The unified framework of some proximal-based decomposition methods for monotone variational inequalities with separable structure

Bingsheng He[1] and **Xiaoming Yuan**[2]

March 22, 2010

Abstract. Some existing decomposition methods for solving a class of variational inequalities (VI) with separable structures are closely related to the classical proximal point algorithm, as their decomposed sub-VIs are regularized by proximal terms. Differing in whether the generated sub-VIs are suitable for parallel computation, these proximal-based methods can be categorized into the parallel decomposition methods and alternating decomposition methods. This paper generalizes these methods and thus presents the unified framework of proximal-based decomposition methods for solving this class of VIs, in both exact and inexact versions. Then, for various special cases of the unified framework, we analyze the respective strategies for fulfilling a condition that ensures the convergence, which are realized by determining appropriate proximal parameters. Moreover, some concrete numerical algorithms for solving this class of VIs are derived. In particular, the inexact version of the unified framework gives rise to some implementable algorithms that allow the involved sub-VIs to be solved under those favorable criteria that are well-developed in the literature of the proximal point algorithm.

Keywords: Variational inequality, proximal point algorithm, decomposition, parallel, alternating.

1 Introduction

Variational inequalities (VI) capture a broad spectrum of applications in diverse fields, see, e.g., [8, 13, 20, 23]. In this paper, we consider the VI with the following separable structure:

$$(x^*, y^*) \in \mathcal{D}, \quad \begin{cases} (x - x^*)^T f(x^*) \ge 0, \\ (y - y^*)^T g(y^*) \ge 0, \end{cases} \qquad \forall\, (x, y) \in \mathcal{D}, \tag{1.1}$$

where

$$\mathcal{D} = \{(x, y) \in \Re^n \,|\, x \in \mathcal{X},\ y \in \mathcal{Y},\ Ax + By = b\}, \tag{1.2}$$

$\mathcal{X}$ and $\mathcal{Y}$ are given nonempty closed convex subsets of $\Re^{n_1}$ and $\Re^{n_2}$, respectively; $A \in \Re^{m\times n_1}$ and $B \in \Re^{m\times n_2}$ are given matrices; $b \in \Re^m$ is a given vector; $f : \mathcal{X} \to \Re^{n_1}$ and $g : \mathcal{Y} \to \Re^{n_2}$ are monotone operators. We refer to [2, 4, 9, 10, 15, 23] for the various applications of (1.1)-(1.2) in some other fields. In particular, (1.1)-(1.2) include the following minimization problem as a special case: $\min\{\theta_1(x) + \theta_2(y) \,|\, (x, y) \in \mathcal{D}\}$, where both $\theta_1(x)$ and $\theta_2(y)$ are differentiable convex functions.

[1]Department of Mathematics, Nanjing University, Nanjing, 210093, China. This author was supported by Jiangsu Province NSF BK2008255, The Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China 708044.Email: hebma@nju.edu.cn

[2]Corresponding author. Department of Mathematics, Hong Kong Baptist University, Hong Kong, China. This author was supported in part by the RGC grant 203009 and the NSFC grant 10701055. Email: xmyuan@hkbu.edu.hk

By attaching a Lagrangian multiplier vector $\lambda \in \Re^m$ to the linear constraint $Ax + By = b$, the VI (1.1)-(1.2) is converted into the following equivalent form:

$$(x^*, y^*, \lambda^*) \in \Omega, \quad \begin{cases} (x - x^*)^T(f(x^*) - A^T\lambda^*) \geq 0, \\ (y - y^*)^T(g(y^*) - B^T\lambda^*) \geq 0, \quad \forall\, (x, y, \lambda) \in \Omega, \\ (\lambda - \lambda^*)^T(Ax^* + By^* - b) \geq 0, \end{cases} \tag{1.3}$$

where

$$\Omega := \mathcal{X} \times \mathcal{Y} \times \Re^m.$$

We denote (1.3) by VI(Ω, F), where

$$u = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} \qquad \text{and} \qquad F(u) = \begin{pmatrix} f(x) - A^T\lambda \\ g(y) - B^T\lambda \\ Ax + By - b \end{pmatrix}. \tag{1.4}$$

The proximal point algorithm (PPA), which was originally proposed by Martinet [22] and concretely developed by Rockafellar [24, 25], is very widely-used in the community of Optimization. In fact, it particularly boosted various efficient numerical algorithms for solving VIs. Applying the PPA to solve VI(Ω, F) gives rise to the following iterative scheme to generate the new iterate $u^{k+1} := \tilde{u}^k$, where $\tilde{u}^k$ is the solution of the following subproblem:

$$(u - \tilde{u}^k)^T\{F(\tilde{u}^k) + \tfrac{1}{\alpha}(\tilde{u}^k - u^k)\} \geq 0, \quad \forall\, u \in \Omega. \tag{1.5}$$

Note that $\frac{1}{\alpha} > 0$ is regarded as the proximal parameter. More specifically, due to the separable structure (1.4), the task of solving the PPA subproblem (1.5) is equivalent to solving the following sub-VIs: Find $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$, such that

$$(x' - \tilde{x}^k)^T\big\{f(\tilde{x}^k) - A^T\tilde{\lambda}^k + \tfrac{1}{\alpha}(\tilde{x}^k - x^k)\big\} \geq 0, \quad \forall\ x' \in \mathcal{X}, \tag{1.6a}$$

$$(y' - \tilde{y}^k)^T\big\{g(\tilde{y}^k) - B^T\tilde{\lambda}^k + \tfrac{1}{\alpha}(\tilde{y}^k - y^k)\big\} \geq 0, \quad \forall\ y' \in \mathcal{Y}, \tag{1.6b}$$

$$(\lambda' - \tilde{\lambda}^k)^T\big\{(A\tilde{x}^k + B\tilde{y}^k - b) + \tfrac{1}{\alpha}(\tilde{\lambda}^k - \lambda^k)\big\} \geq 0, \quad \forall\ \lambda' \in \Re^m. \tag{1.6c}$$

Note that (1.6c) implies that

$$\tilde{\lambda}^k = \lambda^k - \alpha(A\tilde{x}^k + B\tilde{y}^k - b). \tag{1.7}$$

By substituting (1.7) into (1.6a) and (1.6b), we have the equivalent form of (1.6): Find $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$ such that

$$(x' - \tilde{x}^k)^T\big\{f(\tilde{x}^k) - A^T[\lambda^k - \alpha(A\tilde{x}^k + B\tilde{y}^k - b)] + \tfrac{1}{\alpha}(\tilde{x}^k - x^k)\big\} \geq 0, \tag{1.8a}$$

$$(y' - \tilde{y}^k)^T\big\{g(\tilde{y}^k) - B^T[\lambda^k - \alpha(A\tilde{x}^k + B\tilde{y}^k - b)] + \tfrac{1}{\alpha}(\tilde{y}^k - y^k)\big\} \geq 0, \tag{1.8b}$$

$$(\lambda' - \tilde{\lambda}^k)^T\big\{(A\tilde{x}^k + B\tilde{y}^k - b) + \tfrac{1}{\alpha}(\tilde{\lambda}^k - \lambda^k)\big\} \geq 0, \tag{1.8c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$.

Despite that (1.5) is formally decomposed into the sub-VIs (1.8) with lower dimensions, it is in general not implementable to solve the sub-VIs (1.8a) and (1.8b) separably due to the overlap of $\tilde{x}$ and $\tilde{y}$ emerging in these sub-VIs. Hence, the direct application of PPA in the form of (1.8) does not exploit the favorable separable structure of (1.3). Nevertheless, this observation has inspired many efforts to treat (1.8a) and (1.8b) separably by reducing the extent of overlapping of $\tilde{x}$ and $\tilde{y}$.

Thus, some proximal-based decomposition methods for (1.8) have been developed. Depending on the fashion of untying the overlap in (1.8a) and (1.8b), the current proximal-based decomposition methods rooted to (1.8) can be categorized into the parallel decomposition methods and the alternating decomposition methods.

For the proximal-based parallel decomposition methods, the involved sub-VIs are suitable for parallel computation. For example, when the exact version of the proximal-based decomposition method in [3] applied to solve (1.3), the solutions of the following sub-problems constitute the new iterate $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$:

$$(x' - \tilde{x}^k)^T \{ f(\tilde{x}^k) - A^T[\lambda^k - \alpha(Ax^k + By^k - b)] + \tfrac{1}{\alpha}(\tilde{x}^k - x^k) \} \geq 0, \tag{1.9a}$$

$$(y' - \tilde{y}^k)^T \{ g(\tilde{y}^k) - B^T[\lambda^k - \alpha(Ax^k + By^k - b)] + \frac{1}{\alpha}(\tilde{y}^k - y^k) \} \geq 0, \tag{1.9b}$$

$$\tilde{\lambda}^k = \lambda^k - \alpha(A\tilde{x}^k + B\tilde{y}^k - b), \tag{1.9c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$.

Note that the sub-VI (1.9a) does not involve $\tilde{y}^k$ and that (1.9b) is independent on the $\tilde{x}^k$ generated by (1.9a). Hence, (1.9a) and (1.9b) are eligible for parallel computation. We refer to, e.g., [28], for the further development.

On the other hand, opposite to the parallel decomposition methods, it is easy to develop (1.8) in the alternating manner, which solves the involved sub-VIs in the consecutive order. For instance, the following proximal-based alternating decomposition method is a special case of the exact version of the method proposed in [17]: Find $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$ such that

$$(x' - \tilde{x}^k)^T \{ f(\tilde{x}^k) - A^T[\lambda^k - \alpha(A\tilde{x}^k + By^k - b)] + \frac{1}{\alpha}(\tilde{x}^k - x^k) \} \geq 0, \tag{1.10a}$$

$$(y' - \tilde{y}^k)^T \{ g(\tilde{y}^k) - B^T[\lambda^k - \alpha(A\tilde{x}^k + B\tilde{y}^k - b)] + \frac{1}{\alpha}(\tilde{y}^k - y^k) \} \geq 0, \tag{1.10b}$$

$$\tilde{\lambda}^k = \lambda^k - \alpha(A\tilde{x}^k + B\tilde{y}^k - b), \tag{1.10c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$. Note that the sub-VIs (1.10a) and (1.10b) are not eligible for parallel computation, as the solution of (1.10a) is required in the process of solving (1.10b). Nevertheless, although the merit for parallel computation vanishes, alternating decomposition methods inherit well the spirit of the Gauss-Seidel method in the sense that the newly-obtained $\tilde{x}$ via the first sub-VI is used in the second sub-VI to generate $\tilde{y}$. Recall that this particular alternating decomposition method (1.10) is closely-related to the operator-splitting method and it is viewed as the application of PPA to the well-known alternating direction methods, see, e.g., [5, 6, 7, 11, 12, 18, 21, 26, 27, 29].

The main aim of this paper is to present the unified framework of proximal-based decomposition methods for solving (1.3) in both exact and inexact versions. For the inexact version of the unified framework, we extend the well-developed inexact criteria in the literature of PPA to solving the considered VI. Hence, the existing inexact proximal-based decomposition methods are improved. In addition, we pay particular attention to the restrictions on the involved proximal parameter to ensure convergence. With the detailed analysis for some concrete parallel and alternating decomposition methods, we conclude that the alternating decomposition methods generally allow more favorable proximal parameters than the parallel decomposition methods.

The rest of the paper is organized as follows. In Section 2, the unified framework of proximal-based decomposition methods is proposed, in both exact and inexact versions. Section 3 focuses on the restrictions on determining the proximal parameters for various concrete cases of the unified framework. In Section 4, we prove the convergence of the unified framework for both the exact and the inexact versions. Consequently, some implementable algorithms are developed. Finally, some conclusions are drawn in Section 5.

Throughout this paper, we assume that the solution set of (1.3) (denoted by Ω^*) is not empty; the involved sub-VIs are solvable; and that both f and g are monotone mappings.

2 The unified framework of proximal-based decomposition methods

In this section, we present the unified framework of proximal-based decomposition methods for (1.3) that include some existing methods as special cases.

For the given triple $u^k = (x^k, y^k, \lambda^k) \in \Omega$ (or in $\Re^{n_1+n_2+m}$), the following subproblems constitute the main body of the unified framework of proximal-based decomposition methods:

$$\tilde{x}^k \in \mathcal{X}, \;\; (x' - \tilde{x}^k)^T\{f(\tilde{x}^k) - A^T p_x + r(\tilde{x}^k - x^k) - \xi_x^k\} \geq 0, \;\; \forall\, x' \in \mathcal{X}, \tag{2.1a}$$

$$\tilde{y}^k \in \mathcal{Y}, \;\; (y' - \tilde{y}^k)^T\{g(\tilde{y}^k) - B^T p_y + s(\tilde{y}^k - y^k) - \xi_y^k\} \geq 0, \;\; \forall\, y' \in \mathcal{Y}, \tag{2.1b}$$

$$\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k - b), \tag{2.1c}$$

where $r > 0, s > 0$ are proximal parameters of the sub-VIs; $\beta > 0$ is viewed as the penalty parameter to the linear constraints $Ax + By = b$; $\xi_x^k \in \Re^{n_1}$ and $\xi_y^k \in \Re^{n_2}$ are error terms that allow the involved sub-VIs to be solved approximately subject to different inexact criteria; and $p_x \in \Re^m$ and $p_y \in \Re^m$ are terms whose specific choices lead to concrete proximal-based decomposition methods (to be analyzed in detail).

Some remarks regarding (2.1) should be given. First, we observe that for the particular parallel decomposition method (1.9), the generated sub-VIs take $\frac{1}{\alpha}$ as the proximal parameter and α as the penalty parameter to the linear constraints $Ax+By = b$. To study proximal-based decomposition methods in the general setting, we allow these parameter vary in a more general manner, as in [17].

Second, various choices of $p_x \in \Re^m$ and $p_y \in \Re^m$ lead to specific proximal-based parallel or alternating decomposition methods. We concentrate on the following concrete choices of p_x and p_y, two of which include the existing methods in [3, 17] as special cases.

a. The first parallel decomposition method (denoted by PDM1) by setting (p_x, p_y) as:

$$p_x = p_x(x^k, y^k, \lambda^k) = \lambda^k - \beta(Ax^k + By^k - b), \tag{2.2a}$$

$$p_y = p_y(x^k, y^k, \lambda^k) = \lambda^k - \beta(Ax^k + By^k - b). \tag{2.2b}$$

b. The second parallel decomposition method (denoted by PDM2) by setting (p_x, p_y) as:

$$p_x = p_x(\tilde{x}^k, y^k, \lambda^k) = \lambda^k - \beta(A\tilde{x}^k + By^k - b), \tag{2.3a}$$

$$p_y = p_y(x^k, \tilde{y}^k, \lambda^k) = \lambda^k - \beta(Ax^k + B\tilde{y}^k - b). \tag{2.3b}$$

c. The first alternating decomposition method (denoted by ADM1) by setting (p_x, p_y) as

$$p_x = p_x(x^k, y^k, \lambda^k) = \lambda^k - \beta(Ax^k + By^k - b), \tag{2.4a}$$

$$p_y = p_y(\tilde{x}^k, y^k, \lambda^k) = \lambda^k - \beta(A\tilde{x}^k + By^k - b). \tag{2.4b}$$

d. The second alternating decomposition method (denoted by ADM2) by setting (p_x, p_y) as

$$p_x = p_x(\tilde{x}^k, y^k, \lambda^k) = \lambda^k - \beta(A\tilde{x}^k + By^k - b), \tag{2.5a}$$

$$p_y = p_y(\tilde{x}^k, \tilde{y}^k, \lambda^k) = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k - b). \tag{2.5b}$$

It is easy to verify that the proximal-based decomposition methods in [3, 17] are recovered with the particular choices of (2.2) and (2.5), respectively.

For the convenience of proposing the unified framework, we reformulate (2.1) into the more compact form: Find $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$(u' - \tilde{u}^k)^T \{ (F(\tilde{u}^k) + \eta(u^k, \tilde{u}^k)) - (G(u^k - \tilde{u}^k) + \xi^k) \} \geq 0, \quad \forall\, u' \in \Omega, \tag{2.6}$$

where

$$u = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad F(u) = \begin{pmatrix} f(x) - A^T\lambda \\ g(y) - B^T\lambda \\ Ax + By - b \end{pmatrix}, \quad \xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix}, \tag{2.7}$$

$\eta(u^k, \tilde{u}^k) \in R^{n_1+n_2+m}$ and $G \in R^{(n_1+n_2+m)\times(n_1+n_2+m)}$ is positive definite. The respective expressions of $\eta(u^k, \tilde{u}^k)$ and G for some concrete choices of (p_x, p_y) are listed in Table 1 (see also (3.7), (3.20), (3.33) and (3.46)).

Table 1. The specification of $\eta(u^k, \tilde{u}^k)$ and matrix G in (2.6) for different methods

$\begin{pmatrix} p_x \\ p_y \end{pmatrix}$	$\eta(u^k, \tilde{u}^k)$	$G = \begin{pmatrix} G_x & & \\ & G_y & \\ & & G_\lambda \end{pmatrix}$
(2.2)	$\beta \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ 0 \end{pmatrix}$	$\begin{pmatrix} rI_{n_1} & & \\ & sI_{n_2} & \\ & & \frac{1}{\beta}I_m \end{pmatrix}$
(2.3)	$\beta \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ 0 \end{pmatrix}$	$\begin{pmatrix} rI_{n_1} + \beta A^TA & & \\ & sI_{n_2} + \beta B^TB & \\ & & \frac{1}{\beta}I_m \end{pmatrix}$
(2.4)	$\beta \begin{pmatrix} A^TB(y^k - \tilde{y}^k) \\ B^TB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix}$	$\begin{pmatrix} rI_{n_1} - \beta A^TA & & \\ & sI_{n_2} & \\ & & \frac{1}{\beta}I_m \end{pmatrix}$
(2.5)	$\beta \begin{pmatrix} A^TB(y^k - \tilde{y}^k) \\ B^TB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix}$	$\begin{pmatrix} rI_{n_1} & & \\ & sI_{n_2} + \beta B^TB & \\ & & \frac{1}{\beta}I_m \end{pmatrix}$

Third, the error terms ξ_x^k and ξ_y^k in (2.1) allow the sub-VIs to be solved approximately. Recall that the inexact parallel decomposition method in [3] (resp. the inexact alternating decomposition method in [17]) requires the restriction on the inexact criterion that the absolute errors are controlled by a sequence of positive scalars that is square-root-summable (resp. summable). In this paper, we improve these existing inexact criteria in two aspects: relative errors are considered; and the restrictions on the sequence of positive scalars to control the relative errors are relaxed significantly. More specifically, we present the following inexact criteria that are implementable for solving (2.1):

$$\textbf{Inexact Criterion 1:} \quad \|G^{-1}\xi^k\|_G \le \sqrt{\frac{(1-\mu)}{2}}\nu_k\|u^k - \tilde{u}^k\|_G, \quad \text{with } \sum_{k=0}^{\infty}\nu_k^2 < \infty,\ \nu_k > 0 \text{ and } \mu \in (0,1); \tag{2.8}$$

and

$$\textbf{Inexact Criterion 2:} \quad \|G^{-1}\xi^k\|_G \le \frac{\nu}{2}\|u^k - \tilde{u}^k\|_G, \text{ with } \nu \in (0,1). \tag{2.9}$$

Obviously, (2.9) is more relaxed than (2.8), as its relative errors can be uniformly bounded by the constant $\nu/2$. As we will analyze later, however, under Inexact Criterion 1, $\tilde{u}^k$ generated by (2.1) can be directly used as the new iterate u^{k+1}; while under Inexact Criterion 2, $\tilde{u}^k$ generated by (2.1) should be corrected for being the new iterate u^{k+1}.

Now, we are ready to present both the exact and inexact versions of the unified framework of proximal-based decomposition methods for solving (1.3).

The exact version of the unified framework of proximal-based decomposition methods

Step 0. Given $u^k = (x^k, y^k, \lambda^k)$. Set $k = 0$.
Step 1. Generate the new iterate via $u^{k+1} := \tilde{u}^k$, where $\tilde{u}^k$ is solved by (2.1) with $\xi_x^k = 0$ and $\xi_y^k = 0$.
Step 2. If stopping criterion is not satisfied, set $k := k + 1$ and go to step 1.

The inexact version of the unified framework of proximal-based decomposition methods

Step 0. Given $u^k = (x^k, y^k, \lambda^k)$. Set $k = 0$.
Step 1. Obtain $\tilde{u}^k$ via solving (2.1).
Step 2. Generate the new iterate u^{k+1}.
Case 1: If the inexact criterion (2.8) is used for solving (2.1), $u^{k+1} := \tilde{u}^k$.
Case 2: If the inexact criterion (2.9) is used for solving (2.1), $u^{k+1} = u^k - \alpha_k d_k$ where $\alpha_k > 0$ is regarded as the step size along the direction $d_k \in R^{n_1+n_2+m}$; and their specific expressions will be explained in detail later (see (4.13) and (4.18)).
Step 3. If stopping criterion is not satisfied, set $k := k + 1$ and go to step 1.

3 The strategies of determining proximal parameters

In this section, we first propose a condition that is required to ensure the convergence of the proposed unified framework (in both exact and inexact versions), then investigate how to fulfill this condition via choosing appropriate proximal parameters. Hence, the strategy of determining proximal parameters for the proposed unified framework is suggested.

The following lemma motives the condition to ensure the convergence of the proposed unified framework.

Lemma 3.1 *For given $u^k \in \Omega$ (or in $\Re^{n+m}$), let $\tilde{u}^k \in \Omega$ be generated by* (2.1)*. Then, for u^k, $\tilde{u}^k$, $\eta(u^k, \tilde{u}^k)$, G and ξ_k in the compact form* (2.6)*, we have*

$$\begin{aligned} &(u^k - u^*)^T\big(G(u^k - \tilde{u}^k) + \xi^k\big) \\ &\geq \;\; (\tilde{u}^k - u^*)^T\eta(u^k, \tilde{u}^k) + (u^k - \tilde{u}^k)^T\big(G(u^k - \tilde{u}^k) + \xi^k\big), \; \forall u^* \in \Omega^*. \end{aligned} \tag{3.1}$$

Proof. Since $u^* \in \Omega$, set $u' = u^*$ in (2.6), we get

$$(\tilde{u}^k - u^*)^T\big(G(u^k - \tilde{u}^k) + \xi^k\big) \geq (\tilde{u}^k - u^*)^T\big(F(\tilde{u}^k) + \eta(u^k, \tilde{u}^k)\big), \quad \forall u^* \in \Omega^*.$$

Using the monotonicity of F, we have

$$(\tilde{u}^k - u^*)^T F(\tilde{u}^k) \geq (\tilde{u}^k - u^*)^T F(u^*) \geq 0$$

and thus

$$(\tilde{u}^k - u^*)^T\big(G(u^k - \tilde{u}^k) + \xi^k\big) \geq (\tilde{u}^k - u^*)^T\eta(u^k, \tilde{u}^k), \quad \forall u^* \in \Omega^*, \tag{3.2}$$

which implies (3.1) immediately. □

For the exact version of the proposed unified framework (i.e., $\xi^k = 0$), it follows from (3.1) immediately that

$$(u^k - u^*)^T G(u^k - \tilde{u}^k) \geq (\tilde{u}^k - u^*)^T\eta(u^k, \tilde{u}^k) + \|u^k - \tilde{u}^k\|_G^2, \; \forall\, u^* \in \Omega^*,$$

which implies that

$$\begin{aligned} \|\tilde{u}^k - u^*\|_G^2 &= \|(u^k - u^*) - (u^k - \tilde{u}^k)\|_G^2 \\ &= \|u^k - u^*\|_G^2 - 2(u^k - u^*)^T G(u^k - \tilde{u}^k) + \|u^k - \tilde{u}^k\|_G^2 \\ &\overset{(3.3)}{=} \|u^k - u^*\|_G^2 - \|u^k - \tilde{u}^k\|_G^2 - 2(\tilde{u}^k - u^*)^T\eta(u^k, \tilde{u}^k). \end{aligned}$$

Therefore, as long as $\eta(u^k, \tilde{u}^k)$ and G are chosen appropriately such that

$$2|(\tilde{u}^k - u^*)^T\eta(u^k, \tilde{u}^k)| < \|u^k - \tilde{u}^k\|_G^2, \tag{3.3}$$

we have

$$\|\tilde{u}^k - u^*\|_G < \|u^k - u^*\|_G^2,$$

which indicates that the sequence $\{u^k\}$ by taking $u^{k+1} := \tilde{u}^k$ is Fejér monotone to Ω^*. Hence, the convergence of the proposed unified framework (in exact version) is implied, see, e.g., [1].

The requirement (3.3) motivates us to adopt the following condition, which is crucial for ensuring convergence of the proposed unified framework in both exact and inexact versions.

Condition: There exists a constant $\mu \in (0, 1)$ such that

$$|(\tilde{u}^k - u^*)^T\eta(u^k, \tilde{u}^k)| \leq \frac{\mu}{2}\|u^k - \tilde{u}^k\|_G^2, \tag{3.4}$$

where $\tilde{u}^k$ is generated by (2.1), $u^* \in \Omega^*$, $\eta(u^k, \tilde{u}^k)$ and G in (2.6) are listed in Table 1 for different (p_x, p_y).

As shown in Theorems 1-4 , this condition is achievable via choosing appropriate values of the proximal parameters r and s. Therefore, Condition (3.4) also inspires the strategies of determining the proximal parameters r and s in (2.1). As analyzed extensively in [25], the convergence of PPA and its variants requires that the proximal parameter should be bounded below from a positive scalar. However, with this prerequisite satisfied, proximal-based methods with lower bounds of proximal parameters (which means the allowable ranges for proximal parameters are broader) are preferable both theoretically and numerically, see [14] for profound analysis. Note that the value of μ in (2.8) can be taken as exactly that in (3.4) once it is determined in the procedure of fulfilling (3.4). However, the constant ν in (2.9) can be chosen independently on the μ in (3.4).

Therefore, for different cases of the proposed unified framework, it is of particular interest to investigate the allowable ranges of proximal parameters that ensure Condition (3.4). More specifically, with the choices of p_x and p_y specified in (2.2)-(2.5) (hence, the definitions of $\eta(u^k, \tilde{u}^k)$ and G become determinate for each case, as shown in Table 1.), we analyze the respective restrictions on the proximal parameters r and s that ensure Condition (3.4). At the same time, the value of μ is determined accordingly.

3.1 The first proximal-based parallel decomposition method (PDM1)

For the first parallel decomposition method, the pair of (p_x, p_y) is taken as (2.2). We first rewrite (2.1) with the specific (p_x, p_y) into the form of (2.6), and then provide the precise expressions of $\eta(u^k, \tilde{u}^k)$ and G in (2.6)-(2.7).

Substituting the definition of (p_x, p_y) in (2.2) into (2.1), we have $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$ such that

$$(x' - \tilde{x}^k)^T \{ f(\tilde{x}^k) - A^T[\lambda^k - \beta(Ax^k + By^k - b)] + r(\tilde{x}^k - x^k) - \xi_x^k \} \geq 0, \tag{3.5a}$$

$$(y' - \tilde{y}^k)^T \{ g(\tilde{y}^k) - B^T[\lambda^k - \beta(Ax^k + By^k - b)] + s(\tilde{y}^k - y^k) - \xi_y^k \} \geq 0, \tag{3.5b}$$

$$(\lambda' - \tilde{\lambda}^k)^T \{ (A\tilde{x}^k + B\tilde{y}^k - b) + \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \} \geq 0, \tag{3.5c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$. Using the notation of $F(u)$ (see (1.4)), by a manipulation, (3.5) can be written as: find $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ 0 \end{pmatrix} \right\} +$$

$$+ \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} r(\tilde{x}^k - x^k) - \xi_x^k \\ s(\tilde{y}^k - y^k) - \xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \geq 0, \quad \forall u' = (x', y', \lambda') \in \Omega.$$

Therefore, the concrete form of (2.6)-(2.7) for PDM1 is:

$$(u' - \tilde{u}^k)^T \{ (F(\tilde{u}^k) + \eta_I(u^k, \tilde{u}^k)) - (G_I(u^k - \tilde{u}^k) + \xi^k) \} \geq 0, \quad \forall\, u' \in \Omega, \tag{3.6}$$

where

$$\eta_I(u^k, \tilde{u}^k) = \beta \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ 0 \end{pmatrix}, \quad \xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix} \tag{3.7a}$$

and

$$G_I = \begin{pmatrix} rI_{n_1} & & \\ & sI_{n_2} & \\ & & \frac{1}{\beta}I_m \end{pmatrix}. \tag{3.7b}$$

Theorem 3.2 *For given u^k, let $\tilde{u}^k \in \Omega$ be produced by (3.5); $\eta_{_I}(u^k, \tilde{u}^k)$ and G_I be defined in (3.7). If*

$$r > 2\beta\|A^T A\| \qquad \text{and} \qquad s > 2\beta\|B^T B\|, \tag{3.8}$$

then Condition (3.4) is satisfied in the following sense:

$$|(\tilde{u}^k - u^*)^T \eta_{_I}(u^k, \tilde{u}^k)| \le \frac{\mu_I}{2}\|u^k - \tilde{u}^k\|^2_{G_I},$$

with

$$\mu_I = \sqrt{\max\Big\{\frac{2\beta\|A^TA\|}{r}, \frac{2\beta\|B^TB\|}{s}\Big\}} \in (0,1). \tag{3.9}$$

Proof. It follows from $Ax^* + By^* = b$ and $\beta(A\tilde{x}^k + B\tilde{y}^k - b) = \lambda - \tilde{\lambda}^k$ that

$$\begin{aligned}(\tilde{u}^k - u^*)^T \eta_{_I}(u^k, \tilde{u}^k) &= \beta \begin{pmatrix} \tilde{x}^k - x^* \\ \tilde{y}^k - y^* \\ \tilde{\lambda}^k - \lambda^* \end{pmatrix}^T \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ 0 \end{pmatrix} \\ &= (\lambda^k - \tilde{\lambda}^k)^T\big(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\big). \end{aligned} \tag{3.10}$$

By using Cauchy-Schwarz Inequality, we have

$$\begin{aligned} &|(\tilde{u}^k - u^*)^T \eta_{_I}(u^k, \tilde{u}^k)| \\ &= \big|(\lambda^k - \tilde{\lambda}^k)^T\big(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\big)\big| \\ &= \frac{1}{2}\{2|(\lambda^k - \tilde{\lambda}^k)^T A(x^k - \tilde{x}^k)| + 2|(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k)|\} \\ &\le \frac{1}{2}\Big\{\frac{2\beta}{\mu_I}\|A(x^k - \tilde{x}^k)\|^2 + \frac{\mu_I}{2\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\Big\} \\ &\quad + \frac{1}{2}\Big\{\frac{2\beta}{\mu_I}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_I}{2\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\Big\} \\ &= \frac{1}{2}\Big\{\frac{2\beta}{\mu_I}\|A(x^k - \tilde{x}^k)\|^2 + \frac{2\beta}{\mu_I}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_I}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\Big\}. \end{aligned} \tag{3.11}$$

With the μ_I defined in (3.9), we have

$$\frac{2\beta}{\mu_I}\|A(x^k - \tilde{x}^k)\|^2 \le \mu_I r\|x^k - \tilde{x}^k\|^2 \tag{3.12a}$$

and

$$\frac{2\beta}{\mu_I}\|B(y^k - \tilde{y}^k)\|^2 \le \mu_I s\|y^k - \tilde{y}^k\|^2. \tag{3.12b}$$

Combining (3.11), (3.12), and the definition of G_I in (3.7), we obtain (3.9). Hence, Condition (3.4) is satisfied and the proof is complete. □

Remark 3.3 *As analyzed in the Introduction, the proximal-based parallel decomposition method in [3] is a special case of the PDM1 (3.5) with $r = s = \frac{1}{\beta}$. The restriction on β (where $B = I$) proposed in [3] is :*

$$\beta \le \frac{1}{2\max\{\|A\|, 1\}}. \tag{3.13}$$

Corollary 3.4 *If $r = s = \frac{1}{\beta}$ in (3.5) and β satisfies the requirement (3.13), then Condition (3.4) is satisfied in the following sense:*

$$|(\tilde{u}^k - u^*)^T \eta_{_I}(u^k, \tilde{u}^k)| \le \frac{\sqrt{2}}{4} \|u^k - \tilde{u}^k\|^2_{G_{I_0}}, \tag{3.14}$$

where $\eta_{_I}(u^k, \tilde{u}^k)$ is defined identically in (3.7) and

$$G_{I_0} = \frac{1}{\beta} \begin{pmatrix} I_{n_1} & & \\ & I_{n_2} & \\ & & I_m \end{pmatrix}. \tag{3.15}$$

Proof. It follows from (3.11) and $\mu_I = \frac{\sqrt{2}}{2}$ that

$$\begin{aligned} & |(\tilde{u}^k - u^*)^T \eta_{_I}(u^k, \tilde{u}^k)| \\ \le \;& \frac{1}{2}\{\frac{2\beta}{\mu_I}\|A(x^k - \tilde{x}^k)\|^2 + \frac{2\beta}{\mu_I}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_I}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\} \\ = \;& \frac{\sqrt{2}}{4\beta}\{4\beta^2\|A(x^k - \tilde{x}^k)\|^2 + 4\beta^2\|B(y^k - \tilde{y}^k)\|^2 + \|\lambda^k - \tilde{\lambda}^k\|^2\}. \end{aligned} \tag{3.16}$$

From Conditions (3.13), we get $4\beta^2\|A(x^k - \tilde{x}^k)\|^2 \le \|x^k - \tilde{x}^k\|^2$ and $4\beta^2\|B(y^k - \tilde{y}^k)\|^2 \le \|y^k - \tilde{y}^k\|^2$, and thus

$$|(\tilde{u}^k - u^*)^T \eta_{_I}(u^k, \tilde{u}^k)| \;\le\; \frac{\sqrt{2}}{4\beta}\{\|x^k - \tilde{x}^k\|^2 + \|y^k - \tilde{y}^k\|^2 + \|\lambda^k - \tilde{\lambda}^k\|^2\}. \tag{3.17}$$

Assertion (3.14) follows immediately from definition of G_{I_0} in (3.15). □

3.2 The second proximal-based parallel decomposition method (PDM2)

For the second parallel decomposition method, the pair of (p_x, p_y) is taken as (2.3). We first rewrite (2.1) with the specific (p_x, p_y) into the form of (2.6), and then provide the precise expressions of $\eta(u^k, \tilde{u}^k)$ and G in (2.6)-(2.7).

Substituting the definition of (p_x, p_y) in (2.3) into (2.1), we have $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$ such that

$$(x' - \tilde{x}^k)^T\{f(\tilde{x}^k) - A^T[\lambda^k - \beta(A\tilde{x}^k + By^k - b)] + r(\tilde{x}^k - x^k) - \xi_x^k\} \ge 0, \tag{3.18a}$$

$$(y' - \tilde{y}^k)^T\{g(\tilde{y}^k) - B^T[\lambda^k - \beta(Ax^k + B\tilde{y}^k - b)] + s(\tilde{y}^k - y^k) - \xi_y^k\} \ge 0, \tag{3.18b}$$

$$(\lambda' - \tilde{\lambda}^k)^T\{(A\tilde{x}^k + B\tilde{y}^k - b) + \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k)\} \ge 0, \tag{3.18c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$. Using the notation of $F(u)$ (see (1.4)), by a manipulation, (3.18) can be written as: find $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y})^k - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^T B(y^k - \tilde{y}^k) \\ B^T A(x^k - \tilde{x}^k) \\ 0 \end{pmatrix} + \begin{pmatrix} r(\tilde{x}^k - x^k) - \xi_x^k \\ s(\tilde{y}^k - y^k) - \xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \right\} \ge 0,$$

for all $u' \in \Omega$ which is equivalent to

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ 0 \end{pmatrix} \right\} +$$

$$+\begin{pmatrix} x'-\tilde{x}^k \\ y'-\tilde{y}^k \\ \lambda'-\tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} (rI_{n_1}+\beta A^TA)(\tilde{x}^k-x^k)-\xi_x^k \\ (sI_{n_2}+\beta B^TB)(\tilde{y}^k-y^k)-\xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k-\lambda^k) \end{pmatrix} \ge 0,\ \forall u'=(x',y',\lambda')\in\Omega.$$

Therefore, the concrete form of (2.6)-(2.7) for PDM2 is:

$$(u'-\tilde{u}^k)^T\{\big(F(\tilde{u}^k)+\eta_{II}(u^k,\tilde{u}^k)\big)-\big(G_{II}(u^k-\tilde{u}^k)+\xi^k\big)\}\ge 0, \quad \forall\, u'\in\Omega, \tag{3.19}$$

where

$$\eta_{II}(u^k,\tilde{u}^k)=\beta\begin{pmatrix} A^T(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)) \\ B^T(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)) \\ 0 \end{pmatrix}, \quad \xi^k=\begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix} \tag{3.20a}$$

and

$$G_{II}=\begin{pmatrix} rI_{n_1}+\beta A^TA & & \\ & sI_{n_2}+\beta B^TB & \\ & & \frac{1}{\beta}I_m \end{pmatrix}. \tag{3.20b}$$

In fact, the definition of $\eta_{II}(u^k,\tilde{u}^k)$ is identical with that of $\eta_I(u^k,\tilde{u}^k)$ in (3.7). For consistence, we still use the notation of $\eta_{II}(u^k,\tilde{u}^k)$ for PDM2.

Theorem 3.5 *For given u^k, let $\tilde{u}^k\in\Omega$ be produced by* (3.18)*; $\eta_{II}(u^k,\tilde{u}^k)$ and G_{II} be defined in* (3.20)*, If*

$$r>\beta\|A^TA\| \qquad \text{and} \qquad s>\beta\|B^TB\|, \tag{3.21}$$

then Condition (3.4) is satisfied in the following sense:

$$|(\tilde{u}^k-u^*)^T\eta_{II}(u^k,\tilde{u}^k)|\le\frac{\mu_{II}}{2}\|u^k-\tilde{u}^k\|_{G_{II}}^2 \tag{3.22a}$$

with

$$\mu_{II}=\sqrt{\max\Big\{\frac{2\beta\|A^TA\|}{r+\beta\|A^TA\|},\frac{2\beta\|B^TB\|}{s+\beta\|B^TB\|}\Big\}}\in(0,1). \tag{3.22b}$$

Proof. Recall that $\eta_{II}(u^k,\tilde{u}^k)=\eta_I(u^k,\tilde{u}^k)$. As in the proof of (3.10), we have

$$(\tilde{u}^k-u^*)^T\eta_{II}(u^k,\tilde{u}^k)=(\lambda^k-\tilde{\lambda}^k)^T\big(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)\big). \tag{3.23}$$

Therefore, analogous to (3.11), we obtain

$$\begin{aligned} &|(\tilde{u}^k-u^*)^T\eta_{II}(u^k,\tilde{u}^k)| \\ &= \big|(\lambda^k-\tilde{\lambda}^k)^T\big(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)\big)\big| \\ &\le \frac{1}{2}\Big\{\frac{2\beta}{\mu_{II}}\|A(x^k-\tilde{x}^k)\|^2+\frac{2\beta}{\mu_{II}}\|B(y^k-\tilde{y}^k)\|^2+\frac{\mu_{II}}{\beta}\|\lambda^k-\tilde{\lambda}^k\|^2\Big\}. \end{aligned} \tag{3.24}$$

It follows from the value of μ in (3.22) that

$$\frac{2\beta}{\mu_{II}}\|A(x^k-\tilde{x}^k)\|^2\le\mu_{II}\big(r\|x^k-\tilde{x}^k\|^2+\beta\|A(x^k-\tilde{x}^k)\|^2\big)=\mu_{II}\|x^k-\tilde{x}^k\|_{(rI+\beta A^TA)}^2 \tag{3.25a}$$

and

$$\frac{2\beta}{\mu_{II}}\|B(y^k-\tilde{y}^k)\|^2\le\mu_{II}\big(s\|y^k-\tilde{y}^k\|^2+\beta\|B(y^k-\tilde{y}^k)\|^2\big)=\mu_{II}\|y^k-\tilde{y}^k\|_{(sI+\beta B^TB)}^2. \tag{3.25b}$$

Combining (3.24), (3.25), and the definition of G_{II} in (3.20), we obtain

$$|(\tilde{u}^k - u^*)^T \eta_{II}(u^k, \tilde{u}^k)| = \big|(\lambda^k - \tilde{\lambda}^k)^T \big(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\big)\big| \le \frac{\mu_{II}}{2}\|u^k - \tilde{u}^k\|^2_{G_{II}}.$$

Condition (3.4) is satisfied and the proof is complete. □

Remark 3.6 *The proximal-based parallel decomposition method in [19] is a special case of the PDM2* (3.5) *with* $r = s = \frac{1}{\beta}$. *In this case, the restriction on* β *proposed in [19] is:*

$$\beta \le \frac{1}{\sqrt{3}\max\{\|A\|, \|B\|\}}. \tag{3.26}$$

Corollary 3.7 *If* $r = s = \frac{1}{\beta}$ *in* (3.18) *and* β *satisfies the requirement (3.26), then Condition (3.4) is satisfied in the following sense:*

$$|(\tilde{u}^k - u^*)^T \eta_{II}(u^k, \tilde{u}^k)| \le \frac{\sqrt{2}}{4}\|u^k - \tilde{u}^k\|^2_{G_{II_0}}, \tag{3.27}$$

where $\eta_{II}(u^k, \tilde{u}^k)$ *is defined identically in (3.7) and*

$$G_{II_0} = \frac{1}{\beta}\begin{pmatrix} I_{n_1} + \beta A^T A & & \\ & I_{n_2} + \beta B^T B & \\ & & I_m \end{pmatrix}. \tag{3.28}$$

Proof. It follows from (3.24) and $\mu_{II} = \frac{\sqrt{2}}{2}$ that

$$\begin{aligned} & |(\tilde{u}^k - u^*)^T \eta_{II}(u^k, \tilde{u}^k)| \\ \le\ & \frac{1}{2}\Big\{\frac{2\beta}{\mu_{II}}\|A(x^k - \tilde{x}^k)\|^2 + \frac{2\beta}{\mu_{II}}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_{II}}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\Big\} \\ =\ & \frac{\sqrt{2}}{4\beta}\big\{4\beta^2\|A(x^k - \tilde{x}^k)\|^2 + 4\beta^2\|B(y^k - \tilde{y}^k)\|^2 + \|\lambda^k - \tilde{\lambda}^k\|^2\big\}. \end{aligned} \tag{3.29}$$

From Conditions (3.26), we get $4\beta^2\|A(x^k - \tilde{x}^k)\|^2 \le \|x^k - \tilde{x}^k\|^2_{(I+\beta^2 A^T A)}$ and $4\beta^2\|B(y^k - \tilde{y}^k)\|^2 \le \|y^k - \tilde{y}^k\|^2_{(I+\beta^2 B^T B)}$, and thus

$$\begin{aligned} & |(\tilde{u}^k - u^*)^T \eta_{II}(u^k, \tilde{u}^k)| \\ \le\ & \frac{\sqrt{2}}{4\beta}\big\{\|x^k - \tilde{x}^k\|^2_{(I+\beta^2 A^T A)} + \|y^k - \tilde{y}^k\|^2_{(I+\beta^2 B^T B)} + \|\lambda^k - \tilde{\lambda}^k\|^2\big\}. \end{aligned} \tag{3.30}$$

Assertion (3.27) follows the definition of G_{II_0} (3.28) immediately. □

3.3 The first proximal-based alternating decomposition method (ADM1)

For the first alternating decomposition method, the pair of (p_x, p_y) is taken as (2.4). We first rewrite (2.1) with the specific (p_x, p_y) into the form of (2.6), and then provide the precise expressions of $\eta(u^k, \tilde{u}^k)$ and G in (2.6)-(2.7).

Substituting the definition of (p_x, p_y) in (2.4) into (2.1), we have $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$ such that

$$(x' - \tilde{x}^k)^T\big\{f(\tilde{x}^k) - A^T[\lambda^k - \beta(Ax^k + By^k - b)] + r(\tilde{x}^k - x) - \xi^k_x\big\} \ge 0, \tag{3.31a}$$

$$(y' - \tilde{y}^k)^T\{g(\tilde{y}^k) - B^T[\lambda^k - \beta(A\tilde{x}^k + By^k - b)] + s(\tilde{y}^k - y^k) - \xi_y^k\} \geq 0, \tag{3.31b}$$

$$(\lambda' - \tilde{\lambda}^k)^T\{(A\tilde{x}^k + B\tilde{y}^k - b) + \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k)\} \geq 0, \tag{3.31c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$. Using the notation of $F(u)$ (see (1.4)), by a manipulation, (3.31) can be written as: find $\tilde{u} = (\tilde{x}, \tilde{y}, \tilde{\lambda}) \in \Omega$ such that

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^T(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ B^TB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix} \right\} +$$

$$+ \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} r(\tilde{x}^k - x^k) - \xi_x^k \\ s(\tilde{y}^k - y^k) - \xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \geq 0, \ \forall u' = (x', y', \lambda') \in \Omega,$$

which is equivalent to

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^TB(y^k - \tilde{y}^k) \\ B^TB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix} \right\} +$$

$$+ \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} (rI_{n_1} - \beta A^TA)(\tilde{x}^k - x^k) - \xi_x^k \\ s(\tilde{y}^k - y^k) - \xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \geq 0, \ \forall u' = (x', y', \lambda') \in \Omega.$$

Therefore, the concrete form of (2.6)-(2.7) for ADM1 is:

$$(u' - \tilde{u}^k)^T\{(F(\tilde{u}^k) + \eta_{III}(u^k, \tilde{u}^k)) - (G_{III}(u^k - \tilde{u}^k) + \xi^k)\} \geq 0, \quad \forall\, u' \in \Omega, \tag{3.32}$$

where

$$\eta_{III}(u^k, \tilde{u}^k) = \beta \begin{pmatrix} A^TB(y^k - \tilde{y}^k) \\ B^TB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix}, \quad \xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix} \tag{3.33a}$$

and

$$G_{III} = \begin{pmatrix} rI_{n_1} - \beta A^TA & & \\ & sI_{n_2} & \\ & & \frac{1}{\beta}I_m \end{pmatrix}. \tag{3.33b}$$

Theorem 3.8 *For given u, let $\tilde{u}^k \in \Omega$ be produced by (3.31); $\eta_{III}(u^k, \tilde{u}^k)$ and G_{III} be defined in (3.33), If*

$$r > \beta\|A^TA\| \qquad \text{and} \qquad s > \beta\|B^TB\|, \tag{3.34}$$

then Condition (3.4) is satisfied in the following sense:

$$|(\tilde{u}^k - u^*)^T\eta_{III}(u^k, \tilde{u}^k)| \leq \frac{\mu_{III}}{2}\|u^k - \tilde{u}^k\|_{G_{III}}^2 \tag{3.35a}$$

with

$$\mu_{III} = \sqrt{\frac{\beta\|B^TB\|}{s}} \in (0, 1). \tag{3.35b}$$

Proof. First, since F is monotone, we have $(\tilde{u}^k - u^*)^T F(\tilde{u}^k) \geq 0$. Recall the definition of $\eta_{III}(u^k, \tilde{u}^k)$ in (3.33) and the facts $Ax^* + By^* = b$ and $\beta(A\tilde{x}^k + B\tilde{y}^k - b) = \lambda^k - \tilde{\lambda}^k$. Then, we have

$$\begin{aligned}(\tilde{u}^k - u^*)^T \eta_{III}(u^k, \tilde{u}^k) &= \beta \begin{pmatrix} \tilde{x}^k - x^* \\ \tilde{y}^k - y^* \\ \tilde{\lambda}^k - \lambda^* \end{pmatrix}^T \begin{pmatrix} A^T B(y^k - \tilde{y}^k) \\ B^T B(y^k - \tilde{y}^k) \\ 0 \end{pmatrix} \\ &= (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k). \end{aligned} \tag{3.36}$$

By using Cauchy-Schwarz Inequality, we get

$$\begin{aligned}&|(\tilde{u}^k - u^*)^T \eta_{III}(u^k, \tilde{u}^k)| = \left|(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k)\right| \\ &\leq \frac{1}{2}\{\frac{\beta}{\mu_{III}}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_{III}}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\}. \end{aligned} \tag{3.37}$$

Note that under Condition (3.35), we have

$$\frac{\beta}{\mu_{III}}\|B(y^k - \tilde{y}^k)\|^2 \leq \mu_{III} s\|y^k - \tilde{y}^k\|^2. \tag{3.38}$$

Combining (3.38), (3.39) and (3.33), we obtain

$$\begin{aligned}|(\tilde{u}^k - u^*)^T \eta_{III}(u^k, \tilde{u}^k)| &= \left|(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k)\right| \\ &\leq \frac{\mu_{III}}{2}\{s\|y^k - \tilde{y}^k\|^2 + \frac{1}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\} \\ &\leq \frac{\mu_{III}}{2}\|u^k - \tilde{u}^k\|^2_{G_{III}}. \end{aligned}$$

Hence, Condition (3.4) is satisfied and the proof is complete. □

Corollary 3.9 *If* $r = s = \frac{1}{\beta}$ *in* (3.18) *and* β *satisfies the requirement (3.26):*

$$\beta \leq \frac{1}{\sqrt{2}\max\{\|A\|, \|B\|\}}, \tag{3.39}$$

then Condition (3.4) is satisfied in the following sense:

$$|(\tilde{u}^k - u^*)^T \eta_{III}(u^k, \tilde{u}^k)| \leq \frac{\sqrt{2}}{4}\|u^k - \tilde{u}^k\|^2_{G_{III_0}}, \tag{3.40}$$

where $\eta_{III}(u, \tilde{u})$ *is defined in (3.33) and*

$$G_{III_0} = \frac{1}{\beta}\begin{pmatrix} I_{n_1} - \beta^2 A^T A & & \\ & I_{n_2} & \\ & & I_m \end{pmatrix}. \tag{3.41}$$

Proof. It follows from (3.37) and $\mu_{III} = \frac{\sqrt{2}}{2}$ that

$$\begin{aligned}|(\tilde{u}^k - u^*)^T \eta_{III}(u^k, \tilde{u}^k)| &\leq \frac{1}{2}\{\frac{\beta}{\mu_{III}}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_{III}}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\} \\ &= \frac{\sqrt{2}}{4\beta}\{2\beta^2\|B(y^k - \tilde{y}^k)\|^2 + \|\lambda^k - \tilde{\lambda}^k\|^2\}. \end{aligned} \tag{3.42}$$

From Conditions (3.26) we get $2\beta^2\|B(y^k - \tilde{y}^k)\|^2 \leq \|y^k - \tilde{y}^k\|^2$ and thus

$$|(\tilde{u}^k - u^*)^T \eta_{III}(u^k, \tilde{u}^k)| \leq \frac{\sqrt{2}}{4\beta}\{\|y^k - \tilde{y}^k\|^2 + \|\lambda^k - \tilde{\lambda}^k\|^2\}. \tag{3.43}$$

Assertion (3.27) follows from the definition of G_{III_0} (3.41) immediately. □

3.4 The second proximal-based alternating decomposition method (ADM2)

For the first alternating decomposition method, the pair of (p_x, p_y) is taken as (2.5). We first rewrite (2.1) with the specific (p_x, p_y) into the form of (2.6), and then provide the precise expressions of $\eta(u^k, \tilde{u}^k)$ and G in (2.6)-(2.7).

Substituting the definition of (p_x, p_y) in (2.5) into (2.1), we have $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$ such that

$$(x' - \tilde{x}^k)^T \{f(\tilde{x}^k) - A^T[\lambda^k - \beta(A\tilde{x}^k + By^k - b)] + r(\tilde{x}^k - x) - \xi_x^k\} \geq 0, \tag{3.44a}$$

$$(y' - \tilde{y}^k)^T \{g(\tilde{y}^k) - B^T[\lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k - b)] + s(\tilde{y}^k - y^k) - \xi_y^k\} \geq 0, \tag{3.44b}$$

$$(\lambda' - \tilde{\lambda}^k)^T \{(A\tilde{x}^k + B\tilde{y}^k - b) + \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k)\} \geq 0, \tag{3.44c}$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$. Using the notation of $F(u)$ (see (1.4)), by a manipulation, (3.44) can be written as: find $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y})^k - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^T B(y^k - \tilde{y}^k) \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} r(\tilde{x}^k - x^k) - \xi_x^k \\ s(\tilde{y}^k - y) - \xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \right\} \geq 0,$$

for all $(x', y', \lambda') \in \mathcal{X} \times \mathcal{Y} \times \Re^m$. This is equivalent to:

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \beta \begin{pmatrix} A^T B(y^k - \tilde{y}^k) \\ B^T B(y^k - \tilde{y}^k) \\ 0 \end{pmatrix} \right\} +$$

$$+ \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} r(\tilde{x}^k - x^k) - \xi_x^k \\ (sI_{n_2} + \beta B^T B)(\tilde{y}^k - y^k) - \xi_y^k \\ \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \geq 0, \ \forall u' = (x', y', \lambda') \in \Omega.$$

Therefore, the concrete form of (2.6)-(2.7) for ADM2 is:

$$(u' - \tilde{u}^k)^T \{(F(\tilde{u}^k) + \eta_{IV}(u^k, \tilde{u}^k)) - (G_{IV}(u^k - \tilde{u}^k) + \xi^k)\} \geq 0, \quad \forall\, u' \in \Omega, \tag{3.45}$$

where

$$\eta_{IV}(u^k, \tilde{u}^k) = \beta \begin{pmatrix} A^T B(y^k - \tilde{y}^k) \\ B^T B(y^k - \tilde{y}^k) \\ 0 \end{pmatrix}, \quad \xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix} \tag{3.46a}$$

and

$$G_{IV} = \begin{pmatrix} rI_{n_1} & & \\ & sI_{n_2} + \beta B^T B & \\ & & \frac{1}{\beta} I_m \end{pmatrix}. \tag{3.46b}$$

In fact, the definition of $\eta_{IV}(u^k, \tilde{u}^k)$ is identical with that of $\eta_{III}(u^k, \tilde{u}^k)$ in (3.33). For consistence, we still use the notation of $\eta_{IV}(u^k, \tilde{u}^k)$ for ADM2.

Theorem 3.10 *For given u, let $\tilde{u} \in \Omega$ be produced by* (3.31); $\eta_{III}(u,\tilde{u})$ *and* G_{III} *be defined in* (3.46), *If*

$$r > 0 \qquad \text{and} \qquad s > 0, \tag{3.47}$$

then Condition (3.4) is satisfied in the following sense:

$$|(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k)| \le \frac{\mu_{IV}}{2} \|u^k - \tilde{u}^k\|^2_{G_{IV}} \tag{3.48a}$$

with

$$\mu_{IV} = \sqrt{\frac{\beta\|B^TB\|}{s + \beta\|B^TB\|}} \in (0,1). \tag{3.48b}$$

Proof. Recall that $\eta_{IV}(u^k,\tilde{u}^k) = \eta_{III}(u^k,\tilde{u}^k)$. As in the proof of (3.36), we get

$$(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k) = (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k). \tag{3.49}$$

Analogous to (3.37), we have

$$\begin{aligned} |(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k)| &= \big|(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k)\big| \\ &\le \frac{1}{2}\{\frac{\beta}{\mu_{IV}}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_{IV}}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\}. \end{aligned} \tag{3.50}$$

Note that under Condition (3.48), we have

$$\frac{\beta}{\mu_{IV}}\|B(y^k - \tilde{y}^k)\|^2 \le \mu_{IV}\{s\|y^k - \tilde{y}^k\|^2 + \beta\|B(y^k - \tilde{y}^k)\|^2\} = \mu_{IV}\|y^k - \tilde{y}^k\|^2_{(sI+\beta B^TB)}, \tag{3.51}$$

and thus (see the definition of G_{IV} in (3.46))

$$\begin{aligned} |(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k)| &= \big|(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k)\big| \\ &\le \frac{\mu_{IV}}{2}\{\|y^k - \tilde{y}^k\|^2_{(sI+\beta B^TB)} + \frac{1}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\} \\ &\le \frac{\mu_{IV}}{2}\|u^k - \tilde{u}^k\|^2_{G_{IV}}. \end{aligned} \tag{3.52}$$

Hence, Condition (3.4) is satisfied and the proof is complete. □

Corollary 3.11 *If $r = s = \frac{1}{\beta}$ in (3.44) and β satisfies the requirement*

$$\beta \le \frac{1}{\|B\|}, \tag{3.53}$$

then Condition (3.4) is satisfied in the following sense:

$$|(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k)| \le \frac{\sqrt{2}}{4}\|u^k - \tilde{u}^k\|^2_{G_{IV_0}}, \tag{3.54}$$

where $\eta_{IV}(u^k,\tilde{u}^k)$ is defined in (3.46) and

$$G_{IV_0} = \frac{1}{\beta}\begin{pmatrix} I_{n_1} & & \\ & I_{n_2} + \beta^2 B^TB & \\ & & I_m \end{pmatrix}. \tag{3.55}$$

Proof. It follows from (3.52) and $\mu_{IV} = \frac{\sqrt{2}}{2}$ that

$$\begin{aligned} |(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k)| &\le \frac{1}{2}\{\frac{\beta}{\mu_{IV}}\|B(y^k - \tilde{y}^k)\|^2 + \frac{\mu_{IV}}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2\} \\ &= \frac{\sqrt{2}}{4\beta}\{2\beta^2\|B(y^k - \tilde{y}^k)\|^2 + \|\lambda^k - \tilde{\lambda}^k\|^2\}. \end{aligned} \tag{3.56}$$

From Conditions (3.53), we get $2\beta^2\|B(y^k - \tilde{y}^k)\|^2 \le \|y^k - \tilde{y}^k\|^2_{(I_{n_2}+\beta^2 B^T B)}$ and thus

$$|(\tilde{u}^k - u^*)^T \eta_{IV}(u^k, \tilde{u}^k)| \le \frac{\sqrt{2}}{4\beta}\{\|y^k - \tilde{y}^k\|^2_{(I_{n_2}+\beta^2 B^T B)} + \|\lambda^k - \tilde{\lambda}^k\|^2\}. \tag{3.57}$$

Assertion (3.54) follows from the definition of G_{IV_0} (3.55) immediately. □

3.5 Summary

In the last four subsections, we study intensively the respective restrictions on the proximal parameters to enure Condition (3.4), for various cases of the proposed unified framework. We summarize the respective restrictions and the according values of μ in (3.4) in the following table.

Table 2. The restriction on the proximal parameters

$\begin{pmatrix} p_x \\ p_y \end{pmatrix}$	Restrictions on r and s	$\mu \in (0,1)$ in Condition (3.4)
(2.2)	$r > 2\beta\|A^T A\|,\ s > 2\beta\|B^T B\|$	$\mu = \sqrt{\max\{\frac{2\beta\|A^T A\|}{r}, \frac{2\beta\|B^T B\|}{s}\}}$
(2.3)	$r > \beta\|A^T A\|,\ s > \beta\|B^T B\|$	$\mu = \sqrt{\max\{\frac{2\beta\|A^T A\|}{r+\beta\|A^T A\|}, \frac{2\beta\|B^T B\|}{s+\beta\|B^T B\|}\}}$
(2.4)	$r > \beta\|A^T A\|,\ s > \beta\|B^T B\|$	$\mu = \sqrt{\frac{\beta\|B^T B\|}{s}}$
(2.5)	$r > 0, \quad s > 0$	$\mu = \sqrt{\frac{\beta\|B^T B\|}{s+\beta\|B^T B\|}}$

Our analysis indicates that for solving (1.1)-(1.2), the alternating decomposition methods generally allows more favorable lower-bounds of proximal parameters than the parallel decomposition methods do. In particular, among the four cases of the proposed unified framework that have been delineated, the ADM2 admits the most relaxed ranges with smallest lower-bounds of proximal parameters.

4 Convergence

This section proves the convergence of both the exact and the inexact versions of the proposed unified framework, under the Condition (3.4).

4.1 The exact version

Recall that the exact version of the proposed unified framework generates the new iterate u^{k+1} according to:

Algorithm 1:
$u^{k+1} := \tilde{u}^k$,where $\tilde{u}^k$ the solution of (2.1) with $\xi_x^k = 0$ and $\xi_y^k = 0$.

Lemma 4.1 *Let $\{u^k\}$ be the sequence generated by the exact version of the proposed unified framework, and Condition (3.4) is satisfied. Then, we have*

$$\|u^{k+1} - u^*\|_G^2 \le \|u^k - u^*\|_G^2 - (1-\mu)\|u^k - u^{k+1}\|_G^2, \quad \forall u^* \in \Omega^*. \tag{4.1}$$

Proof. From (3.1), Condition (3.4) and $\xi^k = 0$, we get

$$\begin{aligned}
\|\tilde{u}^k - u^*\|_G^2 &= \|(u^k - u^*) - (u^k - \tilde{u}^k)\|_G^2 \\
&= \|u^k - u^*\|_G^2 - 2(u^k - u^*)^T G(u^k - \tilde{u}^k) + \|u^k - \tilde{u}^k\|_G^2 \\
\left(\begin{smallmatrix}\text{using (3.1)} \\ \text{and } \xi^k = 0\end{smallmatrix}\right) &\le \|u^k - u^*\|_G^2 - 2(\tilde{u}^k - u^*)^T \eta(u^k, \tilde{u}^k) - \|u^k - \tilde{u}^k\|_G^2 \\
(\text{using (3.4)}) &\le \|u^k - u^*\|_G^2 - (1-\mu)\|u^k - \tilde{u}^k\|_G^2.
\end{aligned}$$

Recall that $u^{k+1} := \tilde{u}^k$ for the exact version of the proposed unified framework. Thus, the theorem is proved. □

The inequality (4.1) indicates that the sequence $\{u^k\}$ generated by the exact version of the proposed unified framework is Fejér monotone with respect to Ω^* in the G-norm. Hence, the convergence is readily obtained, as in the following theorem (see also [1]).

Theorem 4.2 *The sequence $\{u^k\}$ generated by the exact version of the proposed unified framework converges to a solution of VI(Ω, F).*

Proof. It follows from (4.1) that $\{u^k\}$ is bounded, and it has at least one cluster point. From (2.6), we have $u^{k+1} \in \Omega$ and

$$(u' - u^{k+1})^T\big(F(u^{k+1}) + \eta(u^k, u^{k+1})\big) \ge (u' - u^{k+1})^T G(u^k - u^{k+1}), \quad \forall\, u' \in \Omega.$$

Note that (4.1) implies that $\lim_{k\to\infty} \|u^k - u^{k+1}\| = 0$. On the other hand, it is from Table 1. that there is a constant $C > 0$ (determined by A, B and β), such that

$$\|\eta(u^k, u^{k+1})\| \le C\|u^k - u^{k+1}\|,$$

for all the cases of $\eta(u^k, u^{k+1})$. Thus, we get

$$u^{k+1} \in \Omega, \quad \lim_{k\to\infty} (u' - u^{k+1})^T F(u^{k+1}) \ge 0, \qquad \forall u' \in \Omega.$$

Let u^∞ be a cluster point of $\{u^k\}$. We assume that the subsequence $\{u^{k_j}\}$ converges to u^∞. Then, it follows from the above inequality that

$$u^{k_j} \in \Omega, \qquad \lim_{j\to\infty} (u' - u^{k_j})^T F(u^{k_j}) \ge 0, \qquad \forall u' \in \Omega$$

and consequently

$$u^\infty \in \Omega, \qquad (u' - u^\infty)^T F(u^\infty) \ge 0, \qquad \forall u' \in \Omega.$$

This means that $u^\infty \in \Omega^*$. Because (4.1) is true for any solution point, the sequence $\{u^k\}$ converges to u^∞. Thus, the theorem is proved. □

4.2 The inexact version with Inexact Criterion 1

As well-known, the involved sub-VIs in (2.1) are very expensive (if not impossible) to be solved accurately, no speaking of the little justification of obtaining their exact solutions when they differ from the original VI significantly. Hence, the practical way of treating the involved sub-VIs is to solve them approximately subject to some inexact criteria. As we have analyzed, the proposed unified framework is closely related to the PPA—-both the sub-VIs (2.1a) and (2.1b) are regularized by proximal terms. Therefore, it is necessary to extend the existing inexact criteria in the literature of PPA to the inexact version of the proposed unified framework. Through this extension, some existing inexact proximal-based decomposition methods for solving (1.3) (see, e.g., [3, 17]) will be improved.

Recall that we have proposed the Inexact criterion 1 in (2.8). We refer to [16] for similar inexact criteria in the context of general VI. Under this criterion, the inexact version of the proposed unified framework generates the new iterate u^{k+1} according to:

Algorithm 2:
$u^{k+1} := \tilde{u}^k$, where $\tilde{u}^k$ is the solution of (2.1) under the inexact criterion (2.8).

Lemma 4.3 *Let $\{u^k\}$ be the sequence generated by the inexact version of the proposed unified framework with the inexact criterion (2.8), and the Condition (3.4) is satisfied. Then, we have*

$$(1-\nu_k^2)\|u^{k+1}-u^*\|_G^2 \le \|u^k-u^*\|_G^2 - \frac{1-\mu}{2}\|u^k-u^{k+1}\|_G^2, \quad \forall u^* \in \Omega^*. \tag{4.2}$$

Proof. First, we have

$$\begin{aligned}\|\tilde{u}^k-u^*\|_G^2 &= \|u^k-u^*+\tilde{u}^k-u^k\|_G^2 \\ &= \|u^k-u^*\|_G^2 - 2(u^k-u^*)^T\big(G(u^k-\tilde{u}^k)+\xi^k\big) \\ &\quad +2(u^k-u^*)^T\xi^k + \|u^k-\tilde{u}^k\|_G^2.\end{aligned}$$

Substituting (3.1) in the second term of the right hand side of the above inequality, we get

$$\begin{aligned}\|\tilde{u}^k-u^*\|_G^2 &\le \|u^k-u^*\|_G^2 - 2(\tilde{u}^k-u^*)^T\eta(u^k,\tilde{u}^k) \\ &\quad -2(u^k-\tilde{u}^k)^T\big(G(u^k-\tilde{u}^k)+\xi^k\big) \\ &\quad +2(u^k-u^*)^T\xi^k + \|u^k-\tilde{u}^k\|_G^2 \\ \text{(using (3.4))} &\le \|u^k-u^*\|_G^2 - (1-\mu)\|u^k-\tilde{u}^k\|_G^2 + 2(\tilde{u}^k-u^*)^T\xi^k.\end{aligned}$$

By using Cauchy-Schwarz Inequality,

$$2(\tilde{u}^k-u^*)^T\xi^k \le \nu_k^2\|\tilde{u}^k-u^*\|_G^2 + \frac{1}{\nu_k^2}\|G^{-1}\xi^k\|_G^2,$$

it follows that

$$(1-\nu_k^2)\|\tilde{u}^k-u^*\|_G^2 \le \|u^k-u^*\|_G^2 - (1-\mu)\|u^k-\tilde{u}^k\|_G^2 + \frac{1}{\nu_k^2}\|G^{-1}\xi^k\|_G^2.$$

Since $\sum_{k=0}^{\infty} \nu_k^2 < +\infty$, we have $\lim_{k\to\infty} \nu_k^2 = 0$. Without loss of the generality we can assume that $\nu_k^2 \le \frac{1}{2}$ for all $k \ge 0$. Hence, by using (2.8), we obtain

$$(1-\nu_k^2)\|\tilde{u}^k - u^*\|_G^2 \le \|u^k - u^*\|_G^2 - \frac{1-\mu}{2}\|u^k - \tilde{u}^k\|_G^2. \tag{4.3}$$

Recall that $u^{k+1} := \tilde{u}^k$ for the inexact version of the proposed unified framework with the inexact criterion (2.8), the assertion (4.2) is proved. □.

From $\nu_k^2 \le \frac{1}{2}$, we have

$$1 \le \frac{1}{1-\nu_k^2} \le 1 + 2\nu_k^2. \tag{4.4}$$

By taking $u^{k+1} = \tilde{u}^k$, it follows from (4.2) and (4.4) that

$$\|u^{k+1} - u^*\|_G^2 \le (1+2\nu_k^2)\|u^k - u^*\|_G^2 - \frac{1-\mu}{2}\|u^k - u^{k+1}\|_G^2, \quad \forall u^* \in \Omega^*. \tag{4.5}$$

Comparing with the inequality (4.1) that implies the convergence of the exact version of the unified framework, (4.5) is somewhat weak. However, (4.5) with $\sum_{k=0}^{\infty} \nu_k^2 < \infty$ suffices to ensure the convergence of the inexact version of the proposed unified framework with Inexact Criterion 1, as we show in the following.

Theorem 4.4 *The sequence $\{u^k\}$ generated by the inexact version of the proposed unified framework with the inexact criterion (2.8) converges to a solution of VI(Ω, F).*

Proof. First, we prove that

$$\lim_{k\to\infty} \|u^k - u^{k+1}\| = 0. \tag{4.6}$$

Since $\sum_{k=0}^{\infty} \nu_k^2 < +\infty$, we have

$$\prod_{k=1}^{\infty}(1+\tau\nu_k^2) < \infty, \quad \forall\, \tau > 0. \tag{4.7}$$

A consequent result of (4.5) is

$$\begin{aligned} \|u^{k+1} - u^*\|_G^2 &\le \Pi_{l=0}^{k}(1+2\nu_l^2)\|u^0 - u^*\|^2 \\ &\le \Pi_{l=0}^{\infty}(1+2\nu_l^2)\|u^0 - u^*\|^2 := E, \end{aligned} \tag{4.8}$$

where $E > 0$ is a constant due to the fact that $\Pi_{l=0}^{\infty}(1+2\nu_l^2) < +\infty$ (see (4.7)). Then, it follows from (4.5) and (4.8) that

$$\begin{aligned} \frac{1-\mu}{2}\sum_{k=0}^{\infty}\|u^k - u^{k+1}\|^2 &\overset{(4.5)}{\le} 2\sum_{k=0}^{\infty}\nu_k^2\|u^k - u^*\|^2 + \|u^0 - u^*\|^2 \\ &\overset{(4.8)}{\le} \big(2\sum_{k=0}^{\infty}\nu_k^2 + 1\big)E. \end{aligned} \tag{4.9}$$

By using $\sum_{k=0}^{\infty} \nu_k^2 < +\infty$ and (4.9), we prove (4.6).

Similarly as in the proof of Theorem 4.2, by using (2.8) and (4.6), as in the proof of Theorem 4.2, there is a u^∞, which is a cluster point of $\{u^k\}$ and satisfies

$$u^\infty \in \Omega, \qquad (u' - u^\infty)^T F(u^\infty) \ge 0, \qquad \forall u' \in \Omega.$$

This means that u^∞ is a solution of VI(Ω, F). Since $\lim_{k\to\infty} \nu_k^2 = 0$ and (4.5) is true for any solution point, the sequence $\{u^k\}$ converges to u^∞. □

4.3 The inexact version with Inexact Criterion 2

As analyzed previously, both the exact version and the inexact version with (2.8) adopts directly the solution of (2.1) as the new iterate. However, the inexact criterion (2.1) is still too strengthen, as it requires the involved sub-VIs to be solved with increasing accuracy. Hence, it is interesting to relax it. In particular, motivated by the existing inexact criteria in the literate of PPA that allow the VIs to be solved in a constant accuracy, we have proposed the inexact criterion (2.9) for solving (2.1). This inexact criterion allows the relative errors for solving the involved sub-VIs to be controlled by a constant. Hence, the sub-VIs can be solved easily. As the PPA methods with the similar criteria, the main expense of relaxing the inexact criteria is that the approximate solutions of the involved sub-VIs should be corrected for being the new iterate. Hence, this type of methods are presented in the prediction-correction fashion: the approximate solution of the sub-VIs is viewed as the predictor of the new iterate, and it should be corrected by correction step.

Lemma 4.5 *Let $\{u^k\}$ be the sequence generated by the inexact version of the proposed unified framework with the inexact criterion (2.9), and Condition (3.4) is satisfied. By defining*

$$\varphi(u^k, \tilde{u}^k) := (\tilde{u}^k - u^*)^T \eta(u^k, \tilde{u}^k) + (u^k - \tilde{u}^k)^T \big(G(u^k - \tilde{u}^k) + \xi^k\big), \tag{4.10}$$

we have

$$\varphi(u^k, \tilde{u}^k) \geq \frac{2 - (\mu + \nu)}{2} \|u^k - \tilde{u}^k\|_G^2. \tag{4.11}$$

Proof. Note that $|(u^k - \tilde{u}^k)^T \xi^k| \leq \|u^k - \tilde{u}^k\|_G \cdot \|G^{-1}\xi^k\|_G$. It is from (2.9) that we have

$$(u^k - \tilde{u}^k)^T \big(G(u^k - \tilde{u}^k) + \xi^k\big) \geq \big(1 - \frac{\nu}{2}\big) \|u^k - \tilde{u}^k\|_G^2. \tag{4.12}$$

Then, the assertion (4.11) follows from (4.10), (4.12) and (3.4) immediately. □

4.3.1 The first prediction-correction scheme

The first prediction-correction scheme to implement the inexact version of the proposed unified framework with the inexact criterion (2.9) is the following.

Algorithm 3:

$$u^{k+1} = u^k - \alpha_{Ik}\big((u^k - \tilde{u}^k) + G^{-1}\xi^k\big), \tag{4.13a}$$

where

$$\alpha_{Ik} = \gamma \alpha_{Ik}^*, \qquad \alpha_{Ik}^* = \frac{\varphi(u^k, \tilde{u}^k)}{\|(u^k - \tilde{u}^k) + G^{-1}\xi^k\|_G^2}, \ \gamma \in (0, 2), \tag{4.13b}$$

and $\tilde{u}^k$ is the the solution of (2.1) under the inexact criterion (2.9).

In fact, the rationale of developing the first prediction-correction scheme is easily justified. Recall the Condition (3.4) and that G is positive definite, it is easy to derive that the right-hand-side of (3.1) is nonnegative under the inexact criterion (2.9). This fact implies that $-(u^k - \tilde{u}^k +$

$G^{-1}\xi^k$) is a descent direction of $\frac{1}{2}\|u-u^*\|^2$ at $u=u^k$. Therefore, it is beneficial to generate the new iterate along this direction in the manner of (4.13a).

We now prove that the sequence $\{u^k\}$ generated by the first prediction-correction scheme (4.13) is Fejér monotone to Ω^* under the G-norm. Note that although $\{\tilde{u}^k\}\subset\Omega$, it does not require the sequence $\{u^k\}$ to be contained in Ω.

Lemma 4.6 *Let $\{u^k\}$ be the sequence generated by the first prediction-correction scheme* (4.13). *Then, we have*

$$\|u^{k+1}-u^*\|_G^2 \le \|u^k-u^*\|_G^2 - \gamma(2-\gamma)\alpha_{Ik}^*\varphi(u^k,\tilde{u}^k), \quad \forall\, u^*\in\Omega^*. \tag{4.14}$$

Proof. We define the step size dependent new iterate by

$$u_I(\alpha) = u^k - \alpha\big((u^k-\tilde{u}^k)+G^{-1}\xi^k\big)$$

and consider the difference of $\|u^k-u^*\|_G^2$ and $\|u_I(\alpha)-u^*\|_G^2$. Using (3.1) and (4.10), we get

$$\begin{aligned}
&\|u^k-u^*\|_G^2 - \|u_I(\alpha)-u^*\|_G^2 \\
&\quad= \|u^k-u^*\|_G^2 - \|u^k-u^*-\alpha\big((u^k-\tilde{u}^k)+G^{-1}\xi^k\big)\|_G^2 \\
&\quad= 2\alpha(u^k-u^*)^T\big(G(u^k-\tilde{u}^k)+\xi^k\big) - \alpha^2\|(u^k-\tilde{u}^k)+G^{-1}\xi^k\|_G^2 \\
&\quad\ge 2\alpha\varphi(u^k,\tilde{u}^k) - \alpha^2\|(u^k-\tilde{u}^k)+G^{-1}\xi^k\|_G^2.
\end{aligned}$$

Thus we get

$$\|u^k-u^*\|_G^2 - \|u_I(\alpha)-u^*\|_G^2 \ge q_I(\alpha), \tag{4.15}$$

where

$$q_I(\alpha) = 2\alpha\varphi(u^k,\tilde{u}^k) - \alpha^2\|(u^k-\tilde{u}^k)+G^{-1}\xi^k\|_G^2.$$

Note that $q_I(\alpha)$ is a quadratic function of α which reaches its maximum at α_{Ik}^* in (4.13b) and

$$q_I(\gamma\alpha_{Ik}^*) = \gamma(2-\gamma)\alpha_{Ik}^*\varphi(u^k,\tilde{u}^k). \tag{4.16}$$

Assertion (4.14) follows from (4.15), (4.16) and $u^{k+1}=u_I(\gamma\alpha_{Ik}^*)$ directly. □

Lemma 4.7 *Let α_{Ik}^* be defined in* (4.13b). *If $\nu\le 2\sqrt{1-\mu}$ is additionally requested* [3], *then we have $\alpha_{Ik}^*\ge\frac{1}{2}$.*

Proof Using (4.10) and (3.4), we obtain

$$\begin{aligned}
2\varphi(u^k,\tilde{u}^k) &\ge 2\{(u^k-\tilde{u}^k)^T\big(G(u^k-\tilde{u}^k)+\xi^k\big) - \frac{\mu}{2}\|u^k-\tilde{u}^k\|_G^2\} \\
&= \|u^k-\tilde{u}^k\|_G^2 + 2(u^k-\tilde{u}^k)^T\xi^k + (1-\mu)\|u^k-\tilde{u}^k\|_G^2 \\
&= \|(u^k-\tilde{u}^k)+G^{-1}\xi^k\|_G^2 + \{(1-\mu)\|u^k-\tilde{u}^k\|_G^2 - \|G^{-1}\xi^k\|_G^2\}.
\end{aligned}$$

Since $\nu\le 2\sqrt{1-\mu}$ and (2.9), we have $\|G^{-1}\xi\|_G^2\le(1-\mu)\|u^k-\tilde{u}^k\|_G^2$ and thus

$$2\varphi(u^k,\tilde{u}^k) \ge \|(u^k-\tilde{u}^k)+G^{-1}\xi^k\|_G^2,$$

which states the assertion. □

Lemmas 4.5, 4.6 and 4.7 imply the following result immediately.

[3] The additional request is easy to satisfied. For example, if $\mu=0.8$ is given, $\nu=0.8$ satisfies this condition; when $\mu=0.99$, $\nu=2\sqrt{1-\mu}=0.2\in(0,1)$ is not too small. For $\mu=\sqrt{2}/2$, any $\nu\in(0,1)$ satisfies this request.

Lemma 4.8 *Let $\{u^k\}$ be the sequence generated by the first prediction-correction scheme* (4.13)*. Then for any $u^* \in \Omega^*$ we have*

$$\|u^{k+1} - u^*\|_G^2 \le \|u^k - u^*\|_G^2 - \frac{\gamma(2-\gamma)(2-\mu-\nu)}{4}\|u^k - \tilde{u}^k\|_G^2. \tag{4.17}$$

Now, we are ready to prove the convergence of the first prediction-correction scheme to implement the inexact version of the proposed unified framework with the inexact criterion (2.9).

Theorem 4.9 *The sequence $\{u^k\}$ generated by the first prediction-correction scheme of the inexact version of the proposed unified framework with the inexact criterion (2.9) converges to a solution of VI(Ω, F).*

Proof Due to (4.17), the proof is analogous to that of Theorem 4.2, thus omitted. □

4.3.2 The second prediction-correction scheme

The second prediction-correction scheme to implement the inexact version of the proposed unified framework with the inexact criterion (2.9) is the following.

Algorithm 4:

$$u^{k+1} = P_\Omega\{u^k - \alpha_{II k}\big(F(\tilde{u}^k) + \eta(u^k, \tilde{u}^k)\big)\}, \tag{4.18a}$$

where

$$\alpha_{II k} = \gamma \alpha^*_{II k}, \qquad \alpha^*_{II k} = \frac{\varphi(u^k, \tilde{u}^k)}{\|G(u^k - \tilde{u}^k) + \xi^k\|^2}, \qquad \gamma \in (0, 2). \tag{4.18b}$$

and $\tilde{u}^k$ is the the solution of (2.1) under the inexact criterion (2.9).

We now prove that the sequence $\{u^k\}$ generated by the second prediction-correction scheme (4.18) is Fejér monotone to Ω^* under the Euclidean norm.

Lemma 4.10 *Let $\{u^k\}$ be the sequence generated by the second prediction-correction scheme* (4.18)*. Then, we have*

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \gamma(2-\gamma)\alpha^*_{II k}\varphi(u^k, \tilde{u}^k), \quad \forall\, u^* \in \Omega^*. \tag{4.19}$$

Proof. We define the step size dependent new iterate by

$$u_{II}(\alpha) = P_\Omega\{u^k - \alpha\big(F(\tilde{u}^k) + \eta(u^k, \tilde{u}^k)\big)\} \tag{4.20}$$

and consider the difference of $\|u^k - u^*\|_G^2$ and $\|u_{II}(\alpha) - u^*\|_G^2$. Since $u^* \in \Omega$, it follows that

$$\begin{aligned} \|u_{II}(\alpha) - u^*\|^2 \;\le\; & \|u^k - \alpha\big(F(\tilde{u}^k) + \eta(u^k, \tilde{u}^k)\big) - u^*\|^2 \\ & - \|u^k - \alpha\big(F(\tilde{u}^k) + \eta(u^k, \tilde{u}^k)\big) - u_{II}(\alpha)\|^2. \end{aligned} \tag{4.21}$$

Consequently, we get

$$\begin{aligned}
&\|u-u^*\|^2-\|u_{II}(\alpha)-u^*\|^2\\
&\geq\ \|u^k-u^*\|^2-\|u^k-\alpha\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big)-u^*\|^2\\
&\quad+\|u^k-\alpha\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big)-u_{II}(\alpha)\|^2\\
&=\ \|u^k-u_{II}(\alpha)\|^2+2\alpha(u^k-u^*)^T\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big)\\
&\quad+2\alpha(u_{II}(\alpha)-u)^T\big(F(\tilde u)+\eta(u,\tilde u)\big)\\
&=\ \|u^k-u_{II}(\alpha)\|^2+2\alpha(u_{II}(\alpha)-u^*)^T\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big).
\end{aligned}\tag{4.22}$$

Since $u_{II}(\alpha)\in\Omega$, set $u'=u_{II}(\alpha)$ in (2.6) we obtain

$$(u_{II}(\alpha)-\tilde u^k)^T\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big)\geq(u_{II}(\alpha)-\tilde u^k)^T\big(G(u^k-\tilde u^k)+\xi^k\big).\tag{4.23}$$

Using Condition (2.7) we have

$$\begin{aligned}
&(\tilde u^k-u^*)^T\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big)\\
&\quad\geq\ \varphi(u^k,\tilde u^k)-(u^k-\tilde u^k)^T\big(G(u^k-\tilde u^k)+\xi^k\big).
\end{aligned}\tag{4.24}$$

Adding (4.23) and (4.24), it follows that

$$\begin{aligned}
&(u_{II}(\alpha)-u^*)^T\big(F(\tilde u^k)+\eta(u^k,\tilde u^k)\big)\\
&\quad\geq\ \varphi(u^k,\tilde u^k)+(u_{II}(\alpha)-u^k)^T\big(G(u^k-\tilde u^k)+\xi^k\big)
\end{aligned}\tag{4.25}$$

Substituting (4.25) in the right hand side of (4.22), we get

$$\begin{aligned}
&\|u^k-u^*\|^2-\|u_{II}(\alpha)-u^*\|^2\\
&\geq\ \|u^k-u_{II}(\alpha)\|^2+2\alpha\varphi(u^k,\tilde u^k)+2\alpha(u_{II}(\alpha)-u^k)^T\big(G(u^k-\tilde u^k)+\xi^k\big)\\
&=\ \|u^k-u_{II}(\alpha)-\alpha\big(G(u^k-\tilde u^k)+\xi^k\big)\|^2\\
&\quad+2\alpha\varphi(u^k,\tilde u^k)-\alpha^2\|G(u^k-\tilde u^k)+\xi^k\|^2\\
&\geq\ 2\alpha\varphi(u^k,\tilde u^k)-\alpha^2\|G(u^k-\tilde u^k)+\xi^k\|^2.
\end{aligned}$$

Thus we get

$$\|u^k-u^*\|_G^2-\|u_{II}(\alpha)-u^*\|_G^2\geq q_{II}(\alpha),\tag{4.26}$$

where

$$q_{II}(\alpha)=2\alpha\varphi(u^k,\tilde u^k)-\alpha^2\|G(u^k-\tilde u^k)+\xi^k\|^2.$$

Note that $q_{II}(\alpha)$ is a quadratic function of α which reaches its maximum at α^*_{IIk} in (4.18b) and

$$q_{II}(\gamma\alpha^*_{Ik})=\gamma(2-\gamma)\alpha^*_{IIk}\varphi(u^k,\tilde u^k).\tag{4.27}$$

Assertion (4.14) follows from (4.26), (4.27) and $u^{k+1}=u_{II}(\gamma\alpha^*_{IIk})$ directly. □

Lemma 4.11 *Let α^*_{IIk} be defined in* (4.18b)*. If $\nu\leq2\sqrt{1-\mu}$ is additionally requested, then we have $\alpha^*_{IIk}\geq\frac{1}{2\|G\|}$.*

Proof Note that

$$\|G(u^k-\tilde u^k)+\xi^k\|^2\leq\|G\|\cdot\|(u^k-\tilde u^k)+G^{-1}\xi^k\|_G^2.$$

It follows from (4.13b) and (4.18b) immediately that $\alpha^*_{IIk}\geq\frac{\alpha^*_{Ik}}{\|G\|}$. Recall Lemma 4.7, we prove the assertion. □

Lemmas 4.5, 4.10 and 4.11 imply the following result immediately.

Lemma 4.12 *Let $\{u^k\}$ be the sequence generated by the second prediction-correction scheme* (4.18). *Then, we have*

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \frac{\gamma(2-\gamma)(2-\mu-\nu)}{4\,\|G\|}\|u^k - \tilde{u}^k\|_G^2, \quad \forall\, u^* \in \Omega^*. \tag{4.28}$$

Now, we are ready to prove the convergence of the second prediction-correction scheme to implement the inexact version of the proposed unified framework with the inexact criterion (2.9).

Theorem 4.13 *The sequence $\{u^k\}$ generated by the second prediction-correction scheme of the inexact version of the proposed unified framework with the inexact criterion (2.9) converges to a solution of VI(Ω, F).*

Proof Due to (4.28), the proof is analogous to that of Theorem 4.2, thus omitted. □

5 Conclusions

This paper presents the unified framework of proximal-based decomposition methods for solving a class of monotone variational inequalities (VI) with separable structure. The main task of each iteration consists of solving two sub-VIs that are regularized by the classical proximal terms. Depending on whether approximate solutions of the sub-VIs are allowed, we first present both the exact and inexact versions of the unified framework. Afterwards, we propose a condition to ensure the convergence of the unified framework, and then analyze how to fulfill this condition via choosing appropriate proximal parameters for some special cases of the unified framework. Our theoretical analysis indicates that the proximal-based alternating decomposition methods generally allow more favorable proximal parameters than the proximal-based parallel decomposition methods do. On the other hand, the proposed unified framework results in some new implementable numerical algorithms. In particular, the newly-derived inexact algorithms improve the existing methods in the sense that the involved sub-VIs are allowed to be solved approximately under significantly relaxed criteria.

References

[1] Bauschke, H.H., Combettes, P. L.: A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. Math. Oper. Res. **26(2)**, 248-264 (2001)

[2] Bertsekas, D.P., Tsitsiklis, J.N.: Parallel and Distributed Computation: Numerical Methods. Prentice-Hall, Englewood Cliffs, NJ (1989)

[3] Chen, G., Teboulle, M.: A proximal-based decomposition method for convex minimization problems. Math. Program. **64**, 81-101 (1994)

[4] Dafermos, S.: Taffic equilibrium and variational inequalities. Trans. Sci. **14**, 42-54 (1980)

[5] Eckstein, J.: Some saddle-function splitting methods for convex programming. Optim. Methods Soft. **4**, 75-83 (1994)

[6] Eckstein, J., Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal points algorithm for maximal monotone operators. Math. Program. **55**, 293-318 (1992)

[7] Eckstein, J. , Fukushima, M.: Some reformulation and applications of the alternating directions method of multipliers. In: Hager, W. W. *et al.* eds., Large Scale Optimization: State of the Art, Kluwer Academic Publishers, 115-134 (1994)

[8] Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarity problems. SIAM Rev. **39**, 669-713 (1997)

[9] Fukushima, M.: Application of the alternating directions method of multipliers to separable convex programming problems. Comput. Optim. Appl. **2**, 93-111 (1992)

[10] Fuller, J.D., Chung, W.: Benders decomposition for a class of variational inequalities. Euro. J. Oper. Res. **185**, 76-91 (2008)

[11] Gabay, D.: Application of the method of multipliers to varuational inequalities. In: Fortin, M., Glowinski, R., eds., Augmented Lagrangian methods: Application to the numerical solution of Boundary-Value Problem, North-Holland, Amsterdam, The Netherlands, 299-331 (1983)

[12] Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Comput. Math. Appli. **2**, 17-40 (1976)

[13] Glowinski, R.: Numerical methods for Nonlinear Variational Problems. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo (1984)

[14] Güler, O.: On the convergence of the proximal point algorithm for convex minimization. SIAM J. Con. Optim. **29(2)**, 403-419 (1991)

[15] Han, D. R., Xu, W., Yang, H.: An operator splitting method for variational inequalities with partially unknown mappings. Numer. Math. **111**, 207-237 (2008)

[16] He, B.S.: Inexact implicit methods for monotone general variational inequalities. Math. Program. **86**, 199-217 (1999)

[17] He, B.S, Liao, L.-Z., Han, D.R., Yang, H.: A new inexact alternating directions method for monotone variational inequalities. Math. Program. **92**, 103-118 (2002)

[18] He, B.S, Liao, L.-Z., Wang, S.L.: Self-adaptive operator splitting methods for monotone variational inequalities. Numer. Math. **94 (4)**, 715-737 (2003)

[19] He, N.: Parallel decomposition methods for structured monotone variational inequalities. Master Thesis, Department of Mathematics, Nanjing University, 2009.

[20] Kinderlehrer, D., Stampacchia, G.: A introduction to variational inequalities and their applications. Academic Press, New York (1980)

[21] Kontogiorgis, S., Meyer, R.R.: A variable-penalty alternating directions method for convex optimization. Math. Program. **83**, 29-53 (1998)

[22] Martinet, B.: Regularization d'inequations variationelles par approximations sucessives. Revue Francaise d'Informatique et de RechercheOpérationelle **4**, 154-159 (1970)

[23] Nagurney, A.: Network economics, A variational inequality approach. Kluwer Academics Publishers, Dordrecht (1993)

[24] Rockafellar, R.T.: Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. **1**, 97-116 (1976)

[25] Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Con. Optim. **14**, 877-898 (1976)

[26] Sun, J. and Zhang, S.: A modified alternating direction method for convex quadratically constrained quadratic semidefinite programs. Manuscript, 2009

[27] Tseng, P.: Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming. Math. Program. **48**, 249-263 (1990)

[28] Tseng, P.: Alternating projection-proximal methods for convex programming and variational inequalities. SIAM J. Optim. **7**, 951-965 (1997)

[29] Tseng, P.: A modified forward-backword splitting method for maximal monotone mappings. SIAM J. Con. Optim. **38**, 431-446 (2000)