Fejer-convergent algorithms which accept summable errors, approximated resolvents and the Hybrid Proximal-Extragradient method

Benar F. Svaiter *

Abstract

We prove that a large family of Fejer convergent iterative methods still converges to a solution when summable errors are incorporated to the algorithm. We define approximate resolvents, show that methods based on approximate resolvents fall within the aforementioned family and prove that approximate resolvents are the iteration maps of the hybrid proximal-extragradient method. We prove that the forward-backward splitting method, Tseng's modified forward-backward splitting method and Koreplevich method are all based in particular computations elements in approximate resolvents.

2010 Mathematics Subject Classification: 47H05, 49J52, 47N10.

Keywords: Fejer convergence, approximate resolvent, hybrid proximal-extragradient

1 Introduction

The aim of this work is to prove that convergence under summable errors is a generic property of a large class of Fejer convergent method, to introduce approximate resolvents, to show that approximate resolvents are the iteration map of the Hybrid Proximal Extragradient method and to prove that methods based on approximate resolvents belongs to the aforementioned class of Fejer convergent method.

We also show that the Forward-Backward method, Tseng's Modified Forward-Backward method and Korpelevich's method are all based in particular computations of elements in approximate resolvents, or equivalently, are particular cases of the Hybrid Proximal Extragradient method. In particluar, all these methods fal within the aforementioned class of Fejer convergen methods.

The original contribution of this work are: the definition of a class of Fejer convergent methods which accept summabel errors; the definition of approximate resolvents; a new transportation formula for the ε-enlargement of cocoercive operators and the proof that the Forward-Backwar splitting method is based on a particular computation of points in approximate resolvents/Hybrid Proximal-Extragradient iterations. Beside that, we provide an unifying framework for the Forward-Backward splitting method, Tseng's Modified Forward-Backward method and Korpelevich's method, this extending previous results on [10, 7]. This framework also leads to new algoritms, as exposed on [6].

*IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil (benar@impa.br) Partially supported by CNPq grants 302962/2011-5, 474944/2010-7, FAPERJ grant E-26/102.940/2011 and by PRONEX-Optimization.

This work is organized as follows. In Section 2 we introduce some basic definition and results. In Section 3 we define a class of Fejer convergent methods which accept summable errors in the computation of its iterates. In Section 4 de define approximate resolvents, show that these are the iteration maps of the Hybrid Proximal Extragradient method and prove that' methods based on approximate resolvents fall within the aforementioned class of Fejer convegent methods. In Section 5 we show that the Forward-Backward method is based on approximate resolvents, which is to say that is a particular case of the Hybrid Proximal-Extragrdient method. In Section 6 we reacall that Tseng's Modifed Forward-Backward method is based on approximate resolvents, which is to say that is a particular case of the Hybrid Proximal-Extragrdient method. In Section 7 we recall that Korpelevich's method is based on approximate resolvents, which is to say that is a particular case of the Hybrid Proximal-Extragrdient method. In Section 8

2 Basic definitions and results

In the first part of this section we review the concept and properties of Quasi-Fejer convergence, which will be used in our analysis of a class of Fejer convergent methods which converges even when a summable sequences of errors are incorporated. In the second part we stablish the notation concerning point-to-set maps, which will be used for defining the aforementioned class.

The last part of this section contains the material which will be needed to define approximate resolvents and the Hybrid Proximal-Extragradient (HPE) method, to prove that methods based on approximate resolvents/HPE method belongs to the aforementioned class, and to prove that some well know decomposition methods are based on approximate resolvents/HPE method.

As far as we know, this section contains just one original result, namely, Lemma 2.7 which is a transportation formula for co-coercive operators.

Quasi-Fejer convergence

The concept of Quasi-Fejer convergence was introduce by Ermol'ev[4] in the context of sequences of random variables. We will use a deterministic version of this notion, considered in [5, Definition 4.1]. All this material is standard knowledge and will be included for the sake of completeness. We do no claim to give her any original contribution to this over-studied concept.

Definition 2.1. *Let X be a metrical space. A sequence (x_n) in X is Quasi-Fejer convergent to $\Omega \subset X$ if, for each $x^* \in \Omega$ there exists a non-negative, summable sequence (ρ_n) such that*

$$d(x^*, x_n) \leq d(x^*, x_{n-1}) + \rho_n$$

The next proposition summarizes the properties of Quasi-Fejer convergent sequences in metrical spaces.

Proposition 2.2. *Let X be a metrical space and (x_n) be a sequence in X which is Quasi-Fejer convergent to $\Omega \subset X$ then,*

1. *if Ω is non-empty, then (x_n) is bounded;*
2. *for any $x^* \in \Omega$ there exists $\lim_{n\to\infty} d(x^*, x_n) < \infty$*
3. *if the sequence (x_n) has a cluster point $x^* \in \Omega$ then it converges to such a point.*

Proof. Take $x^* \in \Omega$ and let (ρ_n) be as in Definition 2.1. Then for $n < m$

$$d(x^*, x_m) \leq d(x^*, x_n) + \sum_{i=n+1}^{m} \rho_i$$

Hence

$$\limsup_{m\to\infty} d(x^*, x_m) \leq d(x^*, x_n) + \sum_{i=n+1}^{\infty} \rho_i < \infty,$$

which proves item 1. To prove item 2, note that (ρ_n) is summable and take the $\liminf_{n\to\infty}$ at the right hand side of the first inequality in the above equation. Item 3 follows trivially from item 2. □

Now we recall Opial's Lemma [8], which is useful for analyzing Quasi-Fejer convergence in Hilbert spaces:

Lemma 2.3 (Opial). *If in a Hilbert space X the sequence (x_n) is weakly convergent to x_0, then for any $x \neq x_0$*

$$\liminf_{n\to\infty} \|x_n - x\| > \liminf_{n\to\infty} \|x_n - x_0\|$$

The next result was proved in [9], for the case of a specific sequence generated by an inexact proximal point method, but the proof presented there is quite general, and we summarize it for the sake of completeness.

Proposition 2.4. *If in a Hilbert space X the sequence (x_n) is Quasi-Fejer convergent to $\Omega \subset X$ then, it has at most one weak cluster point in Ω.*

Proof. If $x^* \in \Omega$ is a weak cluster point of (x_n), then there exists a subsequence (x_{n_k}) weakly convergent to x^*. Therefore, using item 2 of Proposition 2.2 and Opial's Lemma we conclude that, for any $x' \in \Omega$, $x' \neq x^*$

$$\lim_{n\to\infty} \|x_n - x'\| = \liminf_{k\to\infty} \|x_{n_k} - x'\| > \liminf_{k\to\infty} \|x_n - x^*\| = \lim_{n\to\infty} \|x_n - x^*\|$$

which trivially implies the desired result. □

Point-to-set operators

Let X, Y be arbitrary sets. A point-to-set map $F : X \rightrightarrows Y$ is a function $F : X \to \wp(Y)$, where $\wp(Y)$ is the power set of Y, that is, the family of all subsets of Y. If $F(x)$ is a singleton, that is, a set with just one element, one says that F is point-to-point. Whenever necessary, we will identify a point-to-point map $F : X \rightrightarrows Y$ whit the unique function $f : X \to Y$ such that $F(x) = \{f(x)\}$ for all $x \in X$,

A point-to-set map $F : X \rightrightarrows Y$ is *L-Lipschitz* if X and Y are normed vector spaces and,

$$\emptyset \neq F(x') \subset \{y + u \mid y \in F(x),\ u \in Y,\ \|u\| \leq L\|x - x'\|\}, \qquad \forall x, x' \in X. \tag{1}$$

Note that if F is point-to-point and it is identified with a function then in the above definition we retrieve the classical definition of a L-Lipschitz continuous function.

Maximal monotone operators and the ε-enlargement

The ε-enlargement of a maximal monotone operators will be used to define approximate resolvents in Section 4. In this section we review the definition of the ε-enlargement and discuss those of its properties which will be used in the analysis and applications of approximate resolvents.

From now on X is a real Hilbert space. A point-to-set operator $T : X \rightrightarrows X$ is *monotone* if

$$\langle x - y, u - v\rangle \geq 0, \forall x, y \in X,\ u \in T(x), v \in T(y)$$

and it is *maximal monotone* if it is monotone and maximal in the family of monotone operators in X with respect to the partial order of the inclusion.

Let $T : X \rightrightarrows X$ be a maximal monotone operator. Recall that the ε-enlargement of T is defined as [1]

$$T^{[\varepsilon]}(x) = \{v \mid \langle x - y, v - u\rangle \geq -\varepsilon\}, \qquad x \in X, \varepsilon \geq 0. \tag{2}$$

Now we will state some elementary of ε-enlargement of T which follows trivially from the above definition and the basic properties of maximal monotone operators. Their proofs can be found in [1, 3, 13].

Proposition 2.5. *Let $T : X \rightrightarrows X$ be maximal monotone. Then*

1. $T = T^{[0]}$*;*
2. *if* $0 \leq \varepsilon_1 \leq \varepsilon_2$ *then* $T^{[\varepsilon_1]}(x) \subset T^{[\varepsilon_2]}(x)$ *for any* $x \in X$*;*
3. $\lambda\left(T^{[\varepsilon]}(x)\right) = (\lambda T)^{[\lambda\varepsilon]}(x)$ *for any* $x \in X$*,* $\varepsilon \geq 0$ *and* $\lambda > 0$*;*
4. *if* $v_k \in T^{[\varepsilon_k]}(x_k)$ *for* $k = 1, 2, \ldots,$ (x_k) *converges weakly to* x*,* (v_k) *converges strongly to* v *and* (ε_k) *converges to* ε *then* $v \in T^{[\varepsilon]}(x)$*;*
5. *if* $T = \partial f$*, where* f *is a proper closed convex function in* X*, then* $\partial_\varepsilon f(x) \subset T^{[\varepsilon]}(x) = (\partial f)^{[\varepsilon]}(x)$ *for any* $x \in X$*,* $\varepsilon \geq 0$*.*

The ε-enlargements of two operators can be "added" as follows. This fact was proved in [1] in a finite dimensional setting, but its extension to Hilbert and Banach spaces are straightforward.

Proposition 2.6. *It* $T_1, T_2 : X \rightrightarrows X$ *are maximal monotone and* $T_1 + T_2$ *is also maximal monotone then, for any* $\varepsilon_1, \varepsilon_2 > 0$ *and* $x \in X$

$$T_1^{[\varepsilon_1]}(x) + T_2^{[\varepsilon_2]}(x) \subset (T_1 + T_2)^{[\varepsilon_1 + \varepsilon_2]}(x)$$

A maximal monotone operator $B : X \to X$ is *α-cocoercive*, For $(\alpha > 0)$ if

$$\langle x - y, Bx - By\rangle \geq \alpha\|Bx - By\|^2, \qquad \forall x, y \in X.$$

There is an interesting "transportation formula" for cocoercive operators. This result was proved in the first draft of [6], but the proof was afterward removed. Whe inted to supress its proof, in case it reapears in that manuscript.

Lemma 2.7. *If* $A : X \to X$ *is α-cocoercive, then for any* $x, z \in X$*,*

$$A(z) \in A^{[\varepsilon]}(x), \qquad \text{with } \varepsilon = \frac{\|x - z\|^2}{4\alpha}$$

Proof. Take $y \in X$. Then

$$\begin{aligned}\langle x-y, Az-Ay\rangle &= \langle x-z, Az-Ay\rangle + \langle z-y, Az-Ay\rangle \\ &\geq \langle x-z, Az-Ay\rangle + \alpha\|Az-Ay\|^2 \\ &\geq -\|x-z\|\|Az-Ay\| + \alpha\|Az-Ay\|^2.\end{aligned}$$

where the first inequality follows form the cocoercivity of A and the second from Cauchy-Schwarz inequality. To end the proof, note that

$$-\|x-z\|\|Az-Ay\| + \alpha\|Az-Ay\|^2 \geq \inf_{t\in\mathbb{R}} \alpha t^2 - \|x-z\|t$$

and compute the value of the left hand-side of this inequality. □

The next result was essentialy proved in [10]. The usefulness of the σ-approximate resolvent follows from the next elementary result, essentially proved in [10, Lemma 2.3, Corollary 4.2].

Lemma 2.8. *Suppose that $T : X \rightrightarrows X$ is maximal monotone, $x \in X$ $\lambda > 0$ and $\sigma \geq 0$. If*

$$\begin{cases} v \in T^{[\varepsilon]}(y), \\ \|\lambda v + y - x\|^2 + 2\lambda\varepsilon \leq \sigma^2\|y-x\|^2, \end{cases} \qquad z = x - \lambda v,$$

then $\|\lambda v\| \leq (1+\sigma)\|y-x\|$, $\|z-y\| \leq \sigma\|y-x\|$ and for any $x^* \in T^{-1}(0)$

$$\begin{aligned}\|x^*-x\|^2 &\geq \|x^*-z\|^2 + \|y-x\|^2 - \left[\|\lambda v + y - x\|^2 + 2\varepsilon\right] \\ &\geq \|x^*-z\|^2 + (1-\sigma^2)\|y-x\|^2.\end{aligned}$$

Proof. Since $\varepsilon \geq 0$, $\|\lambda v + y - x\| \leq \sigma\|y-x\|$ and the two first inequalities of the lemma follows trivially from this inequality, triangle inequality and the definition of z.

To prove the third inequality of the lemma, take $x^* \in T^{-1}(0)$. Direct combination of the algebraic identities

$$\begin{aligned}&\|x^*-x\|^2 \\ &\quad = \|x^*-z\|^2 + 2\langle x^*-y, z-x\rangle + 2\langle y-z, z-x\rangle + \|z-x\|^2 \\ &\quad = \|x^*-z\|^2 + 2\langle x^*-y, z-x\rangle + \|y-x\|^2 - \|y-z\|^2\end{aligned}$$

with the definition of z yields

$$\|x^*-x\|^2 = \|x^*-z\|^2 + 2\lambda\langle x^*-y, -v\rangle + \|y-x\|^2 - \|\lambda v + y - x\|^2$$

Using the inclusions $0 \in T(x^*)$, $v \in T^{[\varepsilon]}(y)$ and definition (2) we conclude that $\langle x^*-y, 0-v\rangle \geq -\varepsilon$. To end the proof, of the third inequality, combine this inequality with the above equations.

The last inequality follows trivially from the third one and the assumptions of the lemma. □

3 A class of Fejer convergent methods

Let X be a Hilbert space and $\Omega \subset X$. We are concerned with iterative methods for solving problem

$$x \in \Omega. \tag{3}$$

which, in their exact or inexact form, generates sequences (x_n) as

$$x_n \in F_n(x_{n-1}) \text{ or } x_n \in F_n(x_{n-1}) + r_n, \qquad n = 1, 2, \dots$$

respectively, where $F_1 : X \rightrightarrows X, F_2 : X \rightrightarrows X, \dots$ are point-to-set maps and $r_1, r_2 \dots$ are errors. The basic elements here are the set Ω and the family of point-to-set maps (F_n).

We will consider two properties of a general family of point-to-set maps $(F_n : X \rightrightarrows X)_{n\in\mathbb{N}}$ with respect to $\Omega \subset X$:

P1: if $\hat{x} \in F_n(x)$ and $x^* \in \Omega$ then

$$\|x^* - \hat{x}\| \leq \|x^* - x\|$$

P2: if $(z_k)_{k\in\mathbb{N}}$ converges weakly to $\bar{z}$, $\hat{z}_k \in F_{n_k}(z_k)$ for $n_1 < n_2 < \cdots$ and for some $w \in \Omega$

$$\lim_{k\to\infty} \|w - z_k\| - \|w - \hat{z}_k\| = 0$$

then $\bar{z} \in \Omega$.

Property **P1** ensures that points in the image of $F_n(x)$ are closer (or no more distant) to Ω than x. Regarding property **P2**, note that (using property **P1**) we have

$$\|w - z_k\| - \|w - \hat{z}_k\| \geq 0.$$

The left hand-side of the above inequality measures the progress of $\hat{z}_k$ toward the solution set, as compared to z_k. Hence, property **P1** ensures that if the progress become "negligible" than the weak limit point of (z_k) belongs to Ω.

Theorem 3.1. *Suppose that $\Omega \subset X$ is non-empty and $(F_n : X \rightrightarrows X)$ is a sequence of point-to-set maps which satisfies conditions* **P1**, **P2** *with respect to Ω.*

If

$$x_n \in F_n(x_{n-1}) + r_n, \qquad \sum \|r_n\| < \infty$$

then (x_n) is Quasi-Fejer convergent to Ω, it converges weakly to some $\bar{x} \in \Omega$ and for any $w \in \Omega$ there exists $\lim_{n\to\infty} \|w - x_n\|$.

Moreover, if $r_n = 0$ for all n, then (x_n) is Fejer-convergent to Ω.

Proof. To simplify the proof, define

$$\hat{x}_n = x_n - r_n.$$

Take an arbitrary $x^* \in \Omega$. Since $\hat{x}_n \in F_n(x_{n-1})$, $\|w - \hat{x}_n\| \leq \|w - x_{n-1}\|$,

$$\|w - x_n\| \leq \|w - \hat{x}_n\| + \|r_n\| \leq \|w - x_{n-1}\| + \|r_n\|$$

and $(x_n)_{n\in\mathbb{N}}$ is Quasi-Fejer convergent to Ω. Therefore this sequence is bounded and there exists $\lim_{n\to\infty}\|w-x_n\|$. In particular

$$\lim_{n,m\to\infty}\|w-x_n\|-\|w-x_m\|=0$$

Since $(x_n)_{n\in\mathbb{N}}$ is bounded it has a weak cluster point, say $\bar{x}$ and there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges weakly to $\bar{x}$. Direct use of the definition of $\hat{x}_n$ yields

$$|\ \|w-x_{n_k}\|-\|w-\hat{x}_{n_k}\|\ |\leq|\ \|w-x_{n_k}\|-\|w-x_{n_k+1}\|\ |+\|r_{n_k+1}\|.$$

Therefore, using **P2**, the two above equations and the inclusion $\hat{x}_n\in F(x_n)$ we conclude that $\bar{x}\in\Omega$. Hence, all weak cluster points of (x_n) belongs to Ω. To end the proof, use Proposition 2.4 □

Note that properties **P1**, **P2** are "inherited" by specializations. Will state formally this result and the proof, being quite trivial, will be omitted

Proposition 3.2. *If* $(F_n:X\rightrightarrows X)$ *is a sequence of point-to-set maps which satisfies conditions* **P1**, **P2** *with respect to* $\Omega\subset X$ *and* $(G_n:X\rightrightarrows X)$ *is a sequence of point-to-set maps such that, for any* $x\in X$

$$G_n(x)\subset F_n(x),\qquad n=1,2,\ldots$$

then $(G_n:X\rightrightarrows X)$ *also satisfies conditions* **P1**, **P2** *with respect to* Ω.

What about compositions? Suppose that (F_n) is a sequence satisfying **P1**, **P2**, and that

$$F_n=G_n\circ X_n$$

where $G_n:X\rightrightarrows Y$, $H_n:Y\rightrightarrows X$ and all G_n's are L-Lipschitz continuous (Y is Hilbert). One may consider a sequences

$$\tilde{x}_n\in H_n(x_{n-1})+r_n,\qquad x_n\in G_n(\tilde{x}_n)+r'_n$$

where (r_n) and (r'_n) are summable. Indeed, since $x_n-r'_n\in G_n(\tilde{x}_n)$, using also (1) we conclude that there exists $\hat{x}_n\in G_n(\tilde{x}_n-r_n)$

$$\|\hat{x}_n-(x_n-r'_n)\|\leq L\ \|r_n\|.$$

Therefore, defining $s_n=x_n-\hat{x}_n$ and noting that $\tilde{x}_n-r_n\in H_n(x_{n-1})$ we conclude that

$$x_n\in F_n(x_{n-1})+s_n,\qquad \sum\|s_n\|\leq\sum L\|r_n\|+\|r'_n\|<\infty.$$

On may also consider compositions of $m+1$ maps

$$F=G_{1,n}\circ G_{2,n}\cdots\circ G_{m,n}\circ H_n$$

adding summable errors in each stage, assuming each $G_{i,n}:Y_i\rightrightarrows Y_{i-1}$ to be L-Lipschitz continuous, $H_n:X\rightrightarrows Y_m$, $Y_0=X$ etc.

4 Approximate resolvents and the Hybrid Proximal-Extragradient Method

In this section first we define σ-approximate resolvents, analyze some of their properties and study conditions under which sequences of σ-approximate resolvents satisfy properties **P1**, **P2**. After that we recall the definition of the Hybrid Proximal Extragradient method and shows that σ-approximate resolvents are the iteration maps of such method. At the end of the section we discuss the incorporation of summable errors to sequences of σ-approximate resolvents and to the Hybrid Proximal Extragradient method.

Recall that the *resolvent* of a maximal monotone operator $T : X \rightrightarrows X$ is defined as

$$J_T(x) = (I + T)^{-1}(x), \qquad x \in X. \tag{4}$$

We shall consider approximations of the resolvent in the following sense.

Definition 4.1. *The* σ-approximate resolvent *of a maximal monotone operator* $T : X \rightrightarrows X$ *is the point-to-set operator* $J_{T,\sigma} : X \rightrightarrows X$

$$J_{T,\sigma}(x) = \left\{ x - v \;\middle|\; \begin{array}{l} \exists \varepsilon \geq 0, y \in H, \\ v \in T^{[\varepsilon]}(y) \\ \|v + y - x\|^2 + 2\varepsilon \leq \sigma^2 \|y - x\|^2 \end{array} \right\}$$

where $\sigma \geq 0$.

First we analyze some elementary properties of approximate resolvent and find a convenient expression for $J_{\lambda T,\sigma}$. In particular, we will show that the σ-approximate resolvent is indeed and extension (in the sense of point-to-set maps) of the classical resolvent.

Proposition 4.2. *Let* $T : X \rightrightarrows X$ *be maximal monotone. Then, for any* $x \in X$,

1. $J_{T,\sigma=0}(x) = \{J_T(x)\}$;
2. *if* $0 \leq \sigma_1 \leq \sigma_2$ *then* $J_{T,\sigma_1}(x) \subset J_{T,\sigma_2}(x)$;
3. *for any* $\lambda > 0$ *and* $\sigma \geq 0$,

$$J_{\lambda T,\sigma}(x) = \left\{ x - \lambda v \;\middle|\; \begin{array}{l} \exists \varepsilon \geq 0, y \in H, \\ v \in T^{[\varepsilon]}(y) \\ \|\lambda v + y - x\|^2 + 2\lambda\varepsilon \leq \sigma^2 \|y - x\|^2 \end{array} \right\}$$

Proof. Items 1, 2 and 3 follow trivially from Definition 4.1 and Proposition 2.5, items 1, 2 and 3. □

Note that in view of item 1 of the above proposition, if point-to-set operators which are point-to-point are identified with functions, we have

$$J_{T,0} = J_T\,.$$

Next theorem is the main result of this section and states that, in some sense, approximate resolvents are "almost as good" as resolvents for finding zeros of maximal monotone operators.

Theorem 4.3. *Suppose that $T : X \rightrightarrows X$ is maximal monotone, $\sigma \in [0,1)$, $\underline{\lambda} > 0$ and (λ_k) is a sequence in $[\underline{\lambda}, \infty)$. Then, the sequence of point-to-set maps*

$$(J_{\lambda_k T,\sigma})_{k\in\mathbb{N}}$$

*satisfies properties **P1**, **P2** with respect to $\Omega = T^{-1}(0)$.*

Proof. Suppose that $\hat{x} \in J_{\lambda_k T,\sigma}(x)$. This means that there exists $y, v \in X$, $\varepsilon \geq 0$ such that

$$\hat{x} = x - \lambda_k v, \quad v \in T^{[\varepsilon]}(x), \quad \|\lambda_k v + y - x\|^2 + 2\lambda\varepsilon \leq \sigma^2\|y - x\|^2.$$

Therefore, using Lemma 2.8 we conclude that for any $x^* \in (\lambda_k T)^{-1}(0) = T^{-1}(0)$,

$$\|x^* - x\|^2 \geq \|x^* - \hat{x}\|^2 + (1-\sigma^2)\|y - x\|^2 \geq \|x^* - \hat{x}\|^2$$

which proves that the family $(J_{\lambda_k T,\sigma})$ satisfies **P1**.

Suppose that (z_k) converges weakly to $\bar{z}$, $\hat{z}_k \in J_{\lambda_{n_k} T,\sigma}(z_k)$, $0 \in T(x^*)$ and

$$\lim_{k\to\infty} \|x^* - z_k\| - \|x^* - \hat{z}_k\| = 0 \tag{5}$$

To simplify the proof, let $\mu_k = \lambda_{n_k} \geq \underline{\lambda}$. For each k there exists $v_k, y_k \in X$, $\varepsilon_k \geq 0$ such that

$$\hat{z}_k = z_k - \mu_k v, \quad v_k \in T^{\varepsilon_k}(z_k), \quad \|\mu_k v_k + y_k - z_k\|^2 + 2\mu_k\varepsilon \leq \sigma^2\|y_k - z_k\|^2. \tag{6}$$

Using again Lemma 2.8 we conclude that

$$\|x^* - z_k\|^2 \geq \|x^* - \hat{z}_k\|^2 + (1-\sigma^2)\|y_k - z_k\|^2.$$

Therefore

$$\begin{aligned}(1-\sigma^2)\|y_k - z_k\|^2 &\leq \|x^* - z_k\|^2 - \|x^* - \hat{z}_k\|^2 \\ &= (\|x^* - z_k\| - \|x^* - \hat{z}_k\|)(\|x^* - z_k\| + \|x^* - \hat{z}_k\|)\end{aligned}$$

Since (z^k) is weakly convergent, it is also bounded. Tanking this fact in to account and using the above equation and (5) we conclude that

$$\lim_{k\to\infty} \|y_k - z_k\| = 0$$

So, (y_k) also converges weakly to $\bar{z}$. Since $\varepsilon_k \geq 0$, using the last relation in (6) we conclude that

$$\mu_k\varepsilon_k \leq \frac{\sigma^2}{2}\|y_k - z_k\|^2, \qquad \|\mu_k v_k\| \leq (1+\sigma)\|y_k - z_k\|$$

Therefore, since (μ_k) is bounded away from 0,

$$\lim_{k\to\infty} \varepsilon_k = 0, \quad \lim_{k\to\infty} v_k = 0$$

and $0 \in T^{[0]}(\bar{z}) = T(\bar{z})$. □

The Hybrid Proximal-Extragradient/Projection methods were introduced in [11, 10, 12]. These methods are variants of the Proximal Pinto method which uses relative error tolerances for accepting inexact solutions of the proximal subproblems. Here we are concerned with the variant introduced in [10], which will be called, from now on, the Hybrid Proximal-Extragradient (HPE) method.

Algorithm: (projection free) HPE method[10]:
Choose $x_0 \in X$, $\sigma \in [0,1)$, $\underline{\lambda} > 0$ and for $k = 1, 2, \dots$
a)
Choose $\lambda_k \geq \underline{\lambda}$ and find/compute $v_k, y_k \in X$ $\varepsilon \geq 0$ such that

$$v_k \in T^{\varepsilon_k}(y_k), \qquad \|\lambda_k v_k + y_k - x_{k-1}\|^2 + 2\lambda_k \varepsilon_k \leq \sigma^2 \|y_k - x_{k-1}\|^2$$

b) Set $x_k = x_{k-1} - \lambda_k v_k$

To generate iteratively sequences by means of approximate resolvents is equivalent to applying the HPE method in the following sense.

Proposition 4.4. *Let $T : X \to X$ be maximal monotone, $\sigma \geq 0$, $\underline{\lambda} > 0$ and (λ_k) be sequence in $[\underline{\lambda}, \infty)$.*

A sequence (x_k) satisfies the recurrent inclusion

$$x_k \in J_{\lambda_k T, \sigma}(x_{k-1}), \qquad k = 1, 2, \dots$$

if and only if there exists sequences (y_k), (v_k), (ε_k) which, together with the sequences (x_k), (λ_k) satisfy steps a) and b) of the HPE method.

Proof. Use Definition 4.1 and Proposition 4.2 item 3. □

Convergence of the HPE method perturbed by a summable sequence of errors was proved directly in [2]. Here we can obtain this result combining Proposition 4.4 with Theorem 4.3.

Corollary 4.5. *If $T : X \rightrightarrows X$ is maximal monotone, $T^{-1}(0) \neq \emptyset$, $\underline{\lambda} > 0$, $\sigma \in [0,1)$, for $k = 1, 2, \dots$*

$$\begin{aligned} &\lambda_k \geq \underline{\lambda} \\ &v_k \in T^{[\varepsilon_k]}(\tilde{x}_k), \ \|\lambda_k v_k + \tilde{x}_k - x_{k-1}\|^2 + 2\lambda_k \varepsilon_k \leq \sigma \|\tilde{x}_k - x_{k-1}\|^2 \\ &x_k = x_{k-1} - \lambda_k v_k + r_k \end{aligned}$$

and $\sum \|r_k\| < \infty$ then (x_k) (and $(\tilde{x}_k)$) converges weakly to a point $\bar{x} \in T^{-1}(0)$

Corollary 4.6. *If $T : X \rightrightarrows X$ is maximal monotone, $T^{-1}(0) \neq \emptyset$, $\overline{\lambda} \geq \underline{\lambda} > 0$, $\sigma \in [0,1)$, for $k = 1, 2, \dots$*

$$\begin{aligned} &\overline{\lambda} \geq \lambda_k \geq \underline{\lambda} \\ &v_k \in T^{[\varepsilon_k]}(\tilde{x}_k), \ \|\lambda_k v_k + \tilde{x}_k - x_{k-1}\|^2 + 2\lambda_k \varepsilon_k \leq \sigma \|\tilde{x}_k - x_{k-1}\|^2 \\ &x_k = x_{k-1} - \lambda_k (v_k + r_k) \end{aligned}$$

and $\sum \|r_k\| < \infty$ then (x_k) (and $(\tilde{x}_k)$) converges weakly to a point $\bar{x} \in T^{-1}(0)$

5 The Forward-Backward splitting method

We will prove in this section that the iteration maps of the Forward-Backward splitting method are specialization or selections of σ-approximate resolvents and the sequence of iteration maps satisfies properties **P1**, **P2**. As a consequence of this result, sequences generated by the inexact Forward-Backward splitting methods with summable errors still converges weakly to solutions of the inclusion problem, if any. Equivalently, the Forward-Backward splitting method is a particular instance of the HPE method.

The Forward-Backward Splitting method solves the inclusion problem

$$0 \in (A + B)x$$

where

h1) $A : X \to X$ is α-cocoercive, $\alpha > 0$;
h2) $B : X \rightrightarrows X$ is maximal monotone.

This method proceed as follows:

Forward-Backward Splitting method
0) Initialization: Choose $0 < \underline{\lambda} \le \bar{\lambda} < 2\alpha$ and $x_0 \in X$;
1) for $k = 1, 2, \ldots$
a) choose $\lambda_k \in [\underline{\lambda}, \bar{\lambda}]$ and define

$$\begin{aligned} x_k &= (I + \lambda_k B)^{-1}(x_k - \lambda_k A(x_{k-1})) \\ &= J_{\lambda_k B} \circ (I - \lambda_k A)\ (x_{k-1}). \end{aligned} \tag{7}$$

Note that the generic iteration map of the Forward-Backward method is

$$J_{\lambda B} \circ (I - \lambda A) \tag{8}$$

whith $\lambda = \lambda_k$ in the k-th iteration.

Lemma 5.1. *If A, B satisfies assumptions* **h1** *and* **h2** *then, for any $\lambda > 0$ and $x \in X$,*

$$J_{\lambda B} \circ (I - \lambda A)(x) \in J_{\lambda(A+B),\sigma}(x).$$

with $\sigma = \sqrt{\lambda/2\alpha}$.

Proof. Take $x \in X$ and let $z = J_{\lambda B} \circ (I - \lambda A)(x)$. This means that

$$a := \lambda^{-1}(x - \lambda A(x) - z) \in A(z).$$

Define $\varepsilon = \|x - z\|^2/(4\alpha)$, $v = a + A(x)$. Using Lemma 2.7 we conclude that $A(x) \in A^{[\varepsilon]}(z)$. Therefore combining this result with these two definitions, the above equation, Proposition 2.6 and Proposition 2.5 item 1 we conclude that

$$\begin{aligned} &v \in (B + A^{[\varepsilon]})(z) \subset (A + B)^{[\varepsilon]}(z), \quad \|\lambda v + z - x\|^2 + 2\lambda\varepsilon = \sigma^2\|z - x\|^2 \\ &z = x - \lambda v, \end{aligned}$$

which proves the lemma. □

Corollary 5.2. *Let A, B be as in* **h1**, **h2** *and* $\underline{\lambda}$, $\bar{\lambda}$ *and* (λ_k), (x_k) *be as in the Forward Backward method. Define*

$$\sigma = \bar{\lambda}/(2\alpha).$$

Then $0 < \sigma < 1$, *for any* $x \in X$

$$J_{\lambda_k B} \circ (I - \lambda_k A)(x) \in J_{\lambda_k (A+B);\sigma}(x), \qquad k = 1, 2, \dots \tag{9}$$

in particular

$$x_k = J_{\lambda_k B} \circ (I - \lambda_k A)(x_k) \in J_{\lambda_k (A+B);\sigma}(x_k), \qquad k = 1, 2, \dots \tag{10}$$

and the sequence of maps $(J_{\lambda_k B} \circ (I - \lambda_k A))$ *satisfies properties* **p1**, **p2** *with respect to* $(A+B)^{-1}(0)$.

Proof. The bounds for σ follows trivially from its definition and the choices for $\underline{\lambda}$, $\bar{\lambda}$ in the Forward-Backward method.

Define $\sigma_k = \lambda_k/(2\alpha)$ for $k = 1, 2, \dots$. Since $\lambda_k \in [\underline{\lambda}, \bar{\lambda}]$, $0 < \sigma_k \leq \sigma$ for all k. Therefore, using also Lemma 5.1 and Proposition 4.2 item 2 we conclude that for any $x \in X$

$$J_{\lambda_k B} \circ (I - \lambda_k A)(x) \in J_{\lambda_k (A+B);\sigma_k}(x) \subset J_{\lambda_k (A+B);\sigma}(x), \quad k = 1, 2, \dots$$

The equality in (10) follows trivially from the definition of the Forward-Backward method while the inclusion follows from the above equation. To end the proof, note that $0 < \underline{\lambda} < \lambda_k$ for all k, and use Theorem 4.3, Proposition 3.2 and the above equation. □

Proposition 5.3. *Let* (λ_k), (x_k) *be sequences generated by the Forward-Backward Splitting method. Define*

$$\sigma = \sqrt{\frac{\bar{\lambda}}{2\alpha}}, \;\; v_k = \lambda_k^{-1}(x_{k-1} - x_k), \;\; \varepsilon_k = \frac{\|x_k - x_{k-1}\|^2}{4\alpha}, \; k = 1, 2, \dots$$

Then $0 < \sigma < 1$ *and for* $k = 1, 2, \dots$

$$\begin{aligned} &v^k \in (B + A)^{[\varepsilon_k]}(x_k), \quad \|\lambda_k v_k + x_k - x_{k-1}\|^2 + 2\lambda_k \varepsilon_k \leq \sigma \|x_k - x_{k-1}\|^2 \\ &x_k = x_{k-1} - \lambda_k v_k. \end{aligned}$$

In particular, the Forward-Backward splitting method above defined is a particular case of the HPE method with $\sigma \in (0, 1)$.

Proof. Se the proofs of Lemma 5.1 and Corollary 5.2. □

6 Tseng's Modified Forward-Backward splitting method

In [10, 7] it was proved that Tseng's Modified Forward-Backward splitting method is a particular case of the HPE method. Here we will cast this result in the framework of approximate resolvents.

We will present examples of application of Theorem 3.1 to the analysis of Tseng's Modified Forward-Backward Splitting Method [14]. It is important to remark that this choice is arbitrary, and we do not pretend to present a new algorithm, because we believe that addition of summable errors to well-know algorithms does not create a new method.

In this section we consider the inclusion problem

$$0 \in (A+B)\,x$$

where

t1) $A : X \to X$ is monotone and L-Lipschitz continuous ($L > 0$);
t2) $B : X \rightrightarrows X$ is maximal monotone.

The exact Tseng's Modified Forward-Backward Splitting method [14] proceeds as follows:
TSENG'S MODIFIED FORWARD-BACKWARD METHOD
Choose $0 < \underline{\lambda} \leq \bar{\lambda} < 1/L$ and $x_0 \in X$;
for $k = 1, 2, \ldots$
a) choose $\lambda_l \in [\underline{\lambda}, \bar{\lambda}]$ and compute

$$y_k = (I + \lambda_k B)^{-1}(x_{k-1} - \lambda_k A(x_{k-1})), \qquad x_k = y_{k-1} - \lambda_k(A(y_k) - A(x_{k-1})).$$

In order to cast this method in the formalism of Section 3 define for $\lambda > 0$

$$H_\lambda : X \to X \times X, \qquad H_\lambda(x) = (x, J_{\lambda B}(x - \lambda A(x))) \tag{11}$$

$$G_\lambda : X \times X \to X, \qquad G_\lambda(x, y) = y - \lambda(A(y) - A(x))) \tag{12}$$

Note that the second component of the (generic) operator H_λ is $J_{\lambda B} \circ (I - \lambda A)$ which is the generic iteration map of the forward backward method in (8). Trivially

$$x_k = G_{\lambda_k} \circ H_{\lambda_k}\,(x_{k-1}). \tag{13}$$

The next two result were essentially proved in [7], in the context of the Hybrid Proximal extragradient method.

Lemma 6.1. *If A, B satisfy assumptions* **t1**, **t2**, *then, for any $\lambda > 0$ and $x \in X$*

$$G_\lambda \circ H_\lambda(x) \in J_{A+B;\sigma}\,(x)$$

for $\sigma = \lambda L$.

Proof. Take $x \in X$ and let

$$y = J_{\lambda B}(x - \lambda A(x)), \qquad z = y + \lambda(A(y) - A(x))$$

Note that $z = G_\lambda \circ H_\lambda(x)$. Using the definition of y we have

$$a := \lambda^{-1}(x - \lambda A(x) - y) \in B(y).$$

Therefore

$$v := a + A(y) \in (A+B)(y), \qquad \begin{aligned}\|\lambda v + x - y\|^2 &= \|\lambda(A(y) - A(x))\|^2 \\ &\leq (\lambda L)^2 \|y - x\|^2\end{aligned}$$

where the inequality follows from assumption **t2)**. To end the proof, note that $z = x - \lambda v$. □

Corollary 6.2. *Let A, B be as in* **t1**, **t2** *and $0 < \underline{\lambda} < \bar{\lambda} < 2\alpha$ and (λ_k), (x_k) be as in Tseng's Modified Forward Backward method. Define*

$$\sigma = \bar{\lambda} L.$$

Then $0 < \sigma < 1$, for any $x \in X$

$$G_{\lambda_k} \circ H_{\lambda_k}(x) \in J_{\lambda_k (A+B);\sigma}(x), \qquad k = 1, 2, \dots \tag{14}$$

in particular

$$x_k = G_{\lambda_k} \circ H_{\lambda_k}(x_{k-1}) \in J_{\lambda_k (A+B);\sigma}(x_{k-1}), \qquad k = 1, 2, \dots \tag{15}$$

and the sequence of maps $(G_{\lambda_k} \circ H_{\lambda_k})$ satisfyes properties **p1**, **p2** *with respect to $(A + B)^{-1}(0)$.*

Proof. The bounds for σ follows trivially from its definition and the choices for $\underline{\lambda}$ and $\bar{\lambda}$ in Tseng's Forward-Backward method.

Define $\sigma_k = \lambda_k L$ for $k = 1, 2, \dots$. Since $\lambda_k \in [\underline{\lambda}, \bar{\lambda}]$, $0 < \sigma_k \le \sigma$ for all k. Therefore, using also Lemma 6.1 and Proposition 4.2 item 2 we conclude that for any $x \in X$

$$J_{\lambda_k B} \circ (I - \lambda_k A)(x) \in J_{\lambda_k (A+B);\sigma_k}(x) \subset J_{\lambda_k (A+B);\sigma}(x), \quad k = 1, 2, \dots$$

The equality in (15) follows trivially from the definition of the Forward-Backward method while the inclusion follows from the above equation. To end the proof, note that $0 < \underline{\lambda} < \lambda_k$ for all k, and use Theorem 4.3, Proposition 3.2 and the above equation. □

Note that for $0 < \lambda \le \bar{\lambda}$, the maps H_λ, G_λ are Lipschitz continuous with constant

$$2 + \bar{\lambda} L, \qquad 1 + 2\bar{\lambda} L$$

respectively. Hence, this method can be perturbed by summable sequences of errors in the evaluations of the resolvents $J_{\lambda_k B}$ and/or in the evaluation of $A(x_k)$, $A(y_k)$ etc, and will still converge weakly to a solution, if any exists.

7 Korpelevich's method

In [7] it was proved that Korpelevich's method, with fixed stepsize, is a particular case of the HPE method. The extension of this result for variable stepsizes is trivial, and here we will analyze such an extension in the framework of approximate resolvents.

In this section we consider the inclusion problem

$$0 \in N_C(x) + B(x)$$

where

k1) $A : X \to X$ is monotone and L-Lipschitz continuous ($L > 0$).
k1) N_C is the normal cone operator of $C \subset X$, a non-emplty closed convex set;

Korpelevich's method
Choose $0 < \underline{\lambda} \le \bar{\lambda} < 1/L$ and $x_0 \in X$;

for $k = 1, 2, \ldots$
a) choose $\lambda_k \in [\underline{\lambda}, \bar{\lambda}]$ and define

$$y_k = P_C(x_{k-1} - \lambda_k F(x_{k-1})), \qquad x_k = P_C(x_{k-1} - \lambda_k F(y_k)) \tag{16}$$

In order to cast this method in the formalism of Section 3 define for $\lambda > 0$

$$H_\lambda : X \to X \times X, \qquad H_\lambda(x) = (x, P_C(x - \lambda A(x)) \tag{17}$$
$$G_\lambda : X \times X \to X, \qquad G_\lambda(x, y) = P_C(x - \lambda A(y)). \tag{18}$$

Observe that since $P_c = J_{\lambda N_C}$, the second component of the (generic) operator H_λ is $J_{\lambda B} \circ (I - \lambda A)$ with $B = N_C$ which is the generic iteration map of the forward backward method in (8) (whit $B = N_C$). Note also that the map H_λ above defined can be obtained from (11) setting $B = N_C$.
map H_λ above defined has an equivalent expression

$$H_\lambda(x) = (x, J_{\lambda N_C}(x - \lambda A(x)))$$

which can be obtained by setting $B = N_C$ in (11) $B = N_C$. Trivially,

$$x_k = G_{\lambda_k} \circ H_{\lambda_k}(x_{k-1}), \qquad k = 1, 2, \ldots \tag{19}$$

The next two result were essentially proved in [7], in the context of the Hybrid Proximal extragradient method.

Lemma 7.1. *If A and C satisfy assumptions* **k1** *and* **k2***, then, for any $\lambda > 0$ and $x \in X$*

$$G_\lambda \circ H_\lambda(x) \in J_{A+N_C;\sigma}(x)$$

for $\sigma = \lambda L$.

Proof. Take $x \in X$ and let

$$y = P_C(x - \lambda A(x)), \quad z = P_C(x - \lambda A(y))$$

Note that $z = G_\lambda \circ H_\lambda(x)$. Define

$$\eta = \frac{1}{\lambda}(x - \lambda A(x) - y),$$
$$\nu = \frac{1}{\lambda}(x - \lambda A(y) - z), \quad \varepsilon = \langle \nu, z - y \rangle, \quad v = \nu + A(y).$$

Trivially, $\eta \in N_C(y)$ and $\nu \in N_C(z) = \partial\delta_C(z)$. Therefore,

$$\nu \in \partial_\varepsilon \delta_C(y) \subset (\partial\delta_C)^{[\varepsilon]}(y) = (N_C)^{[\varepsilon]}(y),$$

and

$$v \in (A + N_C)^{[\varepsilon]}(y), \qquad z = x - \lambda v. \tag{20}$$

Trivially, $\eta \in N_C(y)$. Therefore

$$\begin{aligned}\|\lambda v + y - x\|^2 + 2\lambda\varepsilon &= \|y - z\|^2 + 2\lambda\langle \nu, z - y\rangle \\ &= \|y - z\|^2 + 2\lambda\langle \nu - \eta, z - y\rangle + 2\lambda\langle \eta, z - y\rangle \\ &\leq \|y - z\|^2 + 2\lambda\langle \nu - \eta, z - y\rangle\end{aligned}$$

Direct algebraic manipulations yields

$$\begin{aligned}\|y - z\|^2 + 2\lambda\langle \nu - \eta, z - y\rangle &= \|\lambda(\nu - \eta) + z - y\|^2 - \|\lambda(\nu - \eta)\|^2 \\ &\leq \|\lambda(\nu - \eta) + z - y\|^2 \\ &= \|\lambda(A(x) - A(y))\|^2\end{aligned}$$

Combining the two above equations and using assumption **k1** we conclude that

$$\|\lambda v + y - x\|^2 + 2\lambda\varepsilon \leq (\lambda L)^2\|y - x\|^2$$

The conclusion follows combining this inequality with (20). □

Corollary 7.2. *Let A, N_C be as in* **k1**, **k2** *and $0 < \underline{\lambda} < \bar{\lambda} < 2\alpha$ and (λ_k), (x_k) be as in Korpelevich's method. Define*

$$\sigma = \bar{\lambda} L.$$

Then $0 < \sigma < 1$, for any $x \in X$

$$G_{\lambda_k} \circ H_{\lambda_k}(x) \in J_{\lambda_k(A+B);\sigma}(x), \qquad k = 1, 2, \ldots$$

in particular

$$x_k = G_{\lambda_k} \circ H_{\lambda_k}(x_{k-1}) \in J_{\lambda_k(A+B);\sigma}(x_{k-1}), \qquad k = 1, 2, \ldots$$

and the sequence of maps $(G_{\lambda_k} \circ H_{\lambda_k})$ satisfies properties **p1**, **p2** *with respect to $(A + N_C)^{-1}(0)$.*

Proof. Use the same reasonings as in Corollary 7.2 □

Endowing $X \times X$ with the canonical inner product of Hilbert space products

$$\langle (x, y), (x', y')\rangle = \langle x, x'\rangle + \langle y, y'\rangle$$

it is trivial to check that for $0 < \lambda \leq \bar{\lambda}$, the maps H_λ and G_λ are Lipschitz continuous with constants

$$2 + \bar{\lambda}L, \qquad 1 + \bar{\lambda}L$$

respectively. Hence, one can analyze Korpelevich's method with (summable) errors in the projections and/or evaluations of A etc.

8 Discussion

We defined two properties of *exact* Fejer convergent algorithms which guarantee convergence of the inexact version (of them) using summable errors.

It has been since long recongized that Korpelevich's Method (an may be even the Forward-Backward method) is an "inexact" vesion of the proximal point method. However, the nature and degree of this "inexactness" where not known. We provided formal definition of aproximate solution of the prox by means of the σ-approximate resolvent which, while enconparsing many classical decomposition schemes, also guareantee weak convergence of sequences generated by suchh approximate resolvents (even in the presence of additional summable errors).

References

[1] R. S. Burachik, A. N. Iusem, and B. F. Svaiter. Enlargement of monotone operators with applications to variational inequalities. *Set-Valued Anal.*, 5(2):159–180, 1997.

[2] R. S. Burachik, S. Scheimberg, and B. F. Svaiter. Robustness of the hybrid extragradient proximal-point algorithm. *J. Optim. Theory Appl.*, 111(1):117–136, 2001.

[3] R. S. Burachik and B. F. Svaiter. ϵ-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Anal.*, 7(2):117–132, 1999.

[4] J. M. Ermol′ev. The method of generalized stochastic gradients and stochastic quasi-Fejér sequences. *Kibernetika (Kiev)*, (2):73–83, 1969.

[5] A. N. Iusem, B. F. Svaiter, and M. Teboulle. Entropy-like proximal methods in convex programming. *Math. Oper. Res.*, 19(4):790–814, 1994.

[6] R. D. C. Monteiro and B. F. Svaiter. Iteration-complexity of block-decomposition algorithms and the alternating minimization augmented lagrangian method. Manuscript, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA, August 2010.

[7] R. D. C. Monteiro and B. F. Svaiter. On the complexity of the hybrid proximal extragradient method for the iterates and the ergodic mean. *SIAM J. Optim.*, 20(6):2755–2787, 2010.

[8] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, 73:591–597, 1967.

[9] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization*, 14(5):877–898, 1976.

[10] M. V. Solodov and B. F. Svaiter. A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Anal.*, 7(4):323–345, 1999.

[11] M. V. Solodov and B. F. Svaiter. A hybrid projection-proximal point algorithm. *J. Convex Anal.*, 6(1):59–70, 1999.

[12] M. V. Solodov and B. F. Svaiter. A unified framework for some inexact proximal point algorithms. *Numer. Funct. Anal. Optim.*, 22(7-8):1013–1035, 2001.

[13] B. F. Svaiter. A family of enlargements of maximal monotone operators. *Set-Valued Anal.*, 8(4):311–328, 2000.

[14] P. Tseng. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.*, 38(2):431–446, 2000.