Noname manuscript No.
(will be inserted by the editor)

An Inexact Spingarn's Partial Inverse Method with Applications to Operator Splitting and Composite Optimization

S. Costa Lima · M. Marques Alves

Received: date / Accepted: date

Abstract We propose and study the iteration-complexity of an inexact version of the Spingarn's partial inverse method. Its complexity analysis is performed by viewing it in the framework of the hybrid proximal extragradient (HPE) method, for which pointwise and ergodic iteration-complexity has been established recently by Monteiro and Svaiter. As applications, we propose and analyze the iteration-complexity of an inexact operator splitting algorithm – which generalizes the original Spingarn's splitting method – and of a parallel forward-backward algorithm for multi-term composite convex optimization.

Keywords Inexact proximal point methods · Partial inverse method · Splitting · Composite optimization · Forward-backward · Parallel · Iteration-complexity.

Mathematics Subject Classification (2000) 47H05 · 47J20 · 90C060 · 90C33 · 65K10

1 Introduction

In [1], J. E. Spingarn proposed and analyzed a proximal point type method – called the partial inverse method – for solving the problem of finding a point in the graph of a maximal monotone operator such that the first (primal) variable belongs to a closed subspace and the second (dual) variable belongs to its orthogonal complement. This problem encompasses minimization of convex functions over closed subspaces and inclusion problems given by the sum of finitely many maximal monotone operators. Regarding the latter case, Spingarn also derived an operator splitting method with the distinctive feature of allowing parallel implementations. Spingarn's approach for solving the above mentioned problem consists in recasting it as an inclusion problem for the partial inverse (a concept coined by himself) of the monotone operator involved in the formulation of the problem with respect to the closed subspace. That said, Spingarn's partial inverse method essentially consists of Rockafellar's proximal point method (PPM) applied to this monotone inclusion, which converges either under the assumption of exact computation of the resolvent or under summable error criterion [2]. The hybrid proximal extragradient (HPE) method of Solodov and Svaiter [3] is an inexact version of the Rockafellar's PPM which uses relative error tolerance criterion for solving each proximal subproblem instead of summable error condition. The HPE method has been used for many authors [3–14] as a framework for the design and analysis of several algorithms for monotone inclusion problems, variational inequalities, saddle-point problems and convex optimization. Its iteration-complexity has been established recently by Monteiro and Svaiter [15] and, as a consequence, it has proved the iteration-complexity of various important algorithms in optimization (which use the HPE method as a framework) including Tseng's forward-backward method, Korpelevich extragradient method and the alternating direction method of multipliers (ADMM) [12, 15, 16].

S. Costa Lima
Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Brazil, 88040-900.
Tel.: +55 48 37216560
E-mail: samaraclim@outlook.com

M. Marques Alves (corresponding author)
Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Brazil, 88040-900.
Tel.: +55 48 37213678
Fax: +55 48 37216560
E-mail: maicon.alves@ufsc.br

In this paper, we propose and analyze the iteration-complexity of an inexact version of the Spingarn's partial inverse method in the light of the recent developments in the iteration-complexity of the HPE method. We introduce a notion of approximate solution of the above mentioned Spingarn's problem and prove that our proposed method can be regarded as a special instance of the HPE method applied to this problem and, as a consequence, we obtain pointwise and ergodic iteration-complexity results for our inexact partial inverse method. As applications, we propose and study the iteration-complexity of an inexact operator splitting method for solving monotone inclusions with the sum of finitely many maximal monotone operators as well as of a parallel forward-backward algorithm for multi-term composite convex optimization. We also briefly discuss how a different inexact version of the Spingarn's partial inverse method proposed and studied in [17] is related to our method.

Contents. Section 2 contains two subsections. Subsection 2.1 presents some general results and the basic notation we need in this paper. Subsection 2.2 is devoted to present the iteration-complexity of the HPE method and to briefly discuss the method of [17]. Section 3 presents our main algorithms and its iteration-complexity. Finally, in Section 4 we show how the results of Section 3 can be used to derive an operator splitting method and a parallel forward-backward method for solving multi-term composite convex optimization.

2 Background Materials and Notation

This section contains two subsections. In Subsection 2.1 we present the general notation as well as some basic facts about maximal monotone operators and convex analysis. In Subsection 2.2 we review some important facts about the iteration-complexity of the hybrid proximal extragradient (HPE) method and study some properties of a variant of it.

2.1 General Results and Notation

We denote by $\mathcal{H}$ a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$. For $m \geq 2$, the Hilbert space $\mathcal{H}^m := \mathcal{H} \times \mathcal{H} \times \cdots \times \mathcal{H}$ will be endowed with the inner product $\langle(x_1,\ldots,x_m),(x'_1,\ldots,x'_m)\rangle := \sum_{i=1}^m \langle x_i, x'_i\rangle$ and norm $\|\cdot\| := \sqrt{\langle\cdot,\cdot\rangle}$.

For a set-valued map $S : \mathcal{H} \rightrightarrows \mathcal{H}$, its *graph* and *domain* are taken respectively as $\mathrm{G}(S) = \{(x,v) \in \mathcal{H}\times\mathcal{H} \,:\, v \in S(x)\}$ and $\mathrm{D}(S) = \{x \in \mathcal{H} \,:\, S(x) \neq \emptyset\}$. The *inverse* of S is $S^{-1} : \mathcal{H} \rightrightarrows \mathcal{H}$ such that $v \in S(x)$ if and only if $x \in S^{-1}(v)$. Given $S, S' : \mathcal{H} \rightrightarrows \mathcal{H}$ and $\lambda > 0$ we define $S + S' : \mathcal{H} \rightrightarrows \mathcal{H}$ and $\lambda S : \mathcal{H} \rightrightarrows \mathcal{H}$ by $(S+S')(x) = S(x) + S'(x)$ and $(\lambda S)(x) = \lambda S(x)$ for all $x \in \mathcal{H}$, respectively. Given set-valued maps $S_i : \mathcal{H} \rightrightarrows \mathcal{H}$, for $i = 1,\ldots,m$, we define its product by

$$S_1 \times S_2 \times \cdots \times S_m : \mathcal{H}^m \rightrightarrows \mathcal{H}^m, \quad (x_1,x_2,\ldots,x_m) \mapsto S_1(x_1)\times S_2(x_2)\times\cdots\times S_m(x_m). \tag{1}$$

An operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is *monotone* if

$$\langle v - v', x - x'\rangle \geq 0 \quad \text{whenever} \quad (x,v),(x',v') \in \mathrm{G}(T).$$

It is *maximal monotone* if it is monotone and maximal in the following sense: if $S : \mathcal{H} \rightrightarrows \mathcal{H}$ is monotone and $\mathrm{G}(T) \subset \mathrm{G}(S)$, then $T = S$. The *resolvent* of a maximal monotone operator T is $(T+I)^{-1}$, and $\tilde{z} = (T+I)^{-1}z$ if and only if $z - \tilde{z} \in T(\tilde{z})$. For $T : \mathcal{H} \rightrightarrows \mathcal{H}$ maximal monotone and $\varepsilon \geq 0$, the ε-enlargement of T [18,19] is the operator $T^{\varepsilon} : \mathcal{H} \rightrightarrows \mathcal{H}$ defined by

$$T^{\varepsilon}(x) := \{v \in \mathcal{H} \;\;:\;\; \langle v' - v, x' - x\rangle \geq -\varepsilon \;\; \forall (x',v') \in \mathrm{G}(T)\} \quad \forall x \in \mathcal{H}. \tag{2}$$

Note that $T(x) \subset T^{\varepsilon}(x)$ for all $x \in \mathcal{H}$.

The following summarizes some useful properties of T^{ε} which will be useful in this paper.

Proposition 2.1 *Let $T, S : \mathcal{H} \rightrightarrows \mathcal{H}$ be set-valued maps. Then,*

(a) *if $\varepsilon \leq \varepsilon'$, then $T^{\varepsilon}(x) \subseteq T^{\varepsilon'}(x)$ for every $x \in \mathcal{H}$;*
(b) *$T^{\varepsilon}(x) + S^{\varepsilon'}(x) \subseteq (T+S)^{\varepsilon+\varepsilon'}(x)$ for every $x \in \mathcal{H}$ and $\varepsilon, \varepsilon' \geq 0$;*
(c) *T is monotone if, and only if, $T \subseteq T^0$;*
(d) *T is maximal monotone if, and only if, $T = T^0$;*
(e) *if $f : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ is proper, convex and closed, then $\partial_{\varepsilon} f(x) \subseteq (\partial f)^{\varepsilon}(x)$ for any $\varepsilon \geq 0$ and $x \in \mathcal{H}$.*

Throughout this work we adopt standard notation of convex analysis for subdiferentials, ε-subdiferentials, etc. Moreover, for a closed subspace $V \subseteq \mathcal{H}$ we denote by $V^{\perp}$ its *orthogonal complement* and by P_V and $P_{V^{\perp}}$ the *orthogonal projectors* onto V and $V^{\perp}$, respectively. The *Spingarn's partial inverse* [1] of a set-valued map $S : \mathcal{H} \rightrightarrows \mathcal{H}$ with respect to a closed subspace V of $\mathcal{H}$ is the set-valued operator $S_V : \mathcal{H} \rightrightarrows \mathcal{H}$ whose graph is

$$\mathrm{G}(S_V) := \{(z, v) \in \mathcal{H} \times \mathcal{H} \;:\; P_V(v) + P_{V^{\perp}}(z) \in S(P_V(z) + P_{V^{\perp}}(v))\}. \tag{3}$$

The following lemma will be important for us.

Lemma 2.1 ([17, Lemma 3.1]) *Let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator, $V \subset \mathcal{H}$ a closed subspace and $\varepsilon > 0$. Then,*

$$(T_V)^{\varepsilon} = (T^{\varepsilon})_V.$$

Next we present the transportation formula for ε-enlargements.

Theorem 2.1 ([20, Theorem 2.3]) *Suppose $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone and let $x_\ell, u_\ell \in \mathcal{H}$, $\varepsilon_\ell, \alpha_\ell \in \mathbb{R}_+$, for $\ell = 1, \ldots, k$, be such that*

$$u_\ell \in T^{\varepsilon_\ell}(x_\ell), \quad \ell = 1, \ldots, k, \quad \sum_{\ell=1}^{k} \alpha_\ell = 1,$$

and define

$$x^a := \sum_{\ell=1}^{k} \alpha_\ell \, x_\ell \,, \quad u^a := \sum_{\ell=1}^{k} \alpha_\ell \, u_\ell \,, \quad \varepsilon^a := \sum_{\ell=1}^{k} \alpha_\ell \left[\varepsilon_\ell + \langle x_\ell - x^a, u_\ell - u^a \rangle\right].$$

Then, the following statements hold:

(a) $\varepsilon^a \geq 0$ *and* $u^a \in T^{\varepsilon^a}(x^a)$.
(b) *If, in addition, $T = \partial f$ for some proper, convex and closed function f and $u_\ell \in \partial_{\varepsilon_\ell} f(x_\ell)$ for $\ell = 1, \ldots, k$, then $u^a \in \partial_{\varepsilon^a} f(x^a)$.*

The following results will also be useful in this work.

Lemma 2.2 ([21, Lemma 3.2]) *If $f : \mathcal{H} \to \overline{\mathbb{R}}$ is a proper, closed and convex function and $x, \tilde{x}, v \in \mathcal{H}$ are such that $v \in \partial f(x)$ and $f(\tilde{x}) < \infty$, then $v \in \partial_\varepsilon f(\tilde{x})$ for every $\varepsilon \geq f(\tilde{x}) - f(x) - \langle v, \tilde{x} - x \rangle$.*

Lemma 2.3 *Let $f : \mathcal{H} \to \overline{\mathbb{R}}$ be proper, closed and convex. Then, the following holds for every $\lambda, \varepsilon > 0$:*

(a) $\partial(\lambda f) = \lambda \partial f$;
(b) $\partial_\varepsilon(\lambda f) = \lambda \partial_{\varepsilon/\lambda} f$.

Lemma 2.4 ([22, Lemmas 1.2.3]) *Let $f : \mathcal{H} \to \mathbb{R}$ be convex and continuously differentiable such that there exists a nonnegative constant L satisfying*

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y \in \mathcal{H}.$$

Then,

$$0 \leq f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq \frac{L}{2}\|x - y\|^2 \tag{4}$$

for all $x, \tilde{x} \in \mathcal{H}$.

2.2 On the Hybrid Proximal Extragradient Method

In this subsection we consider the *monotone inclusion problem*

$$0 \in T(z) \tag{5}$$

where $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is a point-to-set maximal monotone operator. We also assume that the solution set $T^{-1}(0)$ of (5) is nonempty. Since problem (5) appears in different fields of applied mathematics including optimization, equilibrium theory and partial differential equations, it is desirable to have efficient numerical schemes to find approximate solutions of it.

An *exact proximal point method* (PPM) iteration for (5) is

$$z_k = (\lambda_k T + I)^{-1} z_{k-1} \tag{6}$$

where z_{k-1} and z_k are the current and new iterate, respectively, and $\lambda_k > 0$ is a sequence of stepsizes. The practical applicability of proximal point algorithms to concrete problems depends on the availability of inexact versions of such methods. In the seminal work [2], Rockafellar proved that if z_k is computed satisfying

$$\|z_k - (\lambda_k T + I)^{-1} z_{k-1}\| \leq \eta_k, \quad \sum_{k=1}^{\infty} \eta_k < \infty, \tag{7}$$

and $\{\lambda_k\}$ is bounded away from zero, then $\{z_k\}$ converges weakly to a solution of (5) – assuming that there exists at least one of them. New inexact versions of the PPM which use *relative error tolerance* to compute approximate solutions have been proposed and intensively studied in the last two decades [3, 12–15, 23, 24]. The key idea behind such methods [3] consists in decoupling (6) as an inclusion and an equation:

$$v_k \in T(z_k), \quad \lambda_k v_k + z_k - z_{k-1} = 0 \tag{8}$$

and in relaxing both of them according to relative error tolerance criteria. The *hybrid proximal extragradient* (HPE) method of [3] has been used in the last few years as a framework for the analysis and development of several algorithms for solving monotone inclusion, saddle-point and convex optimization problems [3–14].

Next we present the HPE method.

Algorithm 1 Hybrid proximal extragradient (HPE) method for (5)

(0) Let $z_0 \in \mathcal{H}$ and $\sigma \in [0, 1[$ be given and set $k = 1$.
(1) Compute $(\tilde{z}_k, v_k, \varepsilon_k) \in \mathcal{H} \times \mathcal{H} \times \mathbb{R}_+$ and $\lambda_k > 0$ such that

$$v_k \in T^{\varepsilon_k}(\tilde{z}_k), \quad \|\lambda_k v_k + \tilde{z}_k - z_{k-1}\|^2 + 2\lambda_k \varepsilon_k \leq \sigma^2 \|\tilde{z}_k - z_{k-1}\|^2. \tag{9}$$

(2) Define

$$z_k = z_{k-1} - \lambda_k v_k, \tag{10}$$

set $k \leftarrow k + 1$ and go to step 1.

Remarks. 1) First note that condition (9) relaxes both the inclusion and the equation in (8). Here, $T^{\varepsilon}(\cdot)$ is the ε-enlargement of T; it has the property that $T(z) \subset T^{\varepsilon}(z)$ (see Subsection 2.1 for details). 2) Instead of $\tilde{z}_k$, the next iterate z_k is defined in (10) as an extragradient step from z_{k-1}. 3) Letting $\sigma = 0$ and using Proposition 2.1(d) we conclude from (9) and (10) that (z_k, v_k) and $\lambda_k > 0$ satisfy (8), i.e., Algorithm 1 is an inexact version of the exact Rockafellar's PPM. 4) Algorithm 1 serves also as a framework for the analysis and development of several numerical schemes for solving concrete instances of (5) (see, e.g., [6, 10–12, 15, 23]); specific strategies for computing $(\tilde{z}_k, v_k, \varepsilon_k)$ and $\lambda_k > 0$ satisfying (9) depends on the particular instance of (5) under consideration.

In the last few years, starting with the paper [15], a lot of research has been done to study and analyze the *iteration-complexity* of the HPE method and its special instances, including Tseng's forward-backward splitting method, Korpelevich extragradient method, ADMM, etc [12, 15, 23]. These iteration-complexity bounds for the HPE method are based on the following termination criterion introduced in [15]: for given tolerances $\rho, \epsilon > 0$, find $\bar{z}, \bar{v} \in \mathcal{H}$ and $\bar{\varepsilon} > 0$ such that $(z, v) := (\bar{z}, \bar{v})$ and $\varepsilon := \bar{\varepsilon}$ satisfy

$$v \in T^{\varepsilon}(z), \quad \|v\| \leq \rho, \quad \varepsilon \leq \epsilon. \tag{11}$$

Using Proposition 2.1(d) we find that if $\rho = \epsilon = 0$ in (11) then $0 \in T(\bar{z})$, i.e., $\bar{z}$ is a solution of (5).

Next we summarize the main results from [15] about *pointwise* and *ergodic* iteration-complexity of the HPE method that we will need in this paper. The *aggregate stepsize sequence* $\{\Lambda_k\}$ and the *ergodic sequences* $\{\tilde{z}_k^a\}$, $\{\tilde{v}_k^a\}$, $\{\varepsilon_k^a\}$ associated to $\{\lambda_k\}$ and $\{\tilde{z}_k\}$, $\{v_k\}$, and $\{\varepsilon_k\}$ are, respectively,

$$\begin{aligned} \Lambda_k &:= \sum_{\ell=1}^{k} \lambda_\ell \,, \\ \tilde{z}_k^a &:= \frac{1}{\Lambda_k} \sum_{\ell=1}^{k} \lambda_\ell \, \tilde{z}_\ell, \quad v_k^a := \frac{1}{\Lambda_k} \sum_{\ell=1}^{k} \lambda_\ell \, v_\ell, \\ \varepsilon_k^a &:= \frac{1}{\Lambda_k} \sum_{\ell=1}^{k} \lambda_\ell (\varepsilon_\ell + \langle \tilde{z}_\ell - \tilde{z}_k^a, v_\ell - v_k^a \rangle). \end{aligned} \tag{12}$$

Theorem 2.2 ([15, Theorem 4.4(a) and 4.7]) *Let $\{\tilde{z}_k\}$, $\{v_k\}$, etc be generated by* Algorithm 1 *and let $\{\tilde{z}_k^a\}$, $\{v_k^a\}$, etc be given in* (12)*. Let also d_0 denote the distance of z_0 to $T^{-1}(0) \neq \emptyset$ and assume that $\underline{\lambda} := \inf \lambda_k > 0$. The following statements hold.*

(a) *For any $k \geq 1$, there exists $i \in \{1, \ldots, k\}$ such that*

$$v_i \in T^{\varepsilon_i}(\tilde{z}_i), \quad \|v_i\| \leq \frac{d_0}{\underline{\lambda}\sqrt{k}} \sqrt{\frac{1+\sigma}{1-\sigma}}, \quad \varepsilon_i \leq \frac{\sigma^2 d_0^2}{2(1-\sigma^2)\underline{\lambda} k} \;;$$

(b) *for any $k \geq 1$,*

$$v_k^a \in T^{\varepsilon_k^a}(\tilde{z}_k^a), \quad \|v_k^a\| \leq \frac{2d_0}{\underline{\lambda} k}, \quad \varepsilon_k^a \leq \frac{2(1 + \sigma/\sqrt{1-\sigma^2}) d_0^2}{\underline{\lambda} k} \,.$$

Remark 2.1 The bounds given in (a) and (b) of Theorem 2.2 are called pointwise and ergodic bounds, respectively. Items (a) and (b) can be used, respectively, to prove that for given tolerances $\rho, \epsilon > 0$ the termination criterion (11) is satisfied in at most

$$\mathcal{O}\left(\max\left\{\left\lceil \frac{d_0^2}{\underline{\lambda}^2 \rho^2} \right\rceil, \left\lceil \frac{d_0^2}{\underline{\lambda} \epsilon} \right\rceil\right\}\right) \text{ and } \mathcal{O}\left(\max\left\{\left\lceil \frac{d_0}{\underline{\lambda} \rho} \right\rceil, \left\lceil \frac{d_0^2}{\underline{\lambda} \epsilon} \right\rceil\right\}\right)$$

iterations, respectively.

The following variant of Algorithm 1 studied in [17] is related to the results of this paper: Let $z_0 \in \mathcal{H}$ and $\hat{\sigma} \in [0, 1[$ be given and iterate for $k \geq 1$,

$$\begin{cases} v_k \in T^{\varepsilon_k}(\tilde{z}_k), \quad \|\lambda_k v_k + \tilde{z}_k - z_{k-1}\|^2 + 2\lambda_k \varepsilon_k \leq \hat{\sigma}^2 \left(\|\tilde{z}_k - z_{k-1}\|^2 + \|\lambda_k v_k\|^2\right), \\ z_k = z_{k-1} - \lambda_k v_k. \end{cases} \tag{13}$$

Remark 2.2 The inequality in (13) is a relative error tolerance proposed in [14] (for a different method); the identity in (13) is the same extragradient step of Algorithm 1. Hence, the method described in (13) can be interpreted as a HPE variant in which a different relative error tolerance is considered in the solution of each subproblem. In what follows in this section we will show that (13) is actually a special instance of Algorithm 1 whenever $\hat{\sigma} \in [0, 1/\sqrt{5}\,[$ and that it may fail to converge if we take $\hat{\sigma} > 1/\sqrt{5}$.

Lemma 2.5 ([14, Lemma 2]) *Suppose $\{z_k\}$, $\{\tilde{z}_k\}$, $\{v_k\}$ and $\{\lambda_k\}$ satisfy the inequality in* (13)*. Then, for every $k \geq 1$,*

$$\frac{1-\theta}{1-\hat{\sigma}^2}\|\tilde{z}_k - z_{k-1}\| \leq \|\lambda_k v_k\| \leq \frac{1+\theta}{1-\hat{\sigma}^2}\|\tilde{z}_k - z_{k-1}\|$$

where

$$\theta := \sqrt{1 - (1-\hat{\sigma}^2)^2}. \tag{14}$$

Proof From the inequality in (13) and the Cauchy-Schwarz inequality we obtain

$$\left(1-\hat{\sigma}^2\right)\|\lambda_k v_k\|^2 - 2\|\tilde{z}_k - z_{k-1}\|\|\lambda_k v_k\| + \left(1-\hat{\sigma}^2\right)\|\tilde{z}_k - z_{k-1}\|^2 \leq 0, \qquad \forall k \geq 1.$$

To finish the proof of the lemma note that (in the above inequality) we have a quadratic function in the term $\|\lambda_k v_k\|$. □

Proposition 2.2 *Let $\{z_k\}$, $\{\tilde{z}_k\}$, $\{v_k\}$, $\{\varepsilon_k\}$ and $\{\lambda_k\}$ be given in* (13) *and assume that $\hat{\sigma} \in [0, 1/\sqrt{5}\,[$. Define, for all $k \geq 1$,*

$$\sigma := \hat{\sigma}\sqrt{1 + \left(\frac{1+\theta}{1-\hat{\sigma}^2}\right)^2}, \tag{15}$$

where $0 \leq \theta < 1$ is given in (14). *Then, $\sigma \geq 0$ belongs to $[0,1[$ and $z_k, \tilde{z}_k, v_k, \varepsilon_k$ and $\lambda_k > 0$ satisfy* (9) *and* (10) *for all $k \geq 1$. As a consequence, the method of* [17] *defined in* (13) *is a special instance of* Algorithm 1 *whenever $\hat{\sigma} \in [0, 1/\sqrt{5}\,[$.*

Proof The assumption $\hat{\sigma} \in [0, 1/\sqrt{5}\,[$, definition (15) and some simple calculations show that $\sigma \in [0,1[$. It follows from (13), (9) and (10) that to finish the proof of the proposition it suffices to prove the inequality in (9). To this end, note that from the second inequality in Lemma 2.5 and (15) we have

$$\begin{aligned}\hat{\sigma}^2\left(\|\tilde{z}_k - z_{k-1}\|^2 + \|\lambda_k v_k\|^2\right) &\leq \hat{\sigma}^2\left(1 + \left(\frac{1+\theta}{1-\hat{\sigma}^2}\right)^2\right)\|\tilde{z}_k - z_{k-1}\|^2 \\ &= \sigma^2\|\tilde{z}_k - z_{k-1}\|^2 \quad \forall k \geq 1,\end{aligned}$$

which in turn gives that the inequality in (9) follows from the one in (13). □

Remark 2.3 Algorithm 1 is obviously a special instance of (13) whenever $\sigma \in [0, 1/\sqrt{5}\,[$ by setting $\hat{\sigma} := \sigma$. Next we will show it is not true in general. Let $T : \mathbb{R} \to \mathbb{R}$ be the maximal monotone operator defined by

$$T(z) := z \quad \forall z \in \mathbb{R}. \tag{16}$$

Assume that $\sigma \in]\sqrt{2/5}, 1[$, take $z_0 = 1$ and define, for all $k \geq 1$,

$$\tilde{z}_k := z_k := \left(1-\sigma^2\right)z_{k-1}, \quad v_k := z_{k-1}, \quad \varepsilon_k := \frac{\sigma^4}{2}|z_{k-1}|^2, \quad \lambda_k := \sigma^2. \tag{17}$$

We will show that $(\tilde{z}_k, v_k, \varepsilon_k)$ and $\lambda_k > 0$ in (17) satisfy (9) but not (13) for any choice of $\hat{\sigma} \in [0, 1/\sqrt{5}\,[$. To this end, we first claim that $v_k \in T^{\varepsilon_k}(\tilde{z}_k)$ for all $k \geq 1$. Indeed, using (16) and (17) we obtain, for all $y \in \mathbb{R}$ and $k \geq 1$,

$$\begin{aligned}(Ty - v_k)(y - \tilde{z}_k) &= (y - z_{k-1})(y - z_{k-1} + \sigma^2 z_{k-1}) \\ &\geq |y - z_{k-1}|^2 - |\sigma^2 z_{k-1}||y - z_{k-1}| \\ &\geq -\frac{|\sigma^2 z_{k-1}|^2}{4} > -\varepsilon_k,\end{aligned}$$

which combined with (2) proves our claim. Moreover, it follows from (17) that

$$\begin{aligned}|\lambda_k v_k + \tilde{z}_k - z_{k-1}|^2 + 2\lambda_k\varepsilon_k &= |\tilde{z}_k - (1-\sigma^2)z_{k-1}|^2 + 2\lambda_k\varepsilon_k \\ &= 2\lambda_k\varepsilon_k \\ &= \sigma^2|\tilde{z}_k - z_{k-1}|^2,\end{aligned} \tag{18}$$

which proves that $(\tilde{z}_k, v_k, \varepsilon_k)$ and $\lambda_k > 0$ satisfy the inequality in (9). The first and second identities in (17) give that they also satisfy (10). Altogether, we have that the iteration defined in (17) is generated by Algorithm 1 for solving (5) with T given in (16). On the other hand, it follows from (17) and the assumption $\sigma > \sqrt{2/5}$ that

$$\begin{aligned}\sigma^2|\tilde{z}_k - z_{k-1}|^2 &= \frac{\sigma^2}{2}\left(|\tilde{z}_k - z_{k-1}|^2 + |\lambda_k v_k|^2\right) \\ &> \frac{1}{5}\left(|\tilde{z}_k - z_{k-1}|^2 + |\lambda_k v_k|^2\right).\end{aligned} \tag{19}$$

Hence, it follows from (18) and (19) that the inequality in (13) can not be satisfied for any choice of $\hat{\sigma} \in [0, 1/\sqrt{5}\,[$ and so the sequence given is (17) is generated by Algorithm 1 but it is not generated by the algorithm described in (13).

Remark 2.4 Next we present an example of a monotone inclusion problem for which an instance of (13) may fail to converge if we take $\hat{\sigma} \in]1/\sqrt{5}, 1[$. To this end, consider problem (5) where the maximal monotone operator $T : \mathbb{R} \to \mathbb{R}$ is defined by

$$T(z) := \alpha z \qquad \forall z \in \mathbb{R}, \tag{20}$$

where

$$\alpha := \frac{2\gamma}{\gamma - 2} + 1, \quad \gamma := \frac{1+\theta}{1-\hat{\sigma}^2}. \tag{21}$$

($\theta > 0$ is defined in (14).) Assuming $\hat{\sigma} \in]1/\sqrt{5}, 1[$ we obtain $5\hat{\sigma}^4 - 6\hat{\sigma}^2 + 1 < 0$, which is clearly equivalent to $\theta > |1 - 2\hat{\sigma}^2|$. Using (21) and the latter inequality we conclude that

$$\gamma > 2, \quad \frac{\alpha\gamma}{\alpha + \gamma} > 2. \tag{22}$$

Now take $z_0 = 1$ and define, for all $k \geq 1$,

$$(\tilde{z}_k, v_k, \varepsilon_k) := \left(\frac{\gamma}{\alpha+\gamma} z_{k-1}, T(\tilde{z}_k), 0\right), \quad \lambda_k := 1, \quad z_k := z_{k-1} - \lambda_k v_k. \tag{23}$$

Direct calculation yields, for all $k \geq 1$,

$$|v_k + \tilde{z}_k - z_{k-1}|^2 = \hat{\sigma}^2 \left(|\tilde{z}_k - z_{k-1}|^2 + |v_k|^2\right), \tag{24}$$

which, in turn, together with (23) imply (13). Using (20) and (23) we find

$$z_k = \left(1 - \frac{\alpha\gamma}{\alpha+\gamma}\right)^k, \quad \forall k \geq 1. \tag{25}$$

Using the second inequality in (22) and the latter identity we easily conclude that $|z_k| \to \infty$ as $k \to \infty$ and so $\{z_k\}$ does not converge to the unique solution $\bar{z} = 0$ of (5).

3 An Inexact Spingarn's Partial Inverse Method

In this section we consider the problem of finding $x, u \in \mathcal{H}$ such that

$$x \in V, \quad u \in V^{\perp} \text{ and } u \in T(x) \tag{26}$$

where $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone and V is a closed subspace of $\mathcal{H}$. We define the solution set of (26) by

$$\mathcal{S}^*(V, T) := \{z \in \mathcal{H} \ : \ \text{there exist } x, u \in \mathcal{H} \text{ satisfying (26) such that } z = x + u\} \tag{27}$$

and assume it is nonempty. Problem (26) encompasses important problems in applied mathematics including minimization of convex functions over closed subspaces, splitting methods for the sum of finitely many maximal monotone operators and decomposition methods in convex optimization [1,17,25, 26]. One of our main goals in this paper is to propose and analyze an inexact partial inverse method for solving (26) in the light of recent developments in the iteration-complexity theory of the HPE method [3, 15], as discussed in Section 2.2. We will show, in particular, that the method proposed in this section generalizes the inexact versions of the Spingarn's partial inverse method for solving (26) proposed in [26] and [17]. The main results of iteration-complexity to find approximate solutions are achieved by analyzing the proposed method in the framework of the HPE method (Algorithm 1).

Regarding the results of iteration-complexity, we will consider the following notion of approximate solution for (26): given tolerances $\rho, \epsilon > 0$, find $\bar{x}, \bar{u} \in \mathcal{H}$ and $\bar{\varepsilon} > 0$ such that $(x, u) = (\bar{x}, \bar{u})$ and $\varepsilon = \bar{\varepsilon}$ satisfy

$$u \in T^{\varepsilon}(x), \quad \max\{\|x - P_V(x)\|, \|u - P_{V^{\perp}}(u)\|\} \leq \rho, \quad \varepsilon \leq \epsilon, \tag{28}$$

where P_V and $P_{V^{\perp}}$ stand for the orthogonal projection onto V and $V^{\perp}$, respectively, and $T^{\varepsilon}(\cdot)$ denotes the ε-enlargement of T (see Section 2.2 for more details on notation). For $\rho = \epsilon = 0$, criterion (28) gives $\bar{x} \in V$, $\bar{u} \in V^{\perp}$ and $\bar{u} \in T(\bar{x})$, i.e., in this case $\bar{x}, \bar{u}$ satisfy (26). Moreover, if $V = \mathcal{H}$ in (26), in which case $P_V = I$ and $P_{V^{\perp}} = 0$, then the criterion (28) coincides with one discussed in Section 2.2 for problem (5) (see (11)).

That said, we next present our inexact version of the Spingarn's partial inverse method for solving (26).

Algorithm 2 An inexact Spingarn's partial inverse method for (26) (I)

(0) Let $x_0 \in \mathcal{H}$ and $\sigma \in [0, 1[$ be given and set $k = 1$.
(1) Compute $(\tilde{x}_k, u_k, \varepsilon_k) \in \mathcal{H} \times \mathcal{H} \times \mathbb{R}_+$ such that

$$u_k \in T^{\varepsilon_k}(\tilde{x}_k), \quad \|u_k + \tilde{x}_k - x_{k-1}\|^2 + 2\varepsilon_k \leq \sigma^2 \|P_V(\tilde{x}_k) + P_{V^\perp}(u_k) - x_{k-1}\|^2. \tag{29}$$

(2) Define

$$x_k = x_{k-1} - \left[P_V(u_k) + P_{V^\perp}(\tilde{x}_k)\right], \tag{30}$$

set $k \leftarrow k + 1$ and go to step 1.

Remarks. 1) Letting $V = \mathcal{H}$ in (26), in which case $P_V = I$ and $P_{V^\perp} = 0$, we obtain that Algorithm 2 coincides with Algorithm 1 with $\lambda_k = 1$ for all $k \geq 1$ for solving (5) (or, equivalently, (26) with $V = \mathcal{H}$). 2) An inexact partial inverse method called *sPIM(ε)* was proposed in [17], Section 4.2, for solving (26). The latter method, with a different notation and scaling factor $\eta = 1$, is given according to the iteration:

$$\begin{cases} u_k \in T^{\varepsilon_k}(\tilde{x}_k), \;\; \|u_k + \tilde{x}_k - x_{k-1}\|^2 + 2\varepsilon_k \leq \hat{\sigma}^2 \left(\|\tilde{x}_k - P_V(x_{k-1})\|^2 + \|u_k - P_{V^\perp}(x_{k-1})\|^2\right), \\ x_k = x_{k-1} - \left[P_V(u_k) + P_{V^\perp}(\tilde{x}_k)\right], \end{cases} \tag{31}$$

where $\hat{\sigma} \in [0, 1[$. The convergence analysis given in [17] for the iteration (31) relies on the fact (proved in the latter reference) that (31) is a special instance of (13) (which we observed in Remark 2.4 may fail to converge if we consider $\hat{\sigma} \in]1/\sqrt{5}, 1[$). Using the fact just mentioned, the last statement in Proposition 2.2 and Proposition 3.3 we conclude that (31) is a special instance of Algorithm 2 whenever $\hat{\sigma} \in [0, 1/\sqrt{5}[$ and it may fail to converge if $\hat{\sigma} > 1/\sqrt{5}$. On the other hand, since, due to Proposition 3.3, Algorithm 2 is a special instance of Algorithm 1, it converges for all $\sigma \in [0, 1[$ (see, e.g., [3, Theorem 3.1]). Note that the difference between sPIM(ε) and Algorithm 2 is the inequality in (29) and (31).

In what follows we will prove iteration-complexity results for Algorithm 2 to obtain approximate solutions of (26), according to (28), as a consequence of the iteration-complexity results from Theorem 2.2. To this end, first let $\{\tilde{x}_k\}$, $\{u_k\}$ and $\{\varepsilon_k\}$ be generated by Algorithm 2 and define the *ergodic* sequences associated to them:

$$\begin{aligned} \tilde{x}_k^a &:= \frac{1}{k}\sum_{\ell=1}^k \tilde{x}_\ell\,, \quad u_k^a := \frac{1}{k}\sum_{\ell=1}^k u_\ell\,, \\ \varepsilon_k^a &:= \frac{1}{k}\sum_{\ell=1}^k \left[\varepsilon_\ell + \langle \tilde{x}_\ell - \tilde{x}_k^a, u_\ell - u_k^a\rangle\right]. \end{aligned} \tag{32}$$

The proof of the next Proposition is given in Subsection 3.1.

Theorem 3.1 *Let* $\{\tilde{x}_k\}$*,* $\{u_k\}$ *and* $\{\varepsilon_k\}$ *be generated by* Algorithm 2 *and let* $\{\tilde{x}_k^a\}$*,* $\{u_k^a\}$ *and* $\{\varepsilon_k^a\}$ *be defined in* (32) *. Let also* $d_{0,V}$ *denote the distance of* x_0 *to the solution set* (27)*. The following statements hold:*

(a) *For any* $k \geq 1$*, there exists* $j \in \{1, \ldots, k\}$ *such that*

$$\begin{aligned} & u_j \in T^{\varepsilon_j}(\tilde{x}_j), \\ & \sqrt{\|\tilde{x}_j - P_V(\tilde{x}_j)\|^2 + \|u_j - P_{V^\perp}(u_j)\|^2} \leq \frac{d_{0,V}}{\sqrt{k}}\sqrt{\frac{1+\sigma}{1-\sigma}}, \quad \varepsilon_j \leq \frac{\sigma^2 d_{0,V}^2}{2(1-\sigma^2)k}\,; \end{aligned} \tag{33}$$

(b) *for any* $k \geq 1$*,*

$$\begin{aligned} & u_k^a \in T^{\varepsilon_k^a}(\tilde{x}_k^a), \\ & \sqrt{\|\tilde{x}_k^a - P_V(\tilde{x}_k^a)\|^2 + \|u_k^a - P_{V^\perp}(u_k^a)\|^2} \leq \frac{2d_{0,V}}{k}, \quad 0 \leq \varepsilon_k^a \leq \frac{2(1 + \sigma/\sqrt{1-\sigma^2})d_{0,V}^2}{k}\,. \end{aligned} \tag{34}$$

Next result, which is a direct consequence of Theorem 3.1(b), gives the iteration-complexity of Algorithm 2 to find $x, u \in \mathcal{H}$ and $\varepsilon > 0$ satisfying the termination criterion (28).

Theorem 3.2 (Iteration-complexity) *Let $d_{0,V}$ denote the distance of x_0 to the solution set* (27) *and let $\rho, \epsilon > 0$ be given torelances. Then,* Algorithm 2 *finds $x, u \in \mathcal{H}$ and $\varepsilon > 0$ satisfying the termination criterion* (28) *in at most*

$$\mathcal{O}\left(\max\left\{\left\lceil\frac{d_{0,V}}{\rho}\right\rceil, \left\lceil\frac{d_{0,V}^2}{\epsilon}\right\rceil\right\}\right) \tag{35}$$

iterations.

We now consider a special instance of Algorithm 2 which will be used in Section 4 to derive operator splitting methods for solving the problem of finding zeroes of a sum of finitely many maximal monotone operators.

Algorithm 3 An inexact Spingarn's partial inverse method for (26) (II)

(0) Let $x_0 \in \mathcal{H}$ and $\sigma \in [0, 1[$ be given and set $k = 1$.
(1) Compute $\tilde{x}_k \in \mathcal{H}$ and $\varepsilon_k \geq 0$ such that

$$u_k := x_{k-1} - \tilde{x}_k \in T^{\varepsilon_k}(\tilde{x}_k), \quad \varepsilon_k \leq \frac{\sigma^2}{2}\|\tilde{x}_k - P_V(x_{k-1})\|^2. \tag{36}$$

(2) Define

$$x_k = P_V(\tilde{x}_k) + P_{V^\perp}(u_k), \tag{37}$$

set $k \leftarrow k + 1$ and go to step 1.

Remarks. 1) Letting $\sigma = 0$ in Algorithm 3 and using Proposition 2.1(d) we obtain from (36) that $x = \tilde{x}_k$ solves the inclusion $0 \in T(x) + x - x_{k-1}$, i.e., $\tilde{x}_k = (T + I)^{-1}x_{k-1}$ for all $k \geq 1$. In other words, if $\sigma = 0$, then Algorithm 3 is the Spingarn's partial inverse method originally presented in [1]. 2) It follows from Proposition 2.1(e) that Algorithm 3 is a generalization to the general setting of inclusions with monotone operators of the *Epsilon-proximal decomposition method scheme (EPDMS)* proposed and studied in [26] for solving convex optimization problems. Indeed, using the identity in (36) we find that the right hand side of the inequality in (36) is equal to $\sigma^2/2\left(\|P_{V^\perp}(\tilde{x}_k)\|^2 + \|P_V(u_k)\|^2\right)$ (cf. EPDMS method in [26], with a different notation). We also mention that no iteration-complexity analysis was performed in [26]. 3) Likewise, letting $V = \mathcal{H}$ in Algorithm 3 and using Proposition 2.1(e) we obtain that Algorithm 3 generalizes the *IPP-CO framework* of [21] (with $\lambda_k := 1$ for all $k \geq 1$), for which iteration-complexity analysis was presented in the latter reference, to the more general setting of inclusions problems with monotone operators.

Proposition 3.1 *The following statements hold true.*

(a) Algorithm 3 *is a special instance of* Algorithm 2.
(b) *The conclusions of Theorem 3.1 and Theorem 3.2 are still valid with* Algorithm 2 *replaced by* Algorithm 3.

Proof (a) Let $\{x_k\}$, $\{\tilde{x}_k\}$, $\{\varepsilon_k\}$ and $\{u_k\}$ be generated by Algorithm 3. Firstly, note that the identity in (36) yields $u_k + \tilde{x}_k - x_{k-1} = 0$ and, consequently,

$$\begin{aligned}\|\tilde{x}_k - P_V(x_{k-1})\|^2 &= \|P_V(\tilde{x}_k - x_{k-1})\|^2 + \|P_{V^\perp}(\tilde{x}_k)\|^2\\ &= \|P_V(\tilde{x}_k - x_{k-1})\|^2 + \|P_{V^\perp}(u_k - x_{k-1})\|^2\\ &= \|P_V(\tilde{x}_k) + P_{V^\perp}(u_k) - x_{k-1}\|^2,\end{aligned}$$

and

$$\begin{aligned}P_V(\tilde{x}_k) + P_{V^\perp}(u_k) &= (\tilde{x}_k - P_{V^\perp}(\tilde{x}_k)) + P_{V^\perp}(u_k)\\ &= (x_{k-1} - u_k) - P_{V^\perp}(\tilde{x}_k) + P_{V^\perp}(u_k)\\ &= x_{k-1} - [P_V(u_k) + P_{V^\perp}(\tilde{x}_k)]\,.\end{aligned}$$

Altogether we obtain (a).

(b) This Item is a direct consequence of (a), Theorem 3.1 and Theorem 3.2. □

Next we observe that Proposition 3.1(b) and the first remark after Algorithm 3 allow us to obtain the iteration-complexity for the Spingarn's partial inverse method.

Proposition 3.2 *Let $d_{0,V}$ denote the distance of x_0 to the solution set* (27) *and consider* Algorithm 3 *with $\sigma = 0$ or, equivalently, the* Spingarn's partial inverse method of [1]. *For given tolerances $\rho, \epsilon > 0$, the latter method finds*

(a) *$x, u \in \mathcal{H}$ such that $u \in T(x)$, $\max\{\|x - P_V(x)\|, \|u - P_{V^\perp}(u)\|\} \leq \rho$ in at most*

$$\mathcal{O}\left(\left\lceil \frac{d_{0,V}^2}{\rho^2} \right\rceil\right) \tag{38}$$

iterations.

(b) *$x, u \in \mathcal{H}$ and $\varepsilon > 0$ satisfying the termination criterion* (28) *in at most a number of iterations given in* (35).

Proof (a) The statement in this item is a direct consequence of Proposition 3.1(b), Theorem 3.1(a) and the fact that $\varepsilon_k = 0$ for all $k \geq 1$ (because $\sigma = 0$ in (36)). (b) Here, the result follows from Proposition 3.1(b) and Theorem 3.2. □

3.1 Proof of Theorem 3.1

The approach adopted in the current section for solving (26) follows the Spingarn's approach [1] which consists in solving the monotone inclusion

$$0 \in T_V(z) \tag{39}$$

where the maximal monotone operator $T_V : \mathcal{H} \rightrightarrows \mathcal{H}$ is the partial inverse of T with respect to the subspace V. In view of (3), we have

$$(T_V)^{-1}(0) = \mathcal{S}^*(V, T), \tag{40}$$

where the latter set is defined in (27). Hence, problem (26) is equivalent to the monotone inclusion problem (39). Before proving Theorem 3.1 we will show that Algorithm 2 can be regarded as a special instance of Algorithm 1 for solving (39).

Proposition 3.3 *Let $\{\tilde{x}_k\}_{k\geq 1}$, $\{u_k\}_{k\geq 1}$, $\{\varepsilon_k\}_{k\geq 1}$ and $\{z_k\}_{k\geq 0}$ be generated by* Algorithm 2. *Define $z_0 = x_0$ and, for all $k \geq 1$,*

$$z_k = x_k, \quad \tilde{z}_k = P_V(\tilde{x}_k) + P_{V^\perp}(u_k), \quad v_k = P_V(u_k) + P_{V^\perp}(\tilde{x}_k). \tag{41}$$

Then, for all $k \geq 1$,

$$\begin{aligned} &v_k \in (T_V)^{\varepsilon_k}(\tilde{z}_k), \quad \|v_k + \tilde{z}_k - z_{k-1}\|^2 + 2\varepsilon_k \leq \sigma^2 \|\tilde{z}_k - z_{k-1}\|^2, \\ &z_k = z_{k-1} - v_k, \end{aligned} \tag{42}$$

i.e., $(\tilde{z}_k, v_k, \varepsilon_k)$ and $\lambda_k := 1$ satisfy (9) *and* (10) *for all $k \geq 1$. As a consequence, the sequences $\{z_k\}_{k\geq 0}$, $\{\tilde{z}_k\}_{k\geq 1}$, $\{v_k\}_{k\geq 1}$ and $\{\varepsilon_k\}_{k\geq 1}$ are generated by* Algorithm 1 *(with $\lambda_k := 1$ for all $k \geq 1$) for solving* (39).

Proof From the inclusion in (29), (3) with $S = T^{\varepsilon_k}$ and Lemma 2.1 we have $P_V(u_k) + P_{V^\perp}(\tilde{x}_k) \in (T_V)^{\varepsilon_k}(P_V(\tilde{x}_k) + P_{V^\perp}(u_k))$ for all $k \geq 1$, which combined with the definitions of $\tilde{z}_k$ and v_k in (41) gives the inclusion in (42). Direct use of (41) and the definition of $\{z_k\}$ yield

$$\begin{aligned} v_k + \tilde{z}_k + z_{k-1} &= u_k + \tilde{x}_k - x_{k-1}, \\ \tilde{z}_k - z_{k-1} &= P_V(\tilde{x}_k) + P_{V^\perp}(u_k) - x_{k-1}, \\ z_{k-1} - v_k &= x_{k-1} - [P_V(u_k) - P_{V^\perp}(\tilde{x}_k)], \end{aligned} \tag{43}$$

which combined with (29), (30) and the definition of $\{z_k\}$ gives the remaining statements in (42). The last statement of the proposition follows from (42) and Algorithm 1's definition. □

Proof of Theorem 3.1. From (41) we obtain

$$\tilde{x}_k = P_V(\tilde{z}_k) + P_{V^\perp}(v_k), \quad u_k = P_V(v_k) + P_{V^\perp}(\tilde{z}_k) \quad \forall k \geq 1. \tag{44}$$

Direct substitution of the latter identities in $\tilde{x}_k^a$ and u_k^a in (32) yields

$$\tilde{x}_k^a = P_V(\tilde{z}_k^a) + P_{V^\perp}(v_k^a), \quad u_k^a = P_V(v_k^a) + P_{V^\perp}(\tilde{z}_k^a) \quad \forall k \geq 1. \tag{45}$$

Using (44) and (45) in the definition of ε_k^a in (32) and the fact that the operators P_V and $P_{V^\perp}$ are self-adjoint and idempotent we find

$$\varepsilon_k^a = \frac{1}{\Lambda_k} \sum_{\ell=1}^{k} \lambda_\ell (\varepsilon_\ell + \langle \tilde{z}_\ell - \tilde{z}_k^a, v_\ell - v_k^a \rangle) \quad \forall k \geq 1, \tag{46}$$

where $\{\varepsilon_k^a\}$ is defined in (32). Now consider the ergodic sequences $\{\Lambda_k\}$, $\{\tilde{z}_k^a\}$ and $\{v_k^a\}$ defined in (12) with $\lambda_k := 1$ for all $k \geq 1$. Let d_0 denote the distance of $z_0 = x_0$ to the solution set $(T_V)^{-1}(0)$ of (39) and note that $d_0 = d_{0,V}$ in view of (40). Based on the above considerations one can use the last statement in Proposition 3.3 and Theorem 2.2 with $\underline{\lambda} := 1$ to conclude that for any $k \geq 1$ there exists $j \in \{1, \ldots, k\}$ such that

$$v_j \in (T_V)^{\varepsilon_j}(\tilde{z}_j), \quad \|v_j\| \leq \frac{d_{0,V}}{\sqrt{k}} \sqrt{\frac{1+\sigma}{1-\sigma}}, \quad \varepsilon_j \leq \frac{\sigma^2 d_{0,V}^2}{2(1-\sigma^2)k}, \tag{47}$$

and

$$v_k^a \in (T_V)^{\varepsilon_k^a}(\tilde{z}_k^a), \quad \|v_k^a\| \leq \frac{2d_{0,V}}{k}, \quad \varepsilon_k^a \leq \frac{2(1+\sigma/\sqrt{1-\sigma^2})d_{0,V}^2}{k}, \tag{48}$$

where $\{\varepsilon_k^a\}$ is given in (32). Using Lemma 2.1, the definition in (3) (for $S = T^{\varepsilon_k}$), (44) and (45) we conclude that the equivalence $v \in (T_V)^\varepsilon(\tilde{z}) \iff v \in (T^\varepsilon)_V(\tilde{z}) \iff u \in T^\varepsilon(\tilde{x})$ holds for $(\tilde{z}, v, \varepsilon) = (\tilde{z}_k, v_k, \varepsilon_k)$ and $(\tilde{x}, u, \varepsilon) = (\tilde{x}_k, u_k, \varepsilon_k)$, and $(\tilde{z}, v, \varepsilon) = (\tilde{z}_k^a, v_k^a, \varepsilon_k^a)$ and $(\tilde{x}, u, \varepsilon) = (\tilde{x}_k^a, u_k^a, \varepsilon_k^a)$, for all $k \geq 1$. As a consequence, the inclusions in (33) and (34) follow from the ones in (47) and (48), respectively. Since (45) gives $v_k^a = P_V(u_k^a) + P_{V^\perp}(\tilde{x}_k^a)$ for all $k \geq 1$, it follows from the definition of $\{v_k\}$ in (41) that $(v, u, \tilde{x}) = (v_k, u_k, \tilde{x}_k)$ and $(v, u, \tilde{x}) = (v_k^a, u_k^a, \tilde{x}_k^a)$ satisfy

$$\|v\|^2 = \|P_V(u)\|^2 + \|P_{V^\perp}(\tilde{x})\|^2 = \|u - P_{V^\perp}(u)\|^2 + \|\tilde{x} - P_V(\tilde{x})\|^2$$

for all $k \geq 1$, which, in turn, gives that the inequalities in (33) and (34) follow from the ones in (47) and (48), respectively. This concludes the proof. □

4 Applications to Operator Splitting and Optimization

In this section we consider the problem of finding $x \in \mathcal{H}$ such that

$$0 \in \sum_{i=1}^{m} T_i(x) \tag{49}$$

where $m \geq 2$ and $T_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone for $i = 1, \ldots, m$. As observed in [1], $x \in \mathcal{H}$ satisfies the inclusion (49) if and only if there exist $u_1, \ldots, u_m \in \mathcal{H}$ such that

$$u_i \in T_i(x) \text{ and } \sum_{i=1}^{m} u_i = 0. \tag{50}$$

That said, we consider the (extended) solution set of (49) – which we assume nonempty – to be defined by

$$\mathcal{S}^*(\Sigma) := \{(z_i)_{i=1}^m \in \mathcal{H}^m \ : \ \exists\, x, u_1, u_2, \ldots, u_m \in \mathcal{H} \text{ satisfying (50)}; \ \ z_i = x + u_i \ \forall i = 1, \ldots, m\}. \tag{51}$$

Due to its importance in solving large-scale problems, numerical schemes for solving (49) use information of each T_i individually instead of using the entire sum [1,17,27,29–31]. In this section, we apply the results of Section 3 to present and study the iteration-complexity of an inexact-version of the Spingarn's operator splitting method [1] for solving (49) and, as a by-product, we obtain the iteration-complexity of the latter method. Moreover, we will apply our results to obtain the iteration-complexity of a parallel forward-backward algorithm for solving multi-term composite convex optimization problems.

To this end, we consider the following notion of approximate solution for (49): given tolerances $\rho, \delta, \epsilon > 0$, find $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m \in \mathcal{H}$, $\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m \in \mathcal{H}$ and $\bar{\varepsilon}_1, \bar{\varepsilon}_2, \ldots, \bar{\varepsilon}_m > 0$ such that $(x_i)_{i=1}^m = (\bar{x}_i)_{i=1}^m$, $(u_i)_{i=1}^m = (\bar{u}_i)_{i=1}^m$ and $(\varepsilon_i)_{i=1}^m = (\bar{\varepsilon}_i)_{i=1}^m$ satisfy

$$\begin{aligned} & u_i \in T_i^{\varepsilon_i}(x_i) \quad \forall i = 1, \ldots, m, \\ & \left\| \sum_{i=1}^m u_i \right\| \leq \rho, \\ & \|x_i - x_\ell\| \leq \delta \quad \forall i, \ell = 1, \ldots, m, \\ & \sum_{i=1}^m \varepsilon_i \leq \epsilon. \end{aligned} \tag{52}$$

For $\rho = \delta = \epsilon = 0$, criterion (52) gives $\bar{x}_1 = \bar{x}_2 = \cdots = \bar{x}_m =: \bar{x}$, $\sum_{i=1}^m \bar{u}_i = 0$ and $\bar{u}_i \in T_i(\bar{x})$ for all $i = 1, \ldots, m$, i.e., in this case $\bar{x}, \bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m$ satisfy (50).

We next present our inexact version of the Spingarn's operator splitting method [1] for solving (49).

Algorithm 4 An inexact Spingarn's operator splitting method for (49)

(0) Let $(x_0, y_{1,0}, \ldots, y_{m,0}) \in \mathcal{H}^{m+1}$ such that $y_{1,0} + \cdots + y_{m,0} = 0$ and $\sigma \in [0, 1[$ be given and set $k = 1$.

(1) For each $i = 1, \ldots, m$, compute $\tilde{x}_{i,k} \in \mathcal{H}$ and $\varepsilon_{i,k} \geq 0$ such that

$$u_{i,k} := x_{k-1} + y_{i,k-1} - \tilde{x}_{i,k} \in T_i^{\varepsilon_{i,k}}(\tilde{x}_{i,k}), \qquad \varepsilon_{i,k} \leq \frac{\sigma^2}{2} \|\tilde{x}_{i,k} - x_{k-1}\|^2. \tag{53}$$

(2) Define

$$x_k = \frac{1}{m} \sum_{i=1}^m \tilde{x}_{i,k}, \qquad y_{i,k} = u_{i,k} - \frac{1}{m} \sum_{\ell=1}^m u_{\ell,k} \ \text{ for } \ i = 1, \ldots, m, \tag{54}$$

set $k \leftarrow k + 1$ and go to step 1.

Remarks. 1) Letting $\sigma = 0$ in Algorithm 4 we obtain the Spingarn's operator splitting method of [1]. 2) In [17], Section 5, an inexact version of the Spingarn's operator splitting method – called *split-sPIM(ε)* – was proposed for solving (49). With a different notation, for $i = 1, \ldots, m$, each iteration of the latter method can be written as:

$$\begin{cases} u_{i,k} \in T_i^{\varepsilon_{i,k}}(\tilde{x}_{i,k}), \\ \|u_{i,k} + \tilde{x}_{i,k} - x_{k-1} - y_{i,k-1}\|^2 + 2\varepsilon_{i,k} \leq \hat{\sigma}^2 \left(\|\tilde{x}_{i,k} - x_{k-1}\|^2 + \|u_{i,k} - y_{i,k-1}\|^2 \right), \\ x_k = x_{k-1} - \dfrac{1}{m} \sum_{i=1}^m u_{i,k}, \quad y_{i,k} = y_{i,k-1} - \tilde{x}_{i,k} + \dfrac{1}{m} \sum_{\ell=1}^m \tilde{x}_{\ell,k} \ \text{ for } \ i = 1, \ldots, m, \end{cases} \tag{55}$$

where $\hat{\sigma} \in [0, 1[$. The convergence analysis of [17] consists in analyzing (55) in the framework of the method described in (13), whose convergence may fail if we take $\hat{\sigma} > 1/\sqrt{5}$, as we observed in Remark 2.4. On the other hand, we will prove in Proposition 4.1 that Algorithm 4 can be regarded as a special instance of Algorithm 3, which converges for all $\sigma \in [0, 1[$ (see Proposition 3.3, Proposition 3.1(b) and [3, Theorem 3.1]). Moreover, we mention that contrary to this work no iteration-complexity analysis is performed in [17].

For each $i = 1, \ldots, m$, let $\{\tilde{x}_{i,k}\}$, $\{u_{i,k}\}$ and $\{\varepsilon_{i,k}\}$ be generated by Algorithm 4 and define the *ergodic* sequences associated to them:

$$\begin{aligned} & \tilde{x}_{i,k}^a := \frac{1}{k} \sum_{\ell=1}^k \tilde{x}_{i,\ell}, \quad u_{i,k}^a := \frac{1}{k} \sum_{\ell=1}^k u_{i,\ell}, \\ & \varepsilon_{i,k}^a := \frac{1}{k} \sum_{\ell=1}^k \left[\varepsilon_{i,\ell} + \langle \tilde{x}_{i,\ell} - \tilde{x}_{i,k}^a, u_{i,\ell} - u_{i,k}^a \rangle \right]. \end{aligned} \tag{56}$$

Next theorem will be proved in Subsection 4.1.

Theorem 4.1 *For each* $i = 1, \ldots, m$, *let* $\{\tilde{x}_{i,k}\}$, $\{u_{i,k}\}$ *and* $\{\varepsilon_{i,k}\}$ *be generated by* Algorithm 4 *and let* $\{\tilde{x}^a_{i,k}\}$, $\{u^a_{i,k}\}$ *and* $\{\varepsilon^a_{i,k}\}$ *be defined in* (56) . *Let also* $d_{0,\Sigma}$ *denote the distance of* $(x_0+y_{1,0}, \ldots, x_0+y_{m,0})$ *to the solution set* (51). *The following statements hold:*

(a) *For any* $k \geq 1$, *there exists* $j \in \{1, \ldots, k\}$ *such that*

$$\begin{aligned} & u_{i,j} \in T_i^{\varepsilon_{i,j}}(\tilde{x}_{i,j}) \quad \forall i = 1, \ldots, m, \\ & \left\| \sum_{i=1}^m u_{i,j} \right\| \leq \frac{\sqrt{m}\, d_{0,\Sigma}}{\sqrt{k}} \sqrt{\frac{1+\sigma}{1-\sigma}}, \\ & \|\tilde{x}_{i,j} - \tilde{x}_{\ell,j}\| \leq \frac{2\, d_{0,\Sigma}}{\sqrt{k}} \sqrt{\frac{1+\sigma}{1-\sigma}} \quad \forall i, \ell = 1, \ldots, m, \\ & \sum_{i=1}^m \varepsilon_{i,j} \leq \frac{\sigma^2 d_{0,\Sigma}^2}{2(1-\sigma^2)k} \ ; \end{aligned} \tag{57}$$

(b) *for any* $k \geq 1$,

$$\begin{aligned} & u^a_{i,k} \in T_i^{\varepsilon^a_{i,k}}(\tilde{x}^a_{i,k}) \quad \forall i = 1, \ldots, m, \\ & \left\| \sum_{i=1}^m u^a_{i,k} \right\| \leq \frac{2\sqrt{m}\, d_{0,\Sigma}}{k}, \\ & \|\tilde{x}^a_{i,k} - \tilde{x}^a_{\ell,k}\| \leq \frac{4\, d_{0,\Sigma}}{k} \quad \forall i, \ell = 1, \ldots, m, \\ & \sum_{i=1}^m \varepsilon^a_{i,k} \leq \frac{2(1+\sigma/\sqrt{1-\sigma^2}) d_{0,\Sigma}^2}{k} \, . \end{aligned} \tag{58}$$

As a consequence of Theorem 4.1(b) we obtain the iteration-complexity of Algorithm 4 to find $x_1, x_2, \ldots, x_m \in \mathcal{H}$, $u_1, u_2, \ldots, u_m \in \mathcal{H}$ and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m > 0$ satisfying the termination criterion (52).

Theorem 4.2 (Iteration-complexity) *Let* $d_{0,\Sigma}$ *denote the distance of* $(x_0 + y_{1,0}, \ldots, x_0 + y_{m,0})$ *to the solution set* (51) *and let* $\rho, \delta, \epsilon > 0$ *be given torelances. Then,* Algorithm 4 *finds* $x_1, x_2, \ldots, x_m \in \mathcal{H}$, $u_1, u_2, \ldots, u_m \in \mathcal{H}$ *and* $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m > 0$ *satisfying the termination criterion* (52) *in at most*

$$\mathcal{O}\left(\max \left\{ \left\lceil \frac{\sqrt{m}\, d_{0,\Sigma}}{\rho} \right\rceil, \left\lceil \frac{d_{0,\Sigma}}{\delta} \right\rceil, \left\lceil \frac{d_{0,\Sigma}^2}{\epsilon} \right\rceil \right\} \right) \tag{59}$$

iterations.

Using the first remark after Algorithm 4 and Theorem 4.1 we also obtain the *pointwise* and *ergodic* iteration-complexity of Spingarn's operator splitting method [1].

Theorem 4.3 (Iteration-complexity) *Let* $d_{0,\Sigma}$ *denote the distance of* $(x_0 + y_{1,0}, \ldots, x_0 + y_{m,0})$ *to the solution set* (51) *and consider* Algorithm 4 *with* $\sigma = 0$ *or, equivalently, the* Spingarn's operator splitting method of [1]. *For given tolerances* $\rho, \delta, \epsilon > 0$, *the latter method finds*

(a) $x_1, x_2, \ldots, x_m \in \mathcal{H}$ *and* $u_1, u_2, \ldots, u_m \in \mathcal{H}$ *such that*

$$\begin{aligned} & u_i \in T_i(x_i) \quad \forall i = 1, \ldots, m, \\ & \left\| \sum_{i=1}^m u_i \right\| \leq \rho, \\ & \|x_i - x_\ell\| \leq \delta \, , \quad \forall i, \ell = 1, \ldots, m, \end{aligned} \tag{60}$$

in at most

$$\mathcal{O}\left(\max \left\{ \left\lceil \frac{m\, d_{0,\Sigma}^2}{\rho^2} \right\rceil, \left\lceil \frac{d_{0,\Sigma}^2}{\delta^2} \right\rceil \right\} \right) \tag{61}$$

iterations.

(b) $x_1, x_2, \ldots, x_m \in \mathcal{H}$, $u_1, u_2, \ldots, u_m \in \mathcal{H}$ *and* $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m > 0$ *satisfying the termination criterion* (52) *in at most the number of iterations given in* (59).

Proof (a) This item follows from Theorem 4.1(a) and the fact that $\varepsilon_{i,k} = 0$ for each $i = 1, \dots, m$ and for all $k \geq 1$ (because $\sigma = 0$ in (53)). (b) This item follows directly from Theorem 4.2. ◻

Applications to optimization. In the remaining part of this section we show how Algorithm 4 and its iteration-complexity results can be used to derive a *parallel forward-backward splitting method* for multi-term composite convex optimization and to study its iteration-complexity. More precisely, consider the minimization problem

$$\min_{x \in \mathcal{H}} \sum_{i=1}^{m} (f_i + \varphi_i)(x) \tag{62}$$

where $m \geq 2$ and the following conditions are assumed to hold for all $i = 1, \dots, m$:

(A.1) $f_i : \mathcal{H} \to \mathbb{R}$ is convex, and differentiable with a L_i-Lipschitz continuous gradient, i.e., there exists $L_i > 0$ such that

$$\|\nabla f_i(x) - \nabla f_i(y)\| \leq L_i \|x - y\| \quad \forall x, y \in \mathcal{H}; \tag{63}$$

(A.2) $\varphi_i : \mathcal{H} \to \overline{\mathbb{R}}$ is proper, convex and closed with an easily computable resolvent $(\lambda \partial \varphi_i + I)^{-1}$, for any $\lambda > 0$;

(A.3) the solution set of (62) is nonempty.

We also assume standard regularity conditions [1] on the functions φ_i which make (62) equivalent to the monotone inclusion problem (49) with $T_i := \nabla f_i + \partial \varphi_i$, for all $i = 1, \dots, m$, i.e., which make it equivalent to the problem of finding $x \in \mathcal{H}$ such that

$$0 \in \sum_{i=1}^{m} (\nabla f_i + \partial \varphi_i)(x). \tag{64}$$

Analogously to (52), we consider the following notion of approximate solution for (62): given tolerances $\rho, \delta, \epsilon > 0$, find $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m \in \mathcal{H}$, $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_m \in \mathcal{H}$ and $\bar{\varepsilon}_1, \bar{\varepsilon}_2, \dots, \bar{\varepsilon}_m > 0$ such that $(x_i)_{i=1}^m = (\bar{x}_i)_{i=1}^m$, $(u_i)_{i=1}^m = (\bar{u}_i)_{i=1}^m$ and $(\varepsilon_i)_{i=1}^m = (\bar{\varepsilon}_i)_{i=1}^m$ satisfy (52) with $T_i^{\varepsilon_i}$ replaced by $\partial_{\varepsilon_i} f_i + \partial \varphi_i$, for each $i = 1, \dots, m$. For $\rho = \delta = \epsilon = 0$, this criterion gives $\bar{x}_1 = \bar{x}_2 = \cdots = \bar{x}_m =: \bar{x}$, $\sum_{i=1}^m \bar{u}_i = 0$ and $\bar{u}_i \in (\nabla f_i + \partial \varphi_i)(\bar{x})$ for all $i = 1, \dots, m$, i.e., in this case $\bar{x}$ solves (64).

We will present a parallel forward-backward method for solving (62) whose iteration-complexity is obtained by regarding it as a special instance of Algorithm 4. Since problem (62) appears in various applications of convex optimization, it turns out that the development of efficient numerical schemes for solving it – specially with $m \geq 2$ very large – is of great importance.

Next is our method for solving (62).

Algorithm 5 A parallel forward-backward splitting method for (62)

(0) Let $(x_0, y_{1,0}, \dots, y_{m,0}) \in \mathcal{H}^{m+1}$ such that $y_{1,0} + \cdots + y_{m,0} = 0$ and $\sigma \in]0, 1[$ be given and set $\lambda = \sigma^2 / \max\{L_i\}_{i=1}^m$ and $k = 1$.

(1) For each $i = 1, \dots, m$, compute

$$\tilde{x}_{i,k} = (\lambda \partial \varphi_i + I)^{-1} (x_{k-1} + y_{i,k-1} - \lambda \nabla f_i(x_{k-1})). \tag{65}$$

(2) Define

$$x_k = \frac{1}{m} \sum_{i=1}^{m} \tilde{x}_{i,k}, \qquad y_{i,k} = y_{i,k-1} + x_k - \tilde{x}_{i,k} \ \text{ for } \ i = 1, \dots, m, \tag{66}$$

set $k \leftarrow k + 1$ and go to step 1.

Remarks. 1) Since in (65) we have a forward step in the direction $-\nabla f_i(x_{k-1})$ and a backward step given by the resolvent of φ_i, Algorithm 5 can be regarded as a parallel variant of the classical forward-backward

[1] see, e.g., [28, Corollary 16.39]

splitting algorithm [32] . 2) For $m=1$ the above method coincides with the forward-backward method of [21], for which the iteration-complexity was studied in the latter reference.

For each $i=1,\dots,m$, let $\{x_k\}$, $\{\tilde x_{i,k}\}$ be generated by Algorithm 5, $\{u_{i,k}\}$ and $\{\varepsilon_{i,k}\}$ be defined in (84) and let $\{\tilde x^a_{i,k}\}$, $\{u^a_{i,k}\}$ and $\{\varepsilon^a_{i,k}\}$ be given in (56). Define, for all $k\geq 1$,

$$\begin{aligned} &u'_{i,k}:=\frac{1}{\lambda}\,u_{i,k}\,,\quad \varepsilon'_{i,k}:=\frac{1}{\lambda}\,\varepsilon_{i,k}\,,\quad {u'}^{a}_{i,k}:=\frac{1}{\lambda}\,u^a_{i,k}\,,\quad {\varepsilon'}^{a}_{i,k}:=\frac{1}{\lambda}\,\varepsilon^a_{i,k}\,,\\ &{\varepsilon''}^{a}_{i,k}:=\frac{1}{k}\sum_{\ell=1}^{k}\Big[\varepsilon'_{i,\ell}+\langle \tilde x_{i,\ell}-\tilde x^a_{i,k},\nabla f_i(x_{\ell-1})-\frac{1}{k}\sum_{s=1}^{k}\nabla f_i(x_{s-1})\rangle\Big]. \end{aligned} \tag{67}$$

Next theorem will be proved in Subsection 4.2.

Theorem 4.4 *For each $i=1,\dots,m$, let $\{\tilde x_{i,k}\}$ be generated by* Algorithm 5 *and $\{\tilde x^a_{i,k}\}$ be given in* (56)*; let $\{u'_{i,k}\}$, $\{\varepsilon'_{i,k}\}$, $\{{u'}^a_{i,k}\}$, $\{{\varepsilon'}^a_{i,k}\}$ and $\{{\varepsilon''}^a_{i,k}\}$ be given in* (67)*. Let also $d_{0,\Sigma}$ denote the distance of $(x_0+y_{1,0},\dots,x_0+y_{m,0})$ to the solution set* (51) *in which $T_i:=\nabla f_i+\partial\varphi_i$ for $i=1,\dots,m$, and define $L_\Sigma:=\max\{L_i\}_{i=1}^m$. The following hold:*

(a) *For any $k\geq 1$, there exists $j\in\{1,\dots,k\}$ such that*

$$\begin{aligned} &u'_{i,j}\in\Big(\partial_{\varepsilon'_{i,j}}f_i+\partial\varphi_i\Big)(\tilde x_{i,j})\quad \forall i=1,\dots,m,\\ &\left\|\sum_{i=1}^m u'_{i,j}\right\|\leq \frac{\sqrt{m}\,L_\Sigma\,d_{0,\Sigma}}{\sigma^2\sqrt{k}}\sqrt{\frac{1+\sigma}{1-\sigma}},\\ &\|\tilde x_{i,j}-\tilde x_{\ell,j}\|\leq \frac{2\,d_{0,\Sigma}}{\sqrt{k}}\sqrt{\frac{1+\sigma}{1-\sigma}}\quad \forall i,\ell=1,\dots,m,\\ &\sum_{i=1}^m\varepsilon'_{i,j}\leq\frac{L_\Sigma\,d^2_{0,\Sigma}}{2(1-\sigma^2)k}\ ; \end{aligned} \tag{68}$$

(b) *for any $k\geq 1$,*

$$\begin{aligned} &{u'}^a_{i,k}\in\Big(\partial_{{\varepsilon''}^a_{i,k}}f_i+\partial_{({\varepsilon'}^a_{i,k}-{\varepsilon''}^a_{i,k})}\varphi_i\Big)(\tilde x^a_{i,k})\quad\forall i=1,\dots,m,\\ &\left\|\sum_{i=1}^m {u'}^a_{i,k}\right\|\leq\frac{2\sqrt{m}\,L_\Sigma\,d_{0,\Sigma}}{\sigma^2 k},\\ &\|\tilde x^a_{i,k}-\tilde x^a_{\ell,k}\|\leq\frac{4d_{0,\Sigma}}{k}\quad\forall i,\ell=1,\dots,m,\\ &\sum_{i=1}^m{\varepsilon'}^a_{i,k}\leq\frac{2(1+\sigma/\sqrt{1-\sigma^2})\,L_\Sigma\,d^2_{0,\Sigma}}{\sigma^2 k}\,. \end{aligned} \tag{69}$$

The following theorem is a direct consequence of Theorem 4.4.

Theorem 4.5 (iteration-complexity) *Let $d_{0,\Sigma}$ denote the distance of $(x_0+y_{1,0},\dots,x_0+y_{m,0})$ to the solution set* (51) *in which $T_i:=\nabla f_i+\partial\varphi_i$, for $i=1,\dots,m$, and let $\rho,\delta,\epsilon>0$ be given torelances. Let $L_\Sigma:=\max\{L_i\}_{i=1}^m$. Then,* Algorithm 5 *finds*

(a) *$x_1,x_2,\dots,x_m\in\mathcal{H}$, $u_1,u_2,\dots,u_m\in\mathcal{H}$ and $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_m>0$ satisfying the termination criterion* (52) *with $T_i^{\varepsilon_i}$ replaced by $\partial_{\varepsilon_i}f_i+\partial\varphi_i$ in at most*

$$\mathcal{O}\left(\max\left\{\left\lceil\frac{m\,L^2_\Sigma\,d^2_{0,\Sigma}}{\rho^2}\right\rceil,\left\lceil\frac{d^2_{0,\Sigma}}{\delta^2}\right\rceil,\left\lceil\frac{L_\Sigma\,d^2_{0,\Sigma}}{\epsilon}\right\rceil\right\}\right) \tag{70}$$

iterations.

(b) *$x_1,x_2,\dots,x_m\in\mathcal{H}$, $u_1,u_2,\dots,u_m\in\mathcal{H}$, $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_m>0$ and $\hat\varepsilon_1,\hat\varepsilon_2,\dots,\hat\varepsilon_m>0$ such that*

$$\begin{aligned} &u_i\in\left(\partial_{\varepsilon_i}f_i+\partial_{\hat\varepsilon_i}\varphi_i\right)(x_i)\quad\forall i=1,\dots,m,\\ &\left\|\sum_{i=1}^m u_i\right\|\leq\rho,\\ &\|x_i-x_\ell\|\leq\delta\,,\quad\forall i,\ell=1,\dots,m,\\ &\sum_{i=1}^m(\varepsilon_i+\hat\varepsilon_i)\leq\epsilon \end{aligned} \tag{71}$$

in at most

$$\mathcal{O}\left(\max\left\{\left\lceil\frac{\sqrt{m}\,L_\Sigma\,d_{0,\Sigma}}{\rho}\right\rceil,\left\lceil\frac{d_{0,\Sigma}}{\delta}\right\rceil,\left\lceil\frac{L_\Sigma\,d_{0,\Sigma}^2}{\epsilon}\right\rceil\right\}\right) \tag{72}$$

iterations.

4.1 Proof of Theorem 4.1

Analogously to Subsection 3.1, in the current section we follow the Spingarn's approach in [1] for solving problem (49) which consists in solving the following inclusion in the product space $\mathcal{H}^m$:

$$\mathbf{0}\in\mathbf{T}_{\mathbf{V}}(\mathbf{z}), \tag{73}$$

where $\mathbf{T}:\mathcal{H}^m\rightrightarrows\mathcal{H}^m$ is the maximal monotone operator defined by

$$\mathbf{T}(x_1,x_2,\dots,x_m):=T_1(x_1)\times T_2(x_2)\times\cdots\times T_m(x_m)\quad\forall(x_1,x_2,\dots,x_m)\in\mathcal{H}^m, \tag{74}$$

and

$$\mathbf{V}:=\{(x_1,x_2,\dots,x_m)\in\mathcal{H}^m\;:\;x_1=x_2=\cdots=x_m\} \tag{75}$$

is a closed subspace of $\mathcal{H}^m$ whose orthogonal complement is

$$\mathbf{V}^\perp=\{(x_1,x_2,\dots,x_m)\in\mathcal{H}^m\;:\;x_1+x_2+\cdots+x_m=0\}. \tag{76}$$

Based on the above observations, we have that problem (49) is equivalent to (26) with $\mathbf{T}$ and $\mathbf{V}$ given in (74) and (75), respectively. Moreover, in this case, the orthogonal projections onto $\mathbf{V}$ and $\mathbf{V}^\perp$ have the explicit formulae:

$$\begin{aligned}P_{\mathbf{V}}(x_1,x_2,\dots,x_m)&=\left(\frac{1}{m}\sum_{i=1}^m x_i,\dots,\frac{1}{m}\sum_{i=1}^m x_i\right),\\ P_{\mathbf{V}^\perp}(x_1,x_2,\dots,x_m)&=\left(x_1-\frac{1}{m}\sum_{i=1}^m x_i,\dots,x_m-\frac{1}{m}\sum_{i=1}^m x_i\right).\end{aligned} \tag{77}$$

Next we show that Algorithm 4 can be regarded as a special instance of Algorithm 3 and, as a consequence, we will obtain that Theorem 4.1 follows from results of Section 3 for Algorithm 3.

Proposition 4.1 *Let* $\{x_k\}_{k\geq 0}$ *and, for each* $i=1,\dots,m$, $\{y_{i,k}\}_{k\geq0}$, $\{\tilde{x}_{i,k}\}_{k\geq1}$, $\{u_{i,k}\}_{k\geq1}$ *and* $\{\varepsilon_{i,k}\}_{k\geq1}$ *be generated by* Algorithm 4. *Consider the sequences* $\{\mathbf{x}_k\}_{k\geq0}$, $\{\tilde{\mathbf{x}}_k\}_{k\geq1}$ *and* $\{\mathbf{u}_k\}_{k\geq1}$ *in* $\mathcal{H}^m$ *and* $\{\varepsilon_k\}_{k\geq1}$ *in* $\mathbb{R}_+$ *where*

$$\begin{aligned}&\mathbf{x}_k:=(x_k+y_{1,k},\dots,x_k+y_{m,k}),\quad\tilde{\mathbf{x}}_k:=(\tilde{x}_{1,k},\dots,\tilde{x}_{m,k}),\\ &\varepsilon_k:=\sum_{i=1}^m\varepsilon_{i,k},\quad\mathbf{u}_k:=(u_{1,k},\dots,u_{m,k}).\end{aligned} \tag{78}$$

Then, for all $k\geq1$,

$$\begin{aligned}&\mathbf{u}_k\in\left(T_1^{\varepsilon_{1,k}}\times\cdots\times T_m^{\varepsilon_{m,k}}\right)(\tilde{\mathbf{x}}_k),\quad\mathbf{u}_k+\tilde{\mathbf{x}}_k-\mathbf{x}_{k-1}=0,\quad\varepsilon_k\leq\frac{\sigma^2}{2}\|\tilde{\mathbf{x}}_k-P_{\mathbf{V}}(\mathbf{x}_{k-1})\|^2,\\ &\mathbf{x}_k=P_{\mathbf{V}}(\tilde{\mathbf{x}}_k)+P_{\mathbf{V}^\perp}(\mathbf{u}_k).\end{aligned} \tag{79}$$

As a consequence of (79), *the sequences* $\{\mathbf{x}_k\}_{k\geq0}$, $\{\tilde{\mathbf{x}}_k\}_{k\geq1}$, $\{\mathbf{u}_k\}_{k\geq1}$ *and* $\{\varepsilon_k\}_{k\geq1}$ *are generated by* Algorithm 3 *for solving* (39) *with* $\mathbf{T}$ *and* $\mathbf{V}$ *given in* (74) *and* (75), *respectively.*

Proof Note that (79) follows directly from (53), (54), (78) and definition (1) (with $S_i=T^{\varepsilon_{i,k}}$ for $i=1,\dots,m$). The last statement of the Proposition is a direct consequence of (79) and Algorithm 3's definition. □

Proof of Theorem 4.1. We start by defining the ergodic sequences associated to the sequences $\{\tilde{\mathbf{x}}_k\}$, $\{\mathbf{u}_k\}$ and $\{\varepsilon_k\}$ in (78):

$$\begin{aligned} \tilde{\mathbf{x}}_k^a &:= \frac{1}{k}\sum_{\ell=1}^{k}\tilde{\mathbf{x}}_\ell\,, \quad \mathbf{u}_k^a := \frac{1}{k}\sum_{\ell=1}^{k}\mathbf{u}_\ell\,, \\ \varepsilon_k^a &:= \frac{1}{k}\sum_{\ell=1}^{k}\left[\varepsilon_\ell + \langle \tilde{\mathbf{x}}_\ell - \tilde{\mathbf{x}}_k^a, \mathbf{u}_\ell - \mathbf{u}_k^a\rangle\right]. \end{aligned} \tag{80}$$

Note that from (27), (51), (74), (75) and (76) we obtain $\mathcal{S}^*(\mathbf{V},\mathbf{T}) = \mathcal{S}^*(\Sigma)$ and, consequently, $d_{0,\mathbf{V}} = d_{0,\Sigma}$. That said, it follows from the last statement in Proposition 4.1, Proposition 3.1(a) and Theorem 3.1 that for any $k \geq 1$, there exists $j \in \{1,\dots,k\}$ such that

$$\begin{aligned} &\mathbf{u}_j \in \left(T_1^{\varepsilon_{1,j}} \times T_2^{\varepsilon_{2,j}} \times \cdots \times T_m^{\varepsilon_{m,j}}\right)(\tilde{\mathbf{x}}_j), \\ &\sqrt{\|\tilde{\mathbf{x}}_j - P_{\mathbf{V}}(\tilde{\mathbf{x}}_j)\|^2 + \|\mathbf{u}_j - P_{\mathbf{V}^\perp}(\mathbf{u}_j)\|^2} \leq \frac{d_{0,\Sigma}}{\sqrt{k}}\sqrt{\frac{1+\sigma}{1-\sigma}}, \quad \varepsilon_j \leq \frac{\sigma^2 d_{0,\Sigma}^2}{2(1-\sigma^2)k}\,, \end{aligned} \tag{81}$$

and

$$\sqrt{\|\tilde{\mathbf{x}}_k^a - P_{\mathbf{V}}(\tilde{\mathbf{x}}_k^a)\|^2 + \|\mathbf{u}_k^a - P_{\mathbf{V}^\perp}(\mathbf{u}_k^a)\|^2} \leq \frac{2d_{0,\Sigma}}{k}, \quad 0 \leq \varepsilon_k^a \leq \frac{2(1+\sigma/\sqrt{1-\sigma^2})d_{0,\Sigma}^2}{k}. \tag{82}$$

In particular, we see that Item (a) of Theorem 4.1 follows from (81), (78) and (77). Note now that from (80), (78) and (56) we obtain, for all $k \geq 1$,

$$\tilde{\mathbf{x}}_k^a = (\tilde{x}_{1,k}^a, \tilde{x}_{2,k}^a, \dots, \tilde{x}_{m,k}^a), \quad \mathbf{u}_k^a = (u_{1,k}^a, u_{2,k}^a, \dots, u_{m,k}^a), \quad \varepsilon_k^a = \sum_{i=1}^{m}\varepsilon_{i,k}^a. \tag{83}$$

Hence, the inequalities in (58) follow from (82), (83) and (77). To finish the proof of the theorem it suffices to show the inclusions in (58) for each $i = 1,\dots,m$ and all $k \geq 1$. To this end, note that for each $i = 1,\dots,m$ the desired inclusion is a direct consequence of the inclusions in (53), the definitions in (56) and Theorem 2.1 (with $T = T_i$ for each $i = 1,\dots,m$). □

4.2 Proof of Theorem 4.4

Next proposition shows that Algorithm 5 is a special instance of Algorithm 4 for solving (49) with $T_i = \nabla(\lambda f_i) + \partial(\lambda\varphi_i)$ for all $i = 1,\dots,m$.

Proposition 4.2 *Let $\{x_k\}_{k\geq 0}$ and, for $i = 1,\dots,m$, $\{y_{i,k}\}_{k\geq 0}$ and $\{\tilde{x}_{i,k}\}_{k\geq 1}$ be generated by* Algorithm 5*. For $i = 1,\dots,m$, consider the sequences $\{u_{i,k}\}_{k\geq 1}$ and $\{\varepsilon_{i,k}\}_{k\geq 1}$ where, for all $k \geq 1$,*

$$\begin{aligned} u_{i,k} &:= x_{k-1} + y_{i,k-1} - \tilde{x}_{i,k}, \\ \varepsilon_{i,k} &:= \lambda\left[f_i(\tilde{x}_{i,k}) - f_i(x_{k-1}) - \langle \nabla f_i(x_{k-1}), \tilde{x}_{i,k} - x_{k-1}\rangle\right]. \end{aligned} \tag{84}$$

Then, for all $k \geq 1$,

$$\nabla(\lambda f_i)(x_{k-1}) \in \partial_{\varepsilon_{i,k}}(\lambda f_i)(\tilde{x}_{i,k}), \tag{85}$$

$$u_{i,k} - \nabla(\lambda f_i)(x_{k-1}) \in \partial(\lambda\varphi_i)(\tilde{x}_{i,k}), \tag{86}$$

$$u_{i,k} \in \left(\partial_{\varepsilon_{i,k}}(\lambda f_i) + \partial(\lambda\varphi_i)\right)(\tilde{x}_{i,k}), \tag{87}$$

$$0 \leq \varepsilon_{i,k} \leq \frac{\sigma^2}{2}\|\tilde{x}_{i,k} - x_{k-1}\|^2, \tag{88}$$

$$x_k \text{ and } y_{i,k} \text{ satisfy } (54). \tag{89}$$

As a consequence of (84)–(89)*, the sequences $\{x_k\}_{k\geq 0}$, $\{y_{i,k}\}_{k\geq 1}$, $\{\tilde{x}_{i,k}\}_{k\geq 0}$, $\{\varepsilon_{i,k}\}_{k\geq 1}$ and $\{u_{i,k}\}_{k\geq 1}$ are generated by* Algorithm 4 *for solving* (49) *with*

$$T_i = \nabla(\lambda f_i) + \partial(\lambda\varphi_i) \quad \forall i = 1,\dots,m.$$

Proof Inclusion (85) follows from Lemma 2.2 with $(f,x,\tilde{x},v,\varepsilon) = (\lambda f_i, x_{k-1}, \tilde{x}_{i,\,k}, \nabla(\lambda f_i)(x_{k-1}), \varepsilon_{i,\,k})$, where $\varepsilon_{i,\,k}$ is given in (84). Inclusion (86) follows from (65), the first identity in (84) and Lemma 2.3(a). Inclusion (87) is a direct consequence of (85) and (86). The inequalities in (88) follow from assumption (A.1), the second identity in (84), Lemma 2.4 and the definition of $\lambda > 0$ in Algorithm 5. The fact that x_k satisfies (54) follows from the first identities in (54) and (66). Direct use of (66) and the assumption that $y_{1,0} + \cdots + y_{m,0} = 0$ in step 0 of Algorithm 5 gives $\sum_{\ell=1}^{m} y_{\ell,\,k} = 0$ for all $k \geq 0$, which, in turn, combined with the second identity in (66) and the first identity in (84) proves that $y_{i,\,k}$ satisfies the second identity in (54). Altogether, we obtain (89). The last statement of the proposition follows from (84)–(89) and Proposition 2.1(b; e). □

Proof of Theorem 4.4. From the last statement of Proposition 4.2, the fact that

$$\left(\sum_{i=1}^{m} [\nabla f_i + \partial \varphi_i]\right)^{-1}(0) = \left(\sum_{i=1}^{m} [\nabla(\lambda f_i) + \partial(\lambda \varphi_i)]\right)^{-1}(0)$$

and Theorem 4.1 we obtain that (57) and (58) hold. As a consequence of the latter fact, (87), (67), Lemma 2.4(b), the fact that $\lambda = \sigma^2/L_\Sigma$ and some direct calculations we obtain (68) and the inequalities in (69). To finish the proof, it suffices to prove the inclusion in (69). To this end, note first that from (85), (56), the last identity in (67), Lemma 2.3(b) and Theorem 2.1(b) we obtain, for each $i = 1, \ldots, m$,

$$\frac{1}{k}\sum_{s=1}^{k} \nabla f_i(x_{s-1}) \in \partial_{\varepsilon''^{\,a}_{i,\,k}} f_i(\tilde{x}^{\,a}_{i,\,k}) \quad \forall k \geq 1. \tag{90}$$

On the other hand, it follows from (86), Lemma 2.3(a), (67), Theorem 2.1(b) and some direct calculations that, for each $i = 1, \ldots, m$,

$$u'^{\,a}_{i,\,k} - \frac{1}{k}\sum_{s=1}^{k} \nabla f_i(x_{s-1}) \in \partial_{\left(\varepsilon'^{\,a}_{i,\,k} - \varepsilon''^{\,a}_{i,\,k}\right)} \varphi_i(\tilde{x}^{\,a}_{i,\,k}) \quad \forall k \geq 1, \tag{91}$$

which, in turn, combined with (90) gives the inclusion in (69). □

5 Conclusions

We proposed and analyzed the iteration-complexity of an inexact version of the Spingar's partial inverse method and, as a consequence, we obtained the iteration-complexity of an inexact version of the Spingarn's operator splitting method as well as of a parallel forward-backward method for multi-term composite convex optimization. We proved that our method falls in the framework of the hybrid proximal extragradient (HPE) method, for which the iteration-complexity has been obtained recently by Monteiro and Svaiter. We also introduced a notion of approximate solution for the Spingarn's problem (which generalizes the one introduced by Monteiro and Svaiter for monotone inclusions) and proved the iteration-complexity for the above mentioned methods based on this notion of approximate solution.

Acknowledgments

The work of S. C. L. was partially supported by CAPES Scholarship no. 201302186. The work of the second author was partially supported by CNPq grants no. 406250/2013-8, 237068/2013-3 and 306317/2014-1.

References

1. Spingarn, J. E.: Partial inverse of a monotone operator. Appl. Math. Optim. 10(3), 247–265 (1983).
2. Rockafellar, R. T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optimization 14(5), 877–898 (1976).
3. Solodov, M.V., Svaiter, B.F.: A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. Set-Valued Anal. 7(4), 323–345 (1999).
4. Boţ, R. I., Csetnek, E. R.: A hybrid proximal-extragradient algorithm with inertial effects. Numer. Funct. Anal. Optim. 36(8), 951–963 (2015).

5. Ceng, L. C., Mordukhovich, B. S., Yao, J. C.: Hybrid approximate proximal method with auxiliary variational inequality for vector optimization. J. Optim. Theory Appl. 146(2), 267–303 (2010).
6. Eckstein, J., Silva, P. J. S.: A practical relative error criterion for augmented Lagrangians. Math. Program. 141(1-2, Ser. A), 319–348 (2013).
7. He, Y., Monteiro, R. D. C.: An accelerated HPE-type algorithm for a class of composite convex-concave saddle-point problems. SIAM J. Optim. 26(1), 29–56 (2016).
8. Iusem, A. N., Sosa, W.: On the proximal point method for equilibrium problems in Hilbert spaces. Optimization 59(8), 1259–1274 (2010).
9. Lotito, P. A., Parente, L. A. Solodov, M. V.: A class of variable metric decomposition methods for monotone variational inclusions. J. Convex Anal. 16(3-4), 857–880 (2009).
10. Monteiro, R. D. C., Ortiz, C., Svaiter, B. F.: Implementation of a block-decomposition algorithm for solving large-scale conic semidefinite programming problems. Comput. Optim. Appl. 57(1), 45–69 (2014).
11. Monteiro, R. D. C., Ortiz, C., Svaiter, B. F.: An adaptive accelerated first-order method for convex optimization. Comput. Optim. Appl. 64(1), 31–73 (2016).
12. Monteiro, R. D. C., Svaiter, B. F.: Iteration-complexity of block-decomposition algorithms and the alternating direction method of multipliers. SIAM J. Optim. 23(1), 475–507 (2013).
13. Solodov, M. V., Svaiter, B. F.: 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).
14. Solodov, M. V., Svaiter, B. F.: A unified framework for some inexact proximal point algorithms. Numer. Funct. Anal. Optim. 22(7-8), 1013–1035 (2001).
15. Monteiro, R. D. C., Svaiter, B. F.: On the complexity of the hybrid proximal extragradient method for the iterates and the ergodic mean. SIAM Journal on Optimization 20, 2755–2787 (2010).
16. Gonçalves, M.L.N., Melo, J.G., Monteiro, R.D.C.: Improved pointwise iteration-complexity of a regularized ADMM and of a regularized non-euclidean HPE framework. Manuscript, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA, January 2016.
17. Burachik, R. S., Sagastizábal, C.A., Scheimberg, S.: An inexact method of partial inverses and a parallel bundle method. Optim. Methods Softw. 21(3), 385–400 (2006).
18. Burachik, R. S., Iusem, A. N., Svaiter, B. F.: Enlargement of monotone operators with applications to variational inequalities. Set-Valued Anal. 5(2), 159–180 (1997).
19. Martinez-Legaz, J.-E., Théra, M.: ϵ-subdifferentials in terms of subdifferentials. Set-Valued Anal. 4(4), 327–332 (1996).
20. Burachik, R. S., Sagastizábal, C. A., Svaiter, B. F.: ϵ-enlargements of maximal monotone operators: theory and applications. In Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods (Lausanne, 1997) volume 22 of Appl. Optim., 25–43. Kluwer Acad. Publ., Dordrecht (1999).
21. Monteiro, R. D. C., Svaiter, B. F.: Convergence rate of inexact proximal point methods with relative error criteria for convex optimization. *Manuscript*, 2010.
22. Nesterov, Y.: Introductory Lectures on Convex Optimization. A Basic Course. Kluwer Academic Publishers, Boston (2004).
23. Monteiro, R. D. C., Svaiter, B. F.: Complexity of variants of Tseng's modified F-B splitting and Korpelevich's methods for hemivariational inequalities with applications to saddle point and convex optimization problems. SIAM Journal on Optimization 21, 1688–1720 (2010).
24. Solodov, M. V., Svaiter, B. F.: A hybrid projection-proximal point algorithm. J. Convex Anal. 6(1), 59–70 (1999).
25. Alghamdi, M. A. , Alotaibi, A., Combettes, P. L., Shahzad, N.: A primal-dual method of partial inverses for composite inclusions. Optim. Lett. 8(8), 2271–2284 (2014).
26. Ouorou, A: Epsilon-proximal decomposition method. Math. Program. 99(1, Ser. A), 89–108 (2004).
27. Bauschke, H. H., Boţ, R. I., Hare, W. L., Moursi, W. M.: Attouch-Théra duality revisited: paramonotonicity and operator splitting. J. Approx. Theory 164(8), 1065–1084 (2012).
28. Bauschke, H. H., Combettes, P. L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2011).
29. Briceño-Arias, L. M., Combettes, P. L.: A monotone + skew splitting model for composite monotone inclusions in duality. SIAM J. Optim. 21(4), 1230–1250 (2011).
30. Eckstein, J., Svaiter, B. F.: A family of projective splitting methods for the sum of two maximal monotone operators. Math. Program. 111(1-2, Ser. B), 173–199 (2008).
31. Eckstein, J., Svaiter, B. F.: General projective splitting methods for sums of maximal monotone operators. SIAM J. Control Optim. 48(2), 787–811 (2009).
32. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems Vol. I. Springer-Verlag, New York (2003).