AMO - Advanced Modeling and Optimization, Volume 13, Number 2, 2011

Convergence of Alternating Directions like Methods for Linearly Constrained Structured Variational Inequalities

Bingsheng He*
Department of Mathematics, Nanjing University, Nanjing, 210093, China.

Abstract

Because of its significant efficiency and easy implementation, alternating direction method (ADM) has attracted wide attention in solving linearly constrained structured convex optimization and variational inequalities. In this paper, we propose the most potential versions of the proximal ADM and investigate their convergence in a uniform framework. The additional proximal term allows us to simplify the sub-problems, and thus the new versions substantially broaden the applicable scope of the alternating direction methods. The convergence is based on the fact that the sequence generated by each different versions approaches to the solution set monotonically in the Fejér sense.

Keywords: alternating direction method, linearly constrained convex programming, separable structure, contraction method

1 Introduction

Variational inequalities (VI) capture a broad spectrum of applications in diverse fields, see, e.g., [7, 12, 15, 26]. In this paper, we consider the VI with the following separable structure:

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

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 [9, 14, 31] for the various applications of (1.1)-(1.2) in other some 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 convex functions.

*This author was supported by the NSFC Grant 10971095 and the NSF of Jiangsu Province Grant BK2008255.
Email: `hebma@nju.edu.cn`
AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

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:

$$(x^*, y^*, \lambda^*) \in \mathcal{W}, \quad \left\{ \begin{array}{l} (x - x^*)^T (f(x^*) - A^T\lambda^*) \geq 0, \\ (y - y^*)^T (g(y^*) - B^T\lambda^*) \geq 0, \\ (\lambda - \lambda^*)^T (Ax^* + By^* - b) \geq 0, \end{array} \right. \quad \forall\, (x, y, \lambda) \in \mathcal{W} := \mathcal{X} \times \mathcal{Y} \times \Re^m. \tag{1.3}$$

We denote (1.3) by SVI$(\mathcal{W}, F)$, where

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

Throughout this paper, we assume that the solution set of SVI$(\mathcal{W}, F)$, denoted by $\mathcal{W}^*$, is non-empty. For solving SVI problems, from a given triplet $w^k = (x^k, y^k, \lambda^k) \in \mathcal{X} \times \mathcal{Y} \times \Re^m$, the *Alternating Directions Methods* (short ADM) [10, 11, 12, 13] produce the new iterate $v^{k+1} = (y^{k+1}, \lambda^{k+1}) \in \mathcal{Y} \times \mathcal{R}^l$ via the following procedure: First, take the solution of problem

$$x \in \mathcal{X}, \quad (x' - x)^T \big\{ f(x) - A^T[\lambda^k - \beta(Ax + By^k - b)] \big\} \geq 0, \quad \forall\ x' \in \mathcal{X}, \tag{1.5a}$$

as x^{k+1}. Then, y^{k+1} is produced by solving

$$y \in \mathcal{Y}, \quad (y' - y)^T \big\{ g(y) - B^T[\lambda^k - \beta(Ax^{k+1} + By - b)] \big\} \geq 0, \quad \forall\ y' \in \mathcal{Y}. \tag{1.5b}$$

Finally, the multipliers are updated by

$$\lambda^{k+1} = \lambda^k - \gamma\beta(Ax^{k+1} + By^{k+1} - b), \tag{1.5c}$$

where $\gamma \in (0, \frac{\sqrt{5}+1}{2})$ is a parameter. In the most of literature about ADM [3, 4, 8, 23], the parameter $\gamma = 1$. A simple form for choosing the parameter β was discussed in [18]. Alternating direction method is well suited to distributed convex optimization and has the benefit that one algorithm could be flexible enough to solve many problems [3].

Some novel and attractive applications of ADM have been discovered very recently, e.g., total-variation regularization problems in image processing [5, 27, 30], ℓ_1-norm minimization in compressive sensing [34], semidefinite programming problems [32], the covariance selection problem and semidefinite least square problem in statistics [20, 37], the sparse and low-rank recovery problem in engineering [24, 36], etc. ADM is also modified to solve convex quadratically constrained quadratic semidefinite programs [29]. Sometimes, solving the subproblems (1.5a) and/or (1.5b) is difficult. The purpose of this paper is to present and study some different modified versions of the alternating direction method in the contraction framework [2, 19].

The rest of the paper is organized as follows. In Section 2, as a preparation for the rest analysis, we review some basic properties of projection mapping and variational inequalities. Section 3 presents the proximal alternating direction method. In section 4, using the rationale of the general framework, we give the updating forms and prove the convergence of the resulting methods. From Section 5 to Section 7, we study different simplifying versions of the alternating directions scheme and give the parallel analysis as in Section 3. Finally, some conclusions are drawn in Section 8. For convenience we use the notations

$$v = \begin{pmatrix} y \\ \lambda \end{pmatrix} \qquad \text{and} \qquad \mathcal{V}^* = \{(y^*, \lambda^*) \,|\, (x^*, y^*, \lambda^*) \in \mathcal{W}^*\}. \tag{1.6}$$

2 Preliminaries

In this section, we summarize some basic properties and related definitions that will be used in the coming analysis and discussions.

2.1 Preliminaries of variational inequalities

Let $\mathcal{W}$ be a nonempty subset of $\Re^l$, F be a continuous mapping from $\Re^l$ to itself. The variational inequality problem, denoted by VI$(\mathcal{W}, F)$, is to find a vector $w^* \in \mathcal{W}$ such that

$$\text{VI}(\mathcal{W}, F) \qquad (w - w^*)^T F(w^*) \geq 0, \quad \forall\, w \in \mathcal{W}.$$

Let G be a $l \times l$ positive definite matrix, we denote $\|w\|_G = \sqrt{w^T G w}$ as the G-norm of vector $w \in \Re^n$. The projection under G-norm will be denoted by $P_{\mathcal{W},G}(\cdot)$. In other words, for given $\bar{w}$,

$$P_{\mathcal{W},G}(\bar{w}) = \text{argmin}\{\|\bar{w} - w\|_G \mid w \in \mathcal{W}\}.$$

From the above definition, it follows that

$$(w - P_{\mathcal{W},G}(w))^T G(\bar{w} - P_{\mathcal{W},G}(w)) \leq 0, \quad \forall\, w \in \Re^l, \forall \bar{w} \in \mathcal{W}. \tag{2.1}$$

Consequently, we have

$$\|P_{\mathcal{W},G}(w) - P_{\mathcal{W},G}(\bar{w})\|_G \leq \|w - \bar{w}\|_G, \qquad \forall\, w, \bar{w} \in \Re^l, \tag{2.2}$$

and

$$\|P_{\mathcal{W},G}(w) - \bar{w}\|_G^2 \leq \|w - \bar{w}\|_G^2 - \|w - P_{\mathcal{W},G}(w)\|_G^2, \qquad \forall\, w \in \Re^l, \forall\, \bar{w} \in \mathcal{W}. \tag{2.3}$$

Definition 2.1. a). F is said to be monotone respect to $\mathcal{W}$ if

$$(w - \bar{w})^T (F(w) - F(\bar{w})) \geq 0, \quad \forall\, w, \bar{w} \in \mathcal{W}.$$

b). F is strongly monotone respect to $\mathcal{W}$ if there exists a constant $\mu > 0$ such that

$$(w - \bar{w})^T (F(w) - F(\bar{w})) \geq \mu \|w - \bar{w}\|^2, \quad \forall\, w, \bar{w} \in \mathcal{W}.$$

We say $VI(\mathcal{W}, F)$ is monotone if the mapping F is monotone.

Lemma 2.2. *Let $G \in \Re^{l\times l}$ be any positive definite matrix. Then w^* is a solution of VI$(\mathcal{W}, F)$ if and only if*

$$w^* = P_{\mathcal{W},G}[w^* - \alpha_k G^{-1} F(w^*)], \qquad \forall\, \alpha > 0. \tag{2.4}$$

Proof. See ([1], pp. 267). $\square$

According to Lemma 2.2, for any positive definite matrix $G \in \Re^{l\times l}$, $p \in \Re^l$ and $\alpha > 0$,

$$w^* = P_{\mathcal{W},G}[w^* - \alpha G^{-1} p] \qquad \Leftrightarrow \qquad w^* \in \mathcal{W}, \quad (w - w^*)^T p \geq 0, \quad \forall\, w \in \mathcal{W}. \tag{2.5}$$

The solution set of a monotone variational inequality is convex (see Theorem 2.3.5 in [6]).

2.2 Concepts concerned with SVI($\mathcal{W}, F$)

The following concepts concern the problem (1.3) in this paper. Let H be a given $m \times m$ positive definite matrix. For monotone VI($\mathcal{W}, F$), we have

$$(\tilde{w} - w^*)^T F(\tilde{w}^k) \geq (\tilde{w} - w^*)^T F(w^*), \quad \forall\, \tilde{w} \in \mathcal{W},\; w^* \in \mathcal{W}^*.$$

Consequently, because $(\tilde{w} - w^*)^T F(w^*) \geq 0$, we obtain

$$(\tilde{w} - w^*)^T F(\tilde{w}) \geq 0, \quad \forall\, \tilde{w} \in \mathcal{W}. \tag{2.6}$$

Throughout this paper, for $y,\, \tilde{y} \in \Re^{n_2}$, we define

Lemma 2.3. *For given w, let $\tilde{w} \in \mathcal{W}$ satisfy*

$$\tilde{w} \in \mathcal{W}, \quad (w' - \tilde{w})^T \big\{ \big(F(\tilde{w}) + \eta(y, \tilde{y})\big) - d(w, \tilde{w}) \big\} \geq 0, \qquad \forall\, w' \in \mathcal{W}, \tag{2.7}$$

where

$$\eta(y, \tilde{y}) = \begin{pmatrix} A^T \\ B^T \\ 0 \end{pmatrix} HB(y - \tilde{y}). \tag{2.8}$$

Then we have

$$(w - w^*)^T d(w, \tilde{w}) \geq \varphi(w, \tilde{w}), \quad \forall\, w^* \in \mathcal{W}^*, \tag{2.9}$$

where

$$\varphi(w, \tilde{w}) = (w - \tilde{w})^T d(w, \tilde{w}) + (y - \tilde{y})^T B^T H(A\tilde{x} + B\tilde{y} - b). \tag{2.10}$$

Proof. For any $w^* \in \mathcal{W}^* \subset \mathcal{W}$, it follows from (2.7) that

$$(\tilde{w} - w^*)^T d(w, \tilde{w}) \geq (\tilde{w} - w^*)^T \{F(\tilde{w}) + \eta(y, \tilde{y})\}, \quad \forall\, w^* \in \mathcal{W}^*. \tag{2.11}$$

In addition, due to (2.6), we have $(\tilde{w} - w^*)^T F(\tilde{w}) \geq 0$. Substituting it in (2.11) and by a manipulation, we get

$$(w - w^*)^T d(w, \tilde{w}) \geq (w - \tilde{w})^T d(w, \tilde{w}) + (\tilde{w} - w^*)^T \eta(y, \tilde{y}), \quad \forall\, w^* \in \mathcal{W}^*. \tag{2.12}$$

For the last term in (2.12), using $Ax^* + By^* = b$, we have

$$(\tilde{w} - w^*)^T \eta(y, \tilde{y}) = (y - \tilde{y})^T B^T H(A\tilde{x} + B\tilde{y} - b), \quad \forall\, w^* \in \mathcal{W}^*. \tag{2.13}$$

Substituting it in the right hand side of (2.12) and using the definition of $\varphi(w, \tilde{w})$, we proved the assertion of this lemma. $\square$

Theorem 2.4. *Let $w, \tilde{w}$ satisfy the conditions in Lemma 2.3 and G be any positive definite matrix. If we take*

$$w(\alpha) = w - \alpha G^{-1} d(w, \tilde{w}), \tag{2.14a}$$

or

$$w(\alpha) = P_{\mathcal{W}, G}\big\{ w - \alpha G^{-1}[F(\tilde{w}) + \eta(y, \tilde{y})] \big\}, \tag{2.14b}$$

then we have

$$\|w - w^*\|_G^2 - \|w(\alpha) - w^*\|_G^2 \geq 2\alpha\varphi(w, \tilde{w}) - \alpha^2 \|G^{-1} d(w, \tilde{w})\|_G^2, \quad \forall w^* \in \mathcal{W}^*. \tag{2.15}$$

Proof. By using (2.14a) and (2.9), we obtain

$$\begin{aligned}
&\|w-w^*\|_G^2-\|w(\alpha)-w^*\|_G^2\\
&\quad= \|w-w^*\|_G^2-\|(w-w^*)-\alpha G^{-1}d(w,\tilde{w}\|_G^2\\
&\quad= 2\alpha(w-w^*)^Td(w,\tilde{w})-\alpha^2\|G^{-1}d(w,\tilde{w})\|_G^2\\
&\quad\geq 2\alpha\varphi(w,\tilde{w})-\alpha^2\|G^{-1}d(w,\tilde{w})\|_G^2. \qquad (2.16)
\end{aligned}$$

Now, we turn to use (2.14b). Since $w^*\in\mathcal{W}$, we have (see (2.3))

$$\|P_{\mathcal{W},G}(\bar{w})-w^*\|_G^2\leq\|\bar{w}-w^*\|_G^2-\|\bar{w}-P_{\mathcal{W},G}(\bar{w})\|_G^2,\quad \forall\bar{w}\in\Re^{n+m}.$$

Setting $\bar{w}=w-\alpha G^{-1}[F(\tilde{w})+\eta(y,\tilde{y})]$ in the above inequality and using $w(\alpha)=P_{\mathcal{W},G}(\bar{w})$, we get

$$\begin{aligned}
\|w(\alpha)-w^*\|_G^2 \leq\ & \|(w-w^*)-\alpha G^{-1}[F(\tilde{w})+\eta(y,\tilde{y})]\|_G^2\\
&-\|(w-w(\alpha))-\alpha G^{-1}[F(\tilde{w})+\eta(y,\tilde{y})]\|_G^2.
\end{aligned}$$

From the above inequality, we obtain

$$\begin{aligned}
&\|w-w^*\|_G^2-\|w(\alpha)-w^*\|_G^2\\
&\quad\geq \|w-w^*\|_G^2-\|(w-w^*)-\alpha G^{-1}[F(\tilde{w})+\eta(y,\tilde{y})]\|_G^2\\
&\qquad+\|(w-w(\alpha))-\alpha G^{-1}[F(\tilde{w})+\eta(y,\tilde{y})]\|_G^2\\
&\quad= \|w-w(\alpha)\|_G^2+2\alpha(w(\alpha)-w^*)^T\big(F(\tilde{w})+\eta(y,\tilde{y})\big). \qquad (2.17)
\end{aligned}$$

Since $w(\alpha)\in\mathcal{W}$, it follows from (2.7) that

$$(w(\alpha)-\tilde{w})^T\big(F(\tilde{w})+\eta(y,\tilde{y})\big)\geq(w(\alpha)-\tilde{w})^Td(w,\tilde{w}).$$

Using $(\tilde{w}-w^*)^TF(\tilde{w})\geq 0$ and (2.13), we have

$$(\tilde{w}-w^*)^T\big(F(\tilde{w})+\eta(y,\tilde{y})\big)\geq(y^k-\tilde{y}^k)^TH(A\tilde{x}+B\tilde{y}-b).$$

Adding the above two inequalities and using the definition of $\varphi(w^k,\tilde{w}^k)$, we obtain

$$(w(\alpha)-w^*)^T\big(F(\tilde{w})+\eta(y,\tilde{y})\big)\geq\varphi(w,\tilde{w})+(w(\alpha)-w)^Td(w,\tilde{w}). \qquad (2.18)$$

Substituting (2.18) in the right hand side of (2.17), we obtain

$$\begin{aligned}
&\|w-w^*\|_G^2-\|\tilde{w}-w^*\|_G^2\\
&\quad\geq \|w-\tilde{w}\|_G^2+2\alpha\varphi(w,\tilde{w})+2\alpha(\tilde{w}-w)^Td(w,\tilde{w})\\
&\quad= \|w-\tilde{w}-\alpha G^{-1}d(w,\tilde{w})\|_G^2+2\alpha\varphi(w,\tilde{w})-\alpha^2\|G^{-1}d(w,\tilde{w})\|_G^2\\
&\quad\geq 2\alpha\varphi(w,\tilde{w})-\alpha^2\|G^{-1}d(w,\tilde{w})\|_G^2. \qquad (2.19)
\end{aligned}$$

The proof of this theorem is complete. □

The right hand side of (2.15) is a quadratic function of α. In the case that $\varphi(w,\tilde{w})>0$, it reaches its maximum at

$$\alpha^*=\frac{\varphi(w,\tilde{w})}{\|G^{-1}d(w,\tilde{w})\|_G^2}.$$

From now on, the matrix M is defined by

$$M = \begin{pmatrix} rI_{n_1} & 0 & 0 \\ 0 & sI_{n_2} + B^T HB & 0 \\ 0 & 0 & H^{-1} \end{pmatrix} \tag{2.20}$$

Note that M $(r, s \geq 0)$ is positive semi-definite. For a proper symmetric matrix M, we use $\|w\|_M^2$ to denote that

$$\|w\|_M^2 = w^T M w,$$

even though M is not positive semi-definite.

3 Proximal alternating direction method scheme

We abuse the notations that have been used in Section 2 without ambiguity. A superscript such as in w^k refers to a specific vector and usually denotes an iteration index. As in the proximal point algorithm [25, 28], we add the proximal term to the subproblems in (1.5). The following scheme was established in [16].

Proximal alternating direction method scheme (ADM-scheme):

1. With available x^k, y^k and λ^k, solve the variational inequality problem

$$x \in \mathcal{X}, \;\; (x' - x)^T \big\{ f(x) - A^T[\lambda^k - H(Ax + By^k - b)] + r(x - x^k) \big\} \geq 0, \; \forall\, x' \in \mathcal{X}, \tag{3.1a}$$

   and denote the solution by $\tilde{x}^k$.

2. With available $\tilde{x}^k, y^k$ and λ^k, solve the variational inequality problem

$$y \in \mathcal{Y}, \;\; (y' - y)^T \big\{ g(y) - B^T[\lambda^k - H(A\tilde{x}^k + By - b)] + s(y - y^k) \big\} \geq 0, \;\; \forall\, y' \in \mathcal{Y}, \tag{3.1b}$$

   and denote the solution by $\tilde{y}^k$.

3. Set

$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b). \tag{3.1c}$$

The scheme (3.1) generates $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{W}$ in an alternating order and thereby adopts the new information whenever possible. The additional $r(x - x^k)$ in (3.1a) (resp. $s(y - y^k)$ in (3.1b)) is the proximal term in the sub-problems. Thus, we named (3.1) *Proximal Alternating Direction Method Scheme*. r, $s \geq 0$ are called the proximal coefficients.

Note that the solution $(\tilde{x}^k, \tilde{y}^k)$ of (3.1a)-(3.1b) satisfies

$$\begin{cases} (x' - \tilde{x}^k)^T \big\{ f(\tilde{x}^k) - A^T \lambda^k + A^T H \big(A\tilde{x}^k + By^k - b \big) + r(\tilde{x}^k - x^k) \big\} \geq 0, \; \forall\, x' \in \mathcal{X}, \\ (y' - \tilde{y}^k)^T \big\{ g(\tilde{y}^k) - B^T \lambda^k + B^T H \big(A\tilde{x}^k + B\tilde{y}^k - b \big) + s(\tilde{y}^k - y^k) \big\} \geq 0, \; \forall\, y' \in \mathcal{Y}. \end{cases} \tag{3.2}$$

Using $\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b)$ and by a manipulation, (3.2) can be rewritten as

$$(\tilde{x}^k, \tilde{y}^k) \in \mathcal{X} \times \mathcal{Y}, \quad \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k + A^T H\big(B(y^k - \tilde{y}^k)\big) \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^T H\big(B(y^k - \tilde{y}^k)\big) \end{pmatrix}$$
$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} & 0 \\ 0 & sI_{n_2} + B^T HB \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \end{pmatrix}, \ \forall (x', y') \in \mathcal{X} \times \mathcal{Y}. \quad (3.3)$$

If $w^k = \tilde{w}^k$, it follows from (3.3) and (3.1c) that

$$\begin{cases} \tilde{x}^k \in \mathcal{X}, & (x' - \tilde{x}^k)^T\{f(\tilde{x}^k) - A^T\tilde{\lambda}^k\} \geq 0, \quad \forall x' \in \mathcal{X}, \\ \tilde{y}^k \in \mathcal{Y}, & (y' - \tilde{y}^k)^T\{g(\tilde{y}^k) - B^T\tilde{\lambda}^k\} \geq 0, \quad \forall y' \in \mathcal{Y}, \\ & A\tilde{x}^k + B\tilde{y}^k - b = 0. \end{cases}$$

Hence, in this case, $\tilde{w}^k$ is a solution of the problem (1.3). In general, we have the following lemma.

Lemma 3.1. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. Then, we have*

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T\big\{\big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big) - d(w^k, \tilde{w}^k)\big\} \geq 0, \qquad \forall\, w' \in \mathcal{W}, \qquad (3.4)$$

where $\eta(y^k, \tilde{y}^k)$ is defined in (2.8),

$$d(w^k, \tilde{w}^k) = M(w^k - \tilde{w}^k), \qquad (3.5)$$

and M is defined in (2.20).

Proof. Since $A\tilde{x}^k + B\tilde{y}^k - b = H^{-1}(\lambda^k - \tilde{\lambda}^k)$, adding the equality

$$(\lambda' - \tilde{\lambda}^k)^T\big(A\tilde{x}^k + B\tilde{y}^k - b\big) = (\lambda' - \tilde{\lambda}^k)^T H^{-1}(\lambda^k - \tilde{\lambda}^k)$$

to (3.3), we obtain $\tilde{w}^k \in \mathcal{W}$ and

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \begin{pmatrix} A^T \\ B^T \\ 0 \end{pmatrix} H\big(B(y^k - \tilde{y}^k)\big) \right\}$$
$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} & 0 & 0 \\ 0 & sI_{n_2} + B^T HB & 0 \\ 0 & 0 & H^{-1} \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix}, \ \forall\, w' \in \mathcal{W}. \quad (3.6)$$

Using the notations of $F(\tilde{w}^k)$, $\eta(y^k, \tilde{y}^k)$ and the matrix M, the above inequality is

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T\big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big) \geq (w' - \tilde{w}^k)^T M(w^k - \tilde{w}^k), \quad \forall\, w' \in \mathcal{W},$$

and the assertion of this lemma is proved. □

Lemma 3.2. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. Then, we have*

$$(w^k - w^*)^T d(w^k, \tilde{w}^k) \geq \varphi(w^k, \tilde{w}^k), \quad \forall\, w^* \in \mathcal{W}^*, \tag{3.7}$$

where

$$\varphi(w^k, \tilde{w}^k) = \|w^k - \tilde{w}^k\|_M^2 + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \tag{3.8}$$

and M is defined in (2.20).

Proof. Note that $(w^k, \tilde{w}^k)$ in Lemma 3.1 satisfies (2.7) in Lemma 2.3. Setting

$$d(w^k, \tilde{w}^k) = M(w^k - \tilde{w}^k) \quad \text{and} \quad H(A\tilde{x}^k + B\tilde{y}^k - b) = \lambda^k - \tilde{\lambda},$$

in Lemma 2.3, we proved the assertion of this lemma. □

For the terms in the right hand side of (3.8), using $H^{-1}(\lambda^k - \tilde{\lambda}^k) = A\tilde{x}^k + B\tilde{y}^k - b$ (see (3.1c), we obtain

$$\begin{aligned} & 2(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \\ = \;& \|B(y^k - \tilde{y}^k) + H^{-1}(\lambda^k - \tilde{\lambda}^k)\|_H^2 - \|B(y^k - \tilde{y}^k)\|_H^2 - \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2 \\ = \;& \|A\tilde{x}^k + By^k - b\|_H^2 - \|B(y^k - \tilde{y}^k)\|_H^2 - \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2. \end{aligned} \tag{3.9}$$

Theorem 3.3. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. Then, we have*

$$\|\tilde{w}^k - w^*\|_M^2 \leq \|w^k - w^*\|_M^2 - \big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2 + \|A\tilde{x}^k + By^k - b\|_H^2\big) \tag{3.10}$$

where M is defined in (2.20).

Proof. Since $(w^k - \tilde{w}^k)^T d(w^k, \tilde{w}^k) = \|w^k - \tilde{w}^k\|_M^2$, it follows from the identity

$$\|w^k - w^*\|_M^2 - \|\tilde{w}^k - w^*\|_M^2 = 2(w^k - w^*)^T M(w^k - \tilde{w}^k) - \|w^k - \tilde{w}^k\|_M^2$$

and Lemma 3.2 that

$$\|w^k - w^*\|_M^2 - \|\tilde{w}^k - w^*\|_M^2 \quad \geq \quad \|w^k - \tilde{w}^k\|_M^2 + 2(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k). \tag{3.11}$$

According to the definition of the matrix M (see (2.20)), the assertion of this theorem is followed from (3.11) and (3.9) directly. □

4 ADMs using different update forms

Based on the same $\tilde{w}^k$ generated by the proximal ADM-scheme (3.1), we consider the following three kinds of update forms for producing the next iterate w^{k+1}.

1. Direct update form: $w^{k+1} = \tilde{w}^k$;

2. Combinative update form: $w^{k+1} = (1 - \alpha_k)w^k + \alpha_k \tilde{w}^k$;

3. Contractive update form:

$$w^{k+1} = w^k - \alpha_k G^{-1}(w^k - \tilde{w}^k),$$

or

$$w^{k+1} = P_{\mathcal{W},G}\big\{w^k - \alpha_k G^{-1}[F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)]\big\},$$

where G is a given positive definite matrix.

By setting the non-negative proximal coefficients $r = 0$ or/and $s = 0$ in (3.1), we obtain different versions of the alternating direction methods. The principal inequalities for the convergence of such methods are followed from (3.10) directly.

The parameters r and s in the different cases

	Case I	Case II	Case III	Case IV
$r,\ s$	$r>0,\ s>0$	$r=0,\ s>0$	$r>0,\ s=0$	$r=0,\ s=0$

4.1 Direct update form based on the ADM-scheme

The alternating direction methods using the direct update form take $w^{k+1} = \tilde{w}^k$ as the next iterate. The assertion in Corollary 4.1 can be found in [16] and [33].

Corollary 4.1. *Let $r,\ s > 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If $w^{k+1} = \tilde{w}^k$, then*

$$\begin{aligned} & r\|x^{k+1}-x^*\|^2 + s\|y^{k+1}-y^*\|^2 + \|B(y^{k+1}-y^*)\|_H^2 + \|\lambda^{k+1}-\lambda^*\|_{H^{-1}}^2 \\ \le\ & \big(r\|x^k-x^*\|^2 + s\|y^k-y^*\|^2 + \|B(y^k-y^*)\|_H^2 + \|\lambda^k-\lambda^*\|_{H^{-1}}^2\big) \\ & -\big(r\|x^k-x^{k+1}\|^2 + s\|y^k-y^{k+1}\|^2 + \|Ax^{k+1}+By^k-b\|_H^2\big). \end{aligned} \tag{4.1}$$

Proof. It follows from (3.10) by setting $w^{k+1} = \tilde{w}^k$. □

Corollary 4.2. *Let $r = 0,\ s > 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector v^k. Setting $v^{k+1} = \tilde{v}^k$, we have*

$$\begin{aligned} & s\|y^{k+1}-y^*\|^2 + \|B(y^{k+1}-y^*)\|_H^2 + \|\lambda^{k+1}-\lambda^*\|_{H^{-1}}^2 \\ \le\ & \big(s\|y^k-y^*\|^2 + \|B(y^k-y^*)\|_H^2 + \|\lambda^k-\lambda^*\|_{H^{-1}}^2\big) \\ & -\big(s\|y^k-y^{k+1}\|^2 + \|Ax^{k+1}+By^k-b\|_H^2\big). \end{aligned} \tag{4.2}$$

Proof. It follows from (4.1) by setting $r = 0$. □

In the following we consider the cases that $s = 0$. First, since $y^k \in \mathcal{Y}$ and $s = 0$, it follows from (3.1b) and (3.1c) that

$$(y^k - \tilde{y})^T\big(g(\tilde{y}^k) - B^T\tilde{\lambda}^k\big) \ge 0.$$

Because $w^{k+1} = \tilde{w}^k$ and $\tilde{y}^k \in \mathcal{Y}$, for the (y^k, λ^k) in the last iteration, we have

$$(\tilde{y} - y^k)^T\big(g(y^k) - B^T\lambda^k\big) \ge 0.$$

Adding the above two inequalities and using the monotonicity of g, we get

$$(y^k - \tilde{y}^k)^T B^T (\lambda^k - \tilde{\lambda}^k) \geq 0.$$

Therefore, by using $(A\tilde{x}^k + B\tilde{y}^k - b) = H^{-1}(\lambda^k - \tilde{\lambda}^k)$, we have

$$\begin{aligned}\|A\tilde{x}^k + By^k - b\|_H^2 &= \|B(y^k - \tilde{y}^k) + (A\tilde{x}^k + B\tilde{y}^k - b)\|_H^2 \\ &= \|B(y^k - \tilde{y}^k) + H^{-1}(\lambda^k - \tilde{\lambda}^k)\|_H^2 \\ &\geq \|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2. \end{aligned} \tag{4.3}$$

Corollary 4.3. *Let $r > 0$, $s = 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. Setting $w^{k+1} = \tilde{w}^k$, we have*

$$\begin{aligned} & r\|x^{k+1} - x^*\|^2 + \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\ & \leq \big(r\|x^k - x^*\|^2 + \|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2\big) \\ & \quad -\big(r\|x^k - x^{k+1}\|^2 + \|B(y^k - y^{k+1})\|_H^2 + \|\lambda^k - \lambda^{k+1}\|_{H^{-1}}^2\big). \end{aligned} \tag{4.4}$$

Proof. It follows from (4.1) and (4.3) directly. □

In the original alternating direction method [10, 12, 13, 23], the parameters $r = s = 0$. The assertion in the following corollary is a special case of Theorem 1 in [22] by setting $\gamma = 1$ and $\beta I = \tilde{\beta} I = H$.

Corollary 4.4. *Let $r = s = 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector v^k. Setting $v^{k+1} = \tilde{v}^k$, we have*

$$\begin{aligned} & \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\ & \leq \big(\|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2\big) \\ & \quad -\big(\|B(y^k - y^{k+1})\|_H^2 + \|\lambda^k - \lambda^{k+1}\|_{H^{-1}}^2\big). \end{aligned} \tag{4.5}$$

Proof. The assertion (4.5) follows from (4.4) by setting $r = 0$. □

Note that (4.4) and (4.5) can be written in a compact form

$$\|w^{k+1} - w^*\|_M^2 \leq \|w^k - w^*\|_M^2 - \|w^k - w^{k+1}\|_M^2,$$

where M is defined in (2.20) with $s = 0$ and $r = s = 0$, respectively. By using the direct update form, $w^{k+1} = \tilde{w}^k$, we have always that (see Theorem 3.3)

$$\|w^{k+1} - w^*\|_M^2 < \|w^k - w^*\|_M^2,$$

whenever $\tilde{w}^k \neq w^k$. Thus, the assertion inequality of the corollaries 4.1–4.4 is called contraction inequality which is essential for the global convergence.

Contraction properties of different versions using direct update form

Cases	Case I	Case II	Case III	Case IV
r, s	$r > 0$, $s > 0$	$r = 0$, $s > 0$	$r > 0$, $s = 0$	$r = 0$, $s = 0$
Contraction inequality	(4.1)	(4.2)	(4.4)	(4.5)

4.2 Combinative update form based on the ADM-scheme

The combinative update form

$$w^{k+1} = (1-\alpha)w^k + \alpha\tilde{w}^k$$

can be rewritten as

$$w^{k+1} = w^k - \alpha(w^k - \tilde{w}^k).$$

For any $w^* \in \mathcal{W}^*$, $M(w-w^*)$ is the gradient of the unknown distance function $\frac{1}{2}\|w-w^*\|_M^2$ at point w. A direction d is called a descent direction of $\frac{1}{2}\|w-w^*\|_M^2$ if and only if the inner-product $\big\langle M(w-w^*),\, d\big\rangle < 0$. Applying Cauchy-Schwarz Inequality to (3.8), we have

$$\begin{aligned}\varphi(w^k,\tilde{w}^k) &= \|w^k-\tilde{w}^k\|_M^2 + (\lambda^k-\tilde{\lambda}^k)^T B(y^k-\tilde{y}^k)\\ &\geq \|w^k-\tilde{w}^k\|_M^2 - \frac{1}{2}\big(\|B(y^k-\tilde{y}^k)\|_H^2 + \|\lambda^k-\tilde{\lambda}^k\|_{H^{-1}}^2\big)\\ &= \frac{1}{2}\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2 + \|w^k-\tilde{w}^k\|_M^2\big). \end{aligned} \tag{4.6}$$

According to Lemma 3.2 and (4.6), $-(w^k-\tilde{w}^k)$ is a descent direction of the unknown distance function $\frac{1}{2}\|w-w^*\|_M^2$. Along this direction, we choose a point which is more closed to the solution set.

Combinative update form based on Proximal ADM-scheme (3.1)**.**
The new iterate w^{k+1} is given by

$$w^{k+1} = w^k - \alpha_k(w^k-\tilde{w}^k), \tag{4.7a}$$

where

$$\alpha_k = \omega\alpha_k^*, \quad \alpha_k^* = \frac{\varphi(w^k,\tilde{w}^k)}{\|w^k-\tilde{w}^k\|_M^2} \quad \text{and} \quad \omega\in(0,2). \tag{4.7b}$$

It follows from (4.6) and (4.7b), $\alpha_k^* \geq 0.5$.

Theorem 4.5. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.7)*, then we have*

$$\|w^{k+1}-w^*\|_M^2 \leq \|w^k-w^*\|_M^2 - \frac{\omega(2-\omega)}{4}\|w^k-\tilde{w}^k\|_M^2, \ \ \forall\, w^*\in\mathcal{W}^*, \tag{4.8}$$

where M is defined in (2.20).

Proof. By using (4.7) and (3.7), we obtain

$$\begin{aligned}&\|w^k-w^*\|_M^2 - \|w^{k+1}-w^*\|_M^2\\ &= \|w^k-w^*\|_M^2 - \|(w^k-w^*) - \omega\alpha_k^*(w^k-\tilde{w}^k)\|_M^2\\ &= 2\omega\alpha_k^*(w^k-w^*)^T M(w^k-\tilde{w}^k) - \omega^2(\alpha_k^*)^2\|w^k-\tilde{w}^k\|_M^2\\ &= 2\omega\alpha_k^*\varphi(w^k,\tilde{w}^k) - \omega^2(\alpha_k^*)^2\|w^k-\tilde{w}^k\|_M^2\\ &\geq \omega(2-\omega)\alpha^*\varphi(w^k,\tilde{w}^k). \end{aligned} \tag{4.9}$$

Using $\alpha_k^* \ge \frac{1}{2}$ and (4.6), it follows that

$$\alpha_k^* \varphi(w^k, \tilde{w}^k) \ge \frac{1}{4}\|w^k - \tilde{w}^k\|_M^2. \tag{4.10}$$

The assertion of this theorem follows from (4.9) and (4.10) immediately. □

The assertion inequality (4.8) is contractive and essential for the global convergence of the combinative update form based on the ADM-scheme. The following corollaries indicate the contraction properties by the different proximal parameter choices.

Corollary 4.6. *Let $r, s > 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.7)*, then for any $w^* \in \mathcal{W}^*$ we have*

$$\begin{aligned} & r\|x^{k+1} - x^*\|^2 + s\|y^{k+1} - y^*\|^2 + \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\ \le\ & \big(r\|x^k - x^*\|^2 + s\|y^k - y^*\|^2 + \|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2\big) \\ & -\frac{\omega(2-\omega)}{4}\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2 + \|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2\big). \end{aligned} \tag{4.11}$$

Corollary 4.7. *Let $r = 0$, $s > 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.7)*, then for any $w^* \in \mathcal{W}^*$ we have*

$$\begin{aligned} & s\|y^{k+1} - y^*\|^2 + \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\ \le\ & \big(s\|y^k - y^*\|^2 + \|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2\big) \\ & -\frac{\omega(2-\omega)}{4}\big(s\|y^k - \tilde{y}^k\|^2 + \|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2\big). \end{aligned} \tag{4.12}$$

Corollary 4.8. *Let $r > 0, s = 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.7)*, then for any $w^* \in \mathcal{W}^*$ we have*

$$\begin{aligned} & r\|x^{k+1} - x^*\|^2 + \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\ \le\ & \big(r\|x^k - x^*\|^2 + \|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2\big) \\ & -\frac{\omega(2-\omega)}{4}\big(r\|x^k - \tilde{x}^k\|^2 + \|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2\big). \end{aligned} \tag{4.13}$$

For $r = s = 0$, Ye and Yuan [35] use the update form (4.7) to produce the new iterate. The result in the following corollary followed from Theorem 4.5 and also can be found in [35].

Corollary 4.9. *Let $r = 0$, $s = 0$ and $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.7)*, then for any $w^* \in \mathcal{W}^*$ we have*

$$\begin{aligned} & \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\ \le\ & \big(\|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2\big) \\ & -\frac{\omega(2-\omega)}{4}\big(\|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2\big). \end{aligned} \tag{4.14}$$

4.3 G-norm contractive form based on the ADM-scheme

For any $w^* \in \mathcal{W}^*$, $G(w - w^*)$ is the gradient of the unknown distance function $\frac{1}{2}\|w - w^*\|_G^2$ at point w. A direction d is called a descent direction of $\frac{1}{2}\|w - w^*\|_G^2$ if and only if the inner-product

$\big\langle G(w-w^*),\, d\big\rangle < 0$. Since (see (3.7) and (4.6))

$$\big\langle G(w-w^*),\, G^{-1}d(w^k,\tilde{w}^k)\big\rangle \ge \varphi(w^k,\tilde{w}^k) \ge \frac{1}{2}\|w^k-\tilde{w}^k\|_M^2,$$

According to Lemma 3.2 and (4.6), $-(w^k-\tilde{w}^k)$ is a descent direction of the unknown distance function $\frac{1}{2}\|w^k-w^*\|_M^2$. Along this direction, we choose a point which is more closed to the solution set.

G-norm contractive update form based on Proximal ADM-scheme (3.1)**.**
The next iterate w^{k+1} is given by

$$w^{k+1} = w^k - \alpha_k G^{-1} d(w^k,\tilde{w}^k), \tag{4.15a}$$

or

$$w^{k+1} = P_{\mathcal{W},\,G}\big\{w^k - \alpha_k G^{-1}[F(\tilde{w}^k)+\eta(y^k,\tilde{y}^k)]\big\}, \tag{4.15b}$$

where

$$\alpha_k = \omega\alpha_k^*, \qquad \alpha_k^* = \frac{\varphi(w^k,\tilde{w}^k)}{\|G^{-1}d(w^k,\tilde{w}^k)\|_G^2} \qquad \text{and} \quad \omega\in(0,2). \tag{4.15c}$$

We have the similar principal contractive inequality as in Theorem 4.5.

Theorem 4.10. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (3.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.15)*, then we have*

$$\|w^{k+1}-w^*\|_G^2 \le \|w^k-w^*\|_G^2 - \frac{\omega(2-\omega)}{4\|G^{-1}\|\cdot\|M\|}\|w^k-\tilde{w}^k\|_M^2. \tag{4.16}$$

where M is defined in (2.20)*.*

Proof. First, it follows from Lemma 3.1 and Theorem 2.4 that

$$\|w^k-w^*\|_G^2 - \|w^{k+1}-w^*\|_G^2 \ge 2\omega\alpha_k^*\varphi(w^k,\tilde{w}^k) - (\omega\alpha_k^*)^2\|G^{-1}d(w^k,\tilde{w}^k)\|_G^2, \;\; \forall\, w^*\in\mathcal{W}^*.$$

By using $\alpha_k^*\|G^{-1}d(w^k,\tilde{w}^k)\|_G^2 = \varphi(w^k,\tilde{w}^k)$, we obtain

$$\|w^k-w^*\|_G^2 - \|w^{k+1}-w^*\|_G^2 \ge \omega(2-\omega)\alpha_k^*\varphi(w^k,\tilde{w}^k), \quad \forall w^*\in\mathcal{W}^*. \tag{4.17}$$

Since $d(w^k,\tilde{w}^k) = M(w^k-\tilde{w}^k)$, it follows that

$$\|G^{-1}d(w^k,\tilde{w}^k)\|_G^2 \le \|G^{-1}\|\cdot\|M(w^k-\tilde{w}^k)\|^2 \le (\|G^{-1}\|\cdot\|M\|)\|w^k-\tilde{w}^k\|_M^2.$$

Consequently, using (4.6) and (4.15c), we have

$$\alpha_k^*\varphi(w^k,\tilde{w}^k) \ge \frac{1}{4(\|G^{-1}\|\cdot\|M\|)}\|w^k-\tilde{w}^k\|_M^2.$$

Substituting it in (4.17), the theorem is proved. □

In the case of M is positive definite, a natural choice of G is $G=M$. Note that in this case the update form (4.15a) (resp. the step size (4.15c)) is the same as (4.7a) (resp. (4.7b)). Since M is block diagonal matrix, the update form (4.15b) is separable in forms

$$x^{k+1} = P_{\mathcal{X}}\{x^k - \alpha_k(\tfrac{1}{r})[f(\tilde{x}^k) - A^T\tilde{\lambda}^k + A^T HB(y^k-\tilde{y}^k)]\}, \tag{4.18a}$$

$$y^{k+1} = P_{\mathcal{Y},(sI+B^THB)}\{y^k - \alpha_k(sI + B^THB)^{-1}[g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^THB(y^k - \tilde{y}^k)]\}, \quad (4.18b)$$

and

$$\lambda^{k+1} = \lambda^k - \alpha_k H(A\tilde{x}^k + B\tilde{y}^k - b). \quad (4.18c)$$

Especially, if $sI + B^THB = (s + h)I$ is a scalar matrix, then (4.18b) becomes

$$y^{k+1} = P_{\mathcal{Y}}\{y^k - \alpha_k(\tfrac{1}{s+h})[g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^THB(y^k - \tilde{y}^k)]\}.$$

When the projections on $\mathcal{X}$ and $\mathcal{Y}$ are easy to be carried out, the update form (4.15b) usually outperforms the update form (4.15a) (see [21]).

All the update forms in this section are based on the Proximal ADM-Scheme (3.1). Sometimes, solving the subproblems in (3.1) is difficult and/or costly. In the following three sections, we will present different simplifying versions of the proximal alternating directions scheme and their related contraction methods.

5 Simplifying version A of the proximal ADM-scheme

In the simplifying version A of the proximal alternating direction method scheme, we substitute the function $H(Ax + By^k - b)$ in (3.1a) (resp. $H(A\tilde{x}^k + By - b)$) in (3.1b)) by $H(Ax^k + By^k - b)$ (resp. $H(A\tilde{x}^k + By^k - b)$).

The simplifying version A of the proximal alternating direction method scheme:

1. With available x^k, y^k and λ^k, solve the variational inequality problem

$$x \in \mathcal{X}, \;\; (x' - x)^T\{f(x) - A^T[\lambda^k - H(Ax^k + By^k - b)] + r(x - x^k)\} \ge 0, \; \forall\, x' \in \mathcal{X}, \quad (5.1a)$$

and denote the solution by $\tilde{x}^k$.

2. With available $\tilde{x}^k, y^k$ and λ^k, solve the variational inequality problem

$$y \in \mathcal{Y}, \;\; (y' - y)^T\{g(y) - B^T[\lambda^k - H(A\tilde{x}^k + By^k - b)] + s(y - y^k)\} \ge 0, \;\; \forall\, y' \in \mathcal{Y}, \quad (5.1b)$$

and denote the solution by $\tilde{y}^k$.

3. Set

$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b). \quad (5.1c)$$

The analysis is parallel as in Section 3. The solution $(\tilde{x}^k, \tilde{y}^k)$ of (5.1a)-(5.1b) satisfies

$$\begin{cases} (x' - \tilde{x}^k)^T\{f(\tilde{x}^k) - A^T\lambda^k + A^TH\big(Ax^k + By^k - b\big) + r(\tilde{x}^k - x^k)\} \ge 0, \; \forall\, x' \in \mathcal{X}, \\ (y' - \tilde{y}^k)^T\{g(\tilde{y}^k) - B^T\lambda^k + B^TH\big(A\tilde{x}^k + By^k - b\big) + s(\tilde{y}^k - y^k)\} \ge 0, \; \forall\, y' \in \mathcal{Y}, \end{cases} \quad (5.2)$$

Using $\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b)$ and by a manipulation, (5.2) can be rewritten as

$$(\tilde{x}^k, \tilde{y}^k) \in \mathcal{X} \times \mathcal{Y}, \quad \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k + A^T HB(y^k - \tilde{y}^k) \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^T HB(y^k - \tilde{y}^k) \end{pmatrix}$$
$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} - A^T HA & 0 \\ 0 & sI_{n_2} \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \end{pmatrix}, \ \forall (x', y') \in \mathcal{X} \times \mathcal{Y}. \quad (5.3)$$

If $w^k = \tilde{w}^k$, it follows from (5.3) and (5.1c) that

$$\begin{cases} \tilde{x}^k \in \mathcal{X}, \quad (x' - \tilde{x}^k)^T \{f(\tilde{x}^k) - A^T\tilde{\lambda}^k\} \geq 0, \quad \forall x' \in \mathcal{X}, \\ \tilde{y}^k \in \mathcal{Y}, \quad (y' - \tilde{y}^k)^T \{g(\tilde{y}^k) - B^T\tilde{\lambda}^k\} \geq 0, \quad \forall y' \in \mathcal{Y}, \\ \qquad A\tilde{x}^k + B\tilde{y}^k - b = 0. \end{cases}$$

Lemma 5.1. *Let $\tilde{w}^k$ be generated by the simplifying version* (5.1) *from the given vector w^k. Then, we have*

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T \big\{ \big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big) - M_A(w^k - \tilde{w}^k) \big\} \geq 0, \quad \forall\, w' \in \mathcal{W}, \qquad (5.4)$$

where $\eta(y^k, \tilde{y}^k)$ is defined in (2.8) *and the matrix*

$$M_A = M - \begin{pmatrix} A^T HA & 0 & 0 \\ 0 & B^T HB & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} rI_{n_1} - A^T HA & 0 & 0 \\ 0 & sI_{n_2} & 0 \\ 0 & 0 & H^{-1} \end{pmatrix}. \qquad (5.5)$$

Proof. The proof is similar as those of Lemma 3.1. In comparison (3.3) and (5.3) we obtain $\tilde{w}^k \in \mathcal{W}$ and

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(\tilde{x}^k) - A^T\tilde{\lambda}^k \\ g(\tilde{y}^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \begin{pmatrix} A^T \\ B^T \\ 0 \end{pmatrix} HB(y^k - \tilde{y}^k) \right\}$$
$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} - A^T HA & 0 & 0 \\ 0 & sI_{n_2} & 0 \\ 0 & 0 & H^{-1} \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \\ (\lambda^k - \tilde{\lambda}^k) \end{pmatrix}, \ \forall w' \in \mathcal{W}. \quad (5.6)$$

Note that (5.6) is obtained by substituting

$$M \quad \text{(see (2.20))} \quad \text{in the right hand side of (3.6) by} \quad M_A \quad \text{(see (5.5))}.$$

In addition, as in (3.6) the left hand side of (5.6) is

$$(w' - \tilde{w}^k)^T \big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big).$$

Therefore, using the notation of the matrix M_A, (5.6) can be expressed as

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T \big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big) \geq (w' - \tilde{w}^k)^T M_A (w^k - \tilde{w}^k), \quad \forall\, w' \in \mathcal{W}.$$

The assertion of this lemma is proved. $\square$

Lemma 5.2. *Let $\tilde{w}^k$ be generated by the simplifying version* (5.1) *from the given vector w^k. Then, we have*

$$(w^k - w^*)^T M_A (w^k - \tilde{w}^k) \ge \varphi^A(w^k, \tilde{w}^k), \quad \forall\, w^* \in \mathcal{W}^*, \tag{5.7}$$

where

$$\varphi^A(w^k, \tilde{w}^k) = \|w^k - \tilde{w}^k\|_{M_A}^2 + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k). \tag{5.8}$$

Proof. The proof is similar as those in Lemma 3.2 and thus is omitted. □

Noticing the difference of the matrix M and M_A (see (2.20) and (5.5)), we get

$$\begin{aligned}\varphi^A(w^k, \tilde{w}^k) \;\; = \;\; & \|w^k - \tilde{w}^k\|_M^2 + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \\ & -\big(\|x^k - \tilde{x}^k\|_{(A^THA)}^2 + \|y^k - \tilde{y}^k\|_{(B^THB)}^2\big).\end{aligned} \tag{5.9}$$

We assume that r, s in the simplifying version (5.1) satisfy the following conditions:

(Conditions A)
$$\|x^k - \tilde{x}^k\|_{(A^THA)}^2 \le \nu r \|x^k - \tilde{x}^k\|^2, \tag{5.10a}$$
and
$$\|y^k - \tilde{y}^k\|_{(B^THB)}^2 \le \nu s \|y^k - \tilde{y}^k\|^2. \tag{5.10b}$$

Lemma 5.3. *Let $\tilde{w}^k$ be generated by the simplifying version* (5.1) *from the given vector w^k and the conditions* (5.10) *be satisfied. Then*

$$\begin{aligned}\varphi^A(w^k, \tilde{w}^k) \;\; \ge \;\; & \frac{1}{2}\|w^k - \tilde{w}^k\|_{M_A}^2 + \frac{1}{2}(1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) \\ & + \frac{1}{2}\|A\tilde{x}^k + By^k - b\|_H^2.\end{aligned} \tag{5.11}$$

Proof. By using (5.8) and (5.9), we have

$$\begin{aligned}& 2\varphi^A(w^k, \tilde{w}^k) - \|w^k - \tilde{w}^k\|_{M_A}^2 \\ & \quad = \;\; \|w^k - \tilde{w}^k\|_{M_A}^2 + 2(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \\ & \quad = \;\; \|w^k - \tilde{w}^k\|_M^2 + 2(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \\ & \qquad\quad -\big(\|x^k - \tilde{x}^k\|_{(A^THA)}^2 + \|y^k - \tilde{y}^k\|_{(B^THB)}^2\big).\end{aligned} \tag{5.12}$$

Under the conditions (5.10), we have

$$\big(\|x^k - \tilde{x}^k\|_{(A^THA)}^2 + \|y^k - \tilde{y}^k\|_{(B^THB)}^2\big) \le \nu\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big). \tag{5.13}$$

In addition, we have (see (3.9))

$$\begin{aligned}& \|w^k - \tilde{w}^k\|_M^2 + 2(\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \\ & \quad = \;\; \big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) + \|A\tilde{x}^k + By^k - b\|_H^2.\end{aligned} \tag{5.14}$$

It follows from (5.12), (5.13) and (5.14) that

$$\begin{aligned}& 2\varphi^A(w^k, \tilde{w}^k) - \|w^k - \tilde{w}^k\|_{M_A}^2 \\ & \quad \ge \;\; (1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) + \|A\tilde{x}^k + By^k - b\|_H^2,\end{aligned}$$

and the assertion of this lemma is proved. □

5.1 Direct update form based on the simplifying version A

In this subsection, the update form takes $w^{k+1} = \tilde{w}^k$ as the new iterate, where $\tilde{w}^k$ is generated by the simplifying version (5.1). Following is the principal convergence inequality

Theorem 5.4. *Let $\tilde{w}^k$ be generated by the simplifying version* (5.1) *from the given vector w^k and the conditions* (5.10) *be satisfied. If the new iterate is updated by $w^{k+1} = \tilde{w}^k$, then for any $w^* \in \mathcal{W}^*$ we have*

$$\begin{aligned}\|w^{k+1} - w^*\|_{M_A}^2 \quad &\le \quad \|w^{k+1} - w^*\|_{M_A}^2 - (1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) \\ &\quad -\|A\tilde{x}^k + By^k - b\|_H^2, \ \ \forall\, w^* \in \mathcal{W}^*,\end{aligned} \tag{5.15}$$

where M_A is defined in (5.5).

Proof. By using Lemmas 5.2 and 5.3, we obtain

$$\begin{aligned}&\|w^k - w^*\|_{M_A}^2 - \|w^{k+1} - w^*\|_{M_A}^2 \\ &\quad = \ 2(w^k - w^*)^T M_A (w^k - \tilde{w}^k) - \|w^k - \tilde{w}^k\|_{M_A}^2 \\ &\quad \ge \ 2\varphi^A(w^k, \tilde{w}^k) - \|w^k - \tilde{w}^k\|_{M_A}^2 \\ &\quad \ge \ (1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) + \|A\tilde{x}^k + By^k - b\|_H^2,\end{aligned}$$

and the theorem is proved. □

5.2 Combinative update form based on the simplifying version A

Similarly as in Subsection 4.2, the combinative update form

$$w^{k+1} = (1-\alpha)w^k + \alpha\tilde{w}^k$$

can be viewed as a descent method. The objective is the unknown distance function $\frac{1}{2}\|w - w^*\|_{M_A}^2$ and the descent direction at w^k is $-(w^k - \tilde{w}^k)$.

Combinative update form based on the simplifying version A.
The next iterate w^{k+1} is given by

$$w^{k+1} = w^k - \alpha_k(w^k - \tilde{w}^k) \tag{5.16a}$$

where

$$\alpha_k = \omega\alpha_k^*, \qquad \alpha_k^* = \frac{\varphi^A(w^k, \tilde{w}^k)}{\|w^k - \tilde{w}^k\|_{M_A}^2} \qquad \text{and} \qquad \omega \in (0,2). \tag{5.16b}$$

Note that due to Lemma 5.3, the α_k^* here in (5.16b) is greater than 0.5.

Theorem 5.5. *Let $\tilde{w}^k$ be generated by the simplifying version* (5.1) *from the given vector w^k and the conditions* (5.10) *be satisfied. If the new iterate w^{k+1} is updated by* (5.16)*, then we have*

$$\begin{aligned}&\|w^{k+1} - w^*\|_{M_A}^2 \\ &\quad \le \ \|w^k - w^*\|_{M_A}^2 - \frac{\omega(2-\omega)}{4}\|w^k - \tilde{w}^k\|_{M_A}^2 \\ &\qquad -\frac{\omega(2-\omega)}{4}\big\{(1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) + \|A\tilde{x}^k + By^k - b\|_H^2\big\}.\end{aligned} \tag{5.17}$$

where M_A is defined in (5.5)*.*

Proof. First, due to (5.16a), Lemma 5.2 and the definition of α_k^*, we obtain

$$\begin{aligned}
&\|w^k - w^*\|_{M_A}^2 - \|w^{k+1} - w^*\|_{M_A}^2 \\
&\quad = 2\omega\alpha_k^*(w^k - w^*)^T M_A (w^k - \tilde{w}^k) - (\omega\alpha_k^*)^2 \|w^k - \tilde{w}^k\|_{M_A}^2 \\
&\quad \geq 2\omega\alpha_k^*\varphi^A(w^k, \tilde{w}^k) - (\omega\alpha_k^*)^2 \|w^k - \tilde{w}^k\|_{M_A}^2 \\
&\quad = \omega(2-\omega)\alpha_k^*\varphi^A(w^k, \tilde{w}^k).
\end{aligned}$$

Using $\alpha_k^* > \frac{1}{2}$ and (5.11) again, the theorem is proved. $\square$

In some applications, we need only simplify one of the subproblems in (5.1), (5.1a) or (5.1b). Thereby we set the related proximal coefficient to zero. As the consequent results of Theorem 5.4, we have the following corollaries.

Corollary 5.6. *Let $\tilde{w}^k$ be generated by a modified simplifying version* (5.1) *in which* (5.1a) *is substituted by*

$$\tilde{x}^k \in \mathcal{X}, \ \ (x' - \tilde{x}^k)^T \big\{ f(\tilde{x}^k) - A^T[\lambda^k - H(A\tilde{x}^k + By^k - b)] \big\} \geq 0, \ \forall\, x' \in \mathcal{X}. \tag{5.18}$$

In addition, let the solution of (5.1b) *satisfy (see* (5.10b)*)*

$$\|y^k - \tilde{y}^k\|_{(B^THB)}^2 \leq \nu s \|y^k - \tilde{y}^k\|^2.$$

If we set $w^{k+1} = \tilde{w}^k$, then for any $w^* \in \mathcal{W}^*$ we have

$$\begin{aligned}
& s\|y^{k+1} - y^*\|^2 + \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\
& \leq \big(s\|y^k - y^*\|^2 + \|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2 \big) \\
& \quad - \big((1-\nu)s\|y^k - \tilde{y}^k\|^2 + \|A\tilde{x}^k + By^k - b\|_H^2 \big).
\end{aligned} \tag{5.19}$$

Proof. It follows from (5.15) by setting $r = 0$. $\square$

Corollary 5.7. *Let $\tilde{w}^k$ be generated by a modified simplifying version* (5.1) *in which* (5.1b) *is substituted by*

$$\tilde{y}^k \in \mathcal{Y}, \ \ (y' - \tilde{y}^k)^T \big\{ g(\tilde{y}^k) - B^T[\lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b)] \big\} \geq 0, \ \ \forall\, y' \in \mathcal{Y}. \tag{5.20}$$

In addition, let the solution of (5.1a) *satisfy (see* (5.10a)*)*

$$\|x^k - \tilde{x}^k\|_{(A^THA)}^2 \leq \nu r \|x^k - \tilde{x}^k\|^2.$$

If we set $w^{k+1} = \tilde{w}^k$, then for any $w^* \in \mathcal{W}^*$ we have

$$\begin{aligned}
& r\|x^{k+1} - x^*\|^2 + \|B(y^{k+1} - y^*)\|_H^2 + \|\lambda^{k+1} - \lambda^*\|_{H^{-1}}^2 \\
& \leq \big(r\|x^k - x^*\|^2 + \|B(y^k - y^*)\|_H^2 + \|\lambda^k - \lambda^*\|_{H^{-1}}^2 \big) \\
& \quad - \big((1-\nu)r\|x^k - \tilde{x}^k\|^2 + \|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2 \big).
\end{aligned} \tag{5.21}$$

Proof. Set $s = 0$ in (5.15) and notice that in this case

$$\|A\tilde{x}^k + By^k - b\|_H^2 = \|B(y^k - \tilde{y}^k)\|_H^2 + \|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2.$$

The assertion is proved. $\square$

5.3 G-norm contractive form based on the simplifying version A

In the case that the projections on $\mathcal{X}$ and $\mathcal{Y}$ are easy to be carried out, similarly as in Subsection 4.3, we consider the G-norm contractive update form.

G-norm contractive update form based on the simplifying version A.
The new iterate w^{k+1} is given by

$$w^{k+1} = w^k - \alpha_k G^{-1} d^A(w^k, \tilde{w}^k), \tag{5.22a}$$

or

$$w^{k+1} = P_{\mathcal{W},G}\big\{w^k - \alpha_k G^{-1}[F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)]\big\}, \tag{5.22b}$$

where

$$\alpha_k = \omega\alpha_k^*, \qquad \alpha_k^* = \frac{\varphi^A(w^k, \tilde{w}^k)}{\|G^{-1}M_A(w^k - \tilde{w}^k)\|_{M_A}^2} \qquad \text{and} \quad \omega \in (0,2). \tag{5.22c}$$

Theorem 5.8. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (5.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (4.15)*, then we have*

$$\begin{aligned}
&\|w^{k+1} - w^*\|_G^2 \\
&\le \ \|w^k - w^*\|_G^2 - \frac{\omega(2-\omega)}{4(\|G^{-1}\|\cdot\|M_A\|)}\|w^k - \tilde{w}^k\|_{M_A}^2 \\
&\quad -\frac{\omega(2-\omega)}{4(\|G^{-1}\|\cdot\|M_A\|)}\big\{(1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big) + \|A\tilde{x}^k + By^k - b\|_H^2\big\}
\end{aligned} \tag{5.23}$$

where M_A is defined in (5.5)*.*

Proof. First, it follows from Lemma 5.1 and Theorem 2.4 that

$$\|w^k - w^*\|_G^2 - \|w^{k+1} - w^*\|_G^2 \ge 2\omega\alpha_k^*\varphi^A(w^k, \tilde{w}^k) - (\omega\alpha_k^*)^2\|G^{-1}d^A(w^k, \tilde{w}^k)\|_G^2, \ \ \forall\, w^* \in \mathcal{W}^*.$$

By using $\alpha_k^*\|G^{-1}d^A(w^k, \tilde{w}^k)\|_G^2 = \varphi^A(w^k, \tilde{w}^k)$, we obtain

$$\|w^k - w^*\|_G^2 - \|w^{k+1} - w^*\|_G^2 \ge \omega(2-\omega)\alpha_k^*\varphi^A(w^k, \tilde{w}^k), \quad \forall w^* \in \mathcal{W}^*. \tag{5.24}$$

Since $d^A(w^k, \tilde{w}^k) = M_A(w^k - \tilde{w}^k)$, it follows that

$$\|G^{-1}d^A(w^k, \tilde{w}^k)\|_G^2 \le \|G^{-1}\|\cdot\|M_A(w^k - \tilde{w}^k)\|^2 \le (\|G^{-1}\|\cdot\|M_A\|)\|w^k - \tilde{w}^k\|_{M_A}^2.$$

Consequently, using (5.22c), we have

$$\alpha_k^* \ge \frac{1}{2(\|G^{-1}\|\cdot\|M_A\|)}.$$

Substituting it in (5.24) and using (5.11), the theorem is proved. □

Since M_A is positive definite, a natural choice of G is $G = M_A$. Note that in this case the update form (5.22a) (resp. the step size (5.22c)) is the same as (5.16a) (resp. (5.16b)). Since M_A is block diagonal matrix, the update form (5.22b) is separable in forms

$$x^{k+1} = P_{\mathcal{X},(rI-ATHA)}\{x^k - \alpha_k(rI - ATHA)^{-1}[f(\tilde{x}^k) - A^T\tilde{\lambda}^k + A^THB(y^k - \tilde{y}^k)]\}, \tag{5.25a}$$

$$y^{k+1} = P_{\mathcal{Y}}\{y^k - \alpha_k(\tfrac{1}{s})[g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^THB(y^k - \tilde{y}^k)]\}, \tag{5.25b}$$

and

$$\lambda^{k+1} = \lambda^k - \alpha_k H(A\tilde{x}^k + B\tilde{y}^k - b). \tag{5.25c}$$

Especially, if $rI - A^THA = (r-h)I$ is a scalar matrix and $r - h > 0$, then (5.25a) becomes

$$x^{k+1} = P_{\mathcal{X}}\{x^k - \alpha_k(\tfrac{1}{r-h})[f(\tilde{x}^k) - A^T\tilde{\lambda}^k + A^THB(y^k - \tilde{y}^k)]\}.$$

6 Simplifying version B of the proximal ADM-scheme

In the simplifying version B of the proximal alternating direction method scheme, we substitute the function $f(x)$ in (3.1a) (resp. $g(y)$) in (3.1b)) by $f(x^k)$ (resp. $g(y^k)$).

The simplifying version B of the proximal alternating direction method scheme:

1. With available x^k, y^k and λ^k, solve the variational inequality problem

$$x \in \mathcal{X}, \ \ (x'-x)^T\big\{f(x^k) - A^T[\lambda^k - H(Ax + By^k - b)] + r(x - x^k)\big\} \ge 0, \ \forall\, x' \in \mathcal{X}, \tag{6.1a}$$

   and denote the solution by $\tilde{x}^k$.

2. With available $\tilde{x}^k, y^k$ and λ^k, solve the variational inequality problem

$$y \in \mathcal{Y}, \ \ (y'-y)^T\big\{g(y^k) - B^T[\lambda^k - H(A\tilde{x}^k + By - b)] + s(y - y^k)\big\} \ge 0, \ \ \forall\, y' \in \mathcal{Y}, \tag{6.1b}$$

   and denote the solution by $\tilde{y}^k$.

3. Set

$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b). \tag{6.1c}$$

The analysis is parallel as in Section 3. The solution $(\tilde{x}^k, \tilde{y}^k)$ of (6.1a)-(6.1b) satisfies

$$\begin{cases} (x' - \tilde{x}^k)^T\big\{f(x^k) - A^T\lambda^k + A^TH\big(A\tilde{x}^k + By^k - b\big) + r(\tilde{x}^k - x^k)\big\} \ge 0, \ \forall\, x' \in \mathcal{X}, \\ (y' - \tilde{y}^k)^T\big\{g(y^k) - B^T\lambda^k + B^TH\big(A\tilde{x}^k + B\tilde{y}^k - b\big) + s(\tilde{y}^k - y^k)\big\} \ge 0, \ \forall\, y' \in \mathcal{Y}. \end{cases} \tag{6.2}$$

Using $\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b)$ and by a manipulation, (6.2) can be rewritten as

$$\begin{aligned} &(\tilde{x}^k, \tilde{y}^k) \in \mathcal{X} \times \mathcal{Y}, \quad \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} f(x^k) - A^T\tilde{\lambda}^k + A^THB(y^k - \tilde{y}^k) \\ g(y^k) - B^T\tilde{\lambda}^k + B^THB(y^k - \tilde{y}^k) \end{pmatrix} \\ &\ge \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} & 0 \\ 0 & sI_{n_2} + B^THB \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \end{pmatrix}, \ \ \forall\,(x', y') \in \mathcal{X} \times \mathcal{Y}. \end{aligned} \tag{6.3}$$

If $w^k = \tilde{w}^k$, it follows from (6.3) and (6.1c) that

$$\begin{cases} \tilde{x}^k \in \mathcal{X}, & (x' - \tilde{x}^k)^T\big\{f(\tilde{x}^k) - A^T\tilde{\lambda}^k\big\} \ge 0, \quad \forall x' \in \mathcal{X}, \\ \tilde{y}^k \in \mathcal{Y}, & (y' - \tilde{y}^k)^T\big\{g(\tilde{y}^k) - B^T\tilde{\lambda}^k\big\} \ge 0, \quad \forall y' \in \mathcal{Y}, \\ & A\tilde{x}^k + B\tilde{y}^k - b = 0. \end{cases}$$

Lemma 6.1. *Let $\tilde{w}^k$ be generated by the simplifying version* (6.1) *from the given vector w^k. Then, we have*

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T \big\{ \big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k) \big) - d^B(w^k, \tilde{w}^k) \big\} \geq 0, \quad \forall\, w' \in \mathcal{W}, \tag{6.4}$$

where $\eta(y^k, \tilde{y}^k)$ is defined in (2.8),

$$d^B(w^k, \tilde{w}^k) = M(w^k - \tilde{w}^k) - \begin{pmatrix} f(x^k) - f(\tilde{x}^k) \\ g(y^k) - g(\tilde{y}^k) \\ 0 \end{pmatrix}, \tag{6.5}$$

and M is defined in (2.20).

Proof. The proof is similar as those of Lemma 3.1. In comparison (3.3) and (6.3) we obtain $\tilde{w}^k \in \mathcal{W}$ and

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

$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} & 0 & 0 \\ 0 & sI_{n_2} + B^THB & 0 \\ 0 & 0 & H^{-1} \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix}, \quad \forall\, w' \in \mathcal{W}. \tag{6.6}$$

Note that the right hand side of (6.6) is $(w' - \tilde{w})^T M(w^k - \tilde{w}^k)$. Adding

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ 0 \end{pmatrix}^T \begin{pmatrix} f(\tilde{x}^k) - f(x^k) \\ g(\tilde{y}^k) - g(y^k) \\ 0 \end{pmatrix}$$

to the both sides of (6.6) and using the notations of $d^B(w^k, \tilde{w}^k)$ and $\eta(y^k, \tilde{y}^k)$, the above inequality is

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T \big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k) \big) \geq (w' - \tilde{w}^k)^T d^B(w^k, \tilde{w}^k), \quad \forall\, w' \in \mathcal{W}$$

and the assertion of this lemma is proved. □

Lemma 6.2. *Let $\tilde{w}^k$ be generated by the simplifying version* (6.1) *from the given vector w^k. Then, we have*

$$(w^k - w^*)^T d^B(w^k, \tilde{w}^k) \geq \varphi^B(w^k, \tilde{w}^k), \quad \forall\, w^* \in \mathcal{W}^*, \tag{6.7}$$

where

$$\varphi^B(w^k, \tilde{w}^k) = (w^k - \tilde{w}^k)^T d^B(w^k, \tilde{w}^k) + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k). \tag{6.8}$$

Proof. The proof is similar as those in Lemma 3.2 and thus is omitted. □

By using the definition of $d^B(w^k, \tilde{w}^k)$, we get

$$\begin{aligned} \varphi^B(w^k, \tilde{w}^k) \;\; = \;\; & \|w^k - \tilde{w}^k\|_M^2 + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) \\ & -\big((x^k - \tilde{x}^k)^T (f(x^k) - f(\tilde{x}^k) + (y^k - \tilde{y}^k)^T (g(y^k) - g(\tilde{y}^k))\big). \end{aligned} \tag{6.9}$$

We assume that r, s in the simplifying version (6.1) satisfy the following conditions:

(Condition B) $$\frac{1}{r}\|f(x^k)-f(\tilde{x}^k)\|^2 \le \nu r\|x^k-\tilde{x}^k\|^2, \tag{6.10a}$$

and

$$\|g(y^k)-g(\tilde{y}^k)\|^2_{\{(sI_{n_2}+B^THB)^{-1}\}} \le \nu s\|y^k-\tilde{y}^k\|^2. \tag{6.10b}$$

Note that under the conditions (6.10) the matrix M is positive definite, and the condition (6.10b) is satisfied when

$$\frac{1}{s}\|g(y^k)-g(\tilde{y}^k)\|^2 \le \nu s\|y^k-\tilde{y}^k\|^2,$$

because B^THB is positive semi-definite.

Lemma 6.3. *Let $\tilde{w}^k$ be generated by the simplifying version* (6.1) *from the given vector w^k and the conditions* (6.10) *be satisfied. Then*

$$\begin{aligned}\varphi^B(w^k,\tilde{w}^k) \;\ge\; & \frac{1}{2}\|M^{-1}d^B(w^k,\tilde{w}^k)\|^2_M + \frac{1}{2}(1-\nu)\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big)\\ & +\frac{1}{2}\|A\tilde{x}^k+By^k-b\|^2_H. \end{aligned}\tag{6.11}$$

Proof. By using (6.8), we have

$$\begin{aligned}&2\varphi^B(w^k,\tilde{w}^k) - \|M^{-1}d^B(w^k,\tilde{w}^k)\|^2_M\\ &= 2(w^k-\tilde{w}^k)^Td^B(w^k,\tilde{w}^k) - \|M^{-1}d^B(w^k,\tilde{w}^k)\|^2_M + 2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k)\\ &= (M^{-1}d^B(w^k,\tilde{w}^k))^TM\{2(w^k-\tilde{w}^k) - M^{-1}d^B(w^k,\tilde{w}^k)\} + 2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k).\\ &= \|w^k-\tilde{w}^k\|^2_M - \begin{pmatrix} f(x^k)-f(\tilde{x}^k)\\ g(y^k)-g(\tilde{y}^k)\end{pmatrix}^T \begin{pmatrix} rI_{n_1} & 0\\ 0 & sI_{n_2}+B^THB\end{pmatrix}^{-1} \begin{pmatrix} f(x^k)-f(\tilde{x}^k)\\ g(y^k)-g(\tilde{y}^k)\end{pmatrix}\\ &\quad +2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k)\end{aligned}\tag{6.12}$$

Under the conditions (6.10), we have

$$\begin{aligned}&\begin{pmatrix} f(x^k)-f(\tilde{x}^k)\\ g(y^k)-g(\tilde{y}^k)\end{pmatrix}^T \begin{pmatrix} rI_{n_1} & 0\\ 0 & sI_{n_2}+B^THB\end{pmatrix}^{-1} \begin{pmatrix} f(x^k)-f(\tilde{x}^k)\\ g(y^k)-g(\tilde{y}^k)\end{pmatrix}\\ &\le \nu\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big).\end{aligned}\tag{6.13}$$

In addition, we have (see (3.9))

$$\begin{aligned}&\|w^k-\tilde{w}^k\|^2_M + 2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k)\\ &= \big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big) + \|A\tilde{x}^k+By^k-b\|^2_H.\end{aligned}\tag{6.14}$$

It follows from (6.12), (6.13) and (6.14) that

$$\begin{aligned}&2\varphi^B(w^k,\tilde{w}^k) - \|G^{-1}d^B(w^k,\tilde{w}^k)\|^2_M\\ &\ge (1-\nu)\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big) + \|A\tilde{x}^k+By^k-b\|^2_H,\end{aligned}$$

and the assertion of this lemma is proved. □

Since the nonlinear function $f(x)$ (res. $g(y)$) is simplified by $f(x^k)$ (resp. $g(y^k)$) in (6.1), it seems that we can not establish the similar principal contractive inequalities as in Section 5 if either $\tilde{w}^k$ or the linear combination of w^k and $\tilde{w}^k$ is taken as the new iterate. However, we can consider the G-norm contractive update form based on the $\tilde{w}^k$ generated by the simplifying version (6.1).

G-norm contractive update form based on the simplifying version B.
The new iterate w^{k+1} is given by

$$w^{k+1} = w^k - \alpha_k G^{-1} d^B(w^k, \tilde{w}^k), \tag{6.15a}$$

or

$$w^{k+1} = P_{\mathcal{W},G}\{w^k - \alpha_k G^{-1}[F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)]\}, \tag{6.15b}$$

where

$$\alpha_k = \omega\alpha_k^*, \qquad \alpha_k^* = \frac{\varphi^B(w^k, \tilde{w}^k)}{\|G^{-1}d^B(w^k, \tilde{w}^k)\|_G^2} \qquad \text{and} \quad \omega \in (0, 2). \tag{6.15c}$$

Theorem 6.4. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (6.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (6.15) *with an $G = I$, then we have*

$$\begin{aligned} &\|w^{k+1} - w^*\|_G^2 \\ \le\ & \|w^k - w^*\|_G^2 - \frac{\omega(2-\omega)}{4(\|M^{1/2}G^{-1}M^{1/2}\|)}\big(\|M^{-1}d^B(w^k, \tilde{w}^k)\|_M^2 + \|A\tilde{x}^k + By^k - b\|_H^2\big) \\ & - \frac{\omega(2-\omega)}{4(\|M^{1/2}G^{-1}M^{1/2}\|)}\big\{(1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big)\big\}, \quad \forall w^* \in \mathcal{W}^*, \end{aligned} \tag{6.16}$$

where M is defined in (2.20).

Proof. First, it follows from Lemma 6.1 and Theorem 2.4 that

$$\|w^k - w^*\|_G^2 - \|w^{k+1} - w^*\|_G^2 \ge 2\omega\alpha_k^*\varphi^B(w^k, \tilde{w}^k) - (\omega\alpha_k^*)^2\|G^{-1}d^B(w^k, \tilde{w}^k)\|_G^2, \ \ \forall\, w^* \in \mathcal{W}^*.$$

By using $\alpha_k^*\|G^{-1}d^B(w^k, \tilde{w}^k)\|_G^2 = \varphi^B(w^k, \tilde{w}^k)$, we obtain

$$\|w^k - w^*\|_G^2 - \|w^{k+1} - w^*\|_G^2 \ge \omega(2-\omega)\alpha_k^*\varphi^B(w^k, \tilde{w}^k), \quad \forall w^* \in \mathcal{W}^*. \tag{6.17}$$

It follows from

$$\|G^{-1}d^B(w^k, \tilde{w}^k)\|_G^2 \le \|M^{1/2}G^{-1}M^{1/2}\| \cdot \|M^{-1}d^B(w^k, \tilde{w}^k)\|_M^2,$$

and (6.11) that

$$\alpha_k^*\varphi^B(w^k, \tilde{w}^k) \ge \frac{1}{4(\|M^{1/2}G^{-1}M^{1/2}\|)}\|M^{-1}d^B(w^k, \tilde{w}^k)\|_M^2.$$

Substituting it in (6.17) the theorem is proved. □

The simplest choices in (6.15) are $G = I$ or $G = M$. For $G = M$, the update form (6.15b) is separable in forms

$$x^{k+1} = P_{\mathcal{X}}\{x^k - \alpha_k(\frac{1}{r})[f(\tilde{x}^k) - A^T\tilde{\lambda}^k + A^T HB(y^k - \tilde{y}^k)]\}, \tag{6.18a}$$

$$y^{k+1} = P_{\mathcal{Y},(sI+B^THB)}\{y^k - \alpha_k(sI + B^THB)^{-1}[g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^THB(y^k - \tilde{y}^k)]\}, \quad (6.18b)$$

and

$$\lambda^{k+1} = \lambda^k - \alpha_k H(A\tilde{x}^k + B\tilde{y}^k - b). \quad (6.18c)$$

Especially, if $sI + B^THB = (s + h)I$ is a scalar matrix and $s + h > 0$, then (6.18b) becomes

$$y^{k+1} = P_{\mathcal{Y}}\{y^k - \alpha_k(\tfrac{1}{s+h})[g(\tilde{y}^k) - B^T\tilde{\lambda}^k + B^THB(y^k - \tilde{y}^k)]\}.$$

7 Simplifying version C of the proximal ADM-scheme

In the simplifying version C of the proximal alternating direction method scheme, we substitute

$$f(x) \qquad \text{and} \qquad H(Ax + By^k - b)$$

in (3.1a) by

$$f(x^k) \qquad \text{and} \qquad H(Ax^k + By^k - b).$$

And, in addition, the

$$g(y) \qquad \text{and} \qquad H(A\tilde{x}^k + By - b)$$

in (3.1b) will be substituted by

$$g(y^k) \qquad \text{and} \qquad H(A\tilde{x}^k + By^k - b).$$

The simplifying version C of the proximal alternating direction method scheme:

1. With available x^k, y^k and λ^k, solve the variational inequality problem

$$x \in \mathcal{X}, \;\; (x'-x)^T\{f(x^k)-A^T[\lambda^k-H(Ax^k+By^k-b)]+r(x-x^k)\} \ge 0, \; \forall\, x' \in \mathcal{X}, \; (7.1a)$$

   and denote the solution by $\tilde{x}^k$.

2. With available $\tilde{x}^k, y^k$ and λ^k, solve the variational inequality problem

$$y \in \mathcal{Y}, \;\; (y'-y)^T\{g(y^k)-B^T[\lambda^k-H(A\tilde{x}^k+By^k-b)]+s(y-y^k)\} \ge 0, \;\; \forall\, y' \in \mathcal{Y}, \; (7.1b)$$

   and denote the solution by $\tilde{y}^k$.

3. Set

$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b). \quad (7.1c)$$

This simplifying version of the Proximal ADM scheme was established in [17]. The analysis is parallel as in Section 3. The solution $(\tilde{x}^k, \tilde{y}^k)$ of (7.1a)-(7.1b) satisfies

$$\begin{cases} (x' - \tilde{x}^k)^T\{f(x^k) - A^T\lambda^k + A^TH\big(Ax^k + By^k - b\big) + r(\tilde{x}^k - x^k)\} \ge 0, \; \forall\, x' \in \mathcal{X}, \\ (y' - \tilde{y}^k)^T\{g(y^k) - B^T\lambda^k + B^TH\big(A\tilde{x}^k + By^k - b\big) + s(\tilde{y}^k - y^k)\} \ge 0, \; \forall\, y' \in \mathcal{Y}. \end{cases} \quad (7.2)$$

Using $\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b)$ and by a manipulation, (7.2) can be rewritten as

$$(\tilde{x}^k, \tilde{y}^k) \in \mathcal{X} \times \mathcal{Y}, \quad \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} f(x^k) - A^T\tilde{\lambda}^k + A^T HB(y^k - \tilde{y}^k) \\ g(y^k) - B^T\tilde{\lambda}^k + B^T HB(y^k - \tilde{y}^k) \end{pmatrix}$$
$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} - A^T HA & 0 \\ 0 & sI_{n_2} \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \end{pmatrix}, \quad \forall\, (x', y') \in \mathcal{X} \times \mathcal{Y}. \tag{7.3}$$

If $w^k = \tilde{w}^k$, it follows from (7.3) and (7.1c) that

$$\begin{cases} \tilde{x}^k \in \mathcal{X}, & (x' - \tilde{x}^k)^T \{f(\tilde{x}^k) - A^T\tilde{\lambda}^k\} \geq 0, \quad \forall x' \in \mathcal{X}, \\ \tilde{y}^k \in \mathcal{Y}, & (y' - \tilde{y}^k)^T \{g(\tilde{y}^k) - B^T\tilde{\lambda}^k\} \geq 0, \quad \forall y' \in \mathcal{Y}, \\ & A\tilde{x}^k + B\tilde{y}^k - b = 0. \end{cases}$$

According to Lemma 2.2, the solution of (7.1a) can be obtained by

$$\tilde{x}^k = P_{\mathcal{X}}\Big\{x^k - \frac{1}{r}\Big(f(x^k) - A^T\big[\lambda^k - H(Ax^k + By^k - b)\big]\Big)\Big\}.$$

Similarly, the solution of (7.1b) can be obtained by

$$\tilde{y}^k = P_{\mathcal{Y}}\Big\{y^k - \frac{1}{s}\Big(g(y^k) - A^T\big[\lambda^k - H(A\tilde{x}^k + By^k - b)\big]\Big)\Big\}.$$

Lemma 7.1. *Let $\tilde{w}^k$ be generated by the simplifying version* (7.1) *from the given vector w^k. Then, we have*

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T \big\{\big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big) - d^C(w^k, \tilde{w}^k)\big\} \geq 0, \quad \forall\, w' \in \mathcal{W}, \tag{7.4}$$

where $\eta(y^k, \tilde{y}^k)$ is defined in (2.8),

$$d^C(w^k, \tilde{w}^k) = M(w^k - \tilde{w}^k) - \begin{pmatrix} f(x^k) - f(\tilde{x}^k) + A^T HA(x^k - \tilde{x}^k) \\ g(y^k) - g(\tilde{y}^k) + B^T HB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix}, \tag{7.5}$$

and the matrix M is defined in (2.20).

Proof. The proof is similar as those of Lemma 3.1. In comparison (3.3) and (7.3) we obtain $\tilde{w}^k \in \mathcal{W}$ and

$$\begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f(x^k) - A^T\tilde{\lambda}^k \\ g(y^k) - B^T\tilde{\lambda}^k \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix} + \begin{pmatrix} A^T \\ B^T \\ 0 \end{pmatrix} HB(y^k - \tilde{y}^k) \right\}$$
$$\geq \begin{pmatrix} x' - \tilde{x}^k \\ y' - \tilde{y}^k \\ \lambda' - \tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} rI_{n_1} - A^T HA & 0 & 0 \\ 0 & sI_{n_2} & 0 \\ 0 & 0 & H^{-1} \end{pmatrix} \begin{pmatrix} x^k - \tilde{x}^k \\ y^k - \tilde{y}^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix}, \quad \forall w' \in \mathcal{W}. \tag{7.6}$$

Note that (7.6) is obtained by substituting

$$f(\tilde{x}^k)(\text{resp. } g(\tilde{y}^k)) \quad \text{in the left hand side of (3.6) by} \quad f(x^k)(\text{resp. } g(y^k))$$

and

$$\begin{pmatrix} rI_{n_1} & 0 & 0 \\ 0 & sI_{n_2} + B^T HB & 0 \\ 0 & 0 & H^{-1} \end{pmatrix}$$

in the right hand side of (3.6) by

$$\begin{pmatrix} rI_{n_1} - A^T HA & 0 & 0 \\ 0 & sI_{n_2} & 0 \\ 0 & 0 & H^{-1} \end{pmatrix}.$$

Using the notations of $d^C(w^k, \tilde{w}^k)$ and $\eta(y^k, \tilde{y}^k)$, the inequality (7.6) is

$$\tilde{w}^k \in \mathcal{W}, \quad (w' - \tilde{w}^k)^T \big(F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)\big) \geq (w' - \tilde{w}^k)^T d^C(w^k, \tilde{w}^k), \quad \forall\, w' \in \mathcal{W},$$

and the assertion of this lemma is proved. $\square$

Lemma 7.2. *Let $\tilde{w}^k$ be generated by the simplifying version C of the proximal ADM scheme* (7.1) *from the given vector w^k. Then, we have*

$$(w^k - w^*)^T d^C(w^k, \tilde{w}^k) \geq \varphi^C(w^k, \tilde{w}^k) \quad \forall\, w^* \in \mathcal{W}^*. \tag{7.7}$$

where

$$\varphi^C(w^k, \tilde{w}^k) = (w^k - \tilde{w}^k)^T d^C(w^k, \tilde{w}^k) + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k). \tag{7.8}$$

Proof. The proof is similar as those in Lemma 3.2 and thus is omitted. $\square$

Using the definition of $d^C(w^k, \tilde{w}^k)$ and the relationship between the matrices M and G_C, we have

$$\begin{aligned} \varphi^C(w^k, \tilde{w}^k) \;=\; & \|w^k - \tilde{w}^k\|_M^2 + (\lambda^k - \tilde{\lambda}^k)^T B(y^k - \tilde{y}^k) - \|A(x^k - \tilde{x}^k)\|_H^2 - \|B(y^k - \tilde{y}^k)\|_H^2 \\ & -\big\{(x^k - \tilde{x}^k)^T (f(x^k) - f(\tilde{x}^k)) + (y^k - \tilde{y}^k)^T (g(y^k) - g(\tilde{y}^k))\big\}. \end{aligned} \tag{7.9}$$

We assume that r, s in simplifying version (7.1) satisfy the following conditions:

(Conditions C)
$$\frac{1}{r}\|(f(x^k) - f(\tilde{x}^k)) + A^T HA(x^k - \tilde{x}^k)\|^2 \leq \nu r \|x^k - \tilde{x}^k\|^2, \tag{7.10a}$$

and

$$\|(g(y^k) - g(\tilde{y}^k)) + B^T HB(y^k - \tilde{y}^k)\|^2_{\{(sI_{n_2} + B^T HB)^{-1}\}} \leq \nu s \|y^k - \tilde{y}^k\|^2. \tag{7.10b}$$

Note under the conditions (7.10) the matrix M is positive definite, and the condition (7.10b) is satisfied when

$$\frac{1}{s}\|(g(y^k) - g(\tilde{y}^k)) + B^T HB(y^k - \tilde{y}^k)\|^2 \leq \nu s \|y^k - \tilde{y}^k\|^2,$$

because H is positive definite.

Lemma 7.3. *Let $\tilde{w}^k$ be generated by the simplifying version C of the proximal ADM-scheme* (7.1) *from the given vector w^k. If the conditions* (7.10) *are satisfied, then*

$$\begin{aligned}\varphi^C(w^k,\tilde{w}^k) \quad \geq \quad & \frac{1}{2}\|G^{-1}d^C(w^k,\tilde{w}^k)\|_M^2 + \frac{1}{2}(1-\nu)\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big)\\ & +\frac{1}{2}\|A\tilde{x}^k + By^k - b\|_H^2. \end{aligned} \tag{7.11}$$

Proof. Because

$$G^{-1}d^C(w^k,\tilde{w}^k) = (w^k-\tilde{w}^k) - G^{-1}\begin{pmatrix} f(x^k)-f(\tilde{x}^k)+A^THA(x^k-\tilde{x}^k)\\ g(y^k)-g(\tilde{y}^k)+B^THB(y^k-\tilde{y}^k)\\ 0\end{pmatrix},$$

we have

$$\begin{aligned}&\|G^{-1}d^C(w^k,\tilde{w}^k)\|_G^2\\ = \quad & \|w^k-\tilde{w}^k\|_M^2 - 2(x^k-\tilde{x}^k)^T\big(f(x^k)-f(\tilde{x}^k)+A^THA(x^k-\tilde{x}^k)\big)\\ & -2(y^k-\tilde{y}^k)^T(g(y^k)-g(\tilde{y}^k)+B^THB(y^k-\tilde{y}^k))\\ & +\frac{1}{r}\|(f(x^k)-f(\tilde{x}^k))+A^THA(x^k-\tilde{x}^k)\|^2\\ & +\|(g(y^k)-g(\tilde{y}^k))+B^THB(y^k-\tilde{y}^k)\|^2_{\{(sI_{n_2}+B^THB)^{-1}\}}. \end{aligned} \tag{7.12}$$

By using (7.9) and (7.12), we obtain

$$\begin{aligned}&2\varphi^C(w^k,\tilde{w}^k) - \|G^{-1}d^C(w^k,\tilde{w}^k)\|_M^2\\ = \quad & \|w^k-\tilde{w}^k\|_M^2 + 2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k) - \frac{1}{r}\|(f(x^k)-f(\tilde{x}^k))+A^THA(x^k-\tilde{x}^k)\|^2\\ & -\|(g(y^k)-g(\tilde{y}^k))+B^THB(y^k-\tilde{y}^k)\|^2_{\{(sI_{n_2}+B^THB)^{-1}\}}. \end{aligned}$$

Applying conditions (7.10) to the above inequality, we obtain

$$\begin{aligned}&2\varphi^C(w^k,\tilde{w}^k) - \|G^{-1}d^C(w^k,\tilde{w}^k)\|_M^2\\ \geq \quad & \|w^k-\tilde{w}^k\|_M^2 + 2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k) - \nu\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big). \end{aligned} \tag{7.13}$$

In addition, we have (see (3.9))

$$\begin{aligned}&\|w^k-\tilde{w}^k\|_M^2 + 2(\lambda^k-\tilde{\lambda}^k)^TB(y^k-\tilde{y}^k)\\ = \quad & \big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big) + \|A\tilde{x}^k+By^k-b\|_H^2. \end{aligned} \tag{7.14}$$

It follows from (7.13) and (7.14) that

$$\begin{aligned}&2\varphi^C(w^k,\tilde{w}^k) - \|G^{-1}d^C(w^k,\tilde{w}^k)\|_M^2\\ \geq \quad & (1-\nu)\big(r\|x^k-\tilde{x}^k\|^2 + s\|y^k-\tilde{y}^k\|^2\big) + \|A\tilde{x}^k+By^k-b\|_H^2, \end{aligned}$$

and the assertion of this lemma is proved. □

Again, since the nonlinear function $f(x)$ (res. $g(y)$) is simplified by $f(x^k)$ (resp. $g(y^k)$) in (7.1), it seems that we can not establish the similar principal contractive inequalities as in Section 5 if either $\tilde{w}^k$ or the linear combination of w^k and $\tilde{w}^k$ is taken as the new iterate. However, we can consider the G-norm contractive update form based on the $\tilde{w}^k$ generated by the simplifying version (7.1).

G-norm contractive update form based on the simplifying version C.
The new iterate w^{k+1} is given by

$$w^{k+1} = w^k - \alpha_k G^{-1} d^C(w^k, \tilde{w}^k), \tag{7.15a}$$

or

$$w^{k+1} = P_{\mathcal{W},G}\big\{w^k - \alpha_k G^{-1}[F(\tilde{w}^k) + \eta(y^k, \tilde{y}^k)]\big\}, \tag{7.15b}$$

where

$$\alpha_k = \omega\alpha_k^*, \qquad \alpha_k^* = \frac{\varphi^C(w^k, \tilde{w}^k)}{\|G^{-1}d^C(w^k, \tilde{w}^k)\|_G^2} \qquad \text{and} \quad \omega \in (0,2). \tag{7.15c}$$

Under the conditions (7.10) the matrix M is positive definite.

Theorem 7.4. *Let $\tilde{w}^k$ be generated by the proximal ADM-scheme* (7.1) *from the given vector w^k. If the new iterate w^{k+1} is updated by* (7.15) *with an $G = I$, then we have*

$$\begin{aligned}
&\|w^{k+1} - w^*\|_G^2 \\
\le\; & \|w^k - w^*\|_G^2 - \frac{\omega(2-\omega)}{4(\|M^{1/2}G^{-1}M^{1/2}\|)}\big(\|M^{-1}d^C(w^k, \tilde{w}^k)\|_M^2 + \|A\tilde{x}^k + By^k - b\|_H^2\big) \\
& - \frac{\omega(2-\omega)}{4(\|M^{1/2}G^{-1}M^{1/2}\|)}\big\{(1-\nu)\big(r\|x^k - \tilde{x}^k\|^2 + s\|y^k - \tilde{y}^k\|^2\big)\big\}, \quad \forall w^* \in \mathcal{W}^*,
\end{aligned} \tag{7.16}$$

where M is defined in (2.20).

Proof. First, it follows from Lemma 7.1 and Theorem 2.4 that

$$\|w^k - w^*\|_G^2 - \|w^{k+1} - w^*\|_G^2 \ge 2\omega\alpha_k^*\varphi^C(w^k, \tilde{w}^k) - (\omega\alpha_k^*)^2\|G^{-1}d^C(w^k, \tilde{w}^k)\|_G^2, \;\; \forall\, w^* \in \mathcal{W}^*.$$

By using $\alpha_k^*\|G^{-1}d^C(w^k, \tilde{w}^k)\|_G^2 = \varphi^C(w^k, \tilde{w}^k)$, we obtain

$$\|w^k - w^*\|_G^2 - \|w^{k+1} - w^*\|_G^2 \ge \omega(2-\omega)\alpha_k^*\varphi^C(w^k, \tilde{w}^k), \quad \forall w^* \in \mathcal{W}^*. \tag{7.17}$$

It follows from

$$\|G^{-1}d^C(w^k, \tilde{w}^k)\|_G^2 \le \|M^{1/2}G^{-1}M^{1/2}\| \cdot \|M^{-1}d^C(w^k, \tilde{w}^k)\|_M^2,$$

and (7.11) that

$$\alpha_k^*\varphi^C(w^k, \tilde{w}^k) \ge \frac{1}{4(\|M^{1/2}G^{-1}M^{1/2}\|)}\|M^{-1}d^C(w^k, \tilde{w}^k)\|_M^2.$$

Substituting it in (7.17) the theorem is proved. □

The simplest choices in (7.15) are $G = I$ or $G = M$. For $G = M$, the update form (7.15b) is the same separable forms (6.18).

8 Conclusions

In order to simplify the sub-problems in the alternating directions based contraction method, this paper presents some proximal-like methods by using the proximal alternating direction method schemes to produce the trial points. The convergence of the different versions of the proposed methods follows from the general framework of the contraction methods [19]. The simplifying versions of the proximal alternating direction based contraction methods substantially broaden the applicable scope of the alternating direction method.

References

[1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[2] E. Blum and W. Oettli, *Mathematische Optimierung, Econometrics and Operations Research XX*, Springer Verlag, 1975.

[3] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, Nov. 2010.

[4] J. Eckstein and M. Fukushima, 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, (1994), 115-134.

[5] E. Esser. Applications of lagrangian-based alternating direction methods and connections to split bregman. *UCLA CAM Report 09-31*, 2009.

[6] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity problems, Volume I*, Springer Series in Operations Research, Springer-Verlag, 2003.

[7] M. C. Ferris and J. S. Pang, Engineering and economic applications of complementarity problems, *SIAM Review* 39 (1997), 669–713.

[8] M. Fukushima, Application of the alternating direction method of multipliers to separable convex programming problems, *Computational Optimization and Applications*, 2 (1992), 93-111.

[9] J. D. Fuller and W. Chung, Benders decomposition for a class of variational inequalities, *European Journal Operational Research* 185 (2008), 76-91.

[10] D. Gabay, 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 (1983), 299-331.

[11] D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximations, *Computational Mathematics with Applications* 2 (1976), 17-40.

[12] R. Glowinski, *Numerical methods for Nonlinear Variational Problems*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[13] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM Studies in Applied Mathematics, Philadelphia, PA, 1989.

[14] D. R. Han, W. Xu and H. Yang, An operator splitting method for variational inequalities with partially unknown mappings, *Numerische Mathematik*, 111 (2008), 207-237.

[15] B.S. He, Inexact implicit methods for monotone general variational inequalities, *Mathematical Programming* 86 (1999), 199-217.

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

[17] B. S. He, L.-Z. Liao and M.-J. Qian, Alternating projection based prediction-correction methods for structured variational inequalities, *Journal of Computational Mathematics*, 24 (2006), 693-710.

[18] B.S. He, H. Yang and S.L. Wang, Alternating directions method with self-adaptive penalty parameters for monotone variational inequalities, *Journal of Optimization Theory and applications*, 106 (2000), 349–368.

[19] B. S. He and M.-H. Xu, A general framework of contraction methods for monotone variational inequalities, *Pacific J. Optimization* 4 (2008), 195-212.

[20] B. S. He, M. H. Xu and X. M. Yuan, Solving large-scale least squares covariance matrix problems by alternating direction methods, *SIAM J. Matrix Analysis and Appl.* 32 (2011), 136-152.

[21] B.S. He, X.M. Yuan and J.J.Z. Zhang, Comparison of two kinds of prediction-correction methods for monotone variational inequalities. *Comput. Optim. Appl.* 27 (2004), 247–267.

[22] B. S. He and H. Yang, Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities, *Operations Research Letters* 23 (1998), 151-161.

[23] S. Kontogiorgis and R. R. Meyer, A variable-penalty alternating directions method for convex optimization, *Mathematical Programming* 83 (1998), 29-53.

[24] Z. C. Lin, M. M. Chen, L. Q. Wu, and Y. Ma. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *manuscript*, 2009.

[25] B. Martinet, Regularization d'inequations variationelles par approximations sucessives, *Revue Francaise d'Informatique et de Recherche Opérationelle* 4 (1970), 154–159.

[26] A. Nagurney, *Network economics, A variational inequality approach*, Kluwer Academics Publishers, Dordrecht, 1993.

[27] M. Ng, P. Weiss, and X. M. Yuan. Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods. *Optimization Online*, 2009.

[28] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization* 14 (1976) 877-898.

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

[30] M. Tao, J. F. Yang and B. S. He, Alternating direction algorithms for total variation deconvolution in image reconstruction. *Optimization Online*, 2009.

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

[32] Zaiwen Wen, Wotao Yin, and Donald Goldfarb. Alternating direction augmented lagrangian methods for semidefinite programming. *Technical Report, IEOR, Columbia University*, 2009.

[33] M. H. Xu, Proximal alternating directions method for structured variational inequalities, *J. of Optimization, Theory and Applications*, 134 (2007) 107-117.

[34] J. F. Yang and Y. Zhang. Alternating direction algorithms for l_1-problems in compressive sensing, TR09-37, CAAM, Rice University. 2009.

[35] C. H. Ye and X. M. Yuan, A descent method for structured monotone variational inequalities, *Optimization Methods and Software*, 22 (2007), 329-338.

[36] X. M. Yuan and J. F. Yang. Sparse and low-rank matrix decomposition via alternating direction methods. *Optimization Online*, 2009.

[37] X. M. Yuan Alternating direction methods for sparse covariance selection. *Optimization Online*, 2009.