Journal of Convex Analysis
Volume 23 (2016), No. 2, 615–629

Zero Duality Gap and Attainment with Possibly Non-Convex Data

Emil Ernst
Aix Marseille Université, CNRS, Centrale Marseille,
I2M, UMR 7373, 13453 Marseille, France
Emil.Ernst@univ-amu.fr

Michel Volle
LMA, Avignon Université,
301 rue Baruch de Spinoza, 84911 Avignon, France
Michel.Volle@univ-avignon.fr

Dedicated to the memory of Jean Jacques Moreau.

Received: November 17, 2014
Revised manuscript received: November 26, 2014
Accepted: November 28, 2014

A newly defined notion of convex closedness regarding a set is used in order to state a necessary and sufficient criterion for the min-sup property in non necessarily convex primal-dual optimization problems, generalizing well-known theorems valid in the convex setting. Our main result is then applied to the classical penalty method.

Keywords: Dual optimization, min-sup property, convex closedness regarding a set, penalty method

2000 Mathematics Subject Classification: 49M30, 49N15, 52A20

1. Introduction and notation.

This note is concerned with the following abstract optimization problem depending on parameters:

$$(P) \qquad \text{minimize } F(x, 0_U) \text{ over all } x \in X,$$

where $F : X \times U \to \overline{\mathbb{R}} := \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$ is an extended-real-valued function, the primal variable x belongs to a set X, while the parameter u lies in a separated locally convex space U. A fruitful method to address problem (P) is to consider the Lagrangian

$$L_F : X \times U^* \to \overline{\mathbb{R}}, \quad L_F(x, u^*) := \inf_{u \in U} \{F(x, u) - \langle u^*, u \rangle_U\}$$

associated to F; here U^* is the topological dual space of U, and $\langle \cdot, \cdot \rangle_U$ is the duality pairing between U^* and U. Indeed, in many applications of practical

ISSN 0944-6532 / \$ 2.50 © Heldermann Verlag

interest the Lagrangian is a more regular function than F, so it is often easier to address, instead of the primal problem (P), the dual problem defined by the following relation

$$(D) \qquad \text{maximize } \inf_{x\in X} L_F(x,u^*) \text{ over all } u^* \in U^*.$$

In the particular case when X is a separated locally convex space, X^* is its topological dual space, and $\langle\cdot,\cdot\rangle_X$ is the duality pairing between X^* and X, the dual problem became

$$(D) \qquad \text{maximize } -F^*(0_{X^*},u^*) \text{ over all } u^* \in U^*,$$

where

$$F^* : X^*\times U^* \to \overline{\mathbb{R}}, \quad F^*(x^*,\, u^*) = \sup_{x\in X,\, u\in U} \langle x^*,\, x\rangle_X + \langle u^*,\, u\rangle_U - F(x,\, u),$$

is the Legendre-Fenchel conjugate of F. The standard terminology and notation we use in the sequel can be explicitly found in the classical monographs [2]–[5], [7] and [11]–[13] for instance.

The aim of this note is to provide a necessary and sufficient condition for the min-sup duality property

$$(MS) \qquad -\infty < \sup(D) = \min(P) \le +\infty,$$

meaning that

$$\exists\, \overline{x}\in X: \quad -\infty < \sup_{u^*\in U^*}\left(\inf_{x\in X} L_F(x,u^*)\right) = F(\overline{x},0_U) \le +\infty; \tag{1}$$

as an application, we investigate the (MS) property in the case of a penalty problem.

1.1. Known results.

The special case when X is a separated locally convex space and $F\in\Gamma(X\times U)$, that is when F is a convex lower semi-continuous and proper function, is well understood, and no less than two different necessary and sufficient conditions for the min-sup duality property are available in the mathematical literature.

Let us first mention the following classical result, which can easily be deduced from [11, Theorem 16’].

Proposition 1.1. *Let X be a separated locally convex space and $F\in\Gamma(X\times U)$ be such that $\inf(P)\in\mathbb{R}$. Then, (MS) holds if and only if the dual value function*

$$\gamma : X^* \to \overline{\mathbb{R}}, \quad \gamma(x^*) := \inf_{u^*\in U^*} \left(F^*(x^*,u^*)\right)$$

is sub-differentiable at 0_{X^*}.

The second available characterization of the min-sup property for convex problems may be achieved by taking the dual statement to [5, Th. 9.1], and reads as follows.

Proposition 1.2. *Let X be a separated locally convex space and $F \in \Gamma(X \times U)$ be such that*

$$(\{0_{X^*}\} \times U^*) \cap \operatorname{dom} F^* \neq \emptyset.$$

Then, (MS) holds if and only if $\mathcal{A}_F$ is closed regarding $\{0_U\} \times \mathbb{R}$.

Here $\mathcal{A}_F := \{(u,t) \in U \times \mathbb{R} : \exists\, x \in X \text{ s.t. } F(x,u) \leq t\}$ is the projection of the epigraph of the perturbation function F onto $U \times \mathbb{R}$, and let us recall from [5] that a set A is closed regarding B if $\overline{A} \cap B = A \cap B$.

Another result of relevance for our study addresses the (MS) property in the particular case of convex penalty problems, that is:

$$F : X \times \mathbb{R} \to \overline{\mathbb{R}}, \qquad F(x,r) := \begin{cases} \phi(x) & d(x,A) \leq r \\ +\infty & d(x,A) > r, \end{cases}$$

where $(X, \|\cdot\|_X)$ is a real normed space, $\phi : X \to \mathbb{R} \cup \{+\infty\}$ is a convex lower semi-continuous function, and $A \subset X$ is convex closed, while $d(x,A)$ denotes the distance between the point $x \in X$ and the set A,

$$d(x,A) = \inf_{y \in A} \|x - y\|_X.$$

As (for details, the reader is referred to Section 3)

$$\inf(P) = \inf_{x \in X} \; F(x,0) = \inf_{x \in A} \; \phi(x),$$

while

$$\sup(D) = \lim_{p \to +\infty} \left(\inf_{x \in X} \left(\phi(x) + p\, d(x,A)\right) \right),$$

the (MS) property means in this case that

$$\lim_{p \to +\infty} \left(\inf_{x \in X} \left(\phi(x) + p\, d(x,A)\right) \right) = \min_{x \in A} \; \phi(x);$$

in other words, the minimum over A of ϕ may be obtained precisely by successively minimizing functions of form $x \to (\phi(x) + p\, d(x,A))$ (that is functions obtained by penalizing the objective function ϕ with the distance to the feasible set A), as p goes to $+\infty$.

The the min-sup property in the case of a convex penalty problem came as a consequence of a well-known result, [6, Theorem 2.6], generalized in [1], and [8].

Proposition 1.3. *Let X be a real normed space, $\phi \in \Gamma(X)$ and A be a convex set such that ϕ is finite and continuous at some point of A, and let us set*

$$F : X \times \mathbb{R} \to \overline{\mathbb{R}}, \qquad F(x,r) := \begin{cases} \phi(x) & d(x,A) \leq r \\ +\infty & d(x,A) > r. \end{cases}$$

Then, $\inf(P) = \sup(D)$*. If, moreover, A is weakly compact, then ϕ achieves its infimum over A, so relation (MS) holds true.*

Several studies have considered the case of non necessarily convex data in its full generality. Let us first state the following easy corollary of [10, Theorem 2.2], which establishes a sufficient, but far from being necessary, condition for the (MS) property.

Proposition 1.4. *Let X be a separated locally convex space and let the perturbation function $F : X \times U \to \overline{\mathbb{R}}$ be such that the value function*

$$m : U \to \overline{\mathbb{R}}, \qquad m(u) := \inf_{x \in X} F(x,u)$$

takes at least a finite value over U. Then, the (MS) property holds true provided that $\mathcal{A}_F$ is convex and closed.

Finally, let us recall Theorem 2 from [12], a duality result which is valid when the underlying set X is not necessarily a vector space.

Proposition 1.5. *Let X be a set, U a normed space, and $F : X \times U \to \overline{\mathbb{R}}$ be such that the value function m is lower semi-continuous at 0_U. Then* $\inf(P) = \max(D)$*, provided that m^*, the Legendre-Fenchel conjugate of m, is Fréchet differentiable at one of its minimum points.*

Moreover, if X is a compact topological space, and F is lower semi-continuous, then the hypothesis that m^* is Fréchet differentiable at one of its minimum points implies that $\min(P) = \max(D)$*, and thus that the (MS) property holds true.*

1.2. Description of our results and plan of the article.

Section 2 addresses the main technical results and concepts of our study. This section relies on a new notion, well-adapted for studying the min-sup property in the general case.

Definition 1.6. Given two subsets A, B of a topological vector space, one says that A is convex closed regarding B if

$$B \cap \overline{\mathrm{co}}\, A = B \cap A,$$

where $\overline{\mathrm{co}}\, A$ denotes the closed convex hull of A.

The main result of this note, Theorem 2.3, proves that the min-sup property holds if and only if the set $\mathcal{A}_F$ is convex closed regarding $\{0_U\} \times \mathbb{R}$. Section 2 closes with Theorem 2.9, a general result, containing as particular cases both Proposition 1.2 and Proposition 1.4.

As an application, Section 3 addresses the classical penalty method in the general setting of a metric space $(X, d(\cdot,\cdot))$ endowed with a topology τ which is coarser than the distance topology. Theorem 3.1 uses the concepts and technics developed in Section 2 to prove that, given a bounded τ-compact set $A \subset X$, and a τ-lower semi-continuous function $\phi : X \to \mathbb{R} \cup \{+\infty\}$ which takes at least one finite value on A, the (MS) property

$$\lim_{p \to +\infty} \left(\inf_{x \in X} (\phi(x) + p\, d(x, A)) \right) = \min_{x \in A} \phi(x)$$

holds if and only the objective function ϕ satisfies the following growth condition: for any $x_0 \in X$ there is some $b > 0$ such that the function $x \to (\phi(x) + b\, d(x, x_0))$ is bounded from below.

When specifying X to a real normed space, and τ to its weak topology, Theorem 3.1 obviously generalizes the second part of Proposition 1.3, removing, at the same time, the blanket condition which requests the continuity of the function ϕ at some point of the set A.

1.3. Notation and requirements.

The following definitions and notation are standard in convex analysis (for details, the reader is referred to the monograph [9]). Having a function $h : U \to \overline{\mathbb{R}}$, we denote by $\operatorname{dom} h := \{u \in U : \ h(u) < +\infty\}$ its effective domain and by h^* its Legendre-Fenchel conjugate,

$$h^* : U^* \to \overline{\mathbb{R}}, \quad h^*(u^*) := \sup\{\langle u^*, u\rangle_U - h(u) : \ u \in U\}.$$

The functions $\overline{h}$ (respectively co h, respectively $\overline{\text{co}}\, h$) are the lower semi-continuous (respectively convex, respectively closed convex) hull of h. In terms of epigraphs, it holds that $\operatorname{epi} \overline{h} = \overline{\operatorname{epi} h}$, $\operatorname{epi} \overline{\text{co}}\, h = \overline{\text{co}}(\operatorname{epi} h)$, and, if $\operatorname{dom} h^* \neq \emptyset$, one has $\overline{\text{co}}\, h = h^{**}$, where h^{**} denotes the Legendre-Fenchel conjugate of h^*:

$$h^{**}(u) = \sup_{u^* \in U^*} \{\langle u^*, u\rangle_U - h^*(u^*)\} \quad \forall u \in U.$$

Let us now focus on an abstract optimization problem depending on parameters. Recall that the value function (also called optimal value function) $m : U \to \overline{\mathbb{R}}$ associated with F is given by $m(u) := \inf_{x\in X} F(x, u)$, and its conjugate satisfies the relation

$$-m^*(u^*) = \inf_{x \in X} L_F(x, u^*) \quad \forall\, u^* \in U^*,$$

together with the so-called weak duality property:

$$-\infty \leq \sup(D) = m^{**}(0_U) \leq m(0_U) = \inf(P) \leq +\infty. \tag{2}$$

Let us denote by $\mathrm{epi}_s\, m$ the strict epigraph of the value function; it then holds that

$$\mathrm{epi}_s(\mathrm{co}\, m) = \mathrm{co}(\mathrm{epi}_s m) \subset \mathrm{co}\,(\mathcal{A}_F) \subset \mathrm{co}(\mathrm{epi}\, m) \subset \mathrm{epi}(\mathrm{co}\, m).$$

Taking the topological closure in the above relations, we obtain:

$$\mathrm{epi}(\overline{\mathrm{co}}\, m) = \overline{\mathrm{co}}\, \mathcal{A}_F. \tag{3}$$

2. A characterization of the min-sup property in the general case.

In the sequel we require the following basic hypothesis already in use in Proposition 1.1:

$$(H_F) \qquad \sup(D) > -\infty.$$

Remark 2.1. Assumption (H_F) means that $\mathrm{dom}\, m^* \neq \emptyset$. Moreover, when X is a separated locally convex space, then (H_F) holds true if and only if

$$(\{0_{X^*}\} \times U^*) \cap \mathrm{dom}\, F^* \neq \emptyset,$$

or, equivalently, if 0_{X^*} belongs to the projection of $\mathrm{dom}\, F^*$ onto X^*.

Remark 2.2. If $\sup(D) = +\infty$, then, by (2), (MS) holds with $\overline{x} \in X$ arbitrary in (1).

Theorem 2.3. *Let X be a set and $F : X \times U \to \overline{\mathbb{R}}$ a function satisfying (H_F). Then, the min-sup property (MS) holds if and only if $\mathcal{A}_F$ is convex closed regarding $\{0_U\} \times \mathbb{R}$.*

Proof of Theorem 2.3. Let us set $d := \sup(D)$. By (H_F) one has that $d \in]-\infty, +\infty]$. Since $\mathrm{dom}\, m^* \neq \emptyset$ (see Remark 2.1), one has $m^{**} = \overline{\mathrm{co}}\, m$ and, by (3),

$$(\{0_U\} \times \mathbb{R}) \cap \overline{\mathrm{co}}\, \mathcal{A}_F = \{0_U\} \times [m^{**}(0_U), +\infty[= \{0_U\} \times [d, +\infty[. \tag{4}$$

Assume first that $d = +\infty$. Then,

$$(\{0_U\} \times \mathbb{R}) \cap \mathcal{A}_F \subset (\{0_U\} \times \mathbb{R}) \cap \overline{\mathrm{co}}\, \mathcal{A}_F = \emptyset,$$

and $\mathcal{A}_F$ is convex closed regarding $\{0_U\} \times \mathbb{R}$. On the other hand, (MS) holds by Remark 2.2, and both assertions in Theorem 2.3 are true.

Assume now that $d \in \mathbb{R}$. By (4) one has

$$(0_U, d) \in (\{0_U\} \times \mathbb{R}) \cap \overline{\mathrm{co}}\, \mathcal{A}_F.$$

So, if $\mathcal{A}_F$ is convex closed regarding $\{0_U\} \times \mathbb{R}$, then $(0_U, d) \in \mathcal{A}_F$ and there exists some point $\overline{x} \in X$ such that $\inf(P) \leq F(\overline{x}, 0_U) \leq d \leq \inf(P)$, and (MS) holds.

Conversely, assume that (MS) holds, and let $r \in \mathbb{R}$ be such that $(0_U, r) \in \overline{\mathrm{co}}\, \mathcal{A}_F$. One has to check that $(0_U, r) \in \mathcal{A}_F$. By (4) one has $r \geq d$, and, by (MS), there exists $\overline{x} \in X$ such that $d = F(\overline{x}, 0_U) \leq r$, that means $(0_U, r) \in \mathcal{A}_F$. $\square$

Corollary 2.4. *Let X be a set and $F : X \times U \to \overline{R}$ satisfying (H_F). Assume that $\overline{\mathcal{A}_F}$ is convex. Then, (MS) holds if and only if $\mathcal{A}_F$ is closed regarding $\{0_U\} \times \mathbb{R}$.*

Proof of Corollary 2.4. Observe that $\overline{\text{co}}\,\mathcal{A}_F = \overline{\text{co}}\,\overline{\mathcal{A}_F} = \overline{\mathcal{A}_F}$. Consequently, $\mathcal{A}_F$ is convex closed regarding $\{0_U\} \times \mathbb{R}$ if and only if $\mathcal{A}_F$ is closed regarding $\{0_U\} \times \mathbb{R}$. □

Remark 2.5. It holds that

$$F \text{ convex} \begin{array}{c} \nearrow \\ \searrow \end{array} \begin{array}{c} \mathcal{A}_F \text{ convex} \\ m \text{ convex} \end{array} \begin{array}{c} \searrow \\ \nearrow \end{array} \overline{m} \text{ convex} \Longleftrightarrow \overline{\mathcal{A}_F} \text{ convex.}$$

Remark 2.6. (MS) duality property without convexity of $\overline{\mathcal{A}_F}$: an example. Take $U = X$ and let $f, g : X \to \overline{\mathbb{R}}$ be such that $\inf_X f = 0$, $g(0_X) = 0$, and $\operatorname{dom} g^* \neq \emptyset$. Define $F(x, u) = f(x) + g(u)$ for any $(x, u) \in X^2$. The corresponding problems are

$$(P) \qquad \min\nolimits_{x \in X} f(x), \qquad\qquad (D) \qquad \max\nolimits_{u^* \in U^*} -g^*(u^*).$$

One has $m = g$ and (H_F) holds. Moreover,

$$\mathcal{A}_F = \{(u, t) \in X \times \mathbb{R} : \exists x \in X \text{ s.t. } f(x) + g(u) \leq t\},$$

$$\overline{\mathcal{A}_F} = \overline{\operatorname{epi} g},$$

$$\overline{\text{co}}\,\mathcal{A}_F = \operatorname{epi} g^{**},$$

$$(\{0_X\} \times \mathbb{R}) \cap \mathcal{A}_F = \begin{cases} \{0_X\} \times [0, +\infty[& \operatorname{argmin} f \neq \emptyset \\ \{0_X\} \times]0, +\infty[& \operatorname{argmin} f = \emptyset, \end{cases}$$

and

$$(\{0_X\} \times \mathbb{R}) \cap \overline{\text{co}}\,\mathcal{A}_F = \{0_X\} \times [g^{**}(0_X), +\infty[.$$

We therefore have

$$\begin{array}{c} \mathcal{A}_F \text{ convex closed regarding } \{0_X\} \times \mathbb{R} \\ \Updownarrow \\ \operatorname{argmin} f \neq \emptyset \text{ and } g^{**}(0_X) = g(0_X) \\ \Updownarrow \\ (MS) \text{ holds.} \end{array}$$

Clearly, the above conditions can be verified without the convexity of f and g: take for example $X = \mathbb{R}$, $f(x) = g(x) = \sqrt{|x|}$.

Corollary 2.7. *If X is a separated locally convex space, $\mathcal{A}_F$ is convex closed regarding $\{0_U\} \times \mathbb{R}$, and (H_F) is satisfied, then*

$$-\infty < \sup(D) = \min_{x \in X} F^{**}(x, 0_U) < +\infty.$$

Proof of Corollary 2.7. For any $(x, u^*) \in X \times U^*$ one has $F^{**}(x, 0_U) \geq -F^*(0_{X^*}, u^*)$, and $\inf_{x \in X} F^{**}(x, 0_U) \geq \sup(D)$. By Theorem 2.3 there exists $\overline{x} \in X$ such that

$$\begin{aligned}-\infty < \sup(D) \leq \inf_{x \in X} F^{**}(x, 0_U) &\leq F^{**}(\overline{x}, 0_U) \\ &\leq F(\overline{x}, 0_U) = \sup(D) \leq +\infty,\end{aligned}$$

and, finally, $\sup(D) = \min_{x \in X} F^{**}(x, 0_U)$. $\square$

Remark 2.8. Since $(F^{**})^* = F^*$, it follows from Theorem 2.3 and Corollary 2.7 that, under the assumption (H_F) one has

$$\begin{gathered}\mathcal{A}_F \text{ closed convex regarding } \{0_U\} \times \mathbb{R} \\ \Downarrow \\ \mathcal{A}_{F^{**}} \text{ closed regarding } \{0_U\} \times \mathbb{R}\end{gathered}$$

a property that may be checked directly by noting that $\mathcal{A}_{F^{**}}$ is convex.

We now consider the case when the min-sup property holds simultaneously for a family of perturbation problems.

Theorem 2.9. *Let X be a set, and $F : X \times U \to \overline{\mathbb{R}}$ a function satisfying (H_F). For any nonempty subset V of U, the following statements are equivalent:*

(i) *for all $v \in V$ it holds that*

$$-\infty < \sup_{u^* \in U^*} \left(-\langle u^*, v \rangle_U + \inf_{x \in X} L_F(x, u^*) \right) = \min_{x \in X} F(x, v),$$

(ii) *$\mathcal{A}_F$ is convex closed regarding $V \times \mathbb{R}$.*

Proof of Theorem 2.9. Observe that (ii) means exactly that $\mathcal{A}_F$ is convex closed regarding $\{v\} \times \mathbb{R}$ for any $v \in V$. So, for any $v \in V$, one has to prove that (i) is equivalent to

$$\mathcal{A}_F \text{ is convex closed regarding } \{v\} \times \mathbb{R}, \tag{5}$$

or, equivalently, to

$$\mathcal{A}_F - \{v\} \times \{0\} \text{ is convex closed regarding } \{0_U\} \times \mathbb{R}.$$

To this end, let us apply Theorem 2.3 to the function $G(x, u) := F(x, u + v)$. We have $\mathcal{A}_G = \mathcal{A}_F - \{v\} \times \{0\}$, $L_G(x, u^*) = L_F(x, u^*) - \langle u^*, v \rangle_U$, and G satisfies the basic assumption (H_G) since F does. According to Theorem 2.3 applied for G, we have that for any $v \in V$, (i) is equivalent to (5). $\square$

Corollary 2.10. *Let X be a set, and $F : X \times U \to \overline{\mathbb{R}}$ be a function satisfying (H_F). The following statements are equivalent:*

(i) *for all $v \in V$ it holds that*

$$-\infty < \sup_{u^* \in U^*} \left(- \langle u^*, v \rangle_U + \inf_{x \in X} L_F(x, u^*) \right) = \min_{x \in X} F(x, v),$$

(ii) *$\mathcal{A}_F$ is convex closed.*

Proof of Corollary 2.10. From Theorem 2.9 we know that (i) is equivalent to the fact that $\mathcal{A}_F$ is convex closed regarding $U \times \mathbb{R}$, that means (ii). □

Remark 2.11. Let us remark that the hypothesis of Proposition 1.4 implies that condition (H_F) holds true. Indeed, if $\mathcal{A}_F$ is convex closed then m is lsc and convex. Taking at least a finite value, it follows that m is proper, so $m = m^{**}$, fact which implies (H_F).

Accordingly, Corollary 2.10 may be seen as a slight improvement of Proposition 1.4 (that is of [10, Th. 2.2]), since it only requires that X is a set, and no longer a separated locally convex space.

3. An application of Theorem 2.3 to the study of the duality gap in penalty method.

Let $(X,\, d(\cdot,\cdot))$ be a metric space, endowed with a topology τ coarser than the topology induced by the distance, $A \subset X$ be a τ-closed set, and $\phi : X \to \mathbb{R} \cup \{+\infty\}$ be a function which takes at least one finite value on A. To minimize ϕ over A, let us introduce one additional real variable, usually refereed to as a penalty parameter, and define the following perturbation function:

$$F : X \times \mathbb{R} \to \overline{\mathbb{R}}, \qquad F(x, r) := \begin{cases} \phi(x) & d(x, A) \le r \\ +\infty & d(x, A) > r. \end{cases}$$

Indeed, since A is τ-closed and τ is coarser than the distance topology, it follows that

$$F(x, 0) = \begin{cases} \phi(x) & x \in A \\ +\infty & x \notin A, \end{cases}$$

so

$$\inf(P) = \inf_{x \in X} \; F(x, 0) = \inf_{x \in A} \; \phi(x),$$

and, as

$$\begin{aligned} -m^*(s) &= \inf_{x \in X,\, r \in \mathbb{R}} (F(x, r) - s\, r) \\ &= \inf_{x \in X} \left(\inf_{r \ge d(x,A)} (\phi(x) - s\, r) \right) \\ &= \begin{cases} -\infty & s > 0 \\ \inf_{x \in X} (\phi(x) - s\, d(x, A)) & s \le 0, \end{cases} \end{aligned}$$

it follows that

$$\sup(D) = \sup_{p\geq 0}\left(\inf_{x\in X}\left(\phi(x)+p\,d(x,A)\right)\right).$$

Since the function $0\leq p\to\inf_{x\in X}\left(\phi(x)+p\,d(x,A)\right)$ is non-decreasing, we have

$$\sup(D) = \lim_{p\to+\infty}\left(\inf_{x\in X}\left(\phi(x)+p\,d(x,A)\right)\right). \tag{6}$$

The aim of the next result is to provide, in the case of a set A which is bounded and τ-compact, a characterization of all the τ-lower semi-continuous functions $\phi : X\to\mathbb{R}\cup\{+\infty\}$ for which the (MS) property

$$\lim_{p\to+\infty}\left(\inf_{x\in X}\left(\phi(x)+p\,d(x,A)\right)\right) = \min_{x\in A}\phi(x)\in\mathbb{R} \tag{7}$$

holds true.

Theorem 3.1. *Let $(X, d(\cdot,\cdot))$ be a metric space, endowed with a topology τ coarser than the topology induced by the distance, and define the optimization problem*

$$F : X\times\mathbb{R}\to\overline{\mathbb{R}},\qquad F(x,r) := \begin{cases}\phi(x) & d(x,A)\leq r\\ +\infty & d(x,A)>r,\end{cases}$$

where $A\subset X$ is a bounded τ-compact set, and $\phi : X\to\mathbb{R}\cup\{+\infty\}$ is a τ-lower semi-continuous function which takes at least one finite value on A. The following statements are equivalent:

i) the property (MS) holds,

ii) the function $x\to\phi(x)+p\,d(x,A)$ is bounded from below for some $p\geq 0$,

iii) for every $x_0\in X$, the function $x\to\phi(x)+b\,d(x,\, x_0)$ is bounded from below for some $b\geq 0$.

Proof of Theorem 3.1. *i)* $\Rightarrow$ *ii)*. From relation (7) it follows that, for p large enough,

$$-\infty < \inf_{x\in X}\left(\phi(x)+p\,d(x,A)\right),$$

fact which proves statement *ii)*.

ii) $\Rightarrow$ *iii)*. As A is bounded, it follows that

$$d(x,\, x_0)\leq R\quad\forall x\in A,$$

for some $R>0$, so

$$d(x,A)\leq d(x,\, x_0)+R\quad\forall x\in X, \tag{8}$$

and

$$d(x,\, x_0)\leq d(x,A)+R\quad\forall x\in X. \tag{9}$$

From relation (8) it yields that

$$\phi(x) + p\, d(x,\, x_0) \geq \phi(x) + p\, d(x, A) - p\, R;$$

the desired conclusion results by noticing that, as the function $x \to \phi(x) + p\, d(x, A)$ is bounded from below, the same holds true for the function $x \to (\phi(x) + p\, d(x, A) - p\, R)$.

iii) $\Rightarrow$ *i*). The set $\mathcal{A}_F$ is given by the following formula:

$$\begin{aligned} \mathcal{A}_F = \; & \{(r, t) \in \mathbb{R}_+ \times \mathbb{R} : \\ & \quad \exists x_{r,t} \in X \text{ s.t. } t \geq \phi(x_{r,t}) \text{ and } d(x_{r,t}, A) \leq r\}\,. \end{aligned} \tag{10}$$

It holds that

$$\mathcal{A}_F \cap (\{0\} \times \mathbb{R}) = \{0\} \times [\inf_A \phi; +\infty[; \tag{11}$$

in order to prove the (MS) property, we have to prove that

$$\overline{\text{co}}\,(\mathcal{A}_F) \cap (\{0\} \times \mathbb{R}) = \{0\} \times [\inf_A \phi; +\infty[. \tag{12}$$

Our proof takes two preliminary steps.

Let us first pick $a, b \geq 0$ such that

$$\phi(x) + a + b\, d(x,\, x_0) \geq 0 \quad \forall x \in X,$$

(such numbers exist since statement *iii*) holds true), and also pick $R > 0$ such that

$$d(x,\, x_0) \leq R \quad \forall x \in A;$$

in virtue of relation (8), we obtain that

$$\phi(x) \geq -(a + bR) - b\, d(x,\, x_0) \quad \forall x \in X.$$

The previous inequality applies in particular to the element $x_{r,t} \in X$ from the relation (10):

$$t \geq \phi(x_{r,t}) \geq -(a + b\, R) - b\, r \quad \forall (r,\, t) \in \mathcal{A}_F. \tag{13}$$

The second step needed in order to determine the set $\overline{\text{co}}\,(\mathcal{A}_F) \cap (\{0\} \times \mathbb{R})$, is to prove that

$$\overline{\mathcal{A}_F} \cap (\{0\} \times \mathbb{R}) = \{0\} \times [\inf_A \phi; +\infty[. \tag{14}$$

To this respect, let us pick $(0, t) \in \overline{\mathcal{A}_F}$; accordingly, there is a sequence $(r_n, t_n)_{n \in \mathbb{N}} \subset \mathcal{A}_F$ converging to $(0, t)$. Thus,

$$\phi(x_{r_n, t_n}) \leq t_n \quad \forall n \in \mathbb{N}, \tag{15}$$

and

$$d(x_{r_n,t_n}, A) \le r_n \quad \forall n \in \mathbb{N}. \tag{16}$$

From relation (16) it easily results that there is a sequence $(y_n)_{n\in\mathbb{N}} \subset A$ such that

$$d(x_{r_n,t_n},\, y_n) \le 2\, r_n, \quad \forall n \in \mathbb{N}. \tag{17}$$

But the set A is τ-compact, so, there is $\overline{y} \in A$, a τ-cluster point of the sequence (y_n). Relation (17), the fact that the sequence (r_n) converges to zero, and the fact that τ is coarser than the distance topology allow us to deduce that $\overline{y}$ is also a τ-cluster point for sequence (x_{r_n,t_n}). The function ϕ being τ-lower semi-continuous, it follows that

$$\limsup_{n\to+\infty} \phi(x_{r_n,t_n}) \ge \phi(\overline{y}) \ge \inf_A \phi, \tag{18}$$

while relation (15) proves that

$$t = \lim_{n\to+\infty} t_n \ge \limsup_{n\to+\infty} \phi(x_{r_n,t_n}). \tag{19}$$

From relations (18) and (19) it yields that

$$\overline{\mathcal{A}_F} \cap (\{0\} \times \mathbb{R}) \subset \{0\} \times [\inf_A \phi; +\infty[;$$

this fact, combined with relation (11), proves the statement (14).

We are finally in a position to determine the set $\overline{\mathrm{co}}\,(\mathcal{A}_F) \cap (\{0\} \times \mathbb{R})$. To this respect, let us pick $(0,t)$ an element from $\overline{\mathrm{co}}\,(\mathcal{A}_F)$; accordingly, $(0,t)$ is the limit of some sequence from $\mathrm{co}\,(\mathcal{A}_F)$, and, in virtue of Carathéodory's theorem, there are three sequences $(r_{i,n}, t_{i,n})_{n\in\mathbb{N}} \subset \mathcal{A}_F$, $i \in \{1,2,3\}$, and three sequences $(\lambda_{i,n})_{n\in\mathbb{N}} \subset [0,\, 1]$, $i \in \{1,2,3\}$, such that

$$\Sigma_{i=1}^3 \lambda_{i,n} = 1 \quad \forall n \in \mathbb{N},$$

and

$$(0,t) = \lim_{n\to+\infty} \left(\Sigma_{i=1}^3 \lambda_{i,n}(r_{i,n}, t_{i,n})\right).$$

Since, on one hand, $0 = \lim_{n\to+\infty} (\Sigma_{i=1}^3\, \lambda_{i,n}\, r_{i,n})$, and, on the other hand, sequences $(\lambda_{i,n}\, r_n)_n$, $i \in \{1,2,3\}$, have only non-negative terms, it follows that all three sequences $(\lambda_{i,n}\, r_{i,n})_n$, $i \in \{1,2,3\}$ converge to zero:

$$0 = \lim_{n\to+\infty} \lambda_{i,n}\, r_{i,n} \quad \forall i \in \{1,2,3\}. \tag{20}$$

By taking, if necessary, subsequences, we may assume that the three sequences $(\lambda_{i,n})_n$ converge, and let λ_i, $i \in \{1,2,3\}$, denote their limits. Of course,

$$\Sigma_{i=1}^3\, \lambda_i = 1. \tag{21}$$

We claim that

$$\liminf_{n\to+\infty} \lambda_{i,n}\, t_{i,n} \geq \ \lambda_i \ \inf_A \phi \quad \forall i \in \{1,2,3\}. \tag{22}$$

To prove our claim, let us pick $j \in \{1,2,3\}$; we shall distinguish two cases.

Case 1: $\lambda_j > 0$. Since the sequence $(\lambda_{j,n})_n$ converges to $\lambda_j > 0$, from relation (20) specified to the case $i = j$, it follows that the sequence $(r_{j,n})_n$ converges to zero. Keeping this fact in mind, let us apply relation (13) to $(r_{j,n},\, t_{j,n})$, and take the limit inferior upon n going to $+\infty$:

$$\begin{aligned}\liminf_{n\to+\infty} t_{j,n} \geq \ & \liminf_{n\to+\infty} (-(a+b\,R) - b\, r_{j,n}) \\ = \ & -(a+b\,R) - b \limsup_{n\to+\infty}(r_{j,n}) = -(a+b\,R).\end{aligned} \tag{23}$$

Let w be a cluster point of the visibly bounded from below sequence $(t_{j,n})_n$. Accordingly, $(0,w)$ is the limit of the sequence $(r_{j,n},\, t_{j,n})_n$, so it belongs to the set $\overline{\mathcal{A}_F}$. Relation (14) implies that

$$w \geq \inf_A \phi. \tag{24}$$

Hence, the sequence $(t_{j,n})_n$ is bounded from below, and each of its cluster points is no lest than $\inf_A \phi$, fact which proves relation (22).

Case 2: $\lambda_j = 0$. Let us apply relation (13) to $(r_{j,n},\, t_{j,n})$, multiply by $\lambda_{j,n}$ its both sides, and take the limit inferior upon n going to $+\infty$; since both sequences $(\lambda_{j,n})_n$ and $(\lambda_{j,n}\, r_{j,n})_n$ converge to zero, we get that

$$\begin{aligned}\liminf_{n\to+\infty} \lambda_{j,n}\, t_{j,n} \geq \ & \liminf_{n\to+\infty} \lambda_{j,n}\,(-(a+b\,R) - b\, r_{j,n}) \\ = \ & -(a+b\,R) \limsup_{n\to+\infty} \ \lambda_{j,n} - b \limsup_{n\to+\infty}(\lambda_{j,n}\, r_{j,n}) = 0,\end{aligned}$$

and relation (22) is proved.

We have thus proved that, for any $i \in \{1,2,3\}$, the limit inferior of the sequence $(\lambda_{i,n}\, t_{i,n})_n$ is no less than the real number $\lambda_i \ \inf_A \phi$; accordingly, by use of relation (21), we deduce that

$$\begin{aligned} t = \ & \lim_{n\to+\infty} \left(\Sigma_{i=1}^3 \ (\lambda_{i,n}\, t_{i,n})\right) \geq \Sigma_{i=1}^3 \left(\liminf_{n\to+\infty} (\lambda_{i,n}\, t_{i,n})\right) \\ & \geq \Sigma_{i=1}^3 \left(\lambda_i \ \inf_A \phi\right) = \left(\Sigma_{i=1}^3 \lambda_i\right) \ \inf_A \phi = \inf_A \phi, \end{aligned}$$

and so

$$\mathrm{co}\left(\overline{\mathcal{A}_F}\right) \cap (\{0\} \times \mathbb{R}) \subset \{0\} \times [\inf_A \phi; +\infty[. \tag{25}$$

The desired relation (12) stems from relations (10) and (25). $\square$

Remark 3.2. Let us notice that standard examples (pick for instance $X = \mathbb{R}$, $A = \{0\}$ and $\phi(x) = -\sqrt{|x|}$) prove that the supremum in the dual problem may not be achieved; Theorem 3.1 provides thus a zero duality gap valid regardless of the existence of the Lagrange multipliers.

Remark 3.3. Condition *iii)* in Theorem 3.1 can be easily reformulated as "there exist $x_0 \in X$ and $b \geq 0$ such that the function $x \mapsto \varphi(x) + \mathrm{bd}(x, x_0)$ is bounded from below" or equivalently, "the function φ admits a globally Lipschitzian minorant".

The following easy application of the previous theorem may be seen as a generalization of the second part of Proposition 1.3, as it removes the blanket assumptions of the continuity of ϕ at some point of A, on one hand, and of the convexity of A, on the other.

Corollary 3.4. *Let $(X, \|\cdot\|_X)$ be a real normed space, and define the optimization problem*

$$F : X \times \mathbb{R} \to \overline{\mathbb{R}}, \qquad F(x, r) := \begin{cases} \phi(x) & d(x, A) \leq r \\ +\infty & d(x, A) > r, \end{cases}$$

where $A \subset X$ is a weakly compact set, and $\phi : X \to \mathbb{R} \cup \{+\infty\}$ is a convex lower semi-continuous function which takes at least one finite value on A. Then the property (MS) holds.

Proof of Corollary 3.4. As ϕ is convex and lower semi-continuous, then ϕ is also weakly lower semi-continuous. Accordingly, ϕ and A fulfills the requirements of Theorem 3.1 when X is a real normed space and τ is the weak topology (so that A is norm-bounded).

Moreover, a standard consequence of the Hahn-Banach Theorem proves that any function from $\Gamma(X)$ is bounded from below by an affine continuous function:

$$\phi(x) \geq \langle x^*, x \rangle + a \tag{26}$$

for some $x^* \in X^*$ and $a \in \mathbb{R}$. Let b be a real number such that $b \geq \|x^*\|_{X^*}$ (as customary, $\|\cdot\|_{X^*}$ denotes the standard norm on X^*); thus

$$b\,\|x\|_X \geq \|x^*\|_{X^*}\,\|x\|_X \geq -\langle x^*,\, x\rangle\,, \tag{27}$$

and we get that

$$\phi(x) + b\,\|x\|_X \geq a.$$

Accordingly, given $x_0 \in X$, it holds that

$$\phi(x) + b\|x - x_0\|_X \geq \phi(x) + b\|x\|_X - b\|x_0\|_X \geq a - b\|x_0\|_X,$$

so the function $\phi(x) + b\|x - x_0\|_X$ is bounded from below. Hence, statement *iii)* from Theorem 3.1 is valid, which means that property min-sup holds true. □

4. Conclusions

The newly introduced notion of convex closedness regarding a given set (Definition 1.6) lies at the basis of a new method in addressing the duality gap for fully non-convex optimization problems depending on parameters.

This method, developed in Sections 2 and 3) had proved herself versatile enough in order to allows us to completely characterize (Theorem 3.1) the min-sup property for the classical penalty method with fully non-convex data, in the general setting of a metric space. Moreover, Corollary 3.4 brings something new even in the case of convex data, as, unlike known results, such as Proposition 1.3, it does not require the continuity of the objective function at some point of the feasible set.

Acknowledgements. The authors would like to thank the referee for his careful reading and helpful remarks.

References

[1] L. Ban, W. Song: Duality gap of the conic convex constrained optimization problems in normed spaces, Math. Program. 119 (2009) 195–214.

[2] V. Barbu, T. Precupanu: Convexity and Optimization in Banach Spaces, 2nd Ed., D. Reidel, Dordrecht (1986).

[3] J. M. Borwein, A. S. Lewis: Partially finite convex programming, Part I: Quasi relative interiors and duality theory, Math. Program. 57 (1992) 15–48.

[4] J. M. Borwein, A. S. Lewis: Convex Analysis and Nonlinear Optimization, 2nd Ed., Springer, New York (2006).

[5] R. Boţ: Conjugate Duality in Convex Optimization, Springer, Berlin (2010).

[6] T. Champion: Duality gap in convex programming, Math. Program. 99 (2004) 487–498.

[7] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, North-Holland, Amsterdam (1976).

[8] E. Ernst, M. Volle: Generalized Courant-Beltrami penalty functions and zero duality gap for conic convex programs, Positivity 17 (2013) 945–964.

[9] J.-J. Moreau: Fonctionelles Convexes, Edizioni del Dipartimento di Ingegneria Civile dell'Università di Roma Tor Vergata, Roma (2003).

[10] T. Precupanu: Closedness conditions for the optimality of a family of nonconvex optimization problems, Math. Operationsforsch. Stat., Ser. Optimization 15 (1984) 339–346.

[11] R. T. Rockafellar: Conjugate Duality and Optimization, Reg. Conf. Ser. in Appl. Math. 16, SIAM, Philadelpia (1974).

[12] V. N. Solov'ev: Duality of nonconvex extremal problems, Sov. Math., Dokl. 42 (1991) 380–383.

[13] C. Zălinescu: Convex Analysis in General Vector Spaces, Word Scientific, Singapore (2002).