Variational Analysis and Appls.
F. Giannessi and A. Maugeri, Eds.
Kluwer Acad. Publ., Dordrecht, 2004

CONVERGENCE AND STABILITY OF A REGULARIZATION METHOD FOR MAXIMAL MONOTONE INCLUSIONS AND ITS APPLICATIONS TO CONVEX OPTIMIZATION

Ya. I. Alber,[1] D. Butnariu[2] and G. Kassay[3]
Faculty of Mathematics, Technion - Israel Institute of Technology, Haifa, Israel;[1] Department of Mathematics, University of Haifa, Haifa, Israel;[2] Faculty of Mathematics, University Babes-Bolyai, Cluj-Napoca, Romania[3]

Abstract: In this paper we study the stability and convergence of a regularization method for solving inclusions $f \in Ax$, where A is a maximal monotone point-to-set operator from a reflexive smooth Banach space X with the Kadec-Klee property to its dual. We assume that the data A and f involved in the inclusion are given by approximations A^k and f^k converging to A and f, respectively, in the sense of Mosco type topologies. We prove that the sequence $x^k = (A^k + \alpha_k J^\mu)^{-1} f^k$ which results from the regularization process converges weakly and, under some conditions, converges strongly to the minimum norm solution of the inclusion $f \in Ax$, provided that the inclusion is consistent. These results lead to a regularization procedure for perturbed convex optimization problems whose objective functions and feasibility sets are given by approximations. In particular, we obtain a strongly convergent version of the generalized proximal point optimization algorithm which is applicable to problems whose feasibility sets are given by Mosco approximations

Key words: Maximal monotone inclusion, Mosco convergence of sets, regularization method, convex optimization problem, generalized proximal point method for optimization.

2000 Mathematics Subject Classification: Primary: 47J06, 47A52, 90C32; Secondary: 47H14, 90C48, 90C25.

1. INTRODUCTION

Let X be a reflexive, strictly convex and smooth Banach space with the Kadec-Klee property (i.e., such that if a sequence $\{x^k\}_{k\in\mathbb{N}}$ in X converges weakly to some $x \in X$, then $\{x^k\}_{k\in\mathbb{N}}$ converges strongly whenever $\lim_{k\to\infty}\|x^k\| = \|x\|$) and let X^* be the dual of X. Given a maximal monotone mapping $A : X \to 2^{X^*}$ and an element $f \in X^*$, we consider the following problem

$$\text{Find } x \in X \text{ such that } f \in Ax. \tag{1}$$

Problems like (1) are often ill-posed in the sense that they may not have solutions, may have infinitely many solutions and/or small data perturbations may lead to significant distortions of the solution sets. A *regularization technique,* whose basic idea can be traced back to Browder [16] and Cruceanu [23], consists of replacing the original problem (1) by the problem

$$\text{Find } z^\alpha \in X \text{ such that } f \in (A + \alpha J^\mu) z^\alpha, \tag{2}$$

where α is a positive real number and $J^\mu : X \to X^*$ is the duality mapping of gauge μ defined by the equations

$$\left\langle J^\mu y, y \right\rangle = \|J^\mu y\|_* \|y\| \text{ and } \|J^\mu y\|_* = \mu\left(\|y\|\right), \tag{3}$$

while $\mu : [0,+\infty) \to [0,+\infty)$ is supposed to be continuous, strictly increasing, having $\mu(0) = 0$ and $\lim_{t\to\infty}\mu(t) = +\infty$. One does so for several reasons. First, since the mapping $A + \alpha J^\mu$ is surjective and $\left(A + \alpha J^\mu\right)^{-1}$ is single valued (cf. [22, Proposition 3.10, p. 165]), the regularized problem (2) has unique solution (even if the inclusion (1) has no solution at all). Second, it follows from [47, p. 129] and [23] that, if $\{\alpha_k\}_{k\in\mathbb{N}}$ is a sequence of positive real numbers and $\lim_{k\to\infty}\alpha_k = 0$ then by solving (2) for $\alpha = \alpha_k$ one finds vectors z^{α_k} converging to a solution of (1) provided that this inclusion is consistent. Third, the operator $\left(A + \alpha J^\mu\right)^{-1}$ is continuous and, therefore, small perturbations of f will not make the vector z^α be far from the theoretical solution $\left(A + \alpha J^\mu\right)^{-1} f$ of (2). In applications it frequently happens that not only f but also the operator A involved in (1) can be approximated but not precisely computed. This naturally leads to the question whether the regularized inclusion (2) is stable, that is, whether by solving instead of the regularized inclusion (2) a sequence of regularized inclusions

$$f^k \in (A^k + \alpha_k J^\mu)x$$

in which $A^k : X \to 2^{X^*}$ are maximal monotone operators approximating A and f^k approximates f, the sequence of corresponding solutions

$$x^k = \left(A^k + \alpha_k J^\mu\right)^{-1} f^k \tag{4}$$

still converges to a solution of (1) when $\lim_{k\to\infty} \alpha_k = 0$ and the original inclusion (1) is consistent. This question was previously considered by Lavrentev [36] who dealt with it in Hilbert spaces under the assumption that A is linear and positive semidefinite, $*\mathrm{Dom}\, A = X$ and $\mu(t) = t/2.$ In Alber [1] the problem appears in a more general context but under the assumption that the operator A is defined on the whole Banach space X.

The main purpose of this paper is to show that if the approximations A^k and f^k satisfy some quite mild requirements, then the answer to the question posed above is affirmative, i.e., the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ defined by (4) converges strongly to the minimal norm solution of (1) as $\alpha_k \to 0$ and provided that (1) has at least one solution. Subsequently, we prove that the stability results we have obtained for the regularization method presented above apply to the resolution of convex optimization problems with perturbed data and, in particular, to produce a strongly convergent version of a proximal point method.

The stability results proved in this work (see Section 2) do not make additional demands on the data of the original inclusion (1) besides the assumption that A is maximal monotone. The conditions under which we prove those results only concern the quality of the approximations A^k and f^k. They ask that either the Mosco weak upper limit (as defined in [45]) or the weak-strong upper limit (introduced in Subsection 2.1 below) of the sequence of sets $\left\{Graph\left(A^k\right)\right\}_{k\in\mathbb{N}}$ be a subset of $\mathrm{Graph}(A)$, the later being a somewhat weaker requirement. Also, they ask for a kind of linkage of the approximative data in the form of the boundedness of the sequence

$$\left\{\alpha_k^{-1}\operatorname{dist}_*(f^k, A^k v^k)\right\}_{k\in\mathbb{N}} \tag{5}$$

for some bounded sequence $\left\{v^k\right\}_{k\in\mathbb{N}}$ in X. If approximants A^k and f^k satisfying these conditions exist, then the inclusion (1) is necessarily consistent, the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ defined by (4) is bounded and its weak accumulation points are solutions of it (see Theorem 2.2 and Corollary 2.3). The main stability results we prove for the proposed regularization scheme are Theorem 2.4 and its Corollary 2.5. They show that if solutions of (1) exist and each of them is the limit of a sequence $\left\{v^k\right\}_{k\in\mathbb{N}}$ such that the sequence (5) converges to zero, then the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ given by (4) converges strongly to the minimal norm solution of (1).

When one has to solve optimization problems like that of finding a vector

$$x^* \in \operatorname{argmin}\left\{F(x) : g_i(x) \le 0,\ i \in I\right\}, \tag{6}$$

where the functions F, $g_i : X \to (-\infty, +\infty]$ are convex and lower semicontinuous, perturbations of data are inherent because of imprecise computations and measurements. Since problems like (6) may happen to be ill-posed, replacing the original data F and g_i by approximations F_k and g_i^k may lead to significant distortions of the solution set. In Section 3 we consider (6) and its perturbations in their subgradient inclusion form. We apply the stability results presented in Section 2 for finding out how "good" the approximative data F_k and g_i^k should be in order to ensure that the vectors x^k resulting from the resolution of the regularized perturbed inclusions strongly approximate solutions of (6). Theorem 3.2 answers this question. It shows that for this to happen it is sufficient that the perturbed data would satisfy the conditions (A) and (B) given in Subsection 3.1. Condition (A) asks for sufficiently uniform point-wise convergence of F_k to F. Condition (B) guarantees weak-strong upper convergence of the feasibility sets of the perturbed problems to the feasibility set of the original problem. Proposition 3.6 provides a tool for verifying the validity of condition (B) in the case of optimization problems with affine constraints as well as in the case of some problems of semidefinite programming.

In Section 4 we consider the question whether or under which conditions the generalized proximal point method for optimization which emerged from the works of Martinet [43], [44], Rockafellar [52] and Censor and Zenios [21] can be forced to converge strongly in infinite dimensional Banach spaces. The origin of this question can be traced back to Rockafellar's work [52]. The relevance of the question emerges from the role of the proximal point method in the construction of augmented Lagrangian algorithms (see [53], [18, Chapter 3] and [30]): in this context a better behaved sequence obtained by regularization of the proximal point method may be of use in order to determine better approximations for a solution of the primal problem. It was shown by Butnariu and Iusem [17] that in smooth uniformly convex Banach spaces the generalized proximal point method converges subsequentially weakly, and sometimes weakly, to solutions of the optimization problem to which it is applied. However, it follows from the work of Güler [28] that the sequences generated by the proximal point method may fail to converge strongly. The generalized proximal point method essentially consists of solving a sequence of perturbed variants of the given convex optimization problem. We apply the results established in Section 3 in order to prove that by regularizing the perturbed problems via the scheme studied in this paper we obtain a sequence $\{(y^k, x^k)\}_{k\in\mathbb{N}}$ in $X \times X$ such that, when the optimization problem is consistent, $\{F(y^k)\}_{k\in\mathbb{N}}$ converges to the optimal value of F and $\{x^k\}_{k\in\mathbb{N}}$ converges strongly to the minimum norm optimal solution of the original optimization problem.

The stability of the regularization scheme represented by (2) was studied before in various settings, but mostly as a way of regularizing variational inequalities involving maximal monotone operators (which, in view of Minty's Theorem, can be also seen as a way of regularizing inclusions involving maximal monotone operators). Mosco [45], [46], Liskovets [39], [40], [41], Ryazantseva [54], Alber and Ryazantseva [6], Alber [2], Alber and Notik [5] have considered the scheme under additional assumptions (not made in our current work) concerning the data A and f (as, for instance, some kind of continuity or that the perturbed operators A^k and A should have the same domains). The stability results they have established usually require Hausdorff metric type convergence conditions for the graphs of A^k. Also under Hausdorff metric type convergence conditions, but with no additional demands on the operator A than its maximal monotonicity, strong convergence of the regularized sequence $\{x^k\}_{k\in\mathbb{N}}$ defined by (4) to the minimal norm solution of (1) was proven by Alber, Butnariu and Ryazantseva in [4]. Recently, weak convergence properties of this regularization scheme were proved by Alber [3] under metric and Mosco type convergence assumptions on the approximants. By contrast, we establish here strong convergence of the regularized sequence $\{x^k\}_{k\in\mathbb{N}}$ by exclusively using variants of Mosco type convergence for the approximants.

The stability of regularization schemes applied to ill-posed problems is a multifaceted topic with multiple applications in various fields as one can see from the monographs of Lions and Magenes [37], Dontchev and Zolezzi [24], Kaplan and Tichatschke [31], Engl, Hanke and Neubauer [27], Showalter [55], and Bonnans and Shapiro [14]. We prove here that the regularization scheme (4) has strong and stable convergence behavior under undemanding conditions and that it can be applied to a large class of convex optimization problems. An interesting topic for further research is to find out whether and under which conditions this regularization scheme works when applied to other problems like, for instance, differential equations [55], inverse problems [27], linearized abstract equations [14, Section 5.1.3.], etc. which, in many circumstances, can be represented as inclusions involving maximal monotone operators. Convergence of the regularization scheme (4) may happen to be slow (as shown by an example given in [4]). Its rate of convergence seems to depend not only on the properties of A^k and f^k but also on the geometry of the Banach space X in which the problem is set. It is an interesting open problem to evaluate the rate of convergence of the regularization scheme discussed in this work in a way similar to that in which such rates were evaluated for alternative regularization methods by Kaplan and Tichatschke [34], [33], [32] and [42]. Such an evaluation may help decide for which type of problems and in which settings application of the regularization scheme (4) is efficient.

The convergence and the reliability under errors of the generalized proximal point method in finite dimensional spaces was systematically studied along the last decade (see [25], [26] and see [29] for a survey on this topic). In infinite dimensional Hilbert spaces repeated attempts were recently made in order to discover how the problem data should be in order to ensure that the generalized proximal point method converges weakly or strongly under error perturbations (see [8], [9], [15], [30]). Projected subgradient type regularization techniques meant to force strong convergence in Hilbert spaces of Rockafellar's classical proximal point algorithm were discovered by Bauschke and Combettes [12, Corollary 6.2] and Solodov and Svaiter [57]. The regularized generalized proximal point method we propose in Section 4 works in non Hilbertian spaces too. It presents an interesting feature which can be easily observed from Theorem 4.2 and Corollary 4.3: if X is uniformly convex, smooth and separable, then by applying the regularized generalized proximal point method (60) one can reduce resolution of optimization problems in spaces of infinite dimension to solving a sequence of optimization problems in spaces of finite dimension whose solutions will necessarily converge strongly to the minimal norm optimum of the original problem.

2. CONVERGENCE AND STABILITY ANALYSIS FOR MAXIMAL MONOTONE INCLUSIONS

2.1 We start our discussion about the stability of the regularization scheme (2) by recalling (see [45, Definition 1.1]) that a sequence $\{S_k\}_{k\in\mathbb{N}}$ of subsets of X is called *convergent* (*in Mosco sense*) if

$$w-\overline{\lim}\ S_k = s-\underline{\lim}\ S_k,$$

where $s-\underline{\lim}\ S_k$ represents the collection of all $y\in X$ which are limits (in the strong convergence sense) of sequences with the property that $x^k \in S_k$ for all $k\in\mathbb{N}$ and $w-\overline{\lim} S_k$ denotes the collection of all $x\in X$ such that there exists a sequence $\{y^k\}_{k\in\mathbb{N}}$ in X converging weakly to x and with the property that there exists a subsequence $\{S_{i_k}\}_{k\in\mathbb{N}}$ of $\{S_k\}_{k\in\mathbb{N}}$ such that $y^k \in S_{i_k}$ for all $k\in\mathbb{N}$. In this case, the set

$$S := s-\underline{\lim}\ S_k = w-\overline{\lim}\ S_k$$

is called the *limit* of $\{S_k\}_{k\in\mathbb{N}}$ and is denoted $S = LimS_k$.

By analogy with Mosco's $w-\overline{\lim}$ we introduce the following notion of limit for sequences of sets contained in $X\times X^*$. This induces a form of graphical convergence for point-to-set mappings from X to X^* which we use in the sequel. For a comprehensive discussion of other notions of convergence of sequences of sets see [13].

Definition. The *weak-strong upper limit of a sequence* $\{U_k\}_{k\in\mathbb{N}}$ of subsets of $X\times X^*$, denoted $ws-\overline{\lim}\ U_k$, is the collection of all pairs $(x,y)\in X\times X^*$ for which there exists a sequence $\{x^k\}_{k\in\mathbb{N}}$ contained in X which converges weakly to x and a sequence $\{y^k\}_{k\in\mathbb{N}}$ contained in X^* which converges strongly to y and such that, for some subsequence $\{U_{i_k}\}_{k\in\mathbb{N}}$ of $\{U_k\}_{k\in\mathbb{N}}$ we have $(x^k, y^k)\in U_{i_k}$ for all $k\in\mathbb{N}$.

It is easy to see that, if $A^k : X\to X^*$, $k\in\mathbb{N}$, is a sequence of point-to-set mappings then the weak-strong upper limit of the sequence $U_k = Graph\left(A^k\right)$, $k\in\mathbb{N}$, is the set U of all pairs $(x,y)\in X\times X^*$ with the property that there exists a sequence $\left\{(x^k, y^k)\right\}_{k\in\mathbb{N}} \subset X \times X^*$ such that $\left\{x^k\right\}_{k\in\mathbb{N}}$ converges weakly to x in X, $\left\{y^k\right\}_{k\in\mathbb{N}}$ converges strongly to y in X^* and, for some subsequence $\left\{A^{i_k}\right\}_{k\in\mathbb{N}}$ of $\left\{A^k\right\}_{k\in\mathbb{N}}$ we have

$$y^k \in A^{i_k}(x^k),\ \forall k \in \mathbb{N}.$$

Therefore, in virtue of [11, Proposition 7.1.2.], the graphical upper limit of the sequence $\left\{A^k\right\}_{k\in\mathbb{N}}$, $\lim^{\#}_{k\to\infty} A^k$, considered in [11, Definition 7.1.1], the weak-strong upper limit $ws - \overline{\lim}\ Graph\left(A^k\right)$ and the Mosco upper limit $w - \overline{\lim} Graph\left(A^k\right)$ are related by

$$Graph\left(\lim\nolimits^{\#}_{k\to\infty} A^k\right) \subseteq ws - \overline{\lim}\ Graph\left(A^k\right) \subseteq w - \overline{\lim}\ Graph\left(A^k\right). \qquad (7)$$

As noted in the Introduction, a goal of this work is to establish convergence and stability of the regularization scheme (4) under undemanding convergence requirements for the approximative data A^k and f^k. As far as we know, the most general result in this respect is that presented in [4 Section 2]. It guarantees convergence and stability of the regularization scheme (4) under the requirement that the maximal monotone operators A^k approximate the maximal monotone operator A in the sense that there exist three functions $a, g, \zeta : \mathbb{R}_+ \to \mathbb{R}_+$, where ζ is strictly increasing and continuous at zero, such that for any $(x, y) \in Graph(A)$ and for any $k \in \mathbb{N}$, there exists a pair $(x^k, y^k) \in Graph\left(A^k\right)$ with the property that

$$\left\|x - x^k\right\| \le a(\|x\|)k^{-1} \text{ and } \left\|y - y^k\right\|_* \le g(\|y\|_*)\zeta(k^{-1}). \qquad (8)$$

Clearly, if this requirement is satisfied, then

$$\mathrm{Graph}(A) \subseteq \mathrm{Graph}\left(\lim\nolimits^{\flat}_{k\to\infty} A^k\right), \qquad (9)$$

where $\lim^{\flat}_{k\to\infty} A^k$ stands for the graphical lower limit of the sequence $\left\{A^k\right\}_{k\in\mathbb{N}}$ (see [11, p. 267]). Since the mappings A^k and A we work with are maximal monotone, Proposition 7.1.7 from [11] applies and, due to (9), it implies that A is exactly the graphical limit of the sequence $\left\{A^k\right\}_{k\in\mathbb{N}}$, that is,

$$A=\lim^{\flat}_{k\to\infty}A^k=\lim^{\#}_{k\to\infty}A^k. \tag{10}$$

In this section we show that convergence and stability of the regularization scheme (4) can be ensured under conditions that are much less demanding than the locally uniform graphical convergence (8). In fact, we prove convergence and stability of the scheme (4) by requiring (see (16) below) less than the graphical convergence (10). This allows us to apply the regularization scheme to a wide class of convex optimization problems as shown in Sections 3 and 4.

All over this paper we denote by $\mu:[0,+\infty)\to[0,+\infty)$ a gauge function with the property that the following limit exists and we have

$$\lim_{t\to\infty}\frac{\mu(t)}{t}>0. \tag{11}$$

The duality mapping of gauge μ is denoted J^{μ}, as usual.

2.2 The next result shows that, under quite mild conditions concerning the mappings A^k and the vectors f^k, the sequence $\left\{x^k\right\}_{k\in N}$ generated in X according to (4) is well defined, bounded and that its weak accumulation points are necessarily solutions of (1).

Theorem. *Let $\{\alpha_k\}_{k\in\mathbb{N}}$ be a bounded sequence of positive real numbers. Suppose that, for each $k\in\mathbb{N}$, the mapping $A^k:X\to 2^{X^*}$ is maximal monotone. Then the following statements are true:*

(i) *The sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ given by* (4) *is well defined;*

(ii) *If there exists a bounded sequence $\left\{v^k\right\}_{k\in\mathbb{N}}$ in X such that the sequence* (5) *is bounded, then the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ is bounded too and has weak accumulation points;*

(iii) *If, in addition to the requirements in (ii), we have that the sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ converges to zero, the sequence $\left\{f^k\right\}_{k\in\mathbb{N}}$ converges weakly to f in X^* and*

$$w-\overline{\lim}\ Graph\left(A^k\right)\subseteq Graph\left(A\right), \tag{12}$$

then the problem (1) *has at least one solution and any weak accumulation point of* $\{x^k\}_{k\in\mathbb{N}}$ *is a solution of it.*

Proof. Since the mappings A^k are maximal monotone it follows that $A^k+\alpha_k J^\mu$ are surjective and $\left(A^k+\alpha_k J^\mu\right)^{-1}$ are single valued. Hence, the sequence $\{x^k\}_{k\in\mathbb{N}}$ is well defined. In order to show that this sequence is bounded, observe that, for each $k\in\mathbb{N}$, there exists a function $h^k\in A^k x^k$ such that

$$f^k=h^k+\alpha_k J^\mu x^k. \tag{13}$$

The sets $A^k v^k$ are nonempty because, otherwise, the sequence (5) would be unbounded. Also, these sets are convex and closed. Hence, for each $k\in\mathbb{N}$ there exists $g^k\in A^k v^k$ such that

$$\left\|g^k-f^k\right\|_*=\operatorname{dist}_*(f^k,A^k v^k). \tag{14}$$

Taking into account that A^k is monotone, we deduce

$$\left\langle h^k-g^k,x^k-v^k\right\rangle\geq 0.$$

Hence,

$$\begin{aligned}\left\langle g^k,x^k-v^k\right\rangle &\leq\left\langle h^k,x^k-v^k\right\rangle=\left\langle f^k-\alpha_k J^\mu x^k,x^k-v^k\right\rangle\\ &=\left\langle f^k,x^k-v^k\right\rangle-\alpha_k\left\langle J^\mu x^k,x^k\right\rangle+\alpha_k\left\langle J^\mu x^k,v^k\right\rangle\\ &=\left\langle f^k,x^k-v^k\right\rangle-\alpha_k\mu\left(\left\|x^k\right\|\right)\left\|x^k\right\|+\alpha_k\left\langle J^\mu x^k,v^k\right\rangle\\ &\leq\left\langle f^k,x^k-v^k\right\rangle-\alpha_k\mu\left(\left\|x^k\right\|\right)\left\|x^k\right\|+\alpha_k\mu\left(\left\|x^k\right\|\right)\left\|v^k\right\|,\end{aligned}$$

where the first equality follows from (13) and the third equality, as well as the last inequality, follows from (3). By consequence,

$$\begin{aligned}\alpha_k\mu\left(\left\|x^k\right\|\right)\left(\left\|x^k\right\|-\left\|v^k\right\|\right) &\leq\left\langle f^k-g^k,x^k-v^k\right\rangle\\ &\leq\left\|f^k-g^k\right\|_*\left\|x^k\right\|+\left\|f^k-g^k\right\|_*\left\|v^k\right\|\end{aligned} \tag{15}$$

for all $k \in \mathbb{N}$. Suppose, by contradiction, that $\{x^k\}_{k\in\mathbb{N}}$ is unbounded. Then, for some subsequence $\{x^{i_k}\}_{k\in\mathbb{N}}$ of it we have $\lim_{k\to\infty} \|x^{i_k}\| = +\infty$. From (15) we deduce that, for sufficiently large k, we have

$$\frac{1}{\|x^{i_k}\|}\mu\left(\|x^{i_k}\|\right)\left(\|x^{i_k}\| - \|v^{i_k}\|\right) \le \frac{1}{\alpha_{i_k}}\|f^{i_k} - g^{i_k}\|_* \left(1 + \frac{\|v^{i_k}\|}{\|x^{i_k}\|}\right),$$

where, according to (14) and the hypothesis, the sequence $\left\{\alpha_k^{-1}\|f^k - g^k\|_*\right\}_{k\in\mathbb{N}}$ is bounded. Taking on both sides of this inequality the upper limit as $k \to \infty$ and taking into account (11), (14) and the boundedness of $\{v^k\}_{k\in\mathbb{N}}$ one gets that the limit on the left hand side is $+\infty$ while that on the right hand side is finite, that is, a contradiction. This shows that $\{x^k\}_{k\in\mathbb{N}}$ is bounded and, since X is reflexive, $\{x^k\}_{k\in\mathbb{N}}$ has weak accumulation points.

Now, assume that $\{\alpha_k\}_{k\in\mathbb{N}}$ converges to zero, $\{f^k\}_{k\in\mathbb{N}}$ converges weakly to f in X^* and (12) also holds. Observe that the sequence $\left\{\|g^k - f^k\|_*\right\}_{k\in\mathbb{N}}$ converges to zero because $\{\alpha_k\}_{k\in\mathbb{N}}$ converges to zero, $M := \sup_{k\in\mathbb{N}} \alpha_k^{-1} \operatorname{dist}_*(f^k, A^k v^k)$ is finite and

$$\|f^k - g^k\|_* \le \alpha_k M,$$

for all $k \in \mathbb{N}$. Consequently, since $\{f^k\}_{k\in\mathbb{N}}$ converges weakly to f, we deduce that $\{g^k\}_{k\in\mathbb{N}}$ converges weakly to f too. Let v be a weak accumulation point of the sequence $\{v^k\}_{k\in\mathbb{N}}$ and denote by $\{v^{i_k}\}_{k\in\mathbb{N}}$ a subsequence of $\{v^k\}_{k\in\mathbb{N}}$ converging weakly to v. Since for any $k \in \mathbb{N}$ we have $(v^{i_k}, g^{i_k}) \in Graph\left(A^{i_k}\right)$, condition (12) implies that $(v, f) \in Graph(A)$, i.e., v is a solution of (1). Let x be a weak accumulation point of $\{x^k\}_{k\in\mathbb{N}}$ and let $\{x^{j_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{x^k\}_{k\in\mathbb{N}}$ which converges weakly to x. Note that for any $z \in X$ we have

$$\begin{aligned}\left|\left\langle z, f-h^k\right\rangle\right| &= \left|\left\langle z, f-f^k\right\rangle + \left\langle z, f^k-h^k\right\rangle\right| \\ &= \left|\left\langle z, f-f^k\right\rangle + \alpha_k\left\langle z, J^\mu x^k\right\rangle\right| \\ &\le \left|\left\langle z, f-f^k\right\rangle\right| + \alpha_k\|z\|\left\|J^\mu x^k\right\|_*,\end{aligned}$$

where the last sum converges to zero as $k\to\infty$. This shows that the sequence $\{h^k\}_{k\in\mathbb{N}}$ converges weakly to f. Hence, the sequence $\left\{(x^{j_k}, h^{j_k})\right\}_{k\in\mathbb{N}}$ converges weakly to (x, f) in $X\times X^*$. Since we also have that $h^{j_k}\in A^{j_k}x^{j_k}$ for all $k\in\mathbb{N}$, condition (12) implies that $(x,f)\in Graph(A)$, that is, x is a solution of (1).

2.3 Condition (12) involved in Theorem 2.2 is difficult to verify in applications as those discussed in Section 3 below. We show next that this condition can be relaxed at the expense of strenghtening the convergence requirements for $\{f^k\}_{k\in\mathbb{N}}$. Note that in view of (7) condition (16) below is weaker than (12). Precisely, we have the following result:

Corollary. *Let $\{\alpha_k\}_{k\in\mathbb{N}}$ be a sequence of positive real numbers converging to zero. Suppose that, for each $k\in\mathbb{N}$, the mapping $A^k: X\to 2^{X^*}$ is maximal monotone and that there exists a bounded sequence $\{v^k\}_{k\in\mathbb{N}}$ in X such that the sequence* (5) *is bounded. Then the sequence $\{x^k\}_{k\in\mathbb{N}}$ given by* (4) *is well defined, bounded and has weak accumulation points. If, in addition, the sequence $\{f^k\}_{k\in\mathbb{N}}$ converges strongly to f in X^* and*

$$ws-\overline{\lim}\ Graph\left(A^k\right)\subseteq Graph\left(A\right), \tag{16}$$

then the problem (1) *has solutions and any weak accumulation point of $\{x^k\}_{k\in\mathbb{N}}$ is a solution of it.*

Proof. Well definedness and boundedness of the sequence $\{x^k\}_{k\in\mathbb{N}}$ results from Theorem 2.2. Exactly as in the proof of Theorem 2.2 we deduce that for each $k\in\mathbb{N}$ there exist $h^k\in A^kx^k$ and $g^k\in A^kv^k$ such that (13) and (14) hold. Observe that the sequence $\{g^k\}_{k\in\mathbb{N}}$ converges strongly to f because of (14) and the boundedness of (5). It remain to show that, under the assumptions that $\{f^k\}_{k\in\mathbb{N}}$ converges strongly to f and (16) holds, any weak accumulation point of $\{x^k\}_{k\in\mathbb{N}}$ is a solution of (1). Let v be a weak accumulation point of the sequence $\{v^k\}_{k\in\mathbb{N}}$ (such a point exists because $\{v^k\}_{k\in\mathbb{N}}$ is bounded and X is reflexive) and denote by $\{v^{i_k}\}_{k\in\mathbb{N}}$ a subsequence of $\{v^k\}_{k\in\mathbb{N}}$ converging weakly to v. Since for all $k\in\mathbb{N}$ we have $(v^{i_k}, g^{i_k})\in Graph\left(A^{i_k}\right)$, condition (16) implies that $(v,f)\in Graph(A)$, i.e., v is a solution of (1). Let x be a weak accumulation point of $\{x^k\}_{k\in\mathbb{N}}$ and let $\{x^{j_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{x^k\}_{k\in\mathbb{N}}$ which converges weakly to x. Note that, according to (13), we have

$$\begin{aligned}\left\|f-h^k\right\|_* &\le \left\|f-f^k\right\|_* + \left\|f^k-h^k\right\|_* \\ &= \left\|f-f^k\right\|_* + \alpha_k \left\|J^\mu x^k\right\|_*,\end{aligned}$$

where the last sum converges to zero as $k\to\infty$, because $\{x^k\}_{k\in\mathbb{N}}$ is bounded (and, hence, so is $\{J^\mu x^k\}_{k\in\mathbb{N}}$) and the sequence $\{f^k\}_{k\in\mathbb{N}}$ converges to f by hypothesis. Therefore, the sequence $\{h^k\}_{k\in\mathbb{N}}$ converges strongly to f. Since we also have that $h^k\in A^k x^k$ for all $k\in\mathbb{N}$, condition (16) implies that $(x,f)\in Graph(A)$, that is, x is a solution of (1). □

2.4 If problem (1) has only one solution (as happens, for instance, when A is strictly monotone), then Theorem 2.2 guarantees weak convergence of the whole sequence $\{x^k\}_{k\in\mathbb{N}}$. However, in general, we do not know whether the whole sequence $\{x^k\}_{k\in\mathbb{N}}$ converges weakly. The next result shows that not only weak convergence, but also strong convergence of $\{x^k\}_{k\in\mathbb{N}}$ to a solution of (1) can be ensured provided that any element of $A^{-1}f$ (the solution set) is the limit of a sequence $\{v^k\}_{k\in\mathbb{N}}$ satisfying (17) below. In view of the remarks in Subsection 2.1, this result improves upon Theorem 2.2 in [4].

Theorem. *Suppose that problem* (1) *has at least one solution and that the sequence of positive real numbers* $\{\alpha_k\}_{k\in\mathbb{N}}$ *converges to zero. If* $A^k : X \to 2^{X^*}$, $k \in \mathbb{N}$, *are maximal monotone operators with the property* (12)*, if* $\{f^k\}_{k\in\mathbb{N}}$ *is a sequence converging weakly to* f *in* X^* *and if, for each* $v \in A^{-1}f$, *there exists a sequence* $\{v^k\}_{k\in\mathbb{N}}$ *which converges strongly to* v *in* X *and such that*

$$0 \in s - \underline{\lim}\ \frac{1}{\alpha_k}\left[A^k v^k - f^k\right], \tag{17}$$

then the sequence $\{x^k\}_{k\in\mathbb{N}}$ *given by* (4) *is well defined and converges strongly to the minimal norm solution of problem* (1).

Proof. The assumption that problem (1) has solutions implies, in our current setting, the existence of a bounded sequence $\{v^k\}_{k\in\mathbb{N}}$ as required by Theorem 2.2. Observe that, since (17) holds, the sequence $\{\alpha_k^{-1}\mathrm{dist}_*(f^k, A^k v^k)\}_{k\in\mathbb{N}}$ converges to zero and, therefore, it is bounded. Hence, one can apply Theorem 2.2 in order to deduce well definedness and boundedness of $\{x^k\}_{k\in\mathbb{N}}$ and the fact that any weak accumulation point of it is a solution of (1). Note that, since A is maximal monotone, A^{-1} is maximal monotone too and, therefore, the set $A^{-1}f$, which is exactly the presumed nonempty solution set of problem (1), is convex and closed. The space X is reflexive and strictly convex and, therefore, the nonempty, convex and closed set $A^{-1}f$ contains a unique minimal norm element $\bar{x}$ (the metric projection of 0 onto the set $A^{-1}f$). We show that the only weak accumulation point of $\{x^k\}_{k\in\mathbb{N}}$ is $\bar{x}$. To this end, let $\{x^{i_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{x^k\}_{k\in\mathbb{N}}$ which converges weakly to some $x \in X$. According to Theorem 2.2, x is necessarily contained in $A^{-1}f$. If $x = 0$, then this is necessarily the minimal norm element of $A^{-1}f$, i.e., $x = \bar{x}$. Suppose that $x \neq 0$. Let v be any other solution of problem (1). By hypothesis, there exists a sequence $\{v^k\}_{k\in\mathbb{N}}$ converging strongly in X to v and such that, for some sequence $\{l^k\}_{k\in\mathbb{N}}$ with $l^k \in A^k v^k$ for each $k \in \mathbb{N}$, we have

$$\lim_{k\to\infty}\frac{1}{\alpha_k}\left(l^k - f^k\right) = 0. \tag{18}$$

Clearly,

$$0 < \|x\| \leq \to \liminf_{k\to\infty}\left\|x^{i_k}\right\|$$

and there exists a subsequence $\left\{x^{j_k}\right\}_{k\in\mathbb{N}}$ of $\left\{x^{i_k}\right\}_{k\in\mathbb{N}}$ such that

$$\liminf_{k\to\infty}\left\|x^{i_k}\right\| = \lim_{k\to\infty}\left\|x^{j_k}\right\|. \tag{19}$$

The subsequence $\left\{x^{j_k}\right\}_{k\in\mathbb{N}}$ is still weakly convergent to x and has

$$0 < \mu(\|x\|) \leq \mu\left(\liminf_{k\to\infty}\left\|x^{i_k}\right\|\right) = \mu\left(\lim_{k\to\infty}\left\|x^{j_k}\right\|\right) = \lim_{k\to\infty}\mu\left(\left\|x^{j_k}\right\|\right), \tag{20}$$

because μ is continuous and increasing (as being a gauge function). For each $k \in \mathbb{N}$, let $h^k \in A^k x^k$ be the function for which (13) is satisfied. These functions exist because $\left\{x^k\right\}_{k\in\mathbb{N}}$ is well defined. Due to the monotonicity of A^k, we have

$$\begin{aligned}
0 &\leq \left\langle h^k - l^k, x^k - v^k\right\rangle = \left\langle f^k - \alpha_k J^\mu x^k - l^k, x^k - v^k\right\rangle \\
&= \left\langle f^k - l^k, x^k - v^k\right\rangle - \alpha_k\left\langle J^\mu x^k, x^k\right\rangle + \alpha_k\left\langle J^\mu x^k, v^k\right\rangle \\
&\leq \left\langle f^k - l^k, x^k - v^k\right\rangle - \alpha_k\mu\left(\left\|x^k\right\|\right)\left\|x^k\right\| + \alpha_k\mu\left(\left\|x^k\right\|\right)\left\|v^k\right\|,
\end{aligned}$$

where the first equality results from (13) and the last inequality follows from (3). This implies

$$\mu\left(\left\|x^k\right\|\right)\left\|x^k\right\| \leq \frac{1}{\alpha_k}\left\langle f^k - l^k, x^k - v^k\right\rangle + \mu\left(\left\|x^k\right\|\right)\left\|v^k\right\|, \tag{21}$$

where the first term of the right hand side converges to zero as $k \to \infty$ because of (18) and because of the boundedness of $\{v^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$. Replacing k by j_k in this inequality, we deduce that for k large enough

$$\left\|x^{j_k}\right\| \le \frac{1}{\mu\left(\left\|x^{j_k}\right\|\right)} \frac{1}{\alpha_k} \left\langle f^{j_k} - l^{j_k}, x^{j_k} - v^{j_k} \right\rangle + \left\|v^{j_k}\right\|. \tag{22}$$

Letting here $k \to \infty$ we get

$$\|x\| \le \lim_{k\to\infty} \left\|x^{j_k}\right\| \le \lim_{k\to\infty} \left\|v^{j_k}\right\| = \|v\|,$$

because $\{v^k\}_{k\in\mathbb{N}}$ converges strongly to v and because of (19). Since v is an arbitrarily chosen solution of problem (1), it follows that $x = \bar{x}$. Hence, the sequence $\{x^k\}_{k\in\mathbb{N}}$ converges weakly to $\bar{x}$.

It remains to show that $\{x^k\}_{k\in\mathbb{N}}$ converges strongly. To this end, observe that, since $\{x^k\}_{k\in\mathbb{N}}$ converges weakly to $\bar{x}$ and since X is a space with the Kadec-Klee property, it is sufficient to show that $\left\{\left\|x^k\right\|\right\}_{k\in\mathbb{N}}$ converges to $\|\bar{x}\|$. In other words, it is sufficient to prove that all convergent subsequences of the bounded sequence $\left\{\left\|x^k\right\|\right\}_{k\in\mathbb{N}}$ converge to $\|\bar{x}\|$. In order to prove that, let $\left\{\left\|x^{p_k}\right\|\right\}_{k\in\mathbb{N}}$ be a convergent subsequence of $\left\{\left\|x^k\right\|\right\}_{k\in\mathbb{N}}$. If $\left\{\left\|x^{p_k}\right\|\right\}_{k\in\mathbb{N}}$ converges to 0, then

$$0 \le \|\bar{x}\| \le \liminf_{k\to\infty} \left\|x^k\right\| \le \lim_{k\to\infty} \left\|x^{p_k}\right\| = 0,$$

that is, $\|\bar{x}\| = \lim_{k\to\infty} \left\|x^{p_k}\right\| = 0.$ Suppose now that

$$\lim_{k\to\infty} \left\|x^{p_k}\right\| = \beta > 0.$$

Then, there exists a positive integer k_0 such that, for all integers $k \ge k_0$, we have $\left\|x^{p_k}\right\| > 0.$ According to (21) this implies that, for $k \ge k_0$, one has

$$\left\|x^{p_k}\right\| \le \frac{1}{\mu\left(\left\|x^{p_k}\right\|\right)} \frac{1}{\alpha_{p_k}} \left\langle f^{p_k} - l^{p_k}, x^{p_k} - v^{p_k} \right\rangle + \left\|v^{p_k}\right\|.$$

Letting $k \to \infty$ in this inequality we get

$$\|\bar{x}\| \le \liminf_{k\to\infty} \|x^k\| \le \lim_{k\to\infty} \|x^{p_k}\| \le \lim_{k\to\infty} \|v^{p_k}\| = \|v\|.$$

Since v is an arbitrarily chosen solution of problem (1) we can take here $v = \bar{x}$ and obtain $\|\bar{x}\| = \lim_{k\to\infty} \|x^{p_k}\|$. This completes the proof. □

2.5 Similarly to Corollary 2.3 ensuring that the weak accumulation points of $\{x^k\}_{k\in N}$ are solutions of (1), we can use Theorem 2.4 in order to prove strong convergence of $\{x^k\}_{k\in N}$ to a solution of (1) when condition (12) is replaced by the weaker requirement (16) but strenghtening the convergence requirements on $\{f^k\}_{k\in N}$.

Corollary. *Suppose that problem* (1) *has solutions and the sequence of positive real numbers* $\{\alpha_k\}_{k\in\mathbb{N}}$ *converges to zero. If* $A^k : X \to 2^{X^*}$, $k\in\mathbb{N}$, *are maximal monotone operators with the property* (16)*, if* $\{f^k\}_{k\in\mathbb{N}}$ *is a sequence converging strongly to* f *in* X^* *and if, for each* $v \in A^{-1}f$, *there exists a sequence* $\{v^k\}_{k\in\mathbb{N}}$ *which converges strongly to* v *in* X *and satisfies* (17)*, then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *given by* (4) *is well defined and converges strongly to the minimal norm solution of problem* (1).

Proof. Well definedness and boundedness of $\{x^k\}_{k\in\mathbb{N}}$ as well as the fact that any weak accumulation point of it is a solution of (1) result from Corollary 2.3. In order to show that $\{x^k\}_{k\in\mathbb{N}}$ converges strongly to the minimal norm solution of the problem one reproduces without modification the arguments made for the same purpose in the proof of Theorem 2.4. □

3. REGULARIZATION OF CONVEX OPTIMIZATION PROBLEMS

3.1 We have noted above that Theorem 2.4 and Corollary 2.5, can be of use in order to prove stability properties of the procedure (4) applied to optimization problems with perturbed data. Such properties are of interest in applications in which the data involved in the optimal solution finding process are affected by computational and/or measurement errors. To make things precise, in what follows $F: X \to (-\infty,+\infty]$ is a lower semicontinuous convex function and Ω is a nonempty, closed convex subset of $\operatorname{Int}(\operatorname{Dom} F)$, the interior of the domain of F. We consider the following optimization problem under the assumption that it has at least one solution:

$$(P) \quad \text{Minimize } F(x) \quad \text{subject to} \quad x \in \Omega. \tag{23}$$

It is not difficult to verify that by solving the following inclusion

$$(P') \quad \text{Find } x \in X \text{ such that } 0 \in Ax,$$

where $A: X \to 2^{X^*}$ is the operator defined by

$$A = \partial F + N_{\Omega,} \tag{24}$$

with ∂F denoting the subdifferential of F and $N_\Omega : X \to 2^{X^*}$ denoting the normal cone operator associated to Ω, that is,

$$N_\Omega(x) = \begin{cases} \left\{h \in X^* : \langle h, z - x\rangle \le 0, \quad \forall z \in \Omega\right\} & \text{if } x \in \Omega, \\ \varnothing & \text{otherwise,} \end{cases} \tag{25}$$

one implicity finds solutions of *(P)*. The operators ∂F and N_Ω are maximal monotone (cf. [51] by taking into account that N_Ω is the subgradient of the indicator function of the set Ω). Consequently, the operator A is maximal monotone too (cf. [50]).

We presume that the function F can not be exactly determined and that, instead, we have a sequence of convex, lower semicontinuous functions $F_k : X \to (-\infty,+\infty]$, $(k \in \mathbb{N})$, such that

$$\operatorname{Dom} F \subseteq \operatorname{Dom} F_k, \quad \forall k \in \mathbb{N}, \tag{26}$$

and which approximates F in the following sense:

Condition (A). *There exists a continuous function* $c:[0,+\infty) \to [0,+\infty)$ *and a sequence of positive real numbers* $\{\delta_k\}_{k\in\mathbb{N}}$ *such that* $\lim_{k\to\infty} \delta_k = 0$ *and*

$$|F_k(x) - F(x)| \le c(\|x\|)\delta_k, \tag{27}$$

whenever $x \in \operatorname{Dom} F$ *and* $k \in \mathbb{N}$.

In real world optimization problems it often happens that the set Ω is defined by a system of inequalities $g_i(x) \le 0, \; i \in I,$ where g_i are convex and lower semicontinuous functions on X. The functions g_i may also be hard to precisely evaluate and, then, determining the set Ω (or determining whether a vector belongs to it or not) is done by using some (still convex and lower semicontinuous) approximations $g_i^k, \; k \in \mathbb{N}$, instead. In other words, one replaces the set Ω by some nonempty closed convex approximations $\Omega_k, \; k \in \mathbb{N}$, of it. In what follows we assume that

$$\Omega_k \subseteq \operatorname{Int}(\operatorname{Dom} F), \quad \forall k \in \mathbb{N}, \tag{28}$$

and that the closed convex sets Ω_k approximate the set Ω in the following sense:

Condition (B). *The next two requirements are satisfied:*

(i) *For any $y \in \Omega$ there exists a sequence $\{y^k\}_{k\in\mathbb{N}}$ which converges strongly to y in X and such that $y^k \in \Omega_k$ for all $k \in \mathbb{N}$;*

(ii) *If $\{z^k\}_{k\in\mathbb{N}}$ is a sequence in X which is weakly convergent and such that for some subsequence $\{\Omega_{i_k}\}_{k\in\mathbb{N}}$ of $\{\Omega_k\}_{k\in\mathbb{N}}$ we have $z^k \in \Omega_{i_k}$ for all $k \in \mathbb{N}$, then there exists a sequence $\{w^k\}_{k\in\mathbb{N}}$ contained in Ω with the property that*

$$\lim_{k\to\infty} \left\| z^k - w^k \right\| = 0.$$

Observe that the requirement ($B(i)$) is equivalent to the condition that $\Omega \subseteq s - \underline{\lim}\ \Omega_k$. The requirement ($B(ii)$) implies that $w - \overline{\lim}\ \Omega_k \subseteq \Omega.$ Taken together, the requirements ($B(i)$) and ($B(ii)$) imply that $\Omega = Lim\ \Omega_k$. It can be verified that the requirement ($B(ii)$) is satisfied whenever there exists a function $b : X \to [0, +\infty)$ which is bounded on bounded sets, and a sequence of positive real numbers $\{\gamma_k\}_{k\in N}$ converging to zero such that for any $k \in N$ and each $z \in \Omega_k$, we have that $\operatorname{dist}(z, \Omega) < b(z)\gamma_k$. The last condition was repeatedly used in the regularization of variational inequalities involving maximal monotone operators (see [8]).

For each $k \in \mathbb{N}$, we associate to problem (23) the problem

(P_k) Minimize $F_k(x)$ subject to $x \in \Omega_k$,

which can be solved by finding solutions of the inclusion

$(P_k^{'})$ Find $x^k \in X$ such that $0 \in A^k x^k$,

where the operator $A^k : X \to 2^{X^*}$ is defined by

$$A^k := \partial F_k + N_{\Omega_k}, \tag{29}$$

and is also maximan monotone. The question is whether under the conditions $(A),(B)$, (11) and presuming that $\{\alpha_k\}_{k\in\mathbb{N}}$ converges to zero, the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated according to (4) for the operators A^k given by (29) and for $f^k = f = 0, k \in \mathbb{N}$, i.e., the sequence

$$x^k := \left(A^k + \alpha_k J^\mu\right)^{-1}(0.) \tag{30}$$

converges strongly to a solution of problem $\left(P^{'}\right)$ and, hence, to a solution of the original optimization problem (P). It should be noted that, since by Asplund's Theorem (see, for istance, [22] we have

$$J^\mu x = \partial\varnothing\left(\|x\|\right) \text{ with } \phi(t) := \int_0^t \mu(\tau),$$

determining the vectors x^k defined by (30) amounts to solving the optimization problem

$$\left(Q_k\right) \quad \text{Minimize } F_k\left(x\right) + \alpha_k \phi\left(\|x\|\right) \quad \text{subject to } x \in \Omega_k \tag{31}$$

By contrast to problem $\left(P_k\right)$ which may have intinitely many solutions, the problem $\left(Q_k\right)$ always has unique solution. Moreover, by choosing $\mu(t) = 2t$ and, thus, $\phi(t) = t^2$, one ensures that the objective function of $\left(Q_k\right)$ is strongly convex and, therefore, the problem $\left(Q_k\right)$ may be better posed and easier to solve than $\left(P_k\right)$.

3.2 We aim now towards giving an answer to the question asked in Subsection 3.1. To this end, when D is a nonempty closed convex subset of X and $x \in X$, we denote by $\text{Proj}_D(x)$ the metric projection of x onto the set D (this exists and it is unique by our hypothesis that the space X is strictly convex and reflexive). The next result shows stability and convergence of the regularization technique when applied to convex optimization problems. For proving it, recall that the objective function F of the problem (P) is assumed to be lower semicontinuous and convex and its domain $\operatorname{Dom} F$ has nonempty interior since $\varnothing \neq \Omega \subseteq \operatorname{Int}(\operatorname{Dom} F)$ - (see Subsection 3.1). Consequently, F is continuous on $\operatorname{Int}(\operatorname{Dom} F)$, for each $x \in \operatorname{Int}(\operatorname{Dom} F)$, we have $\partial F(x) \neq \varnothing$ (cf. [48]) and the right hand sided derivative of F at x, i.e. the function $F^{\circ}(x,\cdot): X \to \mathbb{R}$ given by

$$F^{\circ}(x,d) := \lim_{t \searrow 0} \frac{F(x+td)-F(x)}{t},$$

is a well defined continuous seminorm on X.

Theorem. *Suppose that conditions* (A) *and* (B) *are satisfied. If there exists a sequence* $\{\alpha_k\}_{k\in\mathbb{N}}$ *of positive real numbers converging to zero such that for each optimal solution* v *of* (P), *there exists a sequence* $\{v^k\}_{k\in\mathbb{N}}$ *with the properties that* $v^k \in \Omega_k$ *for all* $k \in \mathbb{N}$ *and*

$$\lim_{k\to\infty} \left\| v^k - v \right\| = 0 = \lim_{k\to\infty} \alpha_k^{-1} \left\| \operatorname{Proj}_{\partial F_k(v^k)+N_{\Omega_k}(v^k)}(0) \right\|_*, \tag{32}$$

then the sequence $\{x^k\}_{k\in\mathbb{N}}$ *given by* (30) *converges strongly to the minimal norm solution of the optimization problem* (P).

Proof. We show that Corollary 2.5 applies to the problems (P') and (P_k'), that is, to the maximal monotone operators A and A^k defined by (24) and (29), respectively, and to the functions $f^k = f = 0$, $(k \in \mathbb{N})$. First, we prove that the condition (16) is satisfied. For this purpose, take $(z,h) \in ws-\overline{\lim}\operatorname{Graph}(A^k)$. Then, there exists a sequence $\{z^k\}_{k\in\mathbb{N}}$ converging weakly to z in X and there exists a sequence $\{h^k\}_{k\in\mathbb{N}}$ converging strongly to h in X^* such that for some subsequence $\{A^{i_k}\}_{k\in\mathbb{N}}$ of $\{A^k\}_{k\in\mathbb{N}}$ we have $(z^k,h^k) \in \operatorname{Graph}(A^{i_k})$ for all $k \in \mathbb{N}$. This means that

$$z^k \in \Omega_{i_k} \text{ and } h^k \in \partial F_{i_k}(z^k) + N_{\Omega_{i_k}}(z^k), \quad \forall k \in \mathbb{N},$$

or, equivalently,

$$z^k \in \Omega_{i_k} \text{ and } h^k = \xi^k + \theta^k,$$

with $\xi^k \in \partial F_{i_k}(z^k)$ and $\theta^k \in N_{\Omega_{i_k}}(z^k)$ for all $k \in \mathbb{N}$. We have to show that

$$z \in \Omega \text{ and } h \in \partial F(z) + N_\Omega(z). \tag{33}$$

The sequence $\{z^k\}_{k\in\mathbb{N}}$ is weakly convergent to z and $z^k \in \Omega_{i_k}$ for all $k \in \mathbb{N}$. Therefore, according to ($B(ii)$), there exists a sequence $\left\{w^k\right\}_{k\in\mathbb{N}} \subseteq \Omega$ such that $\lim_{k\to\infty}\left\|z^k - w^k\right\| = 0.$ Clearly, the sequence $\{w^k\}_{k\in\mathbb{N}}$, converges weakly to z. Since the set Ω is closed and convex, and therefore weakly closed, we obtain that $z \in \Omega$. In order to complete the proof of (33), let $u \in \Omega$ be fixed. According to ($B(i)$), there exists a sequence $\{u^k\}_{k\in\mathbb{N}}$ which converges strongly to u and such that $u^k \in \Omega_k$ for any $k \in \mathbb{N}$. Since $h^k - \theta^k = \xi^k \in \partial F_{i_k}(z^k)$ we deduce

$$\begin{aligned}\left\langle h^k - \theta^k, u^{i_k} - z^k \right\rangle &\le F_{i_k}(u^{i_k}) - F_{i_k}(z^k) \\ &\le | F_{i_k}(u^{i_k}) - F(u^{i_k}) | + | F(z^k) - F_{i_k}(z^k) | \\ &\quad + F(u^{i_k}) - F(z^k) \\ &\le (c(\left\|u^{i_k}\right\|) + c(\left\|z^k\right\|))\delta_{i_k} + F(u^{i_k}) - F(z^k),\end{aligned}$$

where the last inequality results from (27). By consequence,

$$\begin{aligned}\left\langle h^k, u^{i_k} - z^k \right\rangle &\le (c(\left\|u^{i_k}\right\|) + c(\left\|z^k\right\|))\delta_{i_k} + F(u^{i_k}) - F(z^k) \\ &\quad + \left\langle \theta^k, u^{i_k} - z^k \right\rangle,\end{aligned}$$

where the last term on the right hand side of the inequality is nonpositive because $\theta^k \in N_{\Omega_{i_k}}(z^k)$ and $u^{i_k} \in \Omega_{i_k}$ (see (25)). Thus, for any $k \in \mathbb{N}$, we obtain

$$\left\langle h^k, u^{i_k} - z^k \right\rangle \le (c(\left\|u^{i_k}\right\|) + c\left\|z^k\right\|))\delta_{i_k} + F(u^{i_k}) - F(z^k). \tag{34}$$

As noted above, the function F is continuous on $\operatorname{Int}(\operatorname{Dom} F)$. Hence, the sequence $\{F(u^{i_k})\}_{k\in\mathbb{N}}$ converges to $F(u)$. Since F is also convex, it is weakly lower semicontinuous and, then, we have $F(z) \le \lim\ \inf_{k\to\infty} F(z^k)$. Taking lim sup for $k \to \infty$ on both sides of (34), and taking into account that the sequences $\{u^{i_k}\}_{k\in\mathbb{N}}$ and $\{z^k\}_{k\in\mathbb{N}}$ are bounded and that the function c is continuous (see condition (A)), we obtain that

$$\langle h, u - z \rangle \le F(u) - F(z). \tag{35}$$

Since the latter holds for arbitrary $u \in \Omega$, it implies that $h \in \partial F_\Omega(z)$, where $F_\Omega : X \to (-\infty, +\infty]$ is the lower semicontinuous convex function defined by

$$F_\Omega := F + \iota_\Omega,$$

with ι_Ω standing for the indicator function of the set Ω. As noted above, the function F is continuous on the interior of its domain and, thus, is continuous on $\Omega = \operatorname{Dom} F_\Omega$. Hence, applying [48] and observing that $\partial \iota_\Omega = N_\Omega$ (see (25)), we deduce that, for any $x \in X$,

$$\partial F_\Omega(x) = \partial F(x) + \partial \iota_\Omega(x) = \partial F(x) + N_\Omega(x).$$

Consequently,

$$h \in \partial F_\Omega(z) = \partial F(z) + N_\Omega(z)$$

and this completes the proof of (33).

Now observe that, according to (32) and (29), we have that for each solution v of (P) there exists a sequence $\{v^k\}_{k\in\mathbb{N}}$ such that $v^k \in \Omega_k$ for all $k \in \mathbb{N}$ and with the property that

$$\lim_{k\to\infty} \alpha_k^{-1} dist_*(0, A^k v^k) = \lim_{k\to\infty} \alpha_k^{-1} \left\| \operatorname{Proj}_{\partial F_k(v^k) + N_{\Omega_k}(v^k)}(0) \right\|_* = 0,$$

that is, condition (17) is also satisfied. □

3.3 Recall (see Subsection 3.1) that we assume that the problem (P) has optimal solutions. By contrast, some or all problems (P_k) may not have optimal solutions. Theorem 3.2 guarantees existence and convergence of $\{x^k\}_{k\in\mathbb{N}}$ to a solution of (P) with no consistency requirements on the problems (P_k). In our circumstances the functions F_k may not have global minimizers either. The following consequence of Theorem 3.2 may be of use for global minimization of F when some of the problems (P_k) have no optimal solutions.

Corollary. *Suppose that conditions (A) and (B) hold. If there exists a sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ of positive real numbers converging to zero such that for each optimal solution v of (P), there exists a sequence $\{v^k\}_{k\in\mathbb{N}}$ with the properties that $v^k \in \Omega_k$ for all $k \in N$ and*

$$\lim_{k\to\infty}\left\|v^k - v\right\| = 0 = \lim_{k\to\infty}\alpha_k^{-1}\left\|\mathrm{Proj}_{\partial F_k(v^k)}\left(0\right)\right\|_*, \tag{36}$$

then the sequence $\{x^k\}_{k\in\mathbb{N}}$ given by (30) *converges strongly to the minimal norm solution of the optimization problem* (P).

Proof. Note that $0 \in N_{\Omega_k}(v^k)$ for all $k \in \mathbb{N}$ and, therefore, when (36) holds, we have

$$\begin{aligned}\mathrm{dist}_*(0, A^k v^k) &= \inf\left\{\left\|g+\zeta\right\|_* : g \in \partial F_k(v^k) \text{ and } \zeta \in N_{\Omega_k}(v^k)\right\}\\ &\le \inf\left\{\left\|g\right\|_* : g \in \partial F_k(v^k)\right\}\\ &= \left\|Proj_{\partial F_k(v^k)}(0)\right\|_*.\end{aligned}$$

This implies (32) because of (36). □

3.4 If X is a Hilbert space and the functions F and F_k are differentiable on the interior of $\mathrm{Dom}\left(F\right)$, then the condition (36) can be relaxed by taking into account (see [52, Remark 3. p. 890]) that, in this case, we have

$$\left\|\nabla F_k(v^k) + \mathrm{Proj}_{N_{\Omega_k}(v^k)}(-\nabla F_k(v^k))\right\|_* = \left\|\mathrm{Proj}_{T_{\Omega_k}(v^k)}(-\nabla F_k(v^k))\right\|_*, \tag{37}$$

where $T_{\Omega_k}(v^k)$ denotes the tangent cone of Ω_k at the point v^k, that is, the polar cone of $N_{\Omega_k}(v^k)$. Precisely, we have the following result whose proof reproduces without modification the arguments in Theorem 3.2 with the only exception that for showing (17) one uses (37), (39) below, and the equalities

$$\begin{aligned}&\mathrm{dist}_*(0, A^k v^k) = \mathrm{dist}_*(0, \nabla F_k(v^k) + N_{\Omega_k}(v^k))\\ &= \mathrm{dist}_*(-\nabla F_k(v^k), N_{\Omega_k}(v^k))\\ &= \left\|\nabla F_k(v^k) + \mathrm{Proj}_{N_{\Omega_k}(v^k)}(-\nabla F_k(v^k))\right\|_*,\end{aligned} \tag{38}$$

where the first is due to the fact that $0 \in N_{\Omega}(v^k)$ and the second follows from (36).

Corollary. *Suppose that X is a Hilbert space and that conditions (A) and (B) hold. If the functions F and F_k, $k \in N$, are (Gâteaux) differentiable on* $\operatorname{Int}(\operatorname{Dom} F)$ *and if there exists a sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ of positive real numbers converging to zero such that for each optimal solution v of (P), there exists a sequence $\{v^k\}_{k\in\mathbb{N}}$ with the properties that $v^k \in \Omega_k$ for all $k \in N$ and*

$$\lim_{k\to\infty}\left\|v^k - v\right\| = 0 = \lim_{k\to\infty}\alpha_k^{-1}\left\|\operatorname{Proj}_{T_{\Omega_k}(v^k)}(-\nabla F_k(v^k))\right\|_*, \tag{39}$$

then the sequence $\{x^k\}_{k\in\mathbb{N}}$ given by (30) *converges strongly to the minimal norm solution of the optimization problem* (P).

3.5 If $\Omega_k = \Omega$ for all $k \in \mathbb{N}$, then condition (B) is, obviously, satisfied. In this case, if there exists a sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ of positive real numbers converging to zero such that for each solution v of (P) we have

$$\lim_{k\to\infty}\alpha_k^{-1}\left\|\nabla F_k(v) - \nabla F(v)\right\|_* = 0, \tag{40}$$

then (32) holds too. Indeed, if v is a solution of (P), then

$$\left\langle \nabla F(v), u - v\right\rangle = \lim_{t\searrow 0}\frac{F(v + t(u-v)) - F(v)}{t} \ge 0,$$

for any $u \in \Omega$ and this shows that $-\nabla F(v) \in N_\Omega(v)$. Therefore, taking $v^k := v$ for all $k \in \mathbb{N}$ we have

$$\begin{aligned}\left\|\operatorname{Proj}_{\partial F_k(v^k)+N_{\Omega_k}(v^k)}(0)\right\|_* &\le \left\|\nabla F_k(v) + Proj_{N_\Omega(v)}(-\nabla F_k(v))\right\|_* \\ &= \left\|\operatorname{Proj}_{N_\Omega(v)}(-\nabla F_k(v)) - (-\nabla F_k(v))\right\|_* \\ &\le \left\|\nabla F_k(v) - \nabla F(v)\right\|_*,\end{aligned}$$

which together with (40) implies (32). Hence, we have the following result:

Corollary. *Suppose that $\Omega_k = \Omega$ for all $k \in \mathbb{N}$ and condition (A) holds. Assume that the functions F and F_k, $k \in \mathbb{N}$, are* (*Gâteaux*) *differentiable on* $\operatorname{Int}(\operatorname{Dom} F)$. *If there exists a sequence of positive real numbers* *converging to zero such that for each solution* v *of* (P) *condition* (40) *is satisfied, then the sequence* $\{x^k\}_{k\in\mathbb{N}}$ *given by* (30) *converges strongly to the minimal norm solution of the optimization problem* (P).

3.6 Theorem 3.2 shows that perturbed convex optimization problems can be regularized by the method (4). This naturally leads to the question of whether this regularization technique still works when the set Ω is defined by continuous affine constraints and one has to replace the affine constraints of (P) by approximations which are still continuous and affine. We are going to show that this is indeed the case when Ω satisfies a Robinson type regularity condition. In order to do that, let $L(X,X^*)$ be the Banach space of all linear continuous operators $L: X \to X^*$ provided with the norm

$$\|L\|_\infty := \sup\left\{\|Lx\|_* : x \in B_X(0,1)\right\}, \tag{41}$$

where $B_X(u,r)$ stands for the closed ball of center u and radius r in X. Suppose that $L,\ L^k \in L(X,X^*)$, $l,\ l^k \in X^*$ and let $K \subset X^*$ be a nonempty closed convex cone. Suppose that the sets Ω and Ω_k are defined by

$$\Omega := \{x \in X : L(x) + l \in K\} \tag{42}$$

and

$$\Omega_k := \{x \in X : L^k(x) + l^k \in K\}. \tag{43}$$

The set Ω is called *regular* if the point-to-set mapping $x \to L(x) + l - K$ is regular in the sense of Robinson [49, p. 132], that is,

$$0 \in \operatorname{Int}\{L(x) + l - y : x \in X,\ y \in K\}. \tag{44}$$

Taking in the next result $K = \{0\}$ one obtains an answer to the question posed above. The fact is that the proposition we prove below is more general and can be also used in order to guarantee validity of condition (B) for some classes of problems of interest in semidefinite programming. Combined with Theorem 3.2 it implies that if the data involved in the constraints of the perturbed problem (P_k) are strong approximations of the data of (P) and if conditions (A) and (32) hold, then the regularization technique (4) can be applied in order to produce strong approximations of the minimal norm solution of (P).

Proposition. *Suppose that* $L,\ L^k \in L(X,X^*)$ *and* $l,\ l^k \in X^*$ *for any* $k \in \mathbb{N}$. *Let* $K \subset X^*$ *be a nonempty closed convex cone and consider the problems* (P) *and* (P_k) *with the feasibility sets* Ω *and* Ω_k *defined at* (42) *and* (43), *respectively. If the set* Ω *is regular, if the sequence* $\{L^k\}_{k\in\mathbb{N}}$ *converges strongly to* L *in* $L(X,X^*)$ *and if the sequence* $\{l^k\}_{k\in\mathbb{N}}$ *converges strongly to* l *in* X^*, *then condition* (B) *is satisfied.*

Proof. We first prove ($B(ii)$). For this purpose we apply Corollary 4 in [10, p. 133] to the set $M = K - l$ and to the point-to-set mapping $G : X \to 2^{X^*}$ defined by $G(x) = L(x) - M$. This is possible because of the regularity of Ω (see (44)) which guarantees that $0 \in \operatorname{Int} G(X)$. Hence, by observing that $\Omega = G^{-1}(0)$, we deduce that for any $x \in \Omega$ there exists a positive real number $\delta(x)$ such that for any $z \in X$ we have

$$\begin{aligned}\operatorname{dist}(z,\Omega) &= \operatorname{dist}(z, G^{-1}(0)) \le (1+\|z-x\|)\frac{1}{\delta(x)}\operatorname{dist}_*(L(z), M)\\ &= (1+\|z-x\|)\frac{1}{\delta(x)}\operatorname{dist}_*(L(z)+l, K).\end{aligned} \tag{45}$$

Now, let $x \in \Omega$ be fixed and let $\delta := \delta(x)$. If $z \in \Omega_k$, then

$$\begin{aligned}\operatorname{dist}_*(L(z)+l, K) &\le \left\|\left(L(z)+l\right)-\left(L^k(z)+l^k\right)\right\|_*\\ &\le \left\|L(z)-L^k(z)\right\|_* + \left\|l-l^k\right\|_*\\ &\le \left\|L-L^k\right\|_\infty \|z\| + \left\|l-l^k\right\|_*\end{aligned}$$

because of (41). Taking into account (45), it follows that for any $z \in \Omega_k$, we have

$$\begin{aligned}\operatorname{dist}(z,\Omega) &\le \frac{1}{\delta}(1+\|z-x\|)\left[\left\|L-L^k\right\|_\infty\|z\| + \left\|l-l^k\right\|_*\right]\\ &\le \frac{1}{\delta}(1+\|z\|+\|x\|)(\|z\|+1)\max\left\{\left\|L-L^k\right\|_\infty, \left\|l-l^k\right\|_*\right\}.\end{aligned} \tag{46}$$

Consider the bounded on bounded sets function $b : X \to [0,+\infty)$ defined by

$$b(u) := \frac{1}{\delta}(\|u\|+1)(1+\|u\|+\|x\|).$$

By (46) we obtain that, for any $z \in \Omega_k$,

$$\left\|z - \mathrm{Proj}_{\Omega}\left(z\right)\right\| = \mathrm{dist}\left(\mathrm{z},\Omega\right) \le b\left(z\gamma_k\right),$$

where $\gamma_k := \max\left\{\left\|L - L^k\right\|_{\infty}, \left\|l - l^k\right\|_{*}\right\}$ converges to zero as $k \to \infty$. Let $\left\{z^k\right\}_{k\in\mathbb{N}}$ be a weakly convergent sequence in X such that, for some subsequence $\left\{\Omega_{i_k}\right\}_{k\in\mathbb{N}}$ of $\left\{\Omega_k\right\}_{k\in\mathbb{N}}$, we have $z^k \in \Omega_{i_k}$ for all $k \in \mathbb{N}$. According to (47), the vectors $w^k := \mathrm{Proj}_{\Omega}(z^k)$ have the property that

$$\left\|z^k - w^k\right\| \le b(z^k)\gamma_k,$$

where, since $\left\{z^k\right\}_{k\in\mathbb{N}}$ is bounded, the sequence $\left\{b(z^k)\right\}_{k\in\mathbb{N}}$ is bounded too. Hence, $\lim_{k\to\infty}\left\|z^k - w^k\right\| = 0$ and ($B\,(ii)$) is satisfied.

Now we prove that ($B\,(i)$) is also satisfied. To this end, we consider the function $g : X \times \mathbb{N} \to X^*$ defined by

$$g(x,k) = \begin{cases} L\left(x\right) + l & \text{if } k = 0 \\ L^{k-1}\left(x\right) + l^{k-1} & \text{if } k \ge 1. \end{cases}$$

and the point-to-set mapping $\Gamma : X \times \mathbb{N} \to 2^{X^*}$ defined by

$$\Gamma\left(x,k\right) = g\left(x,k\right) - K$$

for all $x \in X$ and $k \ge 0$. Since Ω is regular (see (44)), we have that

$$0 \in \mathrm{Int}\left[\mathrm{Im}\Gamma(\cdot,0)\right].$$

Clearly, $\Omega = \Gamma^{-1}(\cdot,0)(0)$. Let $u^0 \in \Omega$. By Theorem 1 in [49], there exists $\eta > 0$ such that

$$B_{X^*}(0,\eta) \subseteq g(B_X(u^0,1),0) - K, \tag{48}$$

Since $\|L^k - L\|_\infty$ and $\|l^k - l\|_*$ converge to 0, there exists $k_0 \in \mathbb{N}$ such that for any integer $k \ge k_0$ we have

$$\|g(x,0) - g(x,k)\|_* = \|L(x) + l - L^{k-1}(x) - l^{k-1}\|_* \le \frac{\eta}{2}, \quad \forall x \in B_X(u^0,1).$$

This implies that whenever $k \ge k_0$ we have

$$g(B_X(u^0,1),0) \subseteq g\left(B_X(u^0,1),k\right) - K + B_X\left(u^0, \frac{\eta}{2}\right). \tag{49}$$

This and (48) show that the function g satisfies the assumptions in [49, Corollary 2] (with $-K$ instead of K). Consequently, application of this result yields that, for each $x \in \Omega$ and for any integer $k \ge k_0$, the set

$$\Omega_{k+1} = \left\{x \in X : g(x,k) \in K\right\}$$

contains the open ball $B_X\left(0, \frac{\eta}{2}\right)$ and that for any $x \in X$ we have

$$\operatorname{dist}(x,\Omega_k) \le \frac{2}{\eta}\left(1 + \|x - u^0\|\right)\operatorname{dist}_*(0, g(x,k) - K). \tag{50}$$

Note that, if $x \in \Omega,$ then $g(x,0) \in K$ and, therefore, we have

$$\begin{aligned}
\operatorname{dist}_*(0, g(x,k) - K) &= \operatorname{dist}_*(g(x,k), K) \\
&\le \|g(x,k) - g(x,0)\|_* \\
&= \|L^{k-1}(x) + l^{k-1} - L(x) - l\|_* \\
&\le \|L^{k-1} - L\|_\infty \|x\| + \|l^{k-1} - l\|_* \\
&\le \left(\|x\| + 1\right)\max\left\{\|L^{k-1} - L\|_\infty, \|l^{k-1} - l\|_*\right\}.
\end{aligned}$$

This and (50) implies

$$\operatorname{dist}(x,\Omega_k) \le \frac{2}{\eta}\left(1 + \|x\| + \|u^0\|\right)\left(\|x\| + 1\right)\max\left\{\|L^{k-1} - L\|_\infty, \|l^{k-1} - l\|_*\right\}, \tag{51}$$

for any $x \in \Omega$. Define the function $a:[0,+\infty) \to [0,+\infty)$ by

$$a(t) = \frac{2}{\eta}\left(1 + \left\|u^0\right\| + t\right)\left(t+1\right)$$

and the sequence of nonnegative real numbers

$$\beta_k = \max\left\{\left\|L^{k-1} - L\right\|_\infty, \left\|l^{k-1} - l\right\|_*\right\} \quad \forall k \in \mathbb{N}.$$

According to (51) we have

$$\left\|x - \mathrm{Proj}_{\Omega_k}(x)\right\| = \mathrm{dist}(x, \Omega_k) \le a(\|x\|)\beta_k,$$

for all integers $k \ge k_0$ and for all $x \in \Omega$. Since, by hypothesis, $\lim_{k\to\infty} \beta_k = 0,$ condition ($B(i)$) holds. □

3.7 The implementation of the regularization procedure discussed in this work requires computing vectors x^k defined by (30) with operators A^k given by (29). This implicitly means solving problems like (31). In some circumstances, in the regularization process, one can reduce problems placed in infinite dimensional settings to finite dimensional problems for which many efficient techniques of computing solutions are available. This is typically the case of the problem considered in the following example.

Let $X = \ell^p$ with $p \in (1,\infty)$, $q = p(p-1)^{-1}$ and, then, $X^* = \ell^q$. Let $a \in \ell^q \setminus \{0\}$ and $b^j \in \ell^q_+ \setminus \{0\}$, for all $j \in J$, where J is a nonempty set of indices and ℓ^q_+ stands for the subset of ℓ^q consisting of vectors with nonnegative coordinates. For each $j \in J$, let β_j be a nonnegative real number. Consider (P) to be the following optimization problem in ℓ^p :

$$\text{Minimize } F(x) = \langle a, x \rangle \tag{52}$$

over the set

$$\Omega := \{x \in \ell^p_+ : \langle b^j, x \rangle \le \beta_j,\ j \in J\}. \tag{53}$$

We assume that $a = (a_1, ..., a_i, ...)$ has infinitely many coordinates $a_i \ne 0$ and that the problem (P) has optimal solutions. Whenever u is an element in ℓ^p or in ℓ^q, we denote by $u[k]$ the vector in the same space as u obtained by replacing by zero all coordinates u_i of u with $i > k$. With this notations, for each $k \in \mathbb{N}$, let $\alpha_k := \|a - a[k]\|_*^{1/2}$, and observe that $\{\alpha_k\}_{k\in\mathbb{N}}$ is a sequence of positive real numbers which converge to zero as $k \to \infty$. We associate to problem (52)+(53) the perturbed problems (P_k) given by

$$\text{Minimize } F_k(x) = \langle a[k], x\rangle \text{ over } \Omega. \tag{54}$$

Note that, for each $k \in \mathbb{N}$, the problem (54) has optimal solutions because its objective function F_k is bounded from below on Ω by $F^* = \inf\{F(x) : x \in \Omega\}$. Problem ($P$) is ill posed and, therefore, even if one can find an optimal solution y^k for each of the essentially finite dimensional linear programming problems (P_k), the sequences $\{y^k\}_{k\in\mathbb{N}}$ may not converge in ℓ^p or, at best, its weak accumulation points (if any) are optimal solutions of (P).

We apply the regularization method (4) to the problems (P) and (P_k) with the function $\mu(t) = t^{p-1}$. It is easy to see that, in this case, determining the vector x^k defined by (30) reduces to finding the unique optimal solution of the problem

$$\text{Minimize } \langle a[k], x\rangle + \alpha_k \|x\|^p \text{ over } \Omega. \tag{55}$$

Theorem 3.2 applies to problems (P) and (P_k) and guarantees that the sequence $\{x^k\}_{k\in\mathbb{N}}$ converges strongly to the minimal norm solution of (P). Indeed, observe that condition (A) is satisfied because, for any $x \in \ell^p$, we have

$$|F(x) - F_k(x)| \le \|a - a[k]\|_* \|x\| = \alpha_k^2 \|x\|,$$

and condition (B) trivially holds. It remains to prove that (32) holds too, that is, for any optimal solution v of (P) there exists a sequence $\{v^k\}_{k\in\mathbb{N}}$ of vectors in Ω such that

$$\lim_{k\to\infty} \|v^k - v\| = 0 = \lim_{k\to\infty} \alpha_k^{-1} \left\|\mathrm{Proj}_{a[k]+N_\Omega(v^k)}(0)\right\|_*. \tag{56}$$

Take the constant sequence $v^k = v$, ($k \in N$). Then the first equality in (56) holds and, for any $x \in \Omega$, we have $\langle -a, x - v \rangle = \langle a, v \rangle - \langle a, x \rangle \le 0$, showing that $-a \in N_\Omega(v)$. Thus, for each $k \in N$, we obtain that

$$\left\| \mathrm{Proj}_{a[k]+N_\Omega(v^k)}(0) \right\|_* \le \| a[k] - a \|_* = \alpha_k^2$$

and this implies the second equality in (56).

Solving problem (55) can be done by finding the unique optimal solution u^k of the following optimization problem in $\mathbb{R}^k$

$$\text{Minimize } \sum_{i=1}^{k} a_i x_i + a_k \sum_{i=1}^{k} |x_i|^p \quad \text{s.t. } \sum_{i=1}^{k} b_i^j x_i \le \beta_j, \ (j \in J), \ \mathrm{x} \ge 0 \quad (57)$$

and taking $x^k = \left(u_1^k, ..., u_k^0, ...\right)$. Indeed, for any $x \in \Omega$ we have

$$\begin{aligned}\langle a[k], x \rangle + \alpha_k \|x\|^p &= \langle a[k], x[k] \rangle + \alpha_k \|x\|^p \\ &\ge \langle a[k], x[k] \rangle + \alpha_k \|x[k]\|^p \ge \langle a[k], u^k \rangle + \alpha \left\| u^k \right\|^p\end{aligned}$$

where the last inequality holds because $x[k] \in \Omega$ (due to the nonnegativity of the vectors b^j).

4. REGULARIZATION OF A PROXIMAL POINT METHOD

4.1 A question of interest in convex optimization concerns the strong convergence of the *generalized proximal point method* (GPPM for short) which emerged from the works of Martinet [43], [44], Rockafellar [52] and Censor and Zenios [21]. When applied to the consistent problem (P) described in Subsection 3.1 the GPPM produces iterates according to the rule

$$y^0 \in \Omega \text{ and } y^{k+1} := \arg\ \min\left\{ F(x) + \omega_k D_G(x, y^k) : x \in \Omega \right\} \quad (58)$$

with $D_G : \mathrm{Dom}(G) \times \mathrm{Int}(\mathrm{Dom}\ G) \to [0, +\infty)$ defined by

$$D_G(x, y) := G(x) - G(y) - \langle \nabla G(y), x - y \rangle, \quad \forall y \in \mathrm{Int}(\mathrm{Dom} \quad G), \quad (59)$$

where $\{\omega_k\}_{k\in\mathbb{N}}$ is a bounded sequence of positive real numbers and $G : X \to (-\infty,+\infty]$ is a *Bregman function* on Ω, that is, a function satisfying the following conditions:

(i) $\Omega \subseteq \text{Int}(\text{Dom}\, G)$;
(ii) G is Fréchet differentiable on $\text{Int}(\text{Dom}\, G)$;
(iii) G is uniformly convex on bounded subsets of Ω;
(iv) For each $x \in \Omega$, the sets

$$R_\alpha^G(x) = \{ y \in \Omega : D_G(x,y) \le \alpha \}$$

are bounded for all real numbers $\alpha > 0$.

The sequences $\{y^k\}_{k\in\mathbb{N}}$ generated by the GPPM are well defined, bounded and their weak accumulation points are solutions of (P) - cf. [17]. Weak convergence of these sequences can be ensured only when the Bregman function G has very special properties as, for instance, when ∇G is sequentially weakly-to-weak* continuous on Ω - (see [18, Chapter 3]). Strong convergence may not happen at all even when weak convergence does occur. This is in fact the case of the classical proximal point method for optimization which is the particular version of GPPM in Hilbert spaces in which $G = \|\cdot\|^2$ (cf. [28]). The conditions under which the GPPM is known to converge strongly (see [52], [35], [7], [17], [20] and the references therein) are quite restrictive and mostly concern the data of (P) [in contrast to those ensuring weak convergence which mostly concern the Bregman function G whose selection can be done from a relatively large pool of known candidates - cf. [18]]. We are going to prove, by applying Theorem 3.2 and its corollaries, that a regularized version of the GPPM produces sequences which behave better than the sequences $\{y^k\}_{k\in\mathbb{N}}$ associated to (P) by (58). By contrast to the regularization method of GPPM proposed in [57] which, in Hilbert spaces, produces strongly convergent sequences whose limits are the projection of their initial points onto the set of optima of (P), the sequences resulting from the regularized version of GPPM proposed here converge strongly to the minimal norm solution of (P).

4.2 From now on we assume that X is an uniformly convex and uniformly smooth Banach space. We are going to show that in this not necessarily Hilbertian setting, by regularizing GPPM following the technique defined in (4), one obtains a procedure which generates sequences converging strongly to optima of (P). To this end, we denote $G(x) = \|x\|^p$

and $\psi(t) = pt^{p-1}$ for some $p \in (1,+\infty)$. Recall (cf. [19]) that G is a Bregman function and that G' is exactly the duality mapping J^{ψ}. We denote by S the presumed nonempty set of optimal solutions of the problem (P) described in Subsection 3.1.

Theorem. *Let $\{\Omega_k\}_{k\in\mathbb{N}}$ be a sequence of closed convex sets contained in Ω such that ($B(i)$) is satisfied and*

(a) $S \cap \left(\bigcap_{k=0}^{\infty} \Omega_k \right) \neq \varnothing$;

(b) $\operatorname{Lim} (S \cap \Omega_k) = S$.

If $\{\alpha_k\}_{k\in\mathbb{N}}$ and $\{\omega_k\}_{k\in\mathbb{N}}$ are sequences of positive real numbers such that the first converges to zero and the second has the property that $\lim_{k\to\infty} (\omega_{k+1}/\alpha_k) = 0$, then, for any initial point $y^0 \in \Omega_0$, the sequences $\{x^k\}_{k\in\mathbb{N}}$ and $\{y^k\}_{k\in\mathbb{N}}$ generated according to the rule

$$y^k = \arg\ \min\left\{F(x) + \omega_k D_G(x, y^{k-1}) : x \in \Omega_k\right\},$$
$$x^k = \left[\partial F + N_{\Omega_k} + \omega_k (J^{\psi} - J^{\psi} y^k) + \alpha_k J^{\mu}\right]^{-1}(0), \tag{60}$$

are well defined and have the following properties:

(i) The sequence $\{y^k\}_{k\in\mathbb{N}}$ is bounded, the sequence $\{F(y^k)\}_{k\in\mathbb{N}}$ converges and

$$\lim_{k\to\infty} F(y^k) = \inf\{F(y) : y \in \Omega\}; \tag{61}$$

(ii) The sequence $\{x^k\}_{k\in\mathbb{N}}$ converges strongly to a solution of (P).

Proof. For each $k \in \mathbb{N}$, define the functions $E_k, H_k : X \to (-\infty,+\infty]$ by

$$E_k(x) = F(x) + \iota_{\Omega_k}(x),$$

and

$$H_k(x) = E_k(x) + \omega_k D_G(x, y^{k-1}),$$

where ι_{Ω_k} stands for the indicator function of the set Ω_k. The functions E_k and H_k are lower semicontinuous, convex and bounded from below.

According to [18] applied to them we deduce that, for any integer $k \geq 1,$ the vector

$$y^k = \arg\ \min\{H_k(y) : y \in \Omega_k\}.$$

exists and is well defined. Note that this is exactly the vector y^k given by (60) and, thus, the sequence $\{y^k\}_{k\in\mathbb{N}}$ is well defined. Let $\overline{z} \in S \cap \left(\bigcap_{k=0}^{\infty} \Omega_k\right).$ Observe that, for each positive integer $k,$ the vector $\overline{z}$ is also a minimizer of the function E_k over $\Omega_k.$ An argument similar to that in the proof of [18] applied to the functions E_k and H_k shows that

$$D_G(\overline{z}, y^{k-1}) - D_G(\overline{z}, y^k) - D_G(y^k, y^{k-1}) \geq \frac{1}{\omega_k}\left[E_k(y^k) - E_k(\overline{z})\right],$$

for all integers $k \geq 1.$ Thus, if ω is a positive upper bound of the bounded sequence $\{\omega_k\}_{k\in N},$ we get

$$D_G(\overline{z}, y^{k-1}) - D_G(\overline{z}, y^k) - D_G(y^k, y^{k-1}) \geq \frac{1}{\omega}\left[F(y^k) - F(\overline{z})\right] \geq 0, \qquad (62)$$

for all integers $k \geq 1,$ because $y^k \in \Omega_k \subseteq \Omega$ and $\overline{z}$ is a solution of $(P).$ From (62) it can be easily seen that the sequence $\left\{D_G(\overline{z}, y^k)\right\}_{k\in\mathbb{N}}$ is nonincreasing, hence, convergent, and, consequently, that the sequence $\left\{D_G(y^k, y^{k-1})\right\}_{k\in N}$ converges to zero. These and (62) imply that $\lim_{k\to\infty}\left[F(y^k) - F(\overline{z})\right] = 0.$ Hence, (61) is proved. Boundedness of $\{y^k\}_{k\in\mathbb{N}}$ follows from the fact that $\left\{D_G(\overline{z}, y^k)\right\}_{k\in\mathbb{N}}$ is bounded by $\alpha = D_G(\overline{z}, y^0)$ and, then, all y^k are contained in the set $R_\alpha^G(\overline{z})$ which is bounded because, as noted above, G is a Bregman function. Hence, the proof of (i) is complete.

In order to prove (ii), we apply Theorem 3.2 to the problem (P) given at (23) and to the problems (P_k) with the functions $F_k : X \to (-\infty, +\infty]$ given by

$$F_k(x) := F(x) + \omega_{k+1} D_G(x, y^k),$$

where, for each nonnegative integer $k,$ the vector y^k is defined by (60), that is,

$$y^k = \arg \ \min\left\{F_{k-1}(x) : x \in \Omega_k\right\}.$$

Note that F_k is convex, lower semicontinuous and has $\operatorname{Dom} F_k = \operatorname{Dom} F.$ Also, by Asplund's Theorem which shows that ∇G is exactly the duality mapping J^ψ, we obtain

$$\partial F_k(x) := \partial F(x) + \omega_{k+1}(J^\psi x - J^\psi y^k).$$

We associate to each function F_k the maximal monotone operator A^k defined by (29). Observe that the vectors x^k defined by (60) are exactly those given by (30) for this specific operator A^k and, thus, it is well defined. We show next that the operators A^k have the properties required by the hypothesis of Theorem 3.2.

Let $x \in \operatorname{Dom} F.$ Then, for any $k \in \mathbb{N}$, we have

$$\begin{aligned}
F_k(x) - F(x) &= \omega_{k+1} D_G(x, y^k) \\
&= \omega_{k+1}\left(\|x\|^p - \|y^k\|^p - \left\langle J^\psi y^k, x - y^k\right\rangle\right) \\
&= \omega_{k+1}\left(\|x\|^p + (p-1)\|y^k\|^p - \left\langle J^\psi y^k, x\right\rangle\right) \\
&\le \omega_{k+1}\left(\|x\|^p + (p-1)\|y^k\|^p + \|J^\psi y^k\|_* \|x\|\right) \\
&= \omega_{k+1}\left(\|x\|^p + (p-1)\|y^k\|^p + p\|y^k\|^{p-1}\|x\|\right).
\end{aligned}$$

The sequence $\left\{y^k\right\}_{k\in\mathbb{N}}$ is bounded as shown above. Let M be a positive upper bound of the sequence $\left\{\|y^k\|\right\}_{k\in N}.$ Define the continuous function $c:[0,+\infty)\to[0,+\infty)$ by

$$c(t) = t^p + pM^{p-1}t + (p-1)M^{p-1}.$$

Hence, for each $k \in, \mathbb{N}$ we have

$$\left|F_k(x) - F(x)\right| \le \omega_{k+1}\left[\|x\|^p + (p-1)M^p + pM^{p-1}\|x\|\right] = \omega_{k+1}c(\|x\|),$$

showing that condition (A) is satisfied with $\delta_k = \omega_{k+1}$, $k \in \mathbb{N}$. Condition ($B(ii)$) holds in our case because, by hypothesis, all Ω_k are subsets of Ω. So, condition (B) is satisfied.

It remains to show that there exists a such that for any solution $v \in \Omega$ of (P) there exists a sequence $\{v^k\}_{k\in\mathbb{N}}$, which has $v^k \in \Omega_k$ for all $k \in \mathbb{N}$, and satisfies (32). To this end, note that each $S_k := \Omega_k \cap S$ is a nonempty, closed and convex subset of X. By (b), we have

$$S = \operatorname{Lim} S_k. \tag{63}$$

Also, there exists a sequence $\{w^k\}_{k\in N}$ such that, for each $k \in N$, $w^k \in \Omega_k$ and $\lim_{k\to\infty}\left\|w^k - v\right\| = 0.$ Let $v^k = \operatorname{Proj}_{S_k}(w^k).$ Observe that, according to [24, p. 40], the space X is an E-space because it is uniformly convex. Therefore, Theorem 10 in [24, p. 49] combined with (63) imply that

$$\lim_{k\to\infty} v^k = \lim_{k\to\infty} \operatorname{Proj}_{S_k}(w^k) = \operatorname{Proj}_S(v) = v. \tag{64}$$

Now, observe that each $v^k \in S_k$ and, hence, is a minimizer of E_k over X, that is, $0 \in \partial E_k(v^k)$. Similarly to F, the functions F_k are continuous on the interiors of their respective domains and, therefore, they are continuous on Ω_k (see (26) and (28)). By consequence, we have

$$\partial E_k(v^k) = \partial F(v^k) + \partial \iota_{\Omega_k}(v^k) = \partial F(v^k) + N_{\Omega_k}(v^k). \tag{65}$$

From (65) we obtain

$$0 \in \partial F(v^k) + N_{\Omega_k}(v^k). \tag{66}$$

For each $k \in \mathbb{N}$, we have

$$\begin{aligned}\left\|\operatorname{Proj}_{\partial F_k(v^k)+N_{\Omega_k}(v^k)}(0)\right\|_* &= \operatorname{dist}_*(0, \partial F(v^k) + N_{\Omega_k}(v^k) + \omega_{k+1}\left(J^\psi v^k - J^\psi y^k\right)) \\ &= \operatorname{dist}_*(\omega_{k+1}\left(J^\psi y^k - J^\psi v^k\right), \partial F(v^k) + N_{\Omega_k}(v^k)) \\ &\le \omega_{k+1}\left\|J^\psi y^k - J^\psi v^k\right\|_*,\end{aligned}$$

where the last inequality follows from (66). By consequence, taking into account (3), we obtain

$$\left\|\mathrm{Proj}_{\partial F_k(v^k)+N_{\Omega_k}(v^k)}(0)\right\|_* \le \omega_{k+1}\left\|J^{\psi}v^k - J^{\psi}y^k\right\|_*$$
$$\le \omega_{k+1}\left(\left\|J^{\psi}v^k\right\|_* + \left\|J^{\psi}y^k\right\|_*\right)$$
$$= p\omega_{k+1}\left(\left\|v^k\right\|^{p-1} + \left\|y^k\right\|^{p-1}\right),$$

and, thus, we have

$$\left\|\mathrm{Proj}_{\partial F_k(v^k)+N_{\Omega_k}(v^k)}(0)\right\|_* \le p\omega_{k+1}\left(\left\|v^k\right\|^{p-1} + M^{p-1}\right), \quad \forall k \in \mathbb{N}. \tag{67}$$

Let N be an upper bound of the sequence $\left\{\left\|v^k\right\|\right\}_{k\in\mathbb{N}}$ and denote

$$q := p\left(N^{p-1} + M^{p-1}\right).$$

Then, by (67), we have

$$\alpha_k^{-1}\left\|\mathrm{Proj}_{\partial F_k(v^k)+N_{\Omega_k}(v^k)}(0)\right\|_* \le q\frac{\omega_{k+1}}{\alpha_k}, \forall k \in \mathbb{N}. \tag{68}$$

This and (64) imply (32) by hypothesis. According to (68), condition (32) holds too. These show that Theorem 3.2 is applicable to F and to the functions F_k. In turn, Theorem 3.2 implies that the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ converges strongly to the minimal norm minimizer of F over Ω. □

4.3 Verifying the conditions (a) and (b) of Theorem 4.2 may be difficult. In some circumstances the following consequence of Theorem 4.2 may be of use. For instance, if X has a countable system of generators $\left\{e^k\right\}_{k\in\mathbb{N}}$ and the problem (P) is unconstrained (i.e., $\Omega = X$), then using the next result with the sets

$$\Omega_k = \overline{\mathrm{aff}}\{e^i : 0 \le i \le k\},$$

which necessarily satisfy condition (c) below, one reduces the resolution of (P) to solving a sequence of unconstrained problems in spaces of finite dimension whose solutions x^k will necessarily converge strongly to an optimum of (P).

Corollary. *Let* $\{\Omega_k\}_{k\in\mathbb{N}}$ *be a sequence of closed convex subsets of* Ω *such that* ($B(i)$) *is satisfied and one of the following conditions hold:*

(c) $S \subseteq \bigcup_{k=0}^{\infty}\Omega_k$ *and* $\Omega_k \subseteq \Omega_{k+1}$ *for all* $k\in\mathbb{N}$;

(d) $\operatorname{Int} S \neq \varnothing$ *and* $S\cap\Omega_k \neq \varnothing$ *for all* $k\in\mathbb{N}$.

If $\{\alpha_k\}_{k\in\mathbb{N}}$ *and* $\{\omega_k\}_{k\in\mathbb{N}}$ *are sequences of positive real numbers such that the first converges to zero and the second has the property that* $\lim_{k\to\infty}(\omega_{k+1}/\alpha_k)=0,$ *then, for any initial point* $y^0\in\Omega,$ *the sequences* $\{x^k\}_{k\in\mathbb{N}}$ *and* $\{y^k\}_{k\in\mathbb{N}}$ *generated according to the rule* (60) *are well defined and have the properties* (i) *and* (ii) *from Theorem 4.2.*

Proof. Suppose that condition (c) holds. Take any $\bar{z}\in S$. There exists a number $k_0\in\mathbb{N}$ such that $\bar{z}\in\Omega_{k_0}$. Denote $\Omega_0' = \bigcup_{k=0}^{k_0}\Omega_k$ and $\Omega_k' = \Omega_{k_0+k}$ for $k\geq 1$. Applying Lemma 1.2 from [45] we deduce that $\operatorname{Lim}(\Omega_k'\cap S)=S$. Applying Theorem 4.2 to the sets Ω_k' we obtain the result. Now, assume that (d) holds. Then, Lemma 1.4 from [45] guarantees that condition (b) of Theorem 4.2 is satisfied and we can apply that proposition in this case too.

□

4.4 The previous results in this section deal with the case that the sets Ω_k are contained in Ω. If F has bounded level sets

$$L_F(\alpha) := \{x\in X : F(x)\leq\alpha\},$$

then the regularized proximal point method (60) is also stable under outer approximations of Ω.

Proposition. *Let* $\{\Omega_k\}_{k\in\mathbb{N}}$ *be a sequence of closed convex sets contained in* $\operatorname{Int}(\operatorname{Dom} F)$ *such that condition* (B) *is satisfied,* $\Omega=\bigcap_{k=0}^{\infty}\Omega_k$ *and* $\Omega_{k+1}\subseteq\Omega_k$ *for all* $k\in\mathbb{N}$. *Let* $\{\alpha_k\}_{k\in\mathbb{N}}$ *be a sequence of positive real numbers converging to zero. Suppose that for each number* $\alpha\geq 0$ *the level set* $L_F(\alpha)$ *of the objective function* F *is bounded. Then, for any initial point* $y^0\in\Omega,$ *the sequences* $\{x^k\}_{k\in\mathbb{N}}$ *and* $\{y^k\}_{k\in\mathbb{N}}$ *generated according to the rule*

(60) *where the positive numbers* ω_k *are chosen such that* $\lim_{k\to\infty}(\omega_{k+1}/\alpha_k)=0$ *and for some positive number* K,

$$\omega_k D_G(y^0, y^{k-1}) \le K, \tag{69}$$

are well defined and have the properties (i) *and* (ii) *from Theorem 4.2.*

Proof. We use the notations F_k, E_k and H_k introduced in the proof of Theorem 4.2. Observe that, since X is reflexive and the level sets of F are bounded, for any $k \in \mathbb{N}$, there exists

$$z^k \in \arg\ \min\{F(x) : x \in \Omega_k\}.$$

According to [18] applied to E_k and H_k we deduce that, for any integer $k \ge 1$, the vector

$$y^k = \arg\ \min\{H_k(y) : y \in \Omega_k\}.$$

exists. Let $z \in S$ and observe that, for any $k \in \mathbb{N}$,

$$F(z^k) \le F(z^{k+1}) \le F(z), \tag{70}$$

because $\Omega \subseteq \Omega_k$ and $\Omega_{k+1} \subseteq \Omega_k$. This shows that all z^k belong to $L_F(F(z))$ and, therefore, the sequence $\{z^k\}_{k\in\mathbb{N}}$ is bounded. Since X is reflexive, the bounded sequence $\{z^k\}_{k\in\mathbb{N}}$ contains a weakly convergent subsequence $\{z^{i_k}\}_{k\in\mathbb{N}}$. Let z' be the weak limit of $\{z^{i_k}\}_{k\in\mathbb{N}}$. According to [45, Lemma 1.3], we have that

$$\Omega = \bigcap_{k=0}^{\infty} \Omega_k = \operatorname{Lim} \Omega_k$$

and this implies that $z' \in \Omega.$ Hence, by taking (70) into account we get

$$F(z^k) \le F(z) \le F(z').$$

Since F is lower semicontinuous and the sequence $\{F(z^k)\}_{k\in\mathbb{N}}$ is nondecreasing this implies

$$F(z^{'}) \le \lim_{k\to\infty} F(z^{i_k}) = \lim_{k\to\infty} F(z^k) \le F(z) \le F(z^{'}),$$

that is, $F(z^{'}) = F(z)$ showing that $\{F(z^k)\}_{k\in\mathbb{N}}$ converges to the minimal value of F over Ω. Now observe that according to (60) and (69) we have

$$\begin{aligned} F(y^k) &\le F(y^k) + \omega_k D_G(y^k, y^{k-1}) \\ &\le F(y^0) + \omega_k D_G(y^0, y^{k-1}) \\ &\le F(y^0) + K \end{aligned}$$

because $y^0 \in \Omega \subseteq \Omega_k$. This implies that the sequence $\{y^k\}_{k\in\mathbb{N}}$ is bounded because it is contained in $L_F(F(y^0)+K)$. Also according to (60) we have

$$\begin{aligned} 0 &\le F(y^k) - F(z^k) \le \omega_k \left[D_G(z^k, y^{k-1}) - D_G(y^k, y^{k-1}) \right] \\ &= \omega_k \left[\left\|z^k\right\|^p - \left\|y^k\right\|^p + \left\langle J^{\psi} y^{k-1}, y^k - z^k \right\rangle \right], \end{aligned} \tag{71}$$

where the quantity between the square brackets is bounded because both sequences $\{y^k\}_{k\in\mathbb{N}}$ and $\{z^k\}_{k\in\mathbb{N}}$ are bounded as shown above. Note that $\{\omega_k\}_{k\in\mathbb{N}}$ converges to zero. Hence, by (71), we obtain that

$$\lim_{k\to\infty} \left[F(y^k) - F(z^k) \right] = 0,$$

and this proves (i).

For proving (ii) one reproduces without modifications the arguments made for the same purpose in the proof of Theorem 4.2 and keeping in mind that in the current circumstances we have that $S_k := S \cap \Omega_k = S$ for all $k \in \mathbb{N}$. □

ACKNOWLEDGMENTS

The work of Yakov Alber was supported in part by the KAMEA Program of the Israeli Ministry of Absorption. Dan Butnariu gratefully acknowledges the support of the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities (Grant

592/00). The work of Gábor Kassay was partially supported by a grant of the foundation "Sapientia" from Cluj-Napoca, Romania.

REFERENCES

[1] Alber Y., The solution of nonlinear equations with monotone operators in Banach spaces (Russian), *Siberian Mathematical Journal*, **16** (1975), 1-8.

[2] Alber Y., The regularization method for variational inequalities with nonsmooth unbounded operators in Banach space, *Applied Mathematical Letters*, **6** (1993), 63-68.

[3] Alber Y., Stability of the proximal projection algorithm for nonsmooth convex optimization problems with perturbed constraint sets, *preprint* 2001.

[4] Alber Y., Butnariu D. and Ryazantseva, I., Regularization methods for ill-posed inclusions and variational inequalities with domain perturbations, *Journal of Nonlinear and Convex Analysis*, **2** (2001), 53-79.

[5] Alber Y. and Notik A.I., Perturbed unstable variational inequalities with unbounded operators on approximately given sets, *Set-Valued Analysis*, **1** (1993), 393-402.

[6] Alber Y. and Ryazantseva I., Variational inequalities with discontinuous mappings, *Soviet Mathematics Doklady*, ,**25** (1982), 206-210.

[7] Alber Y., Burachik R. and Iusem A.N., A proximal point method for nonsmooth convex optimization problems in Banach spaces, *Abstract and Applied Analysis*, **2** (1997), 100-120.

[8] Alexandre P., The perturbed generalized Tikhonov's algorithm, *Serdica Mathematical Journal*, **25** (1999), 91-102.

[9] Alexandre P., Nguyen V.H. and Tossings P., The perturbed generalized proximal algorithm, *Mathematical Modelling and Numerical Analysis,* **32** (1998), 223-253.

[10] Aubin J.-P. and Ekeland, Applied Nonlinear Analysis, *John-Wiley and Sons*, New York, 1984.

[11] Aubin J.-P. and Frankowska H., Set-Valued Analysis, *Birkhäuser*, Boston, 1990.

[12] Bauschke, H.H. and Combettes P.L., A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Mathematics of Operation Research*, **26** (2000), 248-264.

[13] Beer, G., Topologies on Closed and Closed Convex Sets, *Kluwer Academic Publishers*, Dordrecht, 1993.

[14] Bonnans J.F. and Shapiro A., Perturbation Analysis of Optimization Problems, *Springer Verlag*, 2000.

[15] Brohe M. and Tossings P., Perturbed proximal point algorithm with nonquadratic kernel, *Serdica Mathematical Journal*, **26** (2000), 177-206.

[16] Browder F.E., Existence and approximation of solutions for nonlinear variational inequalities, *Proceedings of the National Academy of Science USA*, **56** (1966), 1080-1086.

[17] Butnariu D. and Iusem A.N., On a proximal point method of optimization in Banach spaces, *Numerical Functional Analysis and Optimization*, **18**, (1998), 723-744.

[18] Butnariu D. and Iusem A.N., Totally Convex Functions for Fixed Point Computation and Infinite Dimensional Optimization, *Kluwer Academic Publishers*, Dordrecht, 2000.

[19] Butnariu D., Iusem A.N. and Resmerita E., Total convexity for the powers of the norm in uniformly convex Banach spaces, *Journal of Convex Analysis* **7** (2000), 319-334.

[20] Butnariu D., Iusem A.N. and Zalinescu C., On uniform convexity, total convexity and the convergence of a proximal point and an outer Bregman projection method in Banach spaces, *Journal of Convex Analysis*, to appear.

[21] Censor Y. and Zenios S., Proximal minimization algorithm with D-functions, *Journal of Optimization Theory and Applications*, **73** (1992) 451-464.

[22] Cioranescu I., Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems, *Kluwer Academic Publishers*, Dordrecht, 1990.

[23] Cruceanu S., Regularization pour les problèmes à operateurs monotones et la mèthode de Galerkine, *Commentationes Mathematicae Universitatis Carolinae* 12 (1971), 1-13.

[24] Dontchev A.L. and Zolezzi T., Well-Posed Optimization Problems, *Springer Verlag*, Berlin, 1991.

[25] Eckstein J., Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, *Mathematics of Operation Research*, **18** (1993), 202-226.

[26] Eckstein J., Approximate iterations in Bregman-function-based proximal algorithms, *Mathematical Programming* **83** (1998), 113-123.

[27] Engl H.W., Hanke M. and Neubauer A., Regularization of Inverse Problems, *Kluwer Academic Publishers*, Dordrecht, 1996.

[28] Güler O., On the convergence of the proximal point algorithm for convex minimization, *SIAM Journal on Control and Optimization,* **29** (1991), 403-419.

[29] Iusem A., Augmented Lagrangian methods and proximal point methods for convex optimization, *Investigacion Operativa* **8** (1999), 11-49.

[30] Iusem A.N. and Otero R.G., Inexact version of proximal point and augmented Lagrangian algorithms in Banach spaces, *Numerical Functional Analysis and Optimization* **22** (2001), 609-640.

[31] Kaplan A. and Tichatschke R., Stable Methods for Ill-Posed Variational Problems, *Akademie Verlag*, 1994.

[32] Kaplan A. and Tichatschke R., Proximal point approach and approximation of variational inequalities, *SIAM Journal on Control and Optimization*, **39** (2000), 1136-1159.

[33] Kaplan A. and Tichatschke R., Proximal methods for variational inequalities with set-valued monotone operators, In: "From Convexity to Nonconvexity" (Edited by R.P. Gilbert, P.D. Panagiotopoulos and P.M. Pardalos), pp. 345-361, *Kluwer Academic Publishers*, Dordrecht, 2001.

[34] Kaplan A. and Tichatschke R., A general view on proximal point methods to variational inequalities in Hilbert spaces - iterative regularization and approximation, *Journal of Nonlinear and Convex Analysis* **2** (2001), 305-332.

[35] Kassay G., The proximal point algorithm for reflexive Banach spaces, *Studia Universitatis "Babes-Bolyai" - Sectia Mathematica* **30** (1985), 9-17.

[36] Lavrentev M.M., On Some Ill-Posed Problems of Mathematical Physics (Russian), *Nauka*, Novosibirsk, 1962.

[37] Lions J.L. and Magenes E., Problèmes aux limites non homogènes et applications, Vol. 1, *Dunod*, Paris, 1968.

[38] Liskovets O. A., Variational Methods for Solving Unstable Problems (Russian), *Nauka i Tekhnika,* Minsk, 1981.

[39] Liskovets O. A., Regularization of problems with discontinuous monotone arbitrarily perturbed operators, *Soviet Mathematics Doklady* **28** (1983), 324-327.

[40] Liskovets O. A., Discrete convergence of elements and operators for ill-posed problems with a monotone operator, *Soviet Mathematics Doklady* **31** (1985), 202-206.

[41] Liskovets O. A., External approximations for the regularization of monotone variational inequalities, *Soviet Mathematics Doklady* **36** (1988), 220-224.

[42] Liu F. and Nashed M.Z., Regularization of nonlinear ill-posed variational inequalities and convergence rates, *Set Valued Analysis* **6** (1998), 313-344.

[43] Martinet B., Régularization d'inéquations variationelles par approximations successive, *Revue Française de Informatique at Recherche Opérationelle* **2** (1970), 154-159.

[44] Martinet B., Algorithmes pour la résolution de problèmes d'optimisation et minimax, Thèse d'état, *Université de Grenoble*, Grenoble, France, 1972.

[45] Mosco U., Convergence of convex sets and solutions of variational inequalities, *Advances in Mathematics* **3** (1969), 510-585.

[46] Mosco U., Perturbations of variational inequalities, in: "Nonlinear Functional Analysis", Proceedings of the Symposium in Pure Mathematics, 18, Part 1, pp. 182-194, *American Mathematical Society*, Providence, R.I., 1970.

[47] Pascali D. and Sburlan S., Nonlinear Mappings of Monotone Type, *Stijthoff&Noordhoff International Publishers*, Alphen aan den Rijn, The Netherlands, 1978.

[48] Phelps R.R., Convex Functions, Monotone Operators and Differentiability, 2nd Edition, *Springer Verlag*, Berlin, 1993.

[49] Robinson S.M., Regularity and stability for convex multivalued functions, *Mathematics of Operations Research* **1** (1976), 130-143

[50] Rockafellar R.T., On the maximality of sums of nonlinear monotone operators, *Transactions of the American Mathematical Society* **149** (1970), 75-88.

[51] Rockafellar R.T., On the maximal monotonicity of subdifferential mappings, *Pacific Journal of Mathematics* **33** (1970), 209-216.

[52] Rockafellar R.T., Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization* **14** (1976), 877-898.

[53] Rockafellar R.T., Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Mathematics of Operation Research* **1** (1976), 97-116.

[54] Ryazantseva I., Variational inequalities with monotone operators on sets which are specified approximately, *USSR Computational Mathematics and Mathematical Physics* **24** (1984), 194-197.

[55] Showalter R.E., Monotone Operators in Banach Space and Nonlinear Partial Differential Equations, *American Mathematical Society*, 1997.

[56] Solodov M.V. and Svaiter B.F., An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions, *Mathematics of Operation Research* **25** (2000), 214-230.

[57] Solodov M.V. and Svaiter B.F., Forcing strong convergence of proximal point iterations in a Hilbert space, *Mathematical Programming* **87** (2000), 189-202.

[58] Yamada I., The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, In: "Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications" (D. Butnariu, Y. Censor and S. Reich, Editors), *Elsevier Science Publishers*, Amsterdam, The Netherlands, 2001, pp. 473-504.