

An inexact interior point proximal method for the variational inequality problem

REGINA S. BURACHIK^{1*}, JURANDIR O. LOPES^{2†}
and GECI J.P. DA SILVA^{3‡}

¹School of Mathematics and Statistics, University of South Australia
Mawson Lakes, S.A. 5095, Australia

²Universidade Federal do Piauí, CCN-UFPI, CP 64.000, 64000-000 Teresina, PI, Brazil

³Universidade Federal de Goiás, IME/UFG, CP 131, 74001-970 Goiânia, GO, Brazil

E-mails: regina.burachik@unisa.edu.au / jurandir@ufpi.br / geci@mat.ufg.br

Abstract. We propose an infeasible interior proximal method for solving variational inequality problems with maximal monotone operators and linear constraints. The interior proximal method proposed by Auslender, Teboulle and Ben-Tiba [3] is a proximal method using a distance-like barrier function and it has a global convergence property under mild assumptions. However, this method is applicable only to problems whose feasible region has nonempty interior. The algorithm we propose is applicable to problems whose feasible region may have empty interior. Moreover, a new kind of inexact scheme is used. We present a full convergence analysis for our algorithm.

Mathematical subject classification: 90C51, 65K10, 47J20, 49J40, 49J52, 49J53.

Key words: maximal monotone operators, outer approximation algorithm, interior point method, global convergence.

1 Introduction

Let $C \subset \mathbb{R}^n$ be a closed and convex set, and $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone point-to-set operator. We consider the *variational inequality problem associated*

#744/07. Received: 27/XI/07. Accepted: 31/VII/08.

*This author acknowledges support by the Australian Research Council Discovery Project Grant DP0556685 for this study.

†Partially supported by CNPq.

‡Partially supported by IM-AGIMB.

with T and C : Find \bar{x} such that there exists $\bar{v} \in T(\bar{x})$ satisfying

$$\langle \bar{v}, y - \bar{x} \rangle \geq 0, \quad \text{for all } y \in C. \quad (1.1)$$

This problem is denoted by $VIP(T, C)$. In the particular case in which T is the subdifferential of a proper, convex and lower semicontinuous function $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$, (1.1) reduces to the *constrained convex optimization problem*: Find \bar{x} such that

$$f(\bar{x}) \leq f(y), \quad \text{for all } y \in C. \quad (1.2)$$

We are concerned with C a polyhedral set on \mathbb{R}^n defined by

$$C := \{x \in \mathbb{R}^n \mid Ax \leq b\}, \quad (1.3)$$

where A is an $m \times n$ real matrix, $b \in \mathbb{R}^m$ and $m \geq n$. Well-known methods for solving $VIP(T, C)$ are the so-called *generalized proximal schemes*, which involve a regularization term that incorporates the constraint set C in such a way that all the subproblems have solutions in the interior of C . For this reason, these methods are also called *interior proximal methods*. Examples of these regularizing functionals are the Bregman distances (see, e.g. [1, 8, 13, 14, 20, 25]), φ -divergences ([26, 5, 15, 18, 19, 27, 28]) and *log-quadratic* regularizations ([3, 4]). Being *interior point methods*, it is a basic assumption that the topological interior of C is nonempty. Otherwise, the iterates are not well-defined. However, a set C as above may usually have empty interior. In order to solve problem (1.2) for an arbitrary set $C \neq \emptyset$ of the kind given in (1.3), Yamasita et al. [29] devised an interior-point scheme in which the subproblems deal with a constraint set C^k given by

$$C^k := \{x \in \mathbb{R}^n \mid Ax \leq b + \delta^k\}, \quad (1.4)$$

where the vectors δ^k have positive coordinates and are such that $\sum_1^\infty \|\delta^k\| < \infty$. So, if $C \neq \emptyset$, it holds $C \subset \text{int } C^k$ and hence a regularizing functional can be associated with the set C^k . Denote by \bar{d}_k the regularization functional proposed in [3, 4] (associated with the set C^k with non-empty interior) and by $\nabla_1 \bar{d}_k$ the derivative of \bar{d}_k with respect to its first argument. The subproblems in [29] find an approximate solution $x^k \in \text{int } C^k$ of the inclusion

$$0 \in \lambda_k \partial_{\varepsilon_k} f(x^k) + \nabla_1 \bar{d}_k(x^k, x^{k-1}),$$

where $\lambda_k > 0$, $\partial_\varepsilon f$ is the ε -subdifferential of f [6]. Yamasita et al. prove in [29] convergence under summability assumptions on the “error” sequences $\{\varepsilon^k\}$ and $\{\delta^k\}$. One drawback of conditions of this kind is that there may be no constructive way to enforce them. Indeed, there exist infinitely many summable sequences, and it is not specified how to choose them at a specific iteration and for the given problem, so as to ensure convergence. From the algorithmic standpoint, one would prefer to have a computable error tolerance condition which is related to the progress of the algorithm at every given step when applied to the given problem. This is one of the main motivations of our approach (see condition (3.10) below), where we choose each ε^k so as to verify a specific condition at each iteration k .

Moreover, we also extend the scheme given in [29] to the more general problem (1.1). Namely, we are concerned with iterations of the kind: Find an approximate solution $x^k \in \text{int } C^k$ of

$$0 \in \lambda_k T^{\varepsilon_k}(x^k) + \nabla_1 \bar{d}_k(x^k, x^{k-1}), \quad (1.5)$$

where $\lambda_k > 0$ and T^ε is an enlargement of the operator T [11, 10]. We impose no summability assumption on the parameters $\{\varepsilon_k\}$. Instead, we define a criterion which can be checked at each iteration. On the other hand, we do need summability of the sequence $\{\delta_k\}$.

Our relative error analysis is inspired by the one given in [12], which yields a more practical framework. The convergence analysis presented in [29] (which considers the optimization problem (1.2)) requires an assumption involving the sequence of iterates generated by the method, and the function f , namely that $\sum_{k=0}^{\infty} |f(x^k) - f(P_C(x^k))| < +\infty$, where P_C stands for the orthogonal projection onto C . We make no assumptions of this kind in our analysis. Another difference between [29] and the present paper is that we allow more degrees of freedom in the definition of the inexact step. See Remark 3.6 for a detailed comparison with [29].

The paper is organized as follows. In Section 2 we give some basic definitions and properties of the family of regularizations, as well as some known results on monotone operators. In the same section, the enlargement T^ε is reviewed, together with its elementary features. In Section 3, we describe the

algorithm, prove its well-definedness and give its inexact version. The convergence analysis is presented in Section 3.1, and in Section 4 we give some conclusions.

2 Basic assumptions and properties

A point-to-set valued map $T: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is an operator which associates with each point $x \in \mathbb{R}^n$ a set (possibly empty) $T(x) \subset \mathbb{R}^n$. The domain and the graph of a point-to-set valued map T are defined as:

$$D(T) := \{x \in \mathbb{R}^n \mid T(x) \neq \emptyset\},$$

$$G(T) := \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \in D(T), v \in T(x)\}.$$

A point-to-set operator T is said to be *monotone* if

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

A monotone operator T is said to be *maximal* when its graph is not properly contained in the graph of any other monotone operator. The well-known result below has been proved in [24, Theorem 1]. Denote by $\text{ir } A$ the relative interior of the set A .

Proposition 2.1. *Let T_1, T_2 maximal monotone operators. If $\text{ir } D(T_1) \cap \text{ir } D(T_2) \neq \emptyset$, then $T_1 + T_2$ is maximal monotone.*

We denote by $\text{dom}(f) = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$ the domain of $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ and by f_∞ the asymptotic function [2, Definition 2.5.1] associated with the function $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$.

It is well-known that the existence of solutions of inclusion (1.5) depends on the properties of the distance \bar{d}_k . For a given distance D , a coercivity property (namely surjectivity of $\nabla_1 D(\cdot, y)$ for y fixed) is required (see, for instance [8, Proposition 3]). The result we need to ensure well-definedness of our scheme, which we state below, is [3, Proposition 3.1], which establishes the desired surjectivity in our particular setting.

Theorem 2.2 ([3, Proposition 3.1]). *Let $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function with $\text{dom}(f)$ open. Assume that f is differentiable on $\text{dom}(f)$ and such that $f_\infty(\xi) = +\infty \forall \xi \neq 0$. Let A be an $m \times n$ matrix with $m \geq n$ and $\text{rank } A = n$, $\tilde{b} \in \mathbb{R}^m$ with $(\tilde{b} - A(\mathbb{R}^n)) \cap \text{dom}(f) \neq \emptyset$, and set $h(x) := f(\tilde{b} - Ax)$. Let $\tilde{T}: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone operator such that $D(\tilde{T}) \cap \text{dom}(h) \neq \emptyset$ and set*

$$U(x) := \begin{cases} \tilde{T}(x) + \nabla h(x) & \text{if } x \in D(\tilde{T}) \cap D(\nabla h), \\ \emptyset & \text{otherwise.} \end{cases}$$

Then $\nabla h(x)$ is onto. Moreover, there exist a solution x of equation $0 \in U(x)$, which is unique if f is strictly convex on its domain.

We describe below the family of regularizations we use. From now on, the function $\varphi: \mathbb{R}_+ \rightarrow (-\infty, \infty]$ is given by

$$\varphi(t) := \mu h(t) + \left(\frac{\nu}{2}\right)(t - 1)^2, \tag{2.1}$$

where h is a closed and proper convex function satisfying the following additional properties:

- (1) h is twice continuously differentiable on $\text{int}(\text{dom } h) = (0, +\infty)$,
- (2) h is strictly convex on its domain,
- (3) $\lim_{t \rightarrow 0^+} h'(t) = -\infty$,
- (4) $h(1) = h'(1) = 0$ and $h''(1) > 0$, and
- (5) for $t > 0$

$$h''(1) \left(1 - \frac{1}{t}\right) \leq h'(t) \leq h''(1)(t - 1). \tag{2.2}$$

Items (1)–(4) and (1)–(5) were used in [4] to define, respectively, the families Φ and Φ_2 . The positive parameters μ, ν shall satisfy the following inequality

$$\nu > \mu h''(1) > 0. \tag{2.3}$$

Note that conditions above and (2.2) imply

$$\mu h''(1) \left(1 - \frac{1}{t}\right) + \nu(t - 1) \leq \varphi'(t) < 2\nu(t - 1), \tag{2.4}$$

therefore $\lim_{t \rightarrow \infty} \varphi'(t) = +\infty$. The generalized distance induced by φ , is denoted by $d_\varphi(x, y)$ and defined as:

$$d_\varphi(x, y) := \begin{cases} \sum_{i=1}^n y_i^2 \varphi(x_i/y_i), & x, y \in \mathbb{R}_{++}^n, \\ +\infty, & \text{otherwise,} \end{cases} \quad (2.5)$$

where $\mathbb{R}_{++}^n := \{z \in \mathbb{R}^n \mid z_i > 0 \ \forall i = 1, \dots, n\}$. Since $\lim_{t \rightarrow \infty} \varphi'(t) = +\infty$ it follows that $[d_\varphi(\cdot, y)]_\infty(\xi) = +\infty, \ \forall \xi \neq 0$. Denoting by ∇_1 the gradient with respect to the first variable, it holds that $[\nabla_1 d_\varphi(x, y)]_i = y_i \varphi'(x_i/y_i)$ for all $i = 1, \dots, n$.

The following lemma has a crucial role in the convergence analysis. Its first part has been established in [3]. Define

$$\theta := v + \rho\mu, \quad \tau := v - \rho\mu \quad \text{and} \quad \rho := h''(1). \quad (2.6)$$

Lemma 2.3. *For all $w, z \in \mathbb{R}_{++}^n$ and $v \in \mathbb{R}_+^n := \{z \in \mathbb{R}^n \mid z_i \geq 0 \ \forall i = 1, \dots, n\}$, it holds that*

- (i) $\langle \nabla_1 d_\varphi(w, z), w - v \rangle \geq \frac{\theta}{2} (\|w - v\|^2 - \|z - v\|^2) + \frac{\tau}{2} \|w - z\|^2;$
- (ii) $\langle v, \nabla_1 d_\varphi(w, z) \rangle \leq \theta \|v\| \|w - z\|.$

Proof. For part (i), see [3, Lemma 3.4]. We proceed to prove (ii). Since $\varphi(t) = \mu h(t) + \frac{v}{2}(t-1)^2$, we have that $\varphi'(t) = \mu h'(t) + v(t-1)$. By (2.2) and (2.6) we get $\varphi'(t) \leq (v + \rho\mu)(t-1)$. Letting $t = \frac{w_i}{z_i}$ and multiplying both sides by $v_i z_i$ yield

$$v_i z_i \varphi' \left(\frac{w_i}{z_i} \right) \leq \theta v_i z_i \left(\frac{w_i}{z_i} - 1 \right),$$

for all $i = 1, \dots, n$. Therefore, $\langle v, \nabla_1 d_\varphi(w, z) \rangle \leq \theta \langle v, w - z \rangle$. Using the Cauchy-Schwartz inequality in the expression above, we get (ii). \square

The result below is known as Hoffman's lemma [16].

Lemma 2.4. *Let $C = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ and $C^k = \{x \in \mathbb{R}^n \mid Ax \leq b + \delta^k\}$ where A is matrix $m \times n$ with $m \geq n$ and $b, \delta^k \in \mathbb{R}^m$. Given $x^k \in C^k$ there exists a constant $\alpha > 0$ such that*

$$\text{dist}(x^k, C) := \inf_{y \in C} \|y - x^k\| = \|p^k - x^k\| \leq \alpha \|\delta^k\|$$

where p^k is projection of x^k in C .

We recall next two technical results on nonnegative sequences of real numbers. The first one was taken from [22, Chapter 2] and the second from [21, Lemma 3.5].

Lemma 2.5. *Let $\{\sigma_k\}$ and $\{\beta_k\}$ be nonnegative sequences of real numbers satisfying:*

- (i) $\sigma_{k+1} \leq \sigma_k + \beta_k$;
- (ii) $\sum_{k=1}^{\infty} \beta_k < \infty$.

Then the sequence $\{\sigma_k\}$ converges.

Lemma 2.6. *Let $\{\lambda_k\}$ be a sequence of positive numbers, and $\{a_k\}$ be a sequence of real numbers. Let $\sigma_k := \sum_{j=1}^k \lambda_j$ and $b_k := \sigma_k^{-1} \sum_{j=1}^k \lambda_j a_j$. If $\sigma_k \rightarrow \infty$, then*

- (i) $\liminf_{k \rightarrow \infty} a_k \leq \liminf_{k \rightarrow \infty} b_k \leq \limsup_{k \rightarrow \infty} b_k \leq \limsup_{k \rightarrow \infty} a_k$;
- (ii) If $\lim_{k \rightarrow \infty} a_k = a < \infty$, then $\lim_{k \rightarrow \infty} b_k = a$.

In our analysis, we relax the inclusion $v^k \in T(x^k)$, by means of an ε -enlargement of the operator T introduced in [10]: Given T a monotone operator, define

$$T^\varepsilon(x) := \{v \in \mathbb{R}^N \mid \langle v - w, x - y \rangle \geq -\varepsilon \forall y \in \mathbb{R}^N, w \in T(y)\}. \quad (2.7)$$

This extension has many useful properties, similar to the ε -subdifferential of a proper closed convex function f . Indeed, when $T = \partial f$, we have

$$\partial_\varepsilon f(x) \subseteq T^\varepsilon(x) = (\partial f)^\varepsilon(x). \quad (2.8)$$

For an arbitrary maximal monotone operator T , the relation $T^0(x) = T(x)$ holds trivially. Furthermore, for $\varepsilon' \geq \varepsilon \geq 0$, we have $T^\varepsilon(x) \subset T^{\varepsilon'}(x)$. In particular, for each $\varepsilon \geq 0$, $T(x) \subset T^\varepsilon(x)$ (see [9, Chapter 5] for a detailed study of the properties of T^ε).

3 The algorithm

In this section, we propose an infeasible interior proximal method for the solution of $VIP(T, C)$ (1.1). To state formally our algorithm, we consider

$$C^k = \{x \in \mathbb{R}^n \mid Ax \leq b + \delta^k\} \text{ where } \delta^k \in \mathbb{R}_{++}^m \text{ and } \sum_{k=1}^{\infty} \|\delta^k\| < \infty,$$

which is considered a perturbation of the original constraint set C . Moreover, if $C \neq \emptyset$, then $C \subset \text{int } C^k \neq \emptyset$ for all k . Since $\delta^k \rightarrow 0$ as $k \rightarrow \infty$, the sequence of sets $\{C^k\}$ converges to the set C . Now, if a_i denotes the row i of the matrix A , for each $x \in C^k$ we consider $y^k(x) = (y_1^k(x), y_2^k(x), \dots, y_m^k(x))^T$, where

$$y_i^k(x) = b_i + \delta_i^k - \langle a_i, x \rangle \quad \text{with } i = 1, 2, \dots, m.$$

Therefore, we have the function $\bar{d}_k : \text{int } C^k \times \text{int } C^k \rightarrow \mathbb{R}$ defined by

$$\bar{d}_k(x, z) = d_\varphi(y^k(x), y^k(z)). \quad (3.1)$$

From the definition of d_φ , for each $x^k \in \text{int } C^k$, $x^{k-1} \in \text{int } C^{k-1}$, we have

$$\nabla_1 \bar{d}_k(x^k, x^{k-1}) = -A^T \nabla_1 d_\varphi(y^k(x^k), y^k(x^{k-1})). \quad (3.2)$$

In the method proposed in [29] for the convex optimization problem (1.2) with C defined as in (1.3), the exact algorithm of the iteration k is given by:

For $\lambda_k > 0$, $\delta^k > 0$ and $(x^{k-1}, y^{k-1}) \in \text{int } C^{k-1} \times \mathbb{R}_{++}^m$, find $(x, y) \in \text{int } C^k \times \mathbb{R}_{++}^m$ and $u \in \mathbb{R}^n$ such that

$$\begin{cases} u \in \partial f(x), \\ \lambda_k u + \nabla_1 \bar{d}_k(x, x^{k-1}) = 0, \\ y - (b - Ax) = \delta^k, \end{cases}$$

where $y \in \mathbb{R}_{++}^m$ can be seen as a slack variable associated to $x \in \text{int } C^k$. The corresponding *inexact* iteration in [29] is given by:

$$\begin{cases} \tilde{u} \in \partial_{\varepsilon_k} f(\tilde{x}), \\ \lambda_k \tilde{u} + \nabla_1 \bar{d}_k(\tilde{x}, x^{k-1}) = 0, \\ \tilde{y} - (b - A\tilde{x}) = \delta^k. \end{cases} \quad (3.3)$$

Following the approach in (3.3), the exact version of our algorithm is obtained replacing ∂f by an arbitrary maximal monotone operator T . Namely, given $\lambda_k > 0, \delta^k > 0$ and $(x^{k-1}, y^{k-1}) \in \text{int } C^{k-1} \times \mathbb{R}_{++}^m$, find $(x, y) \in \text{int } C^k \times \mathbb{R}_{++}^m$ and $u \in \mathbb{R}^n$ such that

$$\begin{cases} u \in T(x), \\ \lambda_k u + \nabla_1 \bar{d}_k(x, x^{k-1}) = 0, \\ y - (b - Ax) = \delta^k. \end{cases} \quad (3.4)$$

A detailed comparison with the method in [29] is given in Remark 3.6.

It is important to guarantee the existence of $(x^k, y^k) \in \text{int } C^k \times \mathbb{R}_{++}^m$ satisfying (3.4). In fact, the next proposition shows that there exists a unique pair $(x^k, y^k) \in \text{int } C^k \times \mathbb{R}_{++}^m$ satisfying (3.4) under the following two assumptions:

(H₁) $\text{int } C \cap \text{int } D(T) \neq \emptyset$;

(H₂) $\text{rank}(A) = n$ (and therefore, A injective).

Proposition 3.1. *Suppose that (H₁) and (H₂) hold. For every $\lambda_k > 0, \delta^k > 0$ and $(x^{k-1}, y^{k-1}) \in \text{int } C^k \times \mathbb{R}_{++}^m$, there exists a unique pair $(x^k, y^k) \in \text{int } C^k \times \mathbb{R}_{++}^m$ satisfying (3.4).*

Proof. Define the operator $\tilde{T}^k(x) := T(x) + N_{C^k}(x) + \lambda_k^{-1} \nabla h(x)$, where $h := \bar{d}_k(\cdot, x^{k-1})$. We prove first that we are in the conditions of Theorem 2.2 for $\tilde{T} := T + N_{C^k}$, $f(\cdot) := d_\varphi(\cdot, y^{k-1})$ and $\tilde{b} := b + \delta^k$. Indeed, the operator $T + N_{C^k}$ is maximal monotone by (H₁) and the fact that $C \subseteq C^k$ (we are using here Proposition 2.1). The function $d_\varphi(\cdot, y^{k-1})$ is by definition convex, proper and differentiable on its (open) domain \mathbb{R}_{++}^m and $[d_\varphi(\cdot, y^{k-1})]_\infty(\xi) = +\infty$,

$\forall \xi \neq 0$. By (H_2) , A has maximal rank. We claim that $(b + \delta^k - A(\mathbb{R}^n)) \cap \text{dom } d_\varphi(\cdot, y^{k-1}) \neq \emptyset$. Indeed, fix $x \in C$. It holds that

$$b + \delta^k - Ax \geq \delta^k > 0, \tag{3.5}$$

and therefore $b + \delta^k - Ax \in \mathbb{R}_{++}^n = \text{dom } d_\varphi(\cdot, y^{k-1})$.

The only hypothesis that remains to be checked is: $D(\tilde{T}) \cap \text{dom}(h) \neq \emptyset$, where $\text{dom}(h) = \text{int } C^k$. Indeed, by (H_1) and by definition of the C^k we get

$$\emptyset \neq C \cap D(T) \subset \text{int } C^k \cap D(T) \subset D(\tilde{T}).$$

Hence $\emptyset \neq C \cap D(\tilde{T}) \subset D(\tilde{T}) \cap \text{int } C^k = D(\tilde{T}) \cap \text{dom}(h)$. So the hypotheses of Theorem 2.2 are satisfied and therefore there exists x^* a solution of the equation

$$0 \in T(x) + N_{C^k}(x) + \lambda_k^{-1} \nabla_1 \bar{d}_k(x, x^{k-1}). \tag{3.6}$$

This solution is unique, because $d_\varphi(\cdot, y^{k-1})$ is strictly convex on its domain. So, there exists $u^k \in T(x^k)$, $v^k \in N_{C^k}(x^k)$ and $z^k = \nabla_1 \bar{d}_k(x^k, x^{k-1}) = \nabla_1 d_\varphi(b + \delta^k - A(x^k), y^{k-1})$ such that

$$0 = u^k + v^k + \lambda_k^{-1} z^k. \tag{3.7}$$

Taking $b + \delta^k - Ax^k =: y^k$ we have that y^k is also unique. Since $y^k \in \mathbb{R}_{++}^m$, it holds that $x^k \in \text{int } C^k$, thus $v^k = 0$. Hence by (3.7) there exists a unique pair $(x^k, y^k) \in \text{int } C^k \times \mathbb{R}_{++}^m$ satisfying

$$\begin{cases} u^k \in T(x^k), \\ u^k + \lambda_k^{-1} \nabla_1 \bar{d}_k(x^k, x^{k-1}) = 0, \\ y^k - (b - Ax^k) = \delta^k, \end{cases}$$

which completes the proof. □

Remark 3.2. We point out that the previous proposition can be established (with essentially the same proof) if we replace (H_1) by the weaker requirement $D(T) \cap \text{int}(C^k) \neq \emptyset$. We will need (H_1) , however, for proving that our iterates converge to a solution (see Theorem 3.11).

To deal with approximations, we relax the inclusion and the equation of the exact system (3.4) in a way similar to (3.3):

$$\begin{cases} \tilde{u} \in T^{\varepsilon_k}(\tilde{x}), \\ \lambda_k \tilde{u} + \nabla_1 \bar{d}_k(\tilde{x}, x^{k-1}) = e^k, \\ \tilde{y} - (b - A\tilde{x}) = \delta^k, \end{cases} \quad (3.8)$$

where T^ε is the enlargement of T given in (2.7). In the exact solution, we have $\varepsilon_k = 0$ and $e^k = 0$. An approximate solution should have ε_k and e^k “small”. Our aim is to use a relative error criteria as the one used in [12] to control the size of ε_k and e^k . The intuitive idea is to perform an extragradient step from x^{k-1} to x , using the direction \tilde{u} (see (3.9)), and then check whether the “error terms” of the iteration, given by $\varepsilon_k + \langle \tilde{u}, \tilde{x} - x \rangle$ and $\|\tilde{y} - y\|$ are small enough with respect to the previous step.

Definition 3.3. *Let $\sigma \in [0, 1)$ and $\gamma > 0$. We say that $(\tilde{x}, \tilde{y}, \tilde{u}, \varepsilon_k)$ in (3.8) is an approximated solution of system (3.4) with tolerance σ and γ if for (x, y) such that*

$$\begin{cases} \lambda_k \tilde{u} + \nabla_1 \bar{d}_k(x, x^{k-1}) = 0, \\ y - (b - Ax) = \delta^k. \end{cases} \quad (3.9)$$

it holds that

$$\lambda_k(\varepsilon_k + \langle \tilde{u}, \tilde{x} - x \rangle) \leq \sigma \frac{\tau}{2} \|y - y^{k-1}\|^2, \quad (3.10)$$

$$\|\tilde{y} - y\| \leq \gamma \|y - y^{k-1}\|, \quad (3.11)$$

where $\tau > 0$ is as in (2.6).

Remark 3.4.

- (i) Since the domain of $d_\varphi(\cdot, y^{k-1})$ is \mathbb{R}_{++}^m , for $\tilde{x}, \tilde{u}, \varepsilon_k$ and x as in Definition 3.3, it holds that $\tilde{x}, x \in \text{int } C^k$.
- (ii) If (x, y, u) verifies (3.4), then $(x, y, u, 0)$ is an approximated solution of system (3.4) with tolerance σ, γ for any $\sigma \in [0, 1)$ and $\gamma > 0$. It is clear that in this case $e^k = 0$. Conversely, if $(\tilde{x}, \tilde{y}, \tilde{u}, \varepsilon_k)$ is an approximated solution of system (3.4) with tolerance $\sigma = 0$ and $\gamma > 0$ arbitrary, then

we must have $\varepsilon_k = 0$ and $(\tilde{x}, \tilde{y}, \tilde{u})$ satisfying (3.4). Indeed, since $\gamma > 0$ is arbitrary, we get $y = \tilde{y}$. Using the fact that A is one-to-one, we get $x = \tilde{x}$. From the fact that $\sigma = 0$, we conclude that $\varepsilon_k = 0$.

- (iii) If (H_1) and (H_2) hold, by Proposition 3.1 the system (3.8) with $e^k = 0$ and $\varepsilon_k = 0$ has a solution. By (ii) of this remark, this solution is also an approximated solution.

We describe below our algorithm, called *Extragradient Algorithm*.

Extragradient Algorithm-EA

Initialize: Take $\bar{\rho}, \bar{\lambda} > 0, \sigma \in [0, 1), \gamma > 0, x^0 \in \mathbb{R}^n$ and $y^0 \in \mathbb{R}_{++}^m$ such that $\delta^0 := y^0 - (b - Ax^0) \in \mathbb{R}_{++}^m$.

Iteration: For $k = 1, 2, \dots$,

Step 1. Take λ_k with $\bar{\rho} \leq \lambda_k \leq \bar{\lambda}$ and $0 < \delta^k < \delta^{k-1}$. Find $(\tilde{x}^k, \tilde{y}^k, \tilde{u}^k, \varepsilon_k)$ an approximated solution of system (3.4) with tolerance σ, γ (i.e., they verify (3.8)).

Step 2. Compute (x^k, y^k) such that

$$\begin{cases} \lambda_k \tilde{u}^k + \nabla_1 \bar{d}_k(x^k, x^{k-1}) & = 0, \\ y^k - (b - Ax^k) & = \delta^k. \end{cases} \quad (3.12)$$

Step 3. Set $k := k + 1$, and return to Step 1.

Remark 3.5. The parameter $\bar{\rho} > 0$ ensures that the information on the original problem is taken into account at each iteration. The requirement $\bar{\rho} > 0$ is standard in the convergence analysis of proximal methods.

Remark 3.6. Our algorithm extends the one in [29]. More precisely, our step coincides with the one in [29] when the following hold.

- (i) $T = \partial f$.
- (ii) $e^k = 0$ in (3.8) (so $x^k = \tilde{x}^k$),
- (iii) Choose $\tilde{u}^k \in \partial_{\varepsilon_k} f(\tilde{x}^k) = \partial_{\varepsilon_k} f(x^k)$, instead of taking \tilde{u}^k in the (potentially) bigger set $(\partial f)^{\varepsilon_k}(\tilde{x}^k)$ as we do in (3.8).

From (ii) and (iii), our step allows more freedom in the choice of the next iterate x^k . As mentioned earlier, a conceptual difference with the method in [29] is the fact that the sequence $\{\varepsilon_k\}$ is chosen in a *constructive way*, so as to ensure convergence. Our choice of each ε_k is related with the progress of the algorithm at every given step when applied to the given problem. It can be seen that, if (i)–(iii) above hold, then (3.10) forces $\sum_{i=1}^{\infty} \lambda_k \varepsilon_k < +\infty$, the latter being an assumption for the convergence results in [29]. Indeed, if (i)–(iii) hold, then $x^k = \tilde{x}^k$ and (3.10) yields $\lambda_k \varepsilon_k \leq \frac{\sigma\tau}{2} \|y^k - y^{k-1}\|^2$. Using now Proposition 3.7, Corollary 3.8 (see the next section) and Lemma 2.5 we obtain $\sum_{i=1}^{\infty} \|y^k - y^{k-1}\|^2 < +\infty$. Therefore, $\sum_{i=1}^{\infty} \lambda_k \varepsilon_k < +\infty$.

3.1 Convergence analysis

In this section, we prove convergence of the Algorithm above. From now on $\{x^k\}$, $\{\tilde{x}^k\}$, $\{\tilde{y}^k\}$, $\{y^k\}$, $\{\tilde{u}^k\}$, $\{\varepsilon_k\}$, $\{\lambda_k\}$ and $\{\delta_k\}$ are sequences generated by EA with approximating criteria (3.10)–(3.11). The main result we shall prove is that the sequence $\{x^k\}$ converges to a solution of $VIP(T, C)$.

The next proposition is essential for the convergence analysis, to show this we need the following further assumptions

(H_3) The solution set S of $VIP(T, C)$ is nonempty.

Proposition 3.7. *Suppose that (H_3) holds and let $\bar{x} \in S$ and $\bar{u} \in T(\bar{x})$. Define $\bar{y} := b - A\bar{x}$. Then, for $k = 1, 2, \dots$,*

$$\begin{aligned} \|y^k - \bar{y}\|^2 &\leq \|y^{k-1} - \bar{y}\|^2 - \frac{\tau}{\theta}(1 - \sigma)\|y^k - y^{k-1}\|^2 \\ &\quad + 2\|\delta^k\| \|y^k - y^{k-1}\| + \alpha \frac{2}{\theta} \lambda_k \|\bar{u}\| \|\delta^k\|, \end{aligned} \quad (3.13)$$

where θ, τ are as in (2.6) and α is as in Lemma 2.4.

Proof. Fix $k > 0$ and take $\tilde{u}^k \in T^{\varepsilon_k}(\tilde{x}^k)$. For all $(x, u) \in G(T)$ we have that

$$\lambda_k \langle x - \tilde{x}^k, u - \tilde{u}^k \rangle \geq -\lambda_k \varepsilon_k.$$

Therefore,

$$\begin{aligned} \lambda_k \langle x - \tilde{x}^k, u \rangle &\geq \lambda_k \langle x - \tilde{x}^k, \tilde{u}^k \rangle - \lambda_k \varepsilon_k \\ &= \lambda_k \langle x - x^k, \tilde{u}^k \rangle + \lambda_k \langle x^k - \tilde{x}^k, \tilde{u}^k \rangle - \lambda_k \varepsilon_k. \end{aligned} \quad (3.14)$$

Using (3.9), (3.10) and (3.12) in the inequality above, we get

$$\lambda_k \langle x - \tilde{x}^k, u \rangle \geq \langle x - x^k, -\nabla_1 \bar{d}_k(x^k, x^{k-1}) \rangle - \sigma \frac{\tau}{2} \|y^k - y^{k-1}\|^2. \quad (3.15)$$

Now, using (3.2) we have

$$\begin{aligned} \langle x - x^k, -\nabla_1 \bar{d}_k(x^k, x^{k-1}) \rangle &= \langle x - x^k, A^T \nabla_1 d_\varphi(y^k, y^{k-1}) \rangle \\ &= \langle A(x - x^k), \nabla_1 d_\varphi(y^k, y^{k-1}) \rangle \\ &= \langle y^k - y, \nabla_1 d_\varphi(y^k, y^{k-1}) \rangle \\ &\quad - \langle \delta^k, \nabla_1 d_\varphi(y^k, y^{k-1}) \rangle, \end{aligned}$$

where $y = b - Ax$. Combining the equality above with (3.15), we get

$$\begin{aligned} \lambda_k \langle x - \tilde{x}^k, u \rangle &\geq \langle y^k - y, \nabla_1 d_\varphi(y^k, y^{k-1}) \rangle \\ &\quad - \langle \delta^k, \nabla_1 d_\varphi(y^k, y^{k-1}) \rangle - \sigma \frac{\tau}{2} \|y^k - y^{k-1}\|^2. \end{aligned}$$

Applying Lemma 2.3 in this inequality yields

$$\begin{aligned} \lambda_k \langle x - \tilde{x}^k, u \rangle &\geq \frac{\theta}{2} (\|y^k - y\|^2 - \|y^{k-1} - y\|^2) \\ &\quad + \frac{\tau}{2} (1 - \sigma) \|y^k - y^{k-1}\|^2 - \theta \|\delta^k\| \|y^k - y^{k-1}\|. \end{aligned} \quad (3.16)$$

The inequality above is valid in particular for $(x, u) := (\bar{x}, \bar{u})$ with $\bar{x} \in S$ and \bar{y} such that $\bar{y} = b - A\bar{x}$. Therefore,

$$\begin{aligned} \lambda_k \langle \bar{x} - \tilde{x}^k, \bar{u} \rangle &\geq \frac{\theta}{2} (\|y^k - \bar{y}\|^2 - \|y^{k-1} - \bar{y}\|^2) \\ &\quad + \frac{\tau}{2} (1 - \sigma) \|y^k - y^{k-1}\|^2 - \theta \|\delta^k\| \|y^k - y^{k-1}\|. \end{aligned} \quad (3.17)$$

On the other hand, for (\bar{x}, \bar{u}) with $\bar{x} \in S$ and $\bar{u} \in T(\bar{x})$, we have that $\langle \bar{x} - x, \bar{u} \rangle \leq 0 \forall x \in C$. Let p^k be the projection of \tilde{x}^k onto C . Since $p^k \in C$, we have that $\langle \bar{x} - p^k, \bar{u} \rangle \leq 0$, and therefore $\langle \bar{x} - \tilde{x}^k, \bar{u} \rangle \leq \langle p^k - \tilde{x}^k, \bar{u} \rangle$. Using the Cauchy-Schwarz inequality and multiplying by $\lambda_k > 0$, we get

$$\lambda_k \langle \bar{x} - \tilde{x}^k, \bar{u} \rangle \leq \lambda_k \|\bar{u}\| \|\tilde{x}^k - p^k\|.$$

By Lemma 2.4 we conclude that

$$\lambda_k \langle \bar{x} - \tilde{x}^k, \bar{u} \rangle \leq \lambda_k \alpha \|\bar{u}\| \|\delta^k\|, \quad (3.18)$$

for some $\alpha > 0$. Combining (3.17) and (3.18), we get

$$\begin{aligned} \|y^k - \bar{y}\|^2 &\leq \|y^{k-1} - \bar{y}\|^2 - \frac{\tau}{\theta}(1 - \sigma)\|y^k - y^{k-1}\|^2 \\ &\quad + 2\|\delta^k\| \|y^k - y^{k-1}\| + \alpha \frac{2}{\theta} \lambda_k \|\bar{u}\| \|\delta^k\|. \end{aligned} \quad (3.19)$$

□

The next corollary guarantees boundedness of the sequence $\{y^k - y^{k-1}\}$.

Corollary 3.8. *Suppose that (H_3) holds, then the sequence $\{\|y^k - y^{k-1}\|\}$ is bounded.*

Proof. Assume the sequence $\{\|y^k - y^{k-1}\|\}$ is unbounded. Then there is a subsequence $\{\|y^k - y^{k-1}\|\}_{k \in K}$ such that $\|y^k - y^{k-1}\| \rightarrow \infty$ for $k \in K$, whereas the complementary subsequence $\{\|y^k - y^{k-1}\|\}_{k \notin K}$ is bounded (note that this complementary subsequence could be finite or even empty). From (3.19), we have

$$\begin{aligned} \|y^k - y^{k-1}\| &\left[\frac{\tau}{\theta}(1 - \sigma)\|y^k - y^{k-1}\| - 2\|\delta^k\| \right] \\ &\leq \|y^{k-1} - \bar{y}\|^2 - \|y^k - \bar{y}\|^2 + \alpha \frac{2}{\theta} \lambda_k \|\bar{u}\| \|\delta^k\|. \end{aligned} \quad (3.20)$$

Summing up the inequalities (3.20) over $k = 1, 2, \dots, n$ gives

$$\begin{aligned} &\sum_{\substack{k=1, \dots, n \\ k \notin K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1 - \sigma)\|y^k - y^{k-1}\| - 2\|\delta^k\| \right] \\ &+ \sum_{\substack{k=1, \dots, n \\ k \in K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta}(1 - \sigma)\|y^k - y^{k-1}\| - 2\|\delta^k\| \right] \\ &\leq \|y^0 - \bar{y}\|^2 - \|y^n - \bar{y}\|^2 + \alpha \frac{2\tau}{\theta} \bar{\lambda} \|\bar{u}\| \sum_{k=1}^n \|\delta^k\| \\ &\leq \|y^0 - \bar{y}\|^2 + \alpha \frac{2\tau}{\theta} \bar{\lambda} \|\bar{u}\| \sum_{k=1}^n \|\delta^k\|. \end{aligned}$$

Set

$$a_n = \sum_{\substack{k=1, \dots, n \\ k \notin K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta} (1 - \sigma) \|y^k - y^{k-1}\| - 2 \|\delta^k\| \right],$$

$$b_n = \sum_{\substack{k=1, \dots, n \\ k \in K}} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta} (1 - \sigma) \|y^k - y^{k-1}\| - 2 \|\delta^k\| \right]$$

and

$$c_n = \|y^0 - \bar{y}\|^2 + \alpha \frac{2\tau}{\theta} \bar{\lambda} \|\bar{u}\| \sum_{k=1}^n \|\delta^k\|.$$

So the above inequality can be re-written as

$$a_n + b_n \leq c_n. \quad (3.21)$$

Because $\sum_{k=1}^n \|\delta^k\| < \infty$ we conclude that $c_n \leq \bar{c}$ for some \bar{c} . We show now that $\{a_n\}$ is bounded below and $\lim_{n \rightarrow \infty} b_n = \infty$, which is a contradiction with (3.21). This will complete the proof. Since $\{\|y^k - y^{k-1}\|\}_{k \notin K}$ is bounded, there is L such that $\|y^k - y^{k-1}\| \leq L$ for all $k \notin K$, then

$$-2L \|\delta^k\| \leq -2\|y^k - y^{k-1}\| \|\delta^k\| \leq \frac{\tau}{\theta} (1 - \sigma) \|y^k - y^{k-1}\|^2 - 2\|y^k - y^{k-1}\| \|\delta^k\|,$$

summing up the inequalities, we have

$$-2L \sum_{\substack{k=1, \dots, n \\ k \notin K}} \|\delta^k\| \leq a_n,$$

it follows that the sequence $\{a_n\}$ is bounded below because $\sum_{k=1}^n \|\delta^k\| < \infty$. Since in K the sequence $\{\|y^k - y^{k-1}\|\}$ is unbounded and $\{\|\delta^k\|\}$ converges to zero, there exists an $k_0 \in K$ such that

$$\frac{\tau}{\theta} (1 - \sigma) \|y^k - y^{k-1}\| - 2\|\delta^k\| \geq L > 0, \text{ for all } k > k_0$$

therefore,

$$\begin{aligned} & \sum_{\substack{k > k_0 \\ k \in K}}^{\infty} \|y^k - y^{k-1}\| \left[\frac{\tau}{\theta} (1 - \sigma) \|y^k - y^{k-1}\| - 2\|\delta^k\| \right] \\ & \geq L \sum_{\substack{k > k_0 \\ k \in K}}^{\infty} \|y^k - y^{k-1}\| = \infty \end{aligned}$$

it follows that $\lim_{n \rightarrow \infty} b_n = \infty$. \square

Remark 3.9. We point out that the requirement $\lambda_k \leq \bar{\lambda}$ used in Corollary 3.8 can be weakened to the assumption $\sum_{k=1}^n \|\delta^k\| \lambda_k < \infty$. We have chosen to use the stronger requirement $\lambda_k \leq \bar{\lambda}_k$ for simplicity of the presentation. The assumption $\lambda_k \leq \bar{\lambda}$ is not used in any of the remaining results.

Corollary 3.10. *Suppose that (H_3) holds. Then, for $\bar{x}, \bar{u}, \bar{y}$ as in Proposition 3.7, it holds that*

- (i) $\{\|y^k - \bar{y}\|\}$ converges (and hence $\{y^k\}$ is bounded);
- (ii) $\lim_k \|y^k - y^{k-1}\| = 0$;
- (iii) $\{\|A(x^k - \bar{x})\|\}$ converges (hence $\{\|x^k - \bar{x}\|\}$ converges and $\{x^k\}$ is bounded);
- (iv) $\lim_k \|\tilde{x}^k - x^k\| = 0$;
- (v) $\{\tilde{x}^k\}$ is bounded.

Proof.

- (i) From (3.13) we have that

$$\|y^k - \bar{y}\|^2 \leq \|y^{k-1} - \bar{y}\|^2 + 2\|\delta^k\| \|y^k - y^{k-1}\| + \alpha \frac{2}{\theta} \lambda_k \|\bar{u}\| \|\delta^k\| \forall k. \quad (3.22)$$

Define

$$\sigma_{k+1} := \|y^k - \bar{y}\|^2 \quad \text{and} \quad \beta_k := 2\|\delta^k\| \|y^k - y^{k-1}\| + \alpha \frac{2}{\theta} \lambda_k \|\bar{u}\| \|\delta^k\|.$$

Since $\{\|y^k - y^{k-1}\|\}$ is bounded by Corollary 3.8 and $\sum_{k=1}^{\infty} \|\delta^k\| < \infty$, then $\sum_{k=1}^{\infty} \|\beta^k\| < \infty$. Therefore the sequences $\{\sigma_k\}$ and $\{\beta_k\}$ are in the conditions of Lemma 2.5. This implies that $\{\|y^k - \bar{y}\|\}$ converges and therefore $\{y^k\}$ is bounded.

(ii) It follows from (i) and Proposition 3.7 that $\sum_{k=1}^{\infty} \|y^k - y^{k-1}\|^2 < \infty$, therefore $\lim_k \|y^k - y^{k-1}\| = 0$.

(iii) Since $y^k - \bar{y} = A(\bar{x} - x^k) + \delta^k$, we get that

$$\|y^k - \bar{y}\| - \|\delta^k\| \leq \|A(\bar{x} - x^k)\| \leq \|y^k - \bar{y}\| + \|\delta^k\|.$$

Being $\{\|y^k - \bar{y}\|\}$ convergent and $\{\|\delta^k\|\}$ convergent to zero, we conclude from the expression above that $\{\|A(\bar{x} - x^k)\|\}$ is also convergent. By (H_2) , the function $u \rightarrow \|u\|_A := \|Au\|$ is a norm in \mathbb{R}^n , and then it follows that $\{\|x^k - \bar{x}\|\}$ converges and therefore $\{x^k\}$ is bounded.

(iv) From (ii) and (3.11), it follows that $\lim_{k \rightarrow \infty} \|\tilde{y}^k - y^k\| = 0$. Therefore $\lim_{k \rightarrow \infty} \|A(\tilde{x}^k - x^k)\| = 0$. Again, the assumptions on A imply that $\lim_{k \rightarrow \infty} \|\tilde{x}^k - x^k\| = 0$. Item (v) Follows from (iii) and (iv). \square

We show below that the sequence $\{x^k\}$ converges to a solution of $VIP(T, C)$. Denote by $\mathcal{A}cc(z^k)$ the set of accumulation points of the sequence $\{z^k\}$.

Theorem 3.11. *Suppose that (H_1) – (H_3) hold. Then $\{x^k\}$ converges to an element of S .*

Proof. By Corollary 3.10(iii) and (iv) then $\mathcal{A}cc(\tilde{x}^k) = \mathcal{A}cc(x^k) \neq \emptyset$. We prove first that every element of $\mathcal{A}cc(\tilde{x}^k) = \mathcal{A}cc(x^k)$ is a solution of $VIP(T, C)$. Indeed, by (3.16), for all $(x, u) \in G(T)$ it holds

$$\begin{aligned} \langle x - \tilde{x}^k, u \rangle &\geq (\lambda_k)^{-1} \left[\frac{\theta}{2} (\|y^k - y\|^2 - \|y^{k-1} - y\|^2) \right. \\ &\quad \left. + (1 - \sigma) \frac{\tau}{2} \|y^k - y^{k-1}\|^2 - \theta \|\delta^k\| \|y^k - y^{k-1}\| \right]. \end{aligned} \quad (3.23)$$

Using Corollary 3.10(i)-(iii), we have that $\{x^k\}$ and $\{y^k\}$ are bounded and $\lim_k \|y^k - y^{k-1}\| = 0$. These facts, together with the identity

$$\|y^k - y\|^2 - \|y^{k-1} - y\|^2 = \|y^k - y^{k-1}\|^2 + 2\langle y^k - y^{k-1}, y^{k-1} - y \rangle,$$

yield

$$\lim_{k \rightarrow \infty} \|y^k - y\|^2 - \|y^{k-1} - y\|^2 = 0.$$

Let $\{\tilde{x}^{k_j}\} \subseteq \{\tilde{x}^k\}$ be a subsequence converging to x^* , we have that

$$\langle x - x^*, u \rangle = \lim_j \langle x - \tilde{x}^{k_j}, u \rangle \geq \liminf_k \langle x - \tilde{x}^k, u \rangle. \quad (3.24)$$

Using the above inequality and the fact that $\lambda_k \geq \bar{\rho} > 0$, we obtain the following expression by taking limits for $k \rightarrow \infty$ in (3.23):

$$\langle x - x^*, u \rangle \geq 0 \quad \forall (x, u) \in G(T). \quad (3.25)$$

By definition, $y^{k_j} = b + \delta^{k_j} - Ax^{k_j}$ with $y^{k_j} > 0$. We know that $\{y^{k_j}\}$ converges to $y^* = b - Ax^*$, with $y^* \geq 0$. Therefore $Ax^* \leq b$. Equivalently, $x^* \in C$. By definition of N_C , we have

$$\langle x - x^*, w \rangle \geq 0 \quad \forall (x, w) \in G(N_C). \quad (3.26)$$

Combining (3.25) and (3.26), we conclude that

$$\langle x - x^*, u + w \rangle \geq 0 \quad \forall (x, u + w) \in G(T + N_C).$$

By (H_1) and Proposition 2.1, $T + N_C$ is maximal monotone. Then the above inequality implies that $0 \in (T + N_C)(x^*)$, i.e, $x^* \in S$. Recall that x^* is also an accumulation point of $\{x^k\}$. Using Corollary 3.10(iii), we have that the sequence $\{\|x^* - x^k\|\}$ is convergent. Since it has a subsequence that converges to zero, the whole sequence $\{\|x^* - x^k\|\}$ must converge to zero. This completes the proof. \square

4 Conclusion

We propose an infeasible interior point method with log-quadratic proximal regularization for solving variational inequality problems. Our algorithm can

start from any point of the space (note that the first δ_0 can be chosen arbitrarily large, as long as the sequence $\{\|\delta_k\|\}$ is summable). We also introduce a relative error analysis which can be checked at each iteration, as the one in [12].

Moreover, our method can be applied even when the interior of C is empty. We show convergence under similar assumptions as those in the classical log-quadratic proximal, where the set C is required to have nonempty interior.

We acknowledge the fact that the scheme we propose is mainly theoretical. We point out that no numerical results are available for the method in [29]. However, from Remark 3.6, our inexact iteration includes the one in [29] as a particular case, and thus our step is likely to be computationally cheaper than that in [29]. A full numerical implementation of our method and a comparison with [29] is a fundamental question, which is the subject of future investigations.

REFERENCES

- [1] A. Auslender and M. Haddou, *An interior proximal method for convex linearly constrained problems and its extension to variational inequalities*. Mathematical Programming, **71** (1995), 77–100.
- [2] A. Auslender and M. Teboulle, *Asymptotic cones and functions in Optimization and Variational Inequalities*. Springer Monographs in Mathematics Springer-Verlag, New York, (2003).
- [3] A. Auslender, M. Teboulle and S. Ben-Tiba, *A logarithmic-quadratic proximal method for variational inequalities*. Computational Optimization and Applications, **12** (1999), 31–40.
- [4] A. Auslender, M. Teboulle and S. Ben-Tiba, *Interior proximal and multiplier methods based on second order homogeneous kernels*. Mathematics of Operations Research, **24** (1999), 645–668.
- [5] A. Ben-Tal and M. Zibulevsky, *Penalty-Barrier Methods for convex programming problems*. SIAM Journal on Optimization, **7** (1997), 347–366.
- [6] A. Brøndsted and R.T. Rockafellar, *On the subdifferentiability of convex functions*. Proceedings of the American Mathematical Society, **16** (1965), 605–611.
- [7] F.E. Browder, *Nonlinear operators and nonlinear equations of evolution in Banach spaces*. Proceedings of Symposia in Pure Mathematics, Journal of the American Mathematical Society, **18** (1976).
- [8] R.S. Burachik and A.N. Iusem, *A generalized proximal point method for the variational inequality problem in a Hilbert space*. SIAM Journal on Optimization, **8** (1998), 197–216.
- [9] R.S. Burachik and A.N. Iusem, *Set-valued mappings and enlargements of monotone operators*. Springer Optimization and its Applications, 8, Springer, New York, (2008).

- [10] R.S. Burachik, A.N. Iusem and B.F. Svaiter, *Enlargements of maximal monotone operators with applications to variational inequalities*. Set Valued Analysis, **5** (1997), 159–180.
- [11] R.S. Burachik, C.A. Sagastizábal and B.F. Svaiter, *Epsilon-Enlargement of Maximal Monotone Operators with Application to Variational Inequalities*, in Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Fukushima, M. and Qi, L. (editors), Kluwer Academic Publishers, (1997).
- [12] R.S. Burachik and B.F. Svaiter, *A relative error tolerance for a family of generalized proximal point methods*. Mathematics of Operations Research, **26** (2001), 816–831.
- [13] Y. Censor, A.N. Iusem and S. Zenios, *An interior point method with Bregman functions for the variational inequality problem with paramonotone operators*. Mathematical Programming, **83** (1998), 113–123.
- [14] J. Eckstein, *Approximate iterations in Bregman-function-based proximal algorithms*. Mathematical Programming, **81** (1998), 373–400.
- [15] P. Eggermont, *Multiplicative iterative algorithms for convex programming*. Linear Algebra and its Applications, **130** (1990), 25–42.
- [16] A.J. Hoffman, *On approximate solutions of systems of linear inequalities*. J. of the National Bureau of Standards, **49** (1952), 263–265.
- [17] A.N. Iusem, *On some properties of paramonotone operators*. Math. Oper. Res., **19** (1994), 790–814.
- [18] A.N. Iusem, B.F. Svaiter and M. Teboulle, *Entropy-like proximal methods in convex programming*. Journal Convex Anal., **5** (1998), 269–278.
- [19] A.N. Iusem and M. Teboulle, *Convergence rate analysis of nonquadratic proximal and augmented Lagrangian methods for convex and linear programming*. Math. of Oper. Res., **20** (1995), 657–677.
- [20] K.C. Kiwiel, *Proximal minimization methods with generalized Bregman functions*. SIAM Journal on Control and Optimization, **35** (1997), 1142–1168.
- [21] B. Lemaire, *On the convergence of some iterative methods for convex minimization*. In Recent Developments in Optimization, R. Durier and C. Michelot (eds), Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, **429** (1995), 252–268.
- [22] B.T. Polyak, *Introduction to optimization*. Optimization Software Inc., New York, (1987).
- [23] D. Pascali and S. Sburlan, *Nonlinear mappings of monotone type*, Ed. Academiei, Bucuresti, Romania, (1978).
- [24] R.T. Rockafellar, *On the maximality of sums of nonlinear monotone operators*. Transactions of the American Mathematical Society, **149** (1970), 75–88.
- [25] M.V. Solodov and B.F. Svaiter, *An Inexact Hybrid Generalized Proximal Point Algorithm and some New Results on the Theory of Bregman Functions*. Mathematics of Operations Research, **25** (2000), 214–230.

- [26] M. Teboulle, *Entropic proximal mappings with applications to nonlinear programming*. Mathematics of Operations Research, **1** (1992), 670–690.
- [27] M. Teboulle, *Convergence of proximal-like algorithms*. SIAM J. Optim., **7** (1997) 1069–1083.
- [28] P. Tseng and D. Bertsekas, *On the convergence of exponential multiplier method for convex programming*. Mathematical Programming, **60** (1993), 1–19.
- [29] N. Yamashita, C. Kanzow, T. Morimoto and M. Fukushima, *An infeasible interior proximal method for convex programming problems with linear constraints*. J. Nonlinear Convex Analysis, **2** (2001), 139–156.