} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)$, $\mathbf{a} = (0, q)$ and $\mathbf{b} = (q, 0)$, $0 < q < 1$, we see that the conditions (130) and (131) hold whenever the parameters μ and q satisfy the inequalities

$$q < \sqrt{1 - \mu^2} < \frac{\sqrt{2}}{2}$$

(e.g., $\mu = \sqrt{3}/2$ and $q = 1/3$). In this case we are led to minimize the function

$$q \max \{z_1, z_2\} + \rho_F(x - z), \quad z \in C. \quad (133)$$

Besides that the restriction $\theta|_C$ is proximally regular at each point $x_0 \in \partial C$. Therefore, by Theorem 4, the minimum $\hat{u}(x)$ of (132) is Clarke regular and the formula (86) is valid for each $x \in \mathfrak{A}$ (see also Proposition 1). Unfortunately, we are

not able to deduce anything about the Fréchet continuous differentiability of $\hat{u}(\cdot)$ near the origin due to the lack of smoothness of the input data F , C and $\theta(\cdot)$.

Acknowledgements. Work is fulfilled in framework of the project "Variational Analysis: Theory and Applications" (PTDC/MAT/111809/2009) financially supported by Fundação para Ciência e Tecnologia (FCT), the Portuguese institutions COMPETE, QREN and the European Regional Development Fund (FEDER).

References

1. Aubin, J.-P. and Cellina, A.: *Differential Inclusions*, Springer, Berlin, 1984.
2. Aubin, J.-P. and Frankowska, H.: *Set-Valued Analysis*, Birkhauser, Boston, 1990.
3. Bardi, M. and Capuzzo-Dolcetta, I.: *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, 1997.
4. Bernard, F., Thibault, L. and Zlateva, N.: Characterizations of prox-regular sets in uniformly convex Banach spaces, *J. Convex Anal.* **13** (2006), 525-559.
5. Bernard, F., Thibault, L. and Zlateva, N.: Prox-regular sets and epigraphs in uniformly convex Banach spaces: various regularities and other properties, *Trans. Amer. Math. Soc.* **363** (2011), 2211-2247.
6. Bounkhel, M., Thibault, L.: On various notions of regularity of sets in nonsmooth analysis, *Nonlin. Anal., Theory, Meth. Appl.* **48** (2002), 223-246.
7. Bressan, A.: Hamilton-Jacobi equations and optimal control. An illustrated tutorial, Trondheim, NTNU, 2001.
8. Canino, A.: On p -convex sets and geodesics, *J. Differ. Equat.* **75** (1988), 118-157.
9. Cardaliaguet, P., Dacorogna, B., Gangbo, W. and Georgy, N.: Geometric restrictions for the existence of viscosity solutions, *Ann. Inst. Henri Poincaré* **16** (1999), 189-220.
10. Clarke, F.H.: *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.
11. Clarke, F.H., Stern, R.J. and Wolenski, P.R.: Proximal smoothness and the lower- C^2 property, *J. Convex Anal.* **2** (1995), 117-144.
12. Clarke, F.H., Ledyev, Yu.S., Stern, R.J. and Wolenski, P.R.: *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.
13. Colombo, G. and Goncharov, V.V.: Variational inequalities and regularity properties of closed sets in Hilbert spaces, *J. Convex Anal.* **8** (2001), 197-221.
14. Colombo, G., Goncharov, V.V. and Mordukhovich, B.S.: Well-posedness of minimal time problems with constant dynamics in Banach spaces, *Set-Valued Var. Anal.* **18** (2010), 349-372.
15. Colombo, G. and Thibault, L.: Prox-regular sets and applications. In D.Y. Gao, D. Motreanu (eds.), *Handbook on Nonconvex Analysis*. International Press, Boston, 2010.
16. Colombo, G. and Wolenski, P.R.: Variational Analysis for a class of minimal time functions in Hilbert spaces, *J. Convex Anal.* **11** (2004), 335-361.
17. Chong, Li: On well posed generalized best approximation problems, *J. Approx. Theory* **107** (2000), 96-108.
18. Ekeland, I.: Nonconvex minimization problems, *Bull. Amer. Math. Soc.* **1** (1979), 443-474.

19. Ekeland, I. and Lebourg, G.: Generic Fréchet differentiability and perturbed optimization problems in Banach spaces, *Trans. Amer. Math. Soc.* **224** (1976), 193-216.
20. Chong, Li and Renxing, Ni: Derivatives of generalized distance functions and existence of generalized nearest points, *J. Approx. Theory* **115** (2002), 44-55.
21. Dal Maso, G., Goncharov, V.V. and Ornelas, A.: A Lipschitz selection from the set of minimizers of a nonconvex functional of the gradient, *Nonlin. Anal., Theory, Meth. Appl.* **37** (1999), 707-717.
22. De Blasi, F.S. and Myjak, J.: On a generalized best approximation problem, *J. Approx. Theory* **94** (1998), 54-72.
23. Federer, H.: Curvature measures, *Trans. Amer. Math. Soc.* **93** (1959), 418-491.
24. Georgiev, P.G.: Submonotone mappings in Banach spaces and applications, *Set-Valued Anal.* **5** (1997), 1-35.
25. Goncharov, V.V. and Pereira, F.F.: Neighbourhood retractions of nonconvex sets in a Hilbert space via sublinear functionals, *J. Convex Anal.* **18** (2011), 1-36.
26. Goncharov, V.V. and Pereira, F.F.: Geometric conditions for regularity in a time-minimum problem with constant dynamics, *J. Convex Anal.* **19** (2012), 631-669.
27. Goncharov, V.V. and Pereira, F.F.: Geometric conditions for regularity of viscosity solution to the simplest Hamilton-Jacobi equation. In D. Hömberg, F. Tröltzsch (eds), *System Modeling and Optimization*, Proc. of the 25th IFIP TC7 Conf., CSMO 2011, Berlin, Springer (2012), 245-254.
28. Hiriart-Urruti, J.-B. Gradients generalises de fonctions marginales, *SIAM J. Contr. Optim.* **16** (1978), 301-316.
29. Ioffe, A.D. and Penot, J.-P.: Subdifferentials of performance functions and calculus of coderivatives of set-valued mappings, *Serdica Math. J.* **22** (1996), 359-384.
30. Kružkov, S.N.: Generalized solutions of the Hamilton-Jacobi equation of the Eikonal type. I., *Math. USSR Sbornik* **27** (1975), 406-446.
31. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I. Basic Theory, Springer, Berlin, 2006.
32. Mordukhovich, B.S., Nam, N.M. and Yen, N.D.: Subgradients of marginal functions in parametric mathematical programming, *Math. Program., Ser. B.* **116** (2009), 369-396.
33. Ngai, H.V. and Penot J.-P.: Approximately convex functions and approximately monotonic operators, *Nonlin. Anal., Theory, Meth. Appl.* **66** (2007), 547-564.
34. Poliquin, R.A., Rockafellar, R.T. and Thibault, L.: Local differentiability of distance functions, *Trans. Amer. Math. Soc.* **352** (2000), 5231-5249.
35. Rockafellar, R.T.: Generalized directional derivatives and subgradients of nonconvex functions, *Can. J. Math.* **32** (1980), 257-280.
36. Rockafellar, R.T. and Wets, R.J.-B.: Variational Analysis, Springer-Verlag, Berlin, 1998.
37. Shapiro, A.: Existence and differentiability of metric projections in Hilbert spaces, *SIAM J. Optimiz.* **4** (1994), 130-141.
38. Thibault, L.: On subdifferentials of optimal value functions, *SIAM J. Contr. Optim.* **29** (1991), 1019-1036.
39. Yiran, He and Kung, Fu Ng: Subdifferentials of a minimum time function in Banach spaces, *J. Math. Anal. Appl.* **321** (2006), 896-910.
40. Wolenski, P. R. and Zhuang, Yu.: Proximal analysis and the minimal time function, *SIAM J. Contr. Optim.* **36** (1998), 1048-1072.

41. Zajicek, L.: A generalization of an Ekeland-Lebourg theorem and the differentiability of distance functions, *Suppl. Rend. del Circolo Mat. Palermo, II* **3** (1984), 403-410.