An inexact method of partial inverses. Application to splitting methods.

Regina S. Burachik*, Claudia Sagastizábal†, and Susana Scheimberg ‡

Abstract

For a maximal monotone operator T on a Hilbert space H and a closed subspace A of H, we consider the problem of finding $(x, y \in T(x))$ satisfying $x \in A$ and $y \in A^{\perp}$. An equivalent formulation of this problem makes use of the partial inverse operator of Spingarn. The resulting generalized equation can be solved by using the proximal point algorithm. We consider instead the use of hybrid proximal methods. Hybrid methods use enlargements of operators, close in spirit to the concept of ε-subdifferentials. We characterize the enlargement of the partial inverse operator in terms of the enlargement of T itself. We present a new algorithm of resolution that combines Spingarn and hybrid methods, we prove for this method global convergence only assuming existence of solutions and maximal monotonicity of T. We also show that, under standard assumptions, the method has a linear rate of convergence. For the important problem of finding a zero of a sum of maximal monotone operators $T_1, \ldots, T_m$, we present a highly paralellizable scheme. Unlike the standard requirements, we only assume maximal monotonicity of each term T_i, as in Spingarn's pioneer's work. We also illustrate the applicability of this approach to some interesting problems.

Keywords. maximal monotone operator, splitting methods, proximal point algorithm, enlargement of a maximal monotone operator, hybrid proximal methods.

---

*Engenharia de Sistemas e Computação, COPPE–UFRJ, CP 68511, 21945-970 Rio de Janeiro-RJ, Brazil. Research supported by PRONEX–Optimization and CNPq Grant 301280/94-0(RE)

†IMPA, Estrada Dona Castorina 110, Jardim Botânico, 22460-320 Rio de Janeiro-RJ, Brazil. Also INRIA, BP 105, 78153 Le Chesnay, France. Research supported by FAPERJ (Brazil) under Grant No.E26/150.205/98. `sagastiz@impa.br`

‡Instituto de Matemática/Engenharia de Sistemas e Computação, COPPE–UFRJ, CP 68511, 21945-970 Rio de Janeiro-RJ, Brazil. Research supported by CNPq Grant 302393/85-4(RN)

AMS Subject Classification: 65K05, 90C25

1 Introduction

Let T be a maximal monotone operator in a Hilbert space H, and consider the *primal-dual* problem:

$$\text{Given a subspace } A \text{ in } H, \text{ find } (x, y \in T(x)) \text{ such that } (x, y) \in A \times A^{\perp}, \tag{1}$$

where $A^{\perp}$ stands for the orthogonal space to A. The well-known problem of finding the zeroes of a maximal monotone operator as well as many optimization and equilibrium problems can be embedded in this setting; in § 5.2 below we give a comprehensive list of applications.

To solve (1), in [21] Spingarn considers the equivalent problem of finding a zero of an associated maximal monotone operator, T_A, called the *partial inverse* of T. For x in H, write $x = x_A + x_{A^{\perp}}$, with $x_A \in A$ and $x_{A^{\perp}} \in A^{\perp}$, the projections onto A and $A^{\perp}$, respectively. Then the partial inverse operator is obtained by swapping the $_{A^{\perp}}$-components of each pair in the graph of T:

$$w \in T_A(z) \quad \Longleftrightarrow \quad w_A + z_{A^{\perp}} \in T(z_A + w_{A^{\perp}}) \,. \tag{2}$$

The importance of this operator stems from the fact that (x, y) solves (1) if and only if

$$0 \in T_A(x + y) \,. \tag{3}$$

Thus, the original problem (1) can be replaced by the problem of finding a solution of the generalized equation (3).

A classical algorithm used to solve (3) is the proximal point method, [11, 10, 15]. Each iteration of this method, when applied to find a zero of a maximal monotone operator F, takes the form:

$$\text{given } \tilde{x} \in H \text{ find } \left(x^+, y^+\right) \text{ such that } \begin{cases} y^+ \in F(x^+) & \text{(a)} \\ 0 = cy^+ + x^+ - \tilde{x} \,, & \text{(b)} \end{cases} \tag{4}$$

where c is a scalar positive parameter, possibly varying along iterations. In order to allow the practical implementation of the method, inexact computations have to be introduced in the analysis. The inexact iterations proposed in [15] consist in substituting (4)(b) by

$$0 = cy^+ + x^+ - \tilde{x} + r \,,$$

for $y^+ \in F(x^+)$ and r a residual error term. The original convergence analysis for this method allows inexact iterates to produce residual norms which form a finite series. In other words, asymptotically (4) will still have to be satisfied exactly. The recently introduced "hybrid" methods of [18, 17, 19], instead, develop an error analysis which does not require increasing accuracy. More precisely, they perform inexact iterations where both (a) and (b) in (4) are relaxed and the error of each iteration does not necessarily decreases along the process.

The *Partial Inverse Method* (PIM) proposed in [21] solves (1) by applying the proximal point method to (3). The error analysis in this work follows the summable error approach of [15]. Our aim is to introduce a relative error scheme of the kind studied in [18, 17, 19] for the Partial Inverse Method. The proximal system (4), written for $F := T_A$, becomes:

$$\text{given } z \in H \text{ find } \left(z^+, w^+\right) \text{ such that } \begin{cases} w^+ \in T_A(z^+) & \text{(a)} \\ 0 = cw^+ + z^+ - z\,. & \text{(b)} \end{cases} \tag{5}$$

Besides from relaxing equation (b) in the proximal system (5), relation (a) could be likewise relaxed, using an outer approximation that somewhat *enlarges* the operator in the inclusion. In [1] an appropriate enlargement for maximal monotone operators was introduced. The enlargement of T_A is denoted by T_A^ε, we recall its definition in (10) below.

Here we shall focus on the approximation criterion of [20]. Specifically, when given $z \in H$

$$\left(\varepsilon, \tilde z, \tilde w\right) \text{ satisfies } \begin{cases} \tilde w \in T_A^\varepsilon(\tilde z) & \text{(a)} \\ \|c\tilde w + \tilde z - z\|^2 + 2c\varepsilon \le \sigma^2(\|c\tilde w\|^2 + \|\tilde z - z\|^2) & \text{(b)}, \end{cases} \tag{6}$$

the triplet $(\varepsilon, \tilde z, \tilde w)$ satisfying (6) is said to be an *approximate solution* of the proximal system (5) with tolerance error $\sigma \in [0, 1)$ and the next iterate is taken as $z^+ = z - c\tilde w$. This rule is constructive, in the sense that it is clearly related to the behaviour of the method around the current point for the problem being solved. Moreover, since along iterations tolerances can be kept bounded away from zero (in fact, they can be even fixed), the *relative* error in (6) need not tend to zero (as opposed to previous inexact proximal methods).

In this paper we consider the resolution of (1) with a method defined by using inexact iterates as in (6).

The paper is organized as follows. We start in Section 2 by recalling the definition and some properties of the partial inverse operator. Section 3 is devoted to the ε-enlargement of T, in particular we characterize the enlargement of the partial inverse operator in terms of the enlargement T^ε itself. In Section 4 we give an inexact method of partial inverses, that we call sPIM(ε). To prove its convergence, we show the equivalence of sPIM(ε) with the hybrid extragradient method of [20]. Finally, in Section 5 we give an important instance of sPIM(ε), that results in a new splitting method. We conclude with some comments on fields of applications and comparisons with other splitting methods.

2 Partial Inverse of a monotone operator

Given a monotone operator T (not necessarily maximal), the *partial inverse of T with respect to A* is defined in [21] by (2). The operator T_A is maximal monotone if and only if T is maximal monotone[21, Proposition 2.1]. The terminology "partial inverse" can be understood by examining the limiting cases $A = H$ and $A = \{0\}$. The former case gives $T_A = T$ while for the latter, T_A is T^{-1}. Also, $T_{A^\perp} = {T_A}^{-1}$.

The *scaled* version of the Partial Inverse Method (sPIM) considered in [21] for $T = \partial f$ and in [9] for arbitrary T consists in applying the proximal point algorithm to solve (3) with T therein replaced by ηT, where η is a positive parameter (note that $(\eta T)_A \neq \eta T_A$). Scaling can be useful for numerical purposes; in particular, [9] gives the optimal scaling factor for accelerating convergence if T is a strongly monotone operator.

Specifically, given $z \in H$, sPIM defines the next iterate z^+ as the first component of the couple (z^+, w^+) satisfying the scaled proximal system

$$\text{given } z \in H \text{ find } \left(z^+, w^+\right) \text{ such that } \begin{cases} w^+ \in (\eta T)_A(z^+) & \text{(a)} \\ 0 = cw^+ + z^+ - z\,. & \text{(b)} \end{cases} \tag{7}$$

We relate proximal images in T_A to elements in T via an intermediate maximal monotone operator S_c defined as follows:

$$w \in S_c(z) \quad \Longleftrightarrow \quad \frac{1}{c}w_A + w_{A^\perp} \in T(z_A + \frac{1}{c}z_{A^\perp})\,. \tag{8}$$

In addition, to establish a link with (7), we use the following system associated to S_c:

$$\text{given } x \in A, y \in A^\perp, \text{ find } \left(x^+, y^+\right) \text{ such that } \begin{cases} y^+ \in S_c(x^+) & \text{(a)} \\ 0 = \eta y^+ + x^+ - (x + \eta y)\,. & \text{(b)} \end{cases} \tag{9}$$

Note that the proximal above is a particular instance of (4), written with $F := \eta S_c$, $c := 1$ and $\tilde{x} := (x + \eta y)$. The connection between systems (7) and (9) is formally expressed in the result below, whose proof is straightforward from the definitions of T_A and S_c:

Proposition 2.1 *Let* $\left(z^+, w^+\right)$ *solve the proximal system* (7) *for a given* $z \in H$. *Then,* $\left(z_A^+ + cw_{A^\perp}^+, \frac{1}{\eta}(z_{A^\perp}^+ + cw_A^+)\right)$ *solves* (9) *for* $x := z_A$ *and* $y := \frac{1}{\eta} z_{A^\perp}$.

Reciprocally, let $\left(x^+, y^+\right)$ *be a solution of* (9) *for given* $(x, y) \in A \times A^\perp$. *Then* $\left(x_A^+ + \eta y_{A^\perp}^+, \frac{1}{c}(x_{A^\perp}^+ + \eta y_A^+)\right)$ *solves the proximal system* (7) *for* $z := x + \eta y$. $\square$

Suppose $c = 1$ and $T = \partial f$, the subdifferential of a proper closed convex function f. The solution of (9) amounts to finding x^+, the proximal point of f at $x + \eta y$, with regularization parameter $1/\eta$. This form of sPIM is called *proximal decomposition* algorithm, see [9].

We now consider how to relax the inclusion (a) in both (7) and (9) by using suitable enlargements of the operators T_A and S_c.

3 Enlargements of maximal monotone operators

It is known that the convergence analysis of bundle methods for nonsmooth optimization heavily relies in the concept of ε-subdifferential. Likewise, for a general maximal monotone operator T, the following enlargement plays an essential role:

$$T^\varepsilon(z) := \{w \in H \mid \langle w - y, z - x\rangle \geq -\varepsilon \text{ for all } (x, y \in T(x))\}, \tag{10}$$

where $\varepsilon \geq 0$ and $z \in H$ are fixed. This enlargement was introduced in [1] and was further studied in [2], [3]. Note that $T^0(z) = T(z) \subset T^\varepsilon(z)$. Among other useful properties, it holds that T^ε is continuous as an operator of both x and ε (unlike T itself).

Since T_A is a maximal monotone operator, its $\varepsilon-$enlargement is well defined. The following lemma characterizes the enlargement of the partial inverse operator, $(T_A)^\varepsilon$, as the partial inverse of the enlargement T^ε.

Lemma 3.1 *Given a maximal monotone operator* T, *consider its partial inverse* T_A *and its enlargement* T^ε *defined in* (2) *and* (10), *respectively. The following holds*

$$(T_A)^\varepsilon = (T^\varepsilon)_A\,.$$

Proof. Take $z \in H$ arbitrary and $w \in (T_A)^\varepsilon(z)$. Using (10) with T replaced by T_A together with (2), we see that $w \in (T_A)^\varepsilon(z)$ if and only if for all x and $y \in T(x)$ it holds that

$$-\varepsilon \leq \langle w - (y_A + x_{A^\perp}), z - (x_A + y_{A^\perp})\rangle .$$

Write w and z in $_A$ and $_{A^\perp}$ components, expand the inner product and use the orthogonality between $_A$ and $_{A^\perp}$ to obtain

$$\begin{aligned}
-\varepsilon &\leq \langle w - (y_A + x_{A^\perp}), z - (x_A + y_{A^\perp})\rangle \\
&= \langle w_A - y_A + w_{A^\perp} - x_{A^\perp}, z_A - x_A + z_{A^\perp} - y_{A^\perp}\rangle \\
&= \langle w_A - y_A, z_A - x_A\rangle + \langle w_{A^\perp} - x_{A^\perp}, z_{A^\perp} - y_{A^\perp}\rangle \\
&= \langle w_A - y_A, z_A - x_A\rangle + \langle z_{A^\perp} - y_{A^\perp}, w_{A^\perp} - x_{A^\perp}\rangle \\
&= \langle w_A - y_A + z_{A^\perp} - y_{A^\perp}, z_A - x_A + w_{A^\perp} - x_{A^\perp}\rangle \\
&= \langle (w_A + z_{A^\perp}) - y, (z_A + w_{A^\perp}) - x\rangle
\end{aligned}$$

Thus, $w \in (T_A)^\varepsilon(z)$ if and only if $w_A + z_{A^\perp} \in T^\varepsilon(z_A + w_{A^\perp})$, which is to say $w \in (T^\varepsilon)_A(z)$. □

An immediate consequence of this result is that the notation T_A^ε can be understood without ambiguity.

We now express $\varepsilon-$enlargements of various operators derived from T in terms of the enlargement T^ε itself. These relations will be useful in the next section. For simplicity of notation, we let $S_c^{\tilde{\varepsilon}}$ denote the operator $(S_c)^{\tilde{\varepsilon}}$.

Lemma 3.2 *Given a maximal monotone operator $T : H \rightrightarrows H$, consider the scaled operator (ηT) for $\eta > 0$ and the intermediate operator S_c defined in (8). The following holds:*

(i) $(\eta T)^\varepsilon = \eta T^{\varepsilon/\eta}$.

(ii) $\tilde{w} \in (\eta T)^\varepsilon_A(\tilde{z})$ *if and only if* $\dfrac{1}{\eta}(\tilde{w}_A + \tilde{z}_{A^\perp}) \in T^{\varepsilon/\eta}(\tilde{z}_A + \tilde{w}_{A^\perp})$.

(iii) $\tilde{y} \in S_c^{\tilde{\varepsilon}}(\tilde{x})$ *if and only if* $\dfrac{1}{c}\tilde{y}_A + \tilde{y}_{A^\perp} \in T^{\tilde{\varepsilon}/c}(\tilde{x}_A + \dfrac{1}{c}\tilde{x}_{A^\perp})$.

Proof. Item (i) is straightforward from the definition (10), written with T replaced by ηT. Item (ii) follows from (2). written with T replaced by $(\eta T)^\varepsilon$ and using (i). For proving (iii), define

$$\tilde{u} := \tilde{x}_A + \frac{1}{c}\tilde{x}_{A^\perp} \quad \text{and} \quad \tilde{v} := \frac{1}{c}\tilde{y}_A + \tilde{y}_{A^\perp} ,$$

and take $(u, v \in T(u))$. Use the orthogonality between ${}_A$ and ${}_{A^\perp}$ and some elementary algebra to write

$$\langle \tilde{v} - v, \tilde{u} - u \rangle \;\; = \;\; \frac{1}{c}\langle \tilde{y} - (cv_A + v_{A^\perp}), \tilde{x} - (u_A + cu_{A^\perp}) \rangle \, . \tag{11}$$

Note that (8) can be also expressed as

$$v \in T(u) \iff cv_A + v_{A^\perp} \in S_c(u_A + cu_{A^\perp}) \tag{12}$$

To show the desired equivalence, combine (11) and (12), as follows. In view of (12), $\tilde{y} \in S_c^{\tilde{\varepsilon}}(\tilde{x})$ if and only if the rightmost expression in (11) is greater than or equal to $-\tilde{\varepsilon}/c$, which is in turn equivalent to $\tilde{v} \in T^{\tilde{\varepsilon}/c}(\tilde{u})$, when using the leftmost expression in (11). $\square$

Finally, for an operator $\mathcal{T}$ defined as the cartesian product of two maximal monotone operators T_1 and T_2, we relate the ε-enlargement of $\mathcal{T}$ with the ε-enlargements of T_1 and T_2. We equip $\mathcal{H} := H \times H$ with the natural inner product

$$\langle (y_1, y_2), (x_1, x_2) \rangle_{\mathcal{H}} := \langle y_1, x_1 \rangle_H + \langle y_2, x_2 \rangle_H \, .$$

Lemma 3.3 *Given two maximal monotone operators $T_1 \, , T_2 : H \rightrightarrows H$, consider the operator $\mathcal{T} : \mathcal{H} \rightrightarrows \mathcal{H}$ defined by their cartesian product, $\mathcal{T} := (T_1, T_2)$. The following hold:*
(i) $T_1^{\varepsilon_1}(z_1) \times T_2^{\varepsilon_2}(z_2) \subset \mathcal{T}^{\varepsilon_1+\varepsilon_2}(z_1, z_2)$ for all $\varepsilon_1 \, , \varepsilon_2 \geq 0$, and
(ii) $\mathcal{T}^{\varepsilon}(z_1, z_2) = \bigcup \{T_1^{\varepsilon_1}(z_1) \times T_2^{\varepsilon_2}(z_2) : \varepsilon_1, \, \varepsilon_2 \geq 0 \, , \varepsilon_1 + \varepsilon_2 \leq \varepsilon\}$ for $\varepsilon \geq 0$.

Proof. In the inner products below we drop subindices ${}_{\mathcal{H}}$ and ${}_H$. To show item (i), let $w_i \in T_i^{\varepsilon_i}(z_i)$ and take any $(x_i \, , y_i \in T_i(x_i))$ for $i = 1, 2$. Then (w_1, w_2) is an element of $\mathcal{T}^{\varepsilon_1+\varepsilon_2}(z_1, z_2)$ because

$$\begin{aligned} \langle (w_1, w_2) - (y_1, y_2), (z_1, z_2) - (x_1, x_2) \rangle \;\; &= \langle w_1 - y_1, z_1 - x_1 \rangle + \langle w_2 - y_2, z_2 - x_2 \rangle \\ &\geq -\varepsilon_1 - \varepsilon_2 \, . \end{aligned}$$

To see item (ii), note that the inclusion "$\supset$" holds by (i). We now prove the opposite inclusion. Given $(w_1, w_2) \in \mathcal{T}^{\varepsilon}(z_1, z_2)$, for any $(x_i \, , y_i \in T_i(x_i))$ it holds that $\langle (w_1, w_2) - (y_1, y_2), (z_1, z_2) - (x_1, x_2) \rangle \geq -\varepsilon$, i.e.,

$$\langle w_1 - y_1, x_1 - z_1 \rangle + \langle w_2 - y_2, x_2 - z_2 \rangle \leq \varepsilon \, .$$

Since the inequality holds for arbitrary $(x_i\,, y_i \in T_i(x_i))$, it also holds for the following suprema:

$$\sup_{(x_1\,, y_1 \in T_1(x_1))} \langle w_1 - y_1, x_1 - z_1 \rangle \quad + \quad \sup_{(x_2\,, y_2 \in T_2(x_2))} \langle w_2 - y_2, x_2 - z_2 \rangle \leq \varepsilon\,. \tag{13}$$

Let $\beta_i(z_i, w_i)$ denote each term above:

$$\beta_i(z_i, w_i) := \sup_{(x_i\,, y_i \in T_i(x_i))} \langle w_i - y_i, x_i - z_i \rangle\,.$$

We claim that $\beta_i \geq 0$ for $i = 1, 2$. Suppose, for the purpose of contradiction, that $\beta_i(z_i, w_i) < 0$, for $i = 1$ or 2. Then there exists $\delta_i > 0$ such that $\beta_i(z_i, w_i) \leq -\delta_i < 0$ i.e., the inner product $\langle w_i - y_i, z_i - x_i \rangle \geq \delta_i > 0$ for all $(x_i\,, y_i \in T_i(x_i))$. Since T_i is maximal, this means that $w_i \in T_i(z_i)$, which in turn implies that $\beta_i(z_i, w_i) \geq 0$, by definition of β_i. We have then a contradiction, so letting $\varepsilon_i := \beta_i(z_i, w_i)$ for $i = 1, 2$, it holds that $\varepsilon_i \geq 0$. Together with (13), the result follows. $\square$

4 An inexact method of partial inverses

We now give an inexact method of partial inverses, that we call sPIM(ε). This method extends PIM, called Algorithm 4.1 in [21]. Each iteration of our algorithm is equivalent to one iteration of the hybrid extragradient method (HEPPM) of [20], applied to find a zero of T_A. We start by giving HEPPM as well as its convergence properties.

4.1 Hybrid proximal methods

Hybrid proximal methods define iterates using relaxed versions of exact proximal systems, like the rule in (6). In particular, for the scaled operator ηT and given $z \in H$, when

$$\left(\varepsilon, \tilde{z}, \tilde{w}\right) \text{ satisfies } \begin{cases} \tilde{w} \in (\eta T)^{\varepsilon}_{A}(\tilde{z}) & \text{(a)} \\ \|c\tilde{w} + \tilde{z} - z\|^2 + 2c\varepsilon \leq \sigma^2 \left(\|c\tilde{w}\|^2 + \|\tilde{z} - z\|^2\right) & \text{(b)}, \end{cases} \tag{14}$$

the triplet $(\varepsilon, \tilde{z}, \tilde{w})$ is said to be an *approximate solution* of the proximal system (7) with tolerance error $\sigma \in [0, 1)$. Note that (z^+, w^+), the solution of the exact proximal system (7) corresponds to $(0, z^+, w^+)$, an approximate solution with error tolerance $\sigma = 0$.

The algorithm proposed in [17] works as follows.

Hybrid Extragradient-Proximal Point Method (HEPPM):
Given the current iterate z^k, consider an approximate solution of (7), i.e., $\left(\varepsilon^k, \tilde{z}^k, \tilde{w}^k\right)$ satisfying (14) written with $c = c_k$ and tolerance error $\sigma_k \in [0, 1)$. The next iterate is computed as $z^{k+1} := z^k - c_k \tilde{w}^k$. The algorithm stops if $z^{k+1} = z^k$. □

The approximation error in (14)(b) corresponds to [20, Definition 1]. This approximation criterion, less restrictive than the original one from [17], is used for developing a unified theory for hybrid methods in [20]. The convergence result is the following.

Theorem 4.1 *[Theorems 6 and 8 in [20]] Consider* HEPPM *and suppose that* $c_k > \overline{c} > 0$ *and* $\sigma_k < \overline{\sigma} \in [0, 1)$. *The following holds:*
(i) If the solution set of (1) *is nonempty, the sequence* $\{z^k\}$ *converges weakly to a solution.*
(ii) If, in addition, T_A^{-1} *is Lipschitz continuous at zero, the sequence converges with linear rate.*
(iii) Finally, if (1) *has no solution, the sequence* $\{z^k\}$ *is unbounded.* □

The proof of item (iii) can be done following [17, Corollary 4.2, item 1].

As observed in [22], the condition for linear convergence can be interpreted as follows. Given two parameters $(a, b) \in A \times A^\perp$, consider the *perturbed* version of (1):

$$\text{Find } (x + b, a + y \in T(x + b)) \text{ such that } (x, y) \in A \times A^\perp . \qquad (1)_{a,b}$$

Then T_A^{-1} is Lipschitz continuous at zero if the solution set of $(1)_{a,b}$ varies in a Lipschitz manner with respect to a and b, for small values of the parameters.

4.2 The algorithm SPIM(ε)

In order to define our algorithm, for the intermediate operator S_c we consider the following approximate solutions with error tolerance σ:

Given $(x, y) \in A \times A^\perp$,

$$\left(\tilde{\varepsilon}, \tilde{x}, \tilde{y}\right) \text{ satisfies } \begin{cases} \tilde{y} \in S_c^{\tilde{\varepsilon}}(\tilde{x}) & \text{(a)} \\ \|\tilde{x} + \eta\tilde{y} - (x + \eta y)\|^2 + 2\eta\tilde{\varepsilon} \leq \sigma^2 \left(\|\eta(\tilde{y} - y)\|^2 + \|\tilde{x} - x\|^2\right) & \text{(b).} \end{cases} \qquad (15)$$

Inexact Partial Inverse Method $\mathrm{sPIM}(\varepsilon)$

Step 0. (Initialization) Choose a scaling parameter $\eta > 0$. Take $(x^0, y^0) \in A \times A^\perp$, $c_0 > 0$ and $\sigma_0 \in [0, 1)$. Set $k := 0$.

Step 1. (Inexact Proximal Iteration) Given $(x^k, y^k) \in A \times A^\perp$, c_k and $\sigma_k \in [0, 1)$, find $\left(\tilde{\varepsilon}_k, \tilde{x}^k, \tilde{y}^k\right)$ satisfying (15) written with (x, y, c, σ) therein replaced by $(x^k, y^k, c_k, \sigma_k)$.

Step 2. (Stopping test) If $\tilde{x}^k = x^k$ and $\tilde{y}^k = y^k$ stop.

Step 3. (Update) Define

$$x^{k+1} := x^k - \eta \tilde{y}^k_A \quad \text{and} \quad y^{k+1} := y^k - \frac{1}{\eta} \tilde{x}^k_{A^\perp}. \tag{16}$$

Let $k := k + 1$, update c_k and σ_k, and loop to Step 1. □

Remark 4.2 Approximate solutions are defined so that solutions of exact proximal systems are just approximate solutions with null error tolerance ($\sigma = 0$). Thus, the Partial Inverse Method from [21] is a particular case of $\mathrm{sPIM}(\varepsilon)$, where for each k the values of $\tilde{\varepsilon}_k$, c_k and σ_k are taken, respectively, as 0, 1, and 0.

Along the lines of Proposition 2.1, we relate proximal systems involving operators ηT and S_c, but now solved inexactly.

Proposition 4.3 *Let* $\left(\varepsilon, \tilde{z}, \tilde{w}\right)$ *solve the proximal system* (14) *for a given* $z \in H$. *Then,* $\left(\frac{1}{\eta} c\varepsilon, \tilde{z}_A + c\tilde{w}_{A^\perp}, \frac{1}{\eta}(\tilde{z}_{A^\perp} + c\tilde{w}_A)\right)$ *solves the proximal system* (15) *for* $x := z_A$ *and* $y := \frac{1}{\eta} z_{A^\perp}$.

Reciprocally, let $\left(\tilde{\varepsilon}, \tilde{x}, \tilde{y}\right)$ *solve the proximal system* (15) *for given* $(x, y) \in A \times A^\perp$. *Then* $\left(\eta\tilde{\varepsilon}/c, \tilde{x}_A + \eta\tilde{y}_{A^\perp}, \frac{1}{c}(\tilde{x}_{A^\perp} + \eta\tilde{y}_A)\right)$ *solves the proximal system* (14) *for* $z := x + \eta y$.

Proof. We just prove the first assertion, the reverse relation can be shown using similar arguments. To show satisfaction of (15)(a) we first use (14)(a) and Lemma 3.2(ii) to write

$$\frac{1}{\eta}(\tilde{w}_A + \tilde{z}_{A^\perp}) \in T^{\varepsilon/\eta}(\tilde{z}_A + \tilde{w}_{A^\perp}).$$

Using Lemma 3.2(iii), written with $\tilde{\varepsilon}$ therein replaced by $c\varepsilon/\eta$:

$$\frac{1}{\eta}(c\tilde{w}_A + \tilde{z}_{A^\perp}) \in S_c^{c\varepsilon/\eta}(\tilde{z}_A + c\tilde{w}_{A^\perp}),$$

i.e., the inclusion (15)(a).

To see that (15)(b) holds, call $\tilde{\varepsilon} := \frac{1}{\eta}c\varepsilon$, $\tilde{x} := \tilde{z}_A + c\tilde{w}_{A^\perp}$ and $\tilde{y} := \frac{1}{\eta}(\tilde{z}_{A^\perp} + c\tilde{w}_A)$. Hence, $\tilde{x} + \eta\tilde{y} = \tilde{z} + c\tilde{w}$. Since $x + \eta y = z$, the left hand side term in (15)(b) is

$$\|\tilde{x} + \eta\tilde{y} - (x + \eta y)\|^2 + 2\eta\tilde{\varepsilon} = \|c\tilde{w} + \tilde{z} - z\|^2 + 2c\varepsilon\,.$$

Using now (14)(b)

$$\begin{aligned}
\|c\tilde{w} + \tilde{z} - z\|^2 + 2c\varepsilon &\le \sigma^2(\|c\tilde{w}\|^2 + \|\tilde{z} - z\|^2) \\
&= \sigma^2[(\|c\tilde{w}_A\|^2 + \|\tilde{z}_{A^\perp} - z_{A^\perp}\|^2) + (\|c\tilde{w}_{A^\perp}\|^2 + \|\tilde{z}_A - z_A\|^2)] \\
&= \sigma^2[\|(c\tilde{w}_A + \tilde{z}_{A^\perp}) - z_{A^\perp}\|^2 + \|(c\tilde{w}_{A^\perp} + \tilde{z}_A) - z_A\|^2] \\
&= \sigma^2[\|\eta\tilde{y} - z_{A^\perp}\|^2 + \|\tilde{x} - z_A\|^2] \\
&= \sigma^2[\|\eta(\tilde{y} - y)\|^2 + \|\tilde{x} - x\|^2],
\end{aligned}$$

where we used the equalities $z_A = x$ and $z_{A^\perp} = \eta y$. □

From Proposition 4.3 it is straightforward that one iteration of sPIM(ε) is equivalent to one iteration of HEPPM:

	HEPPM	sPIM(ε)
current iterate	z	$(x, y) \in A \times A^\perp$
inexact proximal system	$\left(\varepsilon, \tilde{z}, \tilde{w}\right)$ in (14)	$\left(\tilde{\varepsilon}, \tilde{x}, \tilde{y}\right)$ in (15)
update	$z^+ = z - c\tilde{w}$	$(x^+, y^+) = (x, y) - (\eta\tilde{y}_A, \frac{1}{\eta}\tilde{x}_{A^\perp})$

Therefore, the following convergence result is also straightforward, from Theorem 4.1:

Theorem 4.4 *Consider* sPIM(ε)*and suppose that* $c_k > \overline{c} > 0$ *and* $\sigma_k < \overline{\sigma} \in [0, 1)$. *The following holds:*

(i) If the solution set of (1) *is nonempty, the sequence* $\{(x^k, y^k)\}$ *converges weakly to a solution.*

(ii) If, in addition, T_A^{-1} *is Lipschitz continuous at zero, the sequence converges with linear rate.*

(iii) Finally, if (1) *has no solution, the sequence is unbounded.* □

The purpose of introducing (15) and the intermediate operator S_c is to shorten complicated expressions involving c and η that appear in (14) when using $(\eta T)_A$, the partial inverse of the scaled operator (see for instance, Lemma 3.2(iii)). In particular, letting $c = 1$ in (14), since $S_1 = T$ by (8), and using Proposition 4.3, approximate solutions from (14) (that use $(\eta T)_A^\varepsilon$) can be interpreted as approximate solutions from (15) (that use only the enlargement T^ε). The particular choice of $c_k = 1$ for all k in sPIM(ε) is used in the algorithm given in the next section.

5 Deriving a splitting method

Splitting methods are adapted for those problems whose particular structure involves some decomposability. Consider for example the problem of finding a zero of a sum of maximal monotone operators,

$$\text{Find } x \in H \text{ such that } 0 \in T_1(x) + T_2(x) + \ldots + T_n(x)\,, \tag{17}$$

where $T_1\,, T_2\,, \cdots\,, T_n$ are maximal monotone operators on H. Instead of applying a proximal point algorithm directly to $T_1 + T_2 + \ldots + T_n$, it can be simpler to devise a method that makes use of each individual T_i, separately.

5.1 The algorithm

Following [21], note that (17) is equivalent to finding x and $y_i \in T_i(x)$ satisfying $y_1 + y_2 + ... + y_n = 0$. Consider the following subspaces in $\mathcal{H}$, the cartesian product of n copies of H:

$$A := \Big\{ X = (x_1, \ldots, x_n) \in \mathcal{H} : x_1 = \ldots = x_n \Big\}$$

and

$$A^\perp := \Big\{ Y = (y_1, \ldots, y_n) \in \mathcal{H} : y_1 + \ldots + y_n = 0 \Big\}\,.$$

For any $Z \in \mathcal{H}$, the $_A$ and $_{A^\perp}$ components have the explicit form

$$Z_A := \left(\frac{1}{n} \sum_{i=1}^n z_i, \ldots, \frac{1}{n} \sum_{i=1}^n z_i \right) \quad \text{and} \quad Z_{A^\perp} := \left(z_1 - \frac{1}{n} \sum_{i=1}^n z_i, \ldots, z_n - \frac{1}{n} \sum_{i=1}^n z_i \right)\,.$$

With the notation above, we see that (17) is just the following special instance of (1), written for $\mathcal{T} := (T_1, T_2, \ldots, T_n)$ and $\mathcal{H}$:

$$\text{Given a subspace } A \text{ in } \mathcal{H}, \text{ find } \Big(X, Y \in \mathcal{T}(X) \Big) \text{ such that } (X, Y) \in A \times A^\perp\,. \tag{18}$$

The application of sPIM(ε) (with $c_k \equiv 1$ and $\eta = 1$) in this particular setting gives the following decomposition method for solving (17):

Inexact Partial Inverse Splitting Method split-sPIM(ε)

Step 0. (Initialization) For $i = 1, \ldots, n$, choose $y_i^0 \in H$ satisfying $y_1^0 + .. + y_n^0 = 0$. Take $x^0 \in H$ and $\sigma_0 \in [0, 1)$. Set $k := 0$.

Step 1. (Splitting) Given $\sigma_k \in [0, 1)$, x^k, and $y_1^k, \ldots, y_n^k$ satisfying $y_1^k + .. + y_n^k = 0$, for $i = 1, \ldots, n$:

Step 1.i (Inexact Proximal Iteration) find $\left(\tilde{\varepsilon}_{i,k}, \tilde{x}_i^k, \tilde{y}_i^k\right)$ such that

$$\begin{cases} \tilde{y}_i^k \in T_i^{\tilde{\varepsilon}_{i,k}}(\tilde{x}_i^k) & \text{(a)} \\ \|\tilde{x}_i^k + \tilde{y}_i^k - (x^k + y_i^k)\|^2 + 2\tilde{\varepsilon}_{i,k} \leq \sigma_k^2 \left(\|\tilde{y}_i^k - y_i^k\|^2 + \|\tilde{x}_i^k - x^k\|^2\right) & \text{(b).} \end{cases} \tag{19}$$

Step 2. (Stopping test) If $\tilde{x}_i^k = x^k$ and $\tilde{y}_i^k = y_i^k$ for all $i = 1, \ldots, n$, stop.

Step 3. (Update) Define

$$x^{k+1} := x^k - \frac{1}{n}\sum_{j=1}^n \tilde{y}_j^k \quad \text{and} \quad y_i^{k+1} := y_i^k - \tilde{x}_i^k + \frac{1}{n}\sum_{j=1}^n \tilde{x}_j^k \text{ for } i = 1, \ldots, n. \tag{20}$$

Let $k := k + 1$, update σ_k, and loop to Step 1. □

When the tolerance error in the inexact proximal systems (19) is set to zero ($\sigma_k \equiv 0$), split-sPIM(ε) becomes the splitting algorithm proposed by Spingarn in [21]. For positive σ_k, split-sPIM(ε) boils down to the Inexact Partial Inverse Method sPIM(ε), writing the iterates in the expanded Hilbert space $\mathcal{H}$:

Proposition 5.1 *Consider algorithm* sPIM(ε)*, written with* $\eta = 1$ *and* $c_k \equiv 1$*, applied to solve* (18)*. Dropping indices* k*, one iteration of split-*sPIM(ε) *corresponds to one iteration of* sPIM(ε)*, according to the following relations:*

*split-*sPIM(ε)	sPIM(ε)
$x\,, y_i$ *such that* $y_1 + .. + y_n = 0$	$(X, Y) := \Big((x, \ldots, x), (y_1, \ldots, y_n)\Big) \in A \times A^\perp$
$\left(\tilde{\varepsilon}_i, \tilde{x}_i, \tilde{y}_i\right)$ *satisfying* (19)	$\begin{pmatrix} \tilde{\varepsilon} := \sum_i \tilde{\varepsilon}_i \\ \tilde{X} := (\tilde{x}_1, \ldots, \tilde{x}_n) \\ \tilde{Y} := (\tilde{y}_1, \ldots, \tilde{y}_n) \end{pmatrix}$ *satisfying* (15) *for* $S_c = \mathcal{T}$

Proof. To see that $(X, Y) \in A \times A^{\perp}$ is straightforward from the definition of X, Y and A. To show that the given values of $(\tilde{\varepsilon}, \tilde{X}, \tilde{Y})$ solve the inexact proximal system (15), we first use the result in Lemma 3.3(i), to show that (15)(a) holds for $S_c = S_1 = \mathcal{T}$. To verify satisfaction of (15)(b), we add the inequalities (19)(b) for $i = 1, \ldots, n$:

$$\sum_i \|\tilde{x}_i^k + \tilde{y}_i^k - (x^k + y_i^k)\|_H^2 + 2\sum_i \tilde{\varepsilon}_{i,k} \leq \sigma_k^2 \left(\sum_i \|\tilde{y}_i^k - y_i^k\|_H^2 + \sum_i \|\tilde{x}_i^k - x^k\|_H^2 \right) .$$

Since $\mathcal{H}$ is endowed with the natural inner product, the square norm is separable: $\|Z\|_{\mathcal{H}}^2 = \sum_i \|z_i\|_H^2$ for any $Z \in \mathcal{H}$. In particular, the last inequality means that

$$\|\tilde{X} + \tilde{Y} - (X + Y)\|_{\mathcal{H}}^2 + 2\tilde{\varepsilon} \leq \sigma^2 \left(\|\tilde{Y} - Y\|_{\mathcal{H}}^2 + \|\tilde{X} - X\|_{\mathcal{H}}^2 \right)$$

and (15)(b) holds. Finally, since

$$\tilde{Y}_A = \Big(\frac{1}{n} \sum_{j=1}^n \tilde{y}_j, \ldots, \frac{1}{n} \sum_{j=1}^n \tilde{y}_j \Big) \quad \text{and} \quad \tilde{X}_{A^\perp} = \Big(\tilde{x}_1 - \frac{1}{n} \sum_{j=1}^n \tilde{x}_j, \ldots, \tilde{x}_n - \frac{1}{n} \sum_{j=1}^n \tilde{x}_j \Big) ,$$

the update (20) in split-sPIM(ε), written in $\mathcal{H}$, is $(X^+, Y^+) = (X, Y) - (\tilde{Y}_A, \tilde{X}_{A^\perp})$, i.e., the update (16) in sPIM(ε) when η is set to one. □

As a result, the convergence properties of split-sPIM(ε) are those established for sPIM(ε) in Theorem 4.4.

We finish with some comments on changing the values of η and c_k. It is possible to define a scaled version of split-sPIM(ε), when in problem (17) the sum is weighted by given $\eta_1, \ldots, \eta_n > 0$: Find $x \in H$ such that $0 \in \eta_1 T_1(x) + \eta_2 T_2(x) + \ldots + \eta_n T_n(x)$. However, it is not possible to take values of c_k different from 1; as already observed in [22] and [5], keeping $c_k \equiv 1$ is critical to the decomposition scheme.

5.2 Fields of application

The problem of finding a zero of a sum of maximal operators provides a general framework for studying decomposition schemes in convex programming and variational inequalities. In order to illustrate the wide range of applicability of this approach, we briefly sketch how some interesting problems can be regarded as special cases of (17):

Convex Programming: for the problem

$$\min_{x \in C \subset \mathbb{R}^n} f_0(x) \quad \text{subject to} \quad f_i(x) \leq 0 \quad i = 1, \ldots, m ,$$

where C is a nonempty closed and convex set and $f_j : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ for $j = 0, \dots, m$ are closed proper and convex. Define the function $F(x, u_1, \dots, u_m) := f_0(x)$ if $x \in C$ and $f_i(x) \leq u_i$; $F(x,u) := +\infty$ otherwise. The problem is then equivalent to minimizing F on the subspace $A := \mathbb{R}^n \times \{0 \in \mathbb{R}^m\}$. The corresponding optimality conditions are

$$(0,0) \in \partial F(x,u) + N_A(x,u) =: T_1(x,u) + T_2(x,u)\,,$$

where $N_A(x,u)$ is the normal cone to A, i.e., $\{0 \in \mathbb{R}^n\} \times \mathbb{R}^m$ in this case. The application of split-sPIM(ε) to the generalized equation above gives a method related to the proximal method of multipliers [15], see also [22]. Note that, since A is a subspace, $N_A(x,u) = A^\perp$, so it is possible (and equivalent) to apply sPIM(ε).

Stochastic programming: In multistage stochastic programming problems, decisions at a given time depend only on the information available at this time. This structure on the information is called *nonanticipativity condition*; it can be formulated by requiring the policies to belong to certain linear subspace $\mathcal{L}$, [16]. Thus, the optimization problem is again equivalent to minimizing a function (expected cost) on a subspace. The progressive hedging method of [14] is a splitting algorithm applied to solve the optimality conditions of this problem: in (17) T_1 is the subdifferential of the expected cost function, and T_2 the normal cone to $\mathcal{L}$, i.e., $\mathcal{L}^\perp$.

Convex feasibility problems: The problem of finding a point in the intersection of n convex sets C_i can be formulated as (17), where each $T_i(x) = N_{C_i}(x)$. The application of split-sPIM(ε) gives a relaxation-type method that uses primal-dual information.

Monotone complementarity and variational problems: The problem of finding nonnegative orthogonal x and y such that and $y \in T(x)$ can be written as finding the zero of sum of $T_1(x) := T(x)$ and the normal cone defined by $T_2(x) := \{y \leq 0 : \langle y, x\rangle = 0\}$ if $x \geq 0$; and $T_2(x) := \emptyset$ otherwise. More general variational problems arise when the condition "nonnegative and orthogonal x and y" is replaced by $x \in C$, a closed convex set, in which case $T_2(x) := N_C(x)$ if $x \in C$; and $T_2(x) := \emptyset$ otherwise.

Incidentally, note that many of the examples above are formulated as the minimization of a function on a subspace. The generalized equations given by the corresponding optimality conditions have the form (1) and they constitute the core of applications of partial inverse methods. Another important example of the same kind of problem arises when coupling constraints are decomposed by dualization and/or by duplicating variables. Along these lines, it is worth mentioning the "Epsilon-Proximal Decomposition Method" introduced in [12], with an application to multicommodity flow problems in telecommunications. This method is a special case of sPIM(ε)(for which $c = 1$ and $T = \partial f$), with an explicit resolution technique of approximate proximal systems, based on bundle methods.

An advantage of the general framework (17) it that it allows the unified analysis of methods that were developed independently, for various types of problems.

5.3 Comparison with other decomposition methods

There is great variety of monotone splitting methods, forward- and backward- backward methods, extended Rachford methods, Partial Inverse methods, decomposition methods based on the Auxiliary Problem Principle. We refer to [13], [8], [4], [23], [5], [9], [24], mentioning that this is just a sample list of references.

We focus here on the extended Rachford class, see [6], comprising both Peaceman-Rachford and Douglas-Rachford methods.

The relation between the Partial Inverse Method and the Douglas-Rachford splitting operator was explained in [5] and [9]. Specifically, Douglas-Rachford splitting algorithm uses a (rather involved) proximal operator T, derived from T_1 and T_2 in (17). When $T_1 = N_A$ and $T_2 = T$, such proximal operator is the resolvent of the scaled partial inverse. As a result, the partial inverse method of Spingarn PIM applied to (17) is a particular Rachford method.

Inexact versions of Rachford methods only allow for residual error terms that form a finite series, like in [15]. Being a proximal iteration on T, one may attempt to define inexact Rachford methods based on HEPPM. Unfortunately, it does not seems straightforward to relate the enlargement of T with T_1^ε and T_2^ε in order to have also decomposability in the enlargement.

We finish with some comments comparing Rachford methods with split-sPIM(ε), to solve a particular case of monotone complementarity problems. Namely, the affine variational inequality problem, where $T(x)$ is single valued, $T(x) = Mx + q$ for M a positive semidefinite matrix and $q \in H := \mathbb{R}^n$:

$$\text{Find } x \in C \text{ such that } \langle Mx + q, w - x\rangle \geq 0 \text{ for all } w \in C\,.$$

As we already mentioned, a possible splitting is

$$T_1(x) := \{Mx + q\} \quad \text{and} \quad T_2(x) := N_C(x)\,.$$

Other splitting schemes, basically using partitions of the matrix M, can be found in [6, §2.2]. In order to apply split-sPIM(ε), we need to characterize elements in the enlargement T_1^ε and T_2^ε. It is shown in [1] that $T_1^\varepsilon = T_1$ for all $\varepsilon \geq 0$. As for the enlargement of normal cones, it holds that

$$T_2^\varepsilon(x) = N_C^\varepsilon(x) := \{y \in \mathbb{R}^n : \langle y, w - x \rangle \leq \varepsilon \text{ for all } w \in C\} ,$$

the ε-normal set to C, see [7, Def. XI.1.1.3].

Beware that the ε-normal set need no longer be a cone. For example, for linear complementarity problems, C is the positive orthant $\mathbb{R}^n_+$ and

$$N^\varepsilon_{\mathbb{R}^n_+}(x) = \left\{ y \in \mathbb{R}^n_- : \text{ if } x_j > 0 \text{ then } y_j \geq -\frac{\varepsilon}{x_j} \right\} .$$

Altogether, split-sPIM(ε) entails the following computations:
Initialize the algorithm for example with $x \in C$, $y_1 = Mx + q$, $y_2 = -y_1$. Dropping indices k, the splitting step consists of:
Step 1.1: Define $\tilde{x}_1$ as the solution of the system $(I + M)\tilde{x}_1 = x + y_1 - q$, where I is the identity matrix. Then $\tilde{x}_1 + \tilde{y}_1 = x + y_1$ and $\left(0, \tilde{x}_1, \tilde{y}_1\right)$ satisfies (19)(b) written for $i = 1$.
Step 1.2: Find $\left(\tilde{\varepsilon}_2, \tilde{x}_2, \tilde{y}_2 \in N_{C,\tilde{\varepsilon}_2}(\tilde{x}_2)\right)$ satisfying (19)(b) written for $i = 2$.

The update (20) is $x^+ := x - \frac{1}{2}(\tilde{y}_1 + \tilde{y}_2)$, $y_1^+ := y_1 + \frac{1}{2}(\tilde{x}_2 - \tilde{x}_1)$, and $y_2^+ := -y_1^+$.

The rough scheme described above for sPIM(ε) is far from the detailed parallel implementation of Rachford splitting in [6]. Nonetheless, we may still compare some properties of both approaches. From the theoretical point of view, split-sPIM(ε) converges giving both the primal and dual optimal values ($x^\star$ and $y_1^\star = -y_2^\star$ respectively). Rachford splitting provides only the primal optimal value $x^\star$ (cf. [6, Prop. 6]). When one of the operators is single-valued, this is not a crucial difference (for our example the computation of the optimal dual value is straightforward: $y_1^\star := Mx^\star + q$). However, when both $T_1(x^\star)$ and $T_2(x^\star)$ are multi-valued, it can be useful to know the dual value that solves the generalized equation. In addition, as opposed to split-sPIM(ε), Rachford splitting only allows inexact computations, provided the accuracy improves sufficiently along iterations (cf. [6, Prop. 7]).

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, C. A. Sagastizábal, and B. F. Svaiter. ϵ-enlargements of maximal monotone operators: theory and applications. In *Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods (Lausanne, 1997)*, pages 25–43. Kluwer Acad. Publ., Dordrecht, 1999.

[3] R.S. Burachik, C.A. Sagastizábal, and B. F. Svaiter. Bundle methods for maximal monotone operators. In R. Tichatschke and M. Théra, editors, *Ill-posed Variational Problems and Regularization Techniques*, number 477 in Lecture Notes in Economics and Mathematical Systems, pages 49–64. Springer-Verlag Berlin Heidelberg, 1999.

[4] J.-C. Culioli and G. Cohen. Decomposition/coordination algorithms in stochastic optimization. *SIAM J. Control Optim.*, 28(6):1372–1403, 1990.

[5] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Programming*, 55(3, Ser. A):293–318, 1992.

[6] J. Eckstein and M. C. Ferris. Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control. *INFORMS J. Comput.*, 10(2), 1998.

[7] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Number 305-306 in Grund. der math. Wiss. Springer-Verlag, 1993. (two volumes).

[8] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16(6):964–979, 1979.

[9] P. Mahey, S. Oualibouch, and P. D. Tao. Proximal decomposition on the graph of a maximal monotone operator. *SIAM J. Optim.*, 5(2):454–466, 1995.

[10] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Rev. Française Informat. Recherche Opérationnelle*, 4(Ser. R-3):154–158, 1970.

[11] J.J. Moreau. Proximité et dualité dans un espace Hilbertien. *Bulletin de la Société Mathématique de France*, 93:273–299, 1965.

[12] A. Ouorou. Epsilon-proximal decomposition method. Technical report, France Telecom R&D, 2000. Submitted to Mathematical Programming.

[13] G. B. Passty. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.*, 72(2):383–390, 1979.

[14] R. T. Rockafellar and R. J-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16, 1991.

[15] R.T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research*, 1:97–116, 1976.

[16] A. Ruszczyński. Decomposition methods in stochastic programming. In *Mathematical Programming*, volume 79, 1997.

[17] 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.

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

[19] M. V. Solodov and B. F. Svaiter. An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions. *Math. Oper. Res.*, 25(2):214–230, 2000.

[20] M. V. Solodov and B. F. Svaiter. A unified framework for some inexact proximal point algorithms. *Numerical Functional Analysis and Optimization*, 22,1013–1025, 2001.

[21] J. E. Spingarn. Partial inverse of a monotone operator. *Appl. Math. Optim.*, 10(3):247–265, 1983.

[22] J. E. Spingarn. Applications of the method of partial inverses to convex programming: decomposition. *Math. Programming*, 32(2):199–223, 1985.

[23] P. Tseng. Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. *SIAM J. Control Optim.*, 29(1):119–138, 1991.

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