Alternating Direction Method of Multipliers with Variable Metric Indefinite Proximal Terms for Convex Optimization

Yan Gu* and Nobuo Yamashita*

*Graduate School of Informatics, Kyoto University, Kyoto 6068501, Japan.
Email: yan@amp.i.kyoto-u.ac.jp; nobuo@i.kyoto-u.ac.jp.

July 1, 2019

Abstract

This paper studies a proximal alternating direction method of multipliers (ADMM) with variable metric indefinite proximal terms for linearly constrained convex optimization problems. The proximal ADMM plays an important role in many application areas, since the subproblems of the method are easy to solve. Recently, it is reported that the proximal ADMM with a certain fixed indefinite proximal term is faster than that with a positive semidefinite term, and still has the global convergence property. On the other hand, Gu and Yamashita studied a variable metric semidefinite proximal ADMM whose proximal term is generated by the BFGS update. They reported that a slightly indefinite matrix also makes the algorithm work well in their numerical experiments. Motivated by this fact, we consider a variable metric indefinite proximal ADMM, and give sufficient conditions on the proximal terms for the global convergence. Moreover, we propose a new indefinite proximal term based on the BFGS update which can satisfy the conditions for the global convergence.

Keywords: alternating direction method of multipliers, variable metric indefinite proximal term, BFGS update, global convergence, convex optimization

1 Introduction

We consider the following convex composite optimization problem:

\[
\min \{ f(x) + g(y) \mid Ax + By = b, \ x \in \mathbb{R}^n, \ y \in \mathbb{R}^n \},
\]

where \( f: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) and \( g: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) are proper convex functions, \( A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \). Various practical problems of science and engineering, such as machine learning \([33, 43]\), total variation denoising \([38]\) and statistics \([39]\) can be formulated as Problem (1.1). Usually, we say that \( f \) is a loss function and \( g \) is a structured regularization term.

The augmented Lagrangian function of (1.1) is defined as

\[
\mathcal{L}_\beta(x, y, \lambda) := f(x) + g(y) - \langle \lambda, Ax + By - b \rangle + \frac{\beta}{2} \| Ax + By - b \|^2,
\]

where \( \lambda \in \mathbb{R}^m \) is the Lagrangian multiplier for the linear constraints \( Ax + By = b \) in (1.1), and \( \beta \) is a positive scalar. Note that \( \mathcal{L}_\beta: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \).
A number of efficient first-order algorithms have been developed for problem (1.1) including operator splitting methods \[13,31,13,35\], gradient methods \[37,40,41\], primal dual methods \[5,7,17\], etc. One may solve problem (1.1) is the classical augmented Lagrangian method (ALM), which generates the updates

\[
\begin{align*}
(x^{k+1}, y^{k+1}, \lambda^{k+1}) &= \arg\min_{x,y} \mathcal{L}_\beta(x, y, \lambda^k) \\
\mathcal{L}_\beta(x, y, \lambda^k) &= \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b).
\end{align*}
\]

In this case, the vectors \(x^{k+1}\) and \(y^{k+1}\) should be updated at the same time ignoring the separability of the original functions. Generally, the joint minimization problem (1.3) is a challenge to be solved exactly or approximately with a high accuracy. We want to exploit the separability of the objective function to reduce the difficulty. The classical ADMM is one of such methods, and it efficiently solves problem (1.1) \[20,22\]. The convergence analysis for the classical ADMM can be referred to \[4,14,19,20,22\].

Fazel et al. \[18\] proposed a more convenient semi-proximal ADMM by adding proximal terms to subproblems which takes the following scheme:

\[
\begin{align*}
x^{k+1} &= \arg\min_x \mathcal{L}_\beta(x, y^k, \lambda^k) + \frac{1}{2}\|x - x^k\|^2_S, \\
y^{k+1} &= \arg\min_y \mathcal{L}_\beta(x^{k+1}, y, \lambda^k) + \frac{1}{2}\|y - y^k\|^2_T, \\
\lambda^{k+1} &= \lambda^k - \alpha_\beta (Ax^{k+1} + By^{k+1} - b),
\end{align*}
\]

where \(\alpha \in (0, (1 + \sqrt{5})/2\), and \(S, T \succeq 0\). For a vector \(z \in \mathbb{R}^n\) and a semidefinite matrix \(G\), the norm \(\| \cdot \|_G\) is defined by \(\|z\|_G = \sqrt{z^\top Gz}\). In this paper, even if \(G \in \mathbb{R}^{n \times n}\) is not positive semidefinite, we denote \(\|z\|_G = z^\top Gz\) for simplicity.

The proximal ADMM covers the classical ADMM when \(S = T = 0\). When \(S\) and \(T\) are two positive definite matrices and \(\alpha = 1\), this semi-proximal ADMM reduces to the proximal ADMM proposed by Eckstein \[12\]. The proximal ADMM has an advantage that its subproblems are easy to solve, and it also can efficiently handle the multi-block convex optimization problem which is known as block-wise ADMM \[31\]. See \[9,18,30,42\] for more details of the semi-proximal ADMM.

It is well known that the global convergence of the semi-proximal ADMM (1.4) is easier to prove. However, it is not satisfactory in numerical performance. The paper \[10\] mentioned that the proximal matrix \(T\) in (1.4b) could be indefinite if \(\alpha \in (0, 1)\) though it provided no further discussions on theoretical properties. Then Li et al. \[34\] proved the global convergence. He et al. \[29\] proposed a linearized version of ADMM with a positive-definite proximal term. They considered the case that matrix \(S = 0\) and \(\alpha = 1\) in (1.4), and generated the proximal matrix \(T\) as

\[
T = \tau r I - \beta B^\top B \text{ with } r > \beta \|B^\top B\|, \quad \tau \in (0.75, 1).
\]

The proximal matrix \(T\) is not necessarily positive semidefinite. A smaller value \(\tau \in (0.75, 1)\) can ensure the convergence and also give better numerical performance.

How to choose the proximal term is also one of the important research topics for ADMM. The popular proximal term is always chosen as a constant matrix. He et al. \[27\] extended the work to allow the parameters \(\beta\), proximal terms \(T\) and \(S\) to be replaced by some bounded sequences of positive definite matrices \(\{T_k\}\) and \(\{S_k\}\). The resulting ADMM is a variable metric proximal ADMM, which is also closely related to the inexact ADMM \[6,13,15,16,27,44\]. The convergences of such methods have been studied in \[2,23,36\] but a better selection of the sequence \(\{T_k\}\) has not been provided.
Quite recently, Gu and Yamashita [25] proposed to construct a variable positive semi-definite sequence \( \{T_k\} \) with \( T_k = B_k - \nabla^2_{xx} \beta(x, y, \lambda) \) when \( f \) is quadratic. Note that \( M = \nabla^2_{xx} \beta(x, y, \lambda) \) is a constant matrix. They generated \( B_k \) via the BFGS update with respect to \( M \) at every iteration. Gu and Yamashita [26] further extended such a proximal ADMM for more general convex optimization problems with the proximal term generated by the Broyden family update. In these ADMMs, the proximal terms \( T_k \) contain some second order information on the augmented Lagrangian function. The papers [25, 26] report some numerical results for LASSO and L1 regularized logistic regression. The results show that the algorithms can get a solution faster than the general indefinite proximal ADMM whose proximal term is fixed. Another interesting numerical result in [25, 26] is that a variable indefinite sequence via the BFGS update also shows a good performance.

Inspired by the variable metric semi-proximal ADMM [25, 26] and the indefinite proximal ADMM [29], it is worth considering ADMM with a sequence of indefinite proximal matrices. We call the resulting ADMM a variable metric indefinite proximal ADMM (VMIP-ADMM). Throughout our discussion, we always choose the stepsize \( \alpha \) in (1.4c) be 1 as that in [29], which is good enough for such methods in practice and simple for the convergence analysis.

We now introduce the whole update scheme of the VMIP-ADMM:

\[
\begin{align*}
    x^{k+1} &= \arg \min_x \mathcal{L}_\beta(x, y^k, \lambda^k) + \frac{1}{2} \|x - x^k\|_S^2, \\
    y^{k+1} &= \arg \min_y \mathcal{L}_\beta(x^{k+1}, y, \lambda^k) + \frac{1}{2} \|y - y^k\|_{T_k}^2, \\
    \lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b),
\end{align*}
\]

where \( S \) is a fixed positive semi-definite and \( T_k \) is possibly indefinite. Note that the VMIP-ADMM can unify the several existing ADMMs.

- Let \( S = 0, T_k \equiv 0 \), VMIP-ADMM reduces to the classical ADMM;
- Let \( S \) and \( T_k \equiv T \) be positive semidefinite matrices, VMIP-ADMM turns to be the semi-proximal ADMM (1.4);
- Let \( \{T_k\} \) be a positive semidefinite sequence, that is, \( T_k \succeq 0 \) for all \( k \). VMIP-ADMM becomes the variable semi-proximal ADMM;
- Let \( S = 0, T_k \equiv T \) be a positive indefinite matrix, VMIP-ADMM covers the indefinite-proximal ADMM proposed in [29].

We present sufficient conditions on \( \{T_k\} \) for the global convergence of VMIP-ADMM. The proof is followed by the analysis technique in Gu et al. [24], which separated the constant indefinite term “\( T \)” into two semidefinite parts as \( T = T_+ - T_- \). Moreover, we provide a construction of the indefinite term \( T_k \) via the BFGS update. We extend a useful theorem in [24] for a special case when \( y \)-subproblems (1.6b) are unconstrained quadratic programming problems. We construct the \( T_k \) with \( T_k = B_k - M \), where \( M \) is the Hessian matrix of the augmented Lagrangian function (1.2) and \( B_k \) is generated by the BFGS update with respect to \( \tau M \), \( \tau < 1 \). We also show that this construction of \( T_k \) satisfies the above conditions for the global convergence property when \( \tau \in (0.75, 1) \).

The remaining parts of the paper are organized as follows. We first give notations and some preliminaries that will be useful for subsequent analysis in Section 2. Then we present sufficient conditions on
the proximal matrices \( \{ T_k \} \) for the global convergence. In Section 3 we discuss the choices of proximal matrix \( T_k \) that guarantees the global convergence. We also show how to determine the value of \( \tau \). Some conclusions and future works are given in Section 4.

2 Global convergence of the variable metric indefinite proximal ADMM

In this section, we show the global convergence of the variable metric indefinite proximal ADMM (VMIP-ADMM) for problem (1.1). To this end, we first present optimality conditions of problem (1.1) and some useful properties which will be frequently used in our analysis. Then we give sufficient conditions on \( \{ T_k \} \) under which VMIP-ADMM converges globally.

2.1 Optimality conditions for problem (1.1)

Let \( \Omega = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \). The KKT conditions of problem (1.1) are written as:

\[
\begin{align*}
\xi_x^* - A^\top \lambda^* &= 0, \\
\xi_y^* - B^\top \lambda^* &= 0, \\
Ax^* + By^* - b &= 0, \\
\xi_x^* &\in \partial f(x^*), \quad \xi_y^* &\in \partial g(y^*).
\end{align*}
\]

Let \( \Omega^* \) be a set of \((x^*, y^*, \lambda^*)\) satisfying the KKT conditions (2.1).

Throughout this paper, we make the following assumption.

Assumption 2.1. The set \( \Omega^* \) of KKT points is non-empty.

The optimality conditions of subproblems (1.6a) and (1.6b) can be obtained respectively that

\[
(x - x^{k+1})^\top \left( \xi_{x}^{k+1} - A^\top \lambda^k + \beta A^\top (Ax^{k+1} + By^k - b) + S(x^{k+1} - x^k) \right) \geq 0, \quad \forall x \in \mathbb{R}^n,
\]

and

\[
(y - y^{k+1})^\top \left( \xi_{y}^{k+1} - B^\top \lambda^k + \beta B^\top (Ax^{k+1} + By^{k+1} - b) + T_k(y^{k+1} - y^k) \right) \geq 0, \quad \forall y \in \mathbb{R}^n,
\]

where \( \xi_{x}^{k+1} \in \partial f(x^{k+1}) \) and \( \xi_{y}^{k+1} \in \partial g(y^{k+1}) \).

Since \( \lambda^{k+1} = \lambda^k - \beta (Ax^{k+1} + By^{k+1} - b) \) from (1.6c), we have

\[
-A^\top \lambda^k + \beta A^\top (Ax^{k+1} - b) = -A^\top \lambda^{k+1} - \beta A^\top By^{k+1}
\]

and

\[
-B^\top \lambda^k + \beta B^\top (Ax^{k+1} + By^{k+1} - b) = -B^\top \lambda^{k+1}.
\]

Then the above optimality conditions can be written as

\[
(x - x^{k+1})^\top \left( \xi_{x}^{k+1} - A^\top \lambda^{k+1} + \beta A^\top B(y^k - y^{k+1}) + S(x^{k+1} - x^k) \right) \geq 0, \quad \forall x \in \mathbb{R}^n,
\]

and

\[
(y - y^{k+1})^\top \left( \xi_{y}^{k+1} - B^\top \lambda^{k+1} + T_k(y^{k+1} - y^k) \right) \geq 0, \quad \forall y \in \mathbb{R}^n.
\]
2.2 Notations and Conditions on \{T_k\}

We use the following notations throughout this paper:

\[ u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}. \]

Since the subdifferential mappings of the closed proper convex functions \( f \) and \( g \) are maximal monotone, there exist two positive semidefinite matrices \( \Sigma_f \) and \( \Sigma_g \) such that for all \( x, \hat{x} \in \mathbb{R}^n, \xi_x \in \partial f(x) \), and \( \hat{\xi}_x \in \partial f(\hat{x}) \),

\[(x - \hat{x})^\top (\xi_x - \hat{\xi}_x) \geq \|x - \hat{x}\|^2_{\Sigma_f}, \tag{2.4}\]

and for all \( y, \hat{y} \in \mathbb{R}^n, \xi_y \in \partial g(y) \), and \( \hat{\xi}_y \in \partial g(\hat{y}) \),

\[(y - \hat{y})^\top (\xi_y - \hat{\xi}_y) \geq \|y - \hat{y}\|^2_{\Sigma_g}. \tag{2.5}\]

Let \( \Sigma \in \mathbb{R}^{2n \times 2n} \) denote

\[ \Sigma = \begin{pmatrix} \Sigma_f & 0 \\ 0 & \Sigma_g \end{pmatrix}. \]

We first give the conditions for \( S \) and the indefinite proximal sequence \{\( T_k \)\} to guarantee the global convergence.

**Condition 2.1.** The matrix \( S \) in (1.6a) satisfies

(a) \( S + \frac{1}{2} \Sigma_f \preceq 0; \)

(b) \( S + \Sigma_f + \beta A^\top A \succ 0. \)

Moreover, for sequence \{\( T_k \)\} generated in (1.6), there exist a non-negative sequence \{\( \gamma_k \)\} and positive semidefinite sequences \{\( T^+_k \)\} and \{\( T^-_k \)\} such that

(c) \( T_k = T^+_k - T^-_k \) for all \( k \);

(d) \( T_k + \Sigma_g + \beta B^\top B \succeq 0 \) for all \( k \);

(e) \( \frac{1}{1+\gamma_k} T^+_k \leq T^{k+1}_+ \leq (1 + \gamma_k) T^+_k, \) \( \forall k \geq 0, \sum_{k=0}^{\infty} \gamma_k < \infty; \)

(f) \( T_{k+1} + \Sigma_g + \beta B^\top B \succeq (1 + \gamma_k)(T_k + \Sigma_g + \beta B^\top B) \) for all \( k \);

(g) \( \exists c \in (0, 0.5), T_k + \frac{3}{2} \Sigma_g - \frac{3k-1}{2} T^+_k - 2T^-_k + (\frac{3}{4} - \frac{1}{2}c) \beta B^\top B \succeq 0 \) for all \( k \).

Condition (a) and (b) indicate that the proximal matrix \( S \) is allowed to be a slight indefinite but no less than \( -\frac{1}{2} \Sigma_f \). Condition (c) decomposes the indefinite matrix \( T_k \) to two positive semidefinite parts. Note that we require the second part \( T^- \) be fixed. This condition will play an important role in the main analysis. Condition (d) allows \( T_k \) to be indefinite. Condition (e) and (f) are the boundness for positive semi-definite part \( T^+_k \) and indefinite \( T_k \), respectively. Condition (g) is a requirement for global convergence and also an important condition for us to discuss the range of the indefiniteness.
For simplicity, we further define the following matrices. For all $k$,

$$P_k = \begin{pmatrix} S & 0 \\ 0 & T_k \end{pmatrix}, \quad D_k = \begin{pmatrix} S & 0 & 0 \\ 0 & T_k & 0 \\ 0 & 0 & \frac{1}{\beta} I \end{pmatrix}, \quad \text{and} \quad G_k = \begin{pmatrix} S + \Sigma_f & 0 & 0 \\ 0 & T_k + \Sigma_g + \beta B^\top B & 0 \\ 0 & 0 & \frac{1}{\beta} I \end{pmatrix},$$

(2.6)

where $S, T_k$ and $\beta$ are given in (1.6).

Moreover, we also define the following matrices.

$$\Gamma_k = T_k^k + T_-, \quad \forall k,$$

(2.7a)

$$\Lambda_k = -\frac{\gamma_{k-1}}{2} T_k^k - 2 T_- + \Sigma_g, \quad \forall k,$$

(2.7b)

$$\Delta_k = T_k + \frac{3}{2} \Sigma_g - \frac{\gamma_{k-1}}{2} T_k^k - 2 T_- + \left(\frac{3}{4} - \frac{1}{2} c\right) \beta B^\top B, \quad \forall k,$$

(2.7c)

where $\{\gamma_k\}$ is a sequence satisfying Condition 2.1. Note that $\Gamma_k \succeq 0$ for all $k$.

### 2.3 Technical lemmas for convergence analysis of the variable metric indefinite proximal ADMM

In order to show that VMIP-ADMM converges to a solution of (1.1) globally, we first give some properties for the sequence $\{w_k\} = \{(x^k, y^k, \lambda^k)\}$ generated by (1.6).

**Lemma 2.2.** Let $\{w^k\}$ be generated by (1.6). Then, for given $w^* = (x^*, y^*, \lambda^*) \in \Omega^*$, we have

$$(w^{k+1} - w^*)^\top D_k (w^{k+1} - w^k) + \|u^{k+1} - u^*\|_\Sigma^2 \leq \beta (Ax^{k+1} - Ax^*)^\top (By^{k+1} - By^k).$$

(2.8)

**Proof.** By taking $x = x^*$ and $y = y^*$ in the optimality conditions (2.2) and (2.3), respectively, we have

$$(x^{k+1} - x^*)^\top (\xi_x^{k+1} - A^\top \lambda^{k+1} + \beta A^\top B (y^k - y^{k+1}) + S(x^{k+1} - x^k)) \leq 0,$$

and

$$(y^{k+1} - y^*)^\top (\xi_y^{k+1} - B^\top \lambda^{k+1} + T_k (y^{k+1} - y^k)) \leq 0,$$

where $\xi_x^{k+1} \in \partial f(x^{k+1})$ and $\xi_y^{k+1} \in \partial g(y^{k+1})$.

The inequalities are further rearranged as

$$(x^{k+1} - x^*)^\top S(x^{k+1} - x^k) + (x^{k+1} - x^*)^\top (\xi_x^{k+1} - A^\top \lambda^{k+1}) \leq \beta (Ax^{k+1} - Ax^*)^\top (By^{k+1} - By^k)$$

(2.9)

and

$$(y^{k+1} - y^*)^\top T_k (y^{k+1} - y^k) + (y^{k+1} - y^*)^\top (\xi_y^{k+1} - B^\top \lambda^{k+1}) \leq 0.$$  

(2.10)

Moreover, from (2.4)-(2.5) with $x = x^{k+1}, y = y^{k+1}, \hat{x} = x^*$ and $\hat{y} = y^*$, we have

$$(x^{k+1} - x^*)^\top (\xi_x^{k+1} - \xi_x^*) \geq \|x^{k+1} - x^*\|_{\Sigma_f}^2,$$

(2.11)

and

$$(y^{k+1} - y^*)^\top (\xi_y^{k+1} - \xi_y^*) \geq \|y^{k+1} - y^*\|_{\Sigma_g}^2,$$

(2.12)
where $\xi'_x \in \partial f(x^*)$ and $\xi'_y \in \partial g(y^*)$ satisfy the KKT conditions (2.1a) and (2.1b), respectively. It then follows from (2.1a) and (2.11) that

$$
(x^{k+1} - x^*)^\top (\xi'_x^{k+1} - A^\top \lambda^{k+1}) = (x^{k+1} - x^*)^\top (\xi'_x^{k+1} - \xi'_x^*) + (x^{k+1} - x^*)^\top (\xi'_x^* - A^\top \lambda^{k+1}) \\
\geq \|x^{k+1} - x^*\|_{\Sigma_j}^2 + (Ax^{k+1} - Ax^*)^\top (\lambda^* - \lambda^{k+1}).
$$

Combining this inequality and (2.9), we have

$$
(x^{k+1} - x^*)^\top S(x^{k+1} - x^k) + (Ax^{k+1} - Ax^*)^\top (\lambda^* - \lambda^{k+1}) + \|x^{k+1} - x^*\|^2_{\Sigma_j} \\
\leq \beta(Ax^{k+1} - Ax^*)^\top (By^{k+1} - By^k).
$$

In a similar way, we have from (2.1b), (2.10) and (2.12) that

$$
(y^{k+1} - y^*)^\top T_k(y^{k+1} - y^k) + (By^{k+1} - By^*)^\top (\lambda^* - \lambda^{k+1}) + \|y^{k+1} - y^*\|^2_{\Sigma_j} \leq 0.
$$

Rearranging (1.6e), we have $Ax^{k+1} + By^{k+1} - b = \frac{1}{\beta} (\lambda^k - \lambda^{k+1})$. It then follows from (2.14) that

$$
Ax^{k+1} + By^{k+1} - Ax^* - By^* = \frac{1}{\beta} (\lambda^k - \lambda^{k+1}).
$$

Adding (2.13) and (2.14), and recalling the definition of $D_k$ and $\Sigma$, it holds that

$$
(w^{k+1} - w^*)^\top D_k(w^{k+1} - w^k) + \|w^{k+1} - w^*\|^2_{\Sigma} \\
= (x^{k+1} - x^*)^\top S(x^{k+1} - x^k) + (y^{k+1} - y^*)^\top T_k(y^{k+1} - y^k) \\
+ \frac{1}{\beta}(\lambda^{k+1} - \lambda^k)^\top (\lambda^{k+1} - \lambda^*) + \|w^{k+1} - w^*\|^2_{\Sigma} \\
= (x^{k+1} - x^*)^\top S(x^{k+1} - x^k) + (y^{k+1} - y^*)^\top T_k(y^{k+1} - y^k) \\
+ (Ax^{k+1} + By^{k+1} - Ax^* - By^*)^\top (\lambda^* - \lambda^{k+1}) + \|w^{k+1} - w^*\|^2_{\Sigma} \\
\leq \beta(Ax^{k+1} - Ax^*)^\top (By^{k+1} - By^k). \quad \square
$$

The inequality (2.8) in Lemma 2.2 is further rearranged as follows.

**Lemma 2.3.** Let $\{w^k\}$ be generated by (1.6). Then, for given $w^* = (x^*, y^*, \lambda^*) \in \Omega^*$, we have

$$
2(w^{k+1} - w^*)^\top D_k(w^{k+1} - w^k) \\
\leq 2(By^{k+1} - By^*)^\top (\lambda^k - \lambda^{k+1}) - 2\beta(By^{k+1} - By^*)^\top (By^{k+1} - By^k) - 2\|w^{k+1} - w^*\|^2_{\Sigma};
$$

(2.15)

**Proof.** Noting that $Ax^* + By^* - b = 0$, the twice of the right hand of (2.8) is written as

$$
2\beta(Ax^{k+1} - Ax^*)^\top (By^{k+1} - By^k) \\
= 2\beta(Ax^{k+1} + By^* - b + By^{k+1} - By^k)^\top (By^{k+1} - By^k) \\
= 2\beta(Ax^{k+1} + By^{k+1} - b)^\top (By^{k+1} - By^k) - 2\beta(By^{k+1} - By^*)^\top (By^{k+1} - By^k) \\
= 2(By^{k+1} - By^k)^\top (\lambda^k - \lambda^{k+1}) - 2\beta(By^{k+1} - By^*)^\top (By^{k+1} - By^k),
$$

where the last equality follows from (1.6e). Then the assertion is directly obtained from (2.8). \quad \square

7
Next we give a simple but important lemma.

**Lemma 2.4.** For vectors \( a, b \in \mathbb{R}^n \), and symmetric positive semidefinite matrices \( M_1, M_2 \in \mathbb{R}^{n \times n} \), we have that

\[
a^\top M_1 b - a^\top M_2 b \leq \frac{1}{2} a^\top (M_1 + M_2) a + \frac{1}{2} b^\top (M_1 + M_2) b.
\]

(2.16)

**Proof.** For a positive semidefinite matrix \( M_1 \), we have

\[
0 \leq \frac{1}{2} \|a - b\|^2_{M_1} = \frac{1}{2} a^\top M_1 a + \frac{1}{2} b^\top M_1 b - a^\top M_1 b,
\]

which implies

\[
a^\top M_1 b \leq \frac{1}{2} a^\top M_1 a + \frac{1}{2} b^\top M_1 b.
\]

(2.17)

In a similar way for \( M_2 \), we have

\[
-a^\top M_2 b \leq \frac{1}{2} a^\top M_2 a + \frac{1}{2} b^\top M_2 b.
\]

(2.18)

The assertion immediately follows by adding (2.17) and (2.18).

In order to bound \((w^{k+1} - w^*)^\top D_k (w^{k+1} - w^k)\) further, we now give two technical lemmas to estimate upper-bounds for the crossing term \((By^{k+1} - By^k)^\top (\lambda^{k+1} - \lambda^k)\) in (2.15).

**Lemma 2.5.** Let \( \{w^k\} \) be generated by the scheme (1.6). Suppose that the proximal sequence \( \{T_k\} \) satisfies Condition 2.1. Then it holds that

\[
(By^{k+1} - By^k)^\top (\lambda^{k+1} - \lambda^k) \leq \frac{1}{2} \|y^{k+1} - y^k\|^2_{\Gamma_{k-1}} - \frac{1}{2} \|y^{k+1} - y^k\|^2_{\Lambda_k} - \|y^{k+1} - y^k\|^2_{\Lambda_k},
\]

(2.19)

where \( \Gamma_k \) and \( \Lambda_k \) are defined in (2.7).

**Proof.** From the optimality condition (2.3) for \( y^{k+1} \), we can easily derive the optimality condition for \( y^k \) as

\[
(y - y^k)^\top (\xi^k_y - B^\top \lambda + T_{k-1}(y^k - y^{k-1})) \geq 0, \quad \forall y \in \mathbb{R}^n.
\]

(2.20)

Choosing \( y = y^k \) in (2.3), we have

\[
0 \leq (y^k - y^{k+1})^\top (\xi^k_y - B^\top \lambda^{k+1} + T_k(y^{k+1} - y^k)) = (y^{k+1} - y^k)^\top (-\xi^k_y + B^\top \lambda^{k+1} - T_k(y^{k+1} - y^k)).
\]

(2.21)

Moreover, letting \( y = y^{k+1} \) in (2.20), we have

\[
0 \leq (y^{k+1} - y^k)^\top (\xi^k_y - B^\top \lambda^{k+1} + T_{k-1}(y^k - y^{k-1})).
\]

(2.22)

Summing inequalities (2.21) and (2.22), we obtain that

\[
0 \leq - (y^{k+1} - y^k)^\top (\xi^k_y + \xi^k_y) + (By^{k+1} - By^k)^\top (\lambda^{k+1} - \lambda^k) + (y^{k+1} - y^k)^\top T_{k-1}(y^k - y^{k-1}) - \|y^{k+1} - y^k\|^2_{\Gamma_k}.
\]

It then follows from (2.5) that

\[
0 \leq -\|y^{k+1} - y^k\|^2_{\Sigma_y} + (By^{k+1} - By^k)^\top (\lambda^{k+1} - \lambda^k) + (y^{k+1} - y^k)^\top T_{k-1}(y^k - y^{k-1}) - \|y^{k+1} - y^k\|^2_{\Gamma_k}.
\]
which is equivalent to
\[
(By^{k+1} - By^k)\top(\lambda^k - \lambda^{k+1}) \leq -\|y^{k+1} - y^k\|_2^2 + (y^{k+1} - y^k)\top T_{k-1}(y^k - y^{k-1}) - \|y^{k+1} - y^k\|_2^2.
\] (2.23)

Recall that \(T_{k-1} = T_+^{k-1} - T_-\) from (c) in Condition 2.1 and \(T_+^{k-1}, T_- \succeq 0\). Then we have
\[
(y^{k+1} - y^k)\top T_{k-1}(y^k - y^{k-1}) = (y^{k+1} - y^k)\top T_+^{k-1}(y^k - y^{k-1}) - (y^{k+1} - y^k)\top T_-(y^k - y^{k-1})
\leq \frac{1}{2}\|y^{k+1} - y^k\|_2^2 + \frac{1}{2}\|y^{k+1} - y^k\|^2_{T_+^{k-1} + T_-},
\] (2.24)
where the inequality follows from (2.16) with \(a = (y^{k+1} - y^k)\), \(b = (y^k - y^{k-1})\), \(M_1 = T_+^{k-1}\) and \(M_2 = T_-\).

We then have from (2.23) that
\[
(By^{k+1} - By^k)\top(\lambda^k - \lambda^{k+1})
\leq -\|y^{k+1} - y^k\|_2^2 + (y^{k+1} - y^k)\top T_{k-1}(y^k - y^{k-1}) - \|y^{k+1} - y^k\|_2^2
\leq -\|y^{k+1} - y^k\|_2^2 + (y^{k+1} - y^k)\top T_{k-1} + \frac{1}{2}\|y^{k+1} - y^k\|_2^2 + \frac{1}{2}\|y^{k+1} - y^k\|^2_{T_+^{k-1} + T_-} - \|y^{k+1} - y^k\|_2^2
\leq -\|y^{k+1} - y^k\|_2^2 + (y^{k+1} - y^k)\top T_{k-1} + \frac{1}{2}\|y^{k+1} - y^k\|_2^2 + \frac{1}{2}\|y^{k+1} - y^k\|^2_{T_+^{k-1} + T_-} - \|y^{k+1} - y^k\|_2^2
= \frac{1}{2}\|y^{k+1} - y^k\|_2^2 + \frac{1}{2}\|y^{k+1} - y^k\|_2^2 + \frac{1}{2}\|y^{k+1} - y^k\|_2^2
= \frac{1}{2}\|y^{k+1} - y^k\|_{T_{k-1}} - \frac{1}{2}\|y^{k+1} - y^k\|_{1_{\lambda_k}} - \|y^{k+1} - y^k\|_{L_k}^2,
\]
where the second inequality follows from \(T_k = T_+^k - T_-\) and (2.24), the third inequality follows from Condition 2.1 (d), and the last equality is from the definitions (2.7a) and (2.7b). Then it shows the assertion (2.19).

Besides Lemma 2.5, we can derive another estimation for \((By^{k+1} - By^k)\top(\lambda^k - \lambda^{k+1})\), whose proof is similar to that in [29] Lemma 4.4.

**Lemma 2.6.** Let \(\{w^k\}\) be generated by the scheme (1.6). Then, for any \(c \in (0, 0.5)\), it holds that
\[
(By^{k+1} - By^k)\top(\lambda^k - \lambda^{k+1}) \leq \left(\frac{1}{4} + \frac{1}{2}c\right)\beta\|By^{k+1} - By^k\|^2 + (1 - c)\frac{1}{\beta}\|\lambda^{k+1} - \lambda^k\|^2.
\] (2.25)

**Proof.** See [29] Lemma 4.4. \(\square\)

Based on the above two lemmas for \((By^{k+1} - By^k)\top(\lambda^k - \lambda^{k+1})\), we can further bound \((w^{k+1} - w^*)\top D_k(w^{k+1} - w^k)\) in (2.15) of Lemma 2.3

**Lemma 2.7.** Let \(\{w^k\}\) be generated by (1.6). Suppose that the proximal sequence \(\{T_k\}\) satisfies Condition 2.7. Then, for given \(w^* = (x^*, y^*, \lambda^*) \in \Omega^*\), we have
\[
2(w^{k+1} - w^*)\top D_k(w^{k+1} - w^k)
\leq \frac{1}{2}\|y^{k+1} - y^k\|_2 + \frac{1}{2}\|y^{k+1} - y^k\|_2^2 - \|y^{k+1} - y^k\|_2^2 + \left(\frac{1}{4} + \frac{1}{2}c\right)\beta\|By^{k+1} - By^k\|^2
\leq \frac{1}{2}\|\lambda^{k+1} - \lambda^k\|^2 - 2\beta(By^{k+1} - By^k)\top (By^{k+1} - By^k) - 2\|u^{k+1} - u^*\|_2^2.
\] (2.26)
Proof. The term $2(By^{k+1} - By^k)\top (\lambda^k - \lambda^{k+1})$ in inequality (2.15) can be bounded by the above lemmas (2.19) and (2.25), and then the assertion is obtained. \hfill \Box

2.4 Global Convergence of the variable metric indefinite proximal ADMM

In this subsection we show the global convergence based on the results in the previous subsection and Condition 2.1. Firstly, we obtain the following contractive result, which will play a key role in proving the convergence of (1.6).

Lemma 2.8. Let $w^* = (x^*, y^*, \lambda^*) \in \Omega^*$, and let $\{w^k\}$ be generated by the scheme (1.6). Suppose that the proximal sequence $\{T_k\}$ satisfies Condition 2.7. Then we have

$$
\|w^k - w^*\|^2_{\Delta_k} + \frac{1}{2}\|y^{k-1} - y^{k}\|^2_{\Delta_{k-1}} - \left(\|w^{k+1} - w^*\|^2_{\Gamma_k} + \frac{1}{2}\|y^{k+1} - y^{k}\|^2_{\Sigma_k}\right)
\geq \|x^{k+1} - x^*\|^2_{S + \frac{\alpha}{2} \Sigma_y} + \|y^{k+1} - y^*\|^2_{\Delta_k} + \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2,
$$

(2.27)

where $\Gamma_k$ and $\Delta_k$ are given in (2.7).

Proof. By the identity $\|a + b\|^2 = \|a\|^2 - \|b\|^2 + 2(a + b)^\top b$, we get

$$
\|By^{k+1} - By^*\|^2 = \|By^k - By^* + By^{k+1} - By^k\|^2
= \|By^k - By^*\|^2 - \|By^{k+1} - By^k\|^2 + 2(By^{k+1} - By^*)^\top (By^{k+1} - By^k).
$$

Moreover,

$$
\|w^{k+1} - w^*\|^2_{D_k} = \|w^k - w^* + w^{k+1} - w^k\|^2_{D_k}
= \|w^k - w^*\|^2_{D_k} - \|w^{k+1} - w^*\|^2_{D_k} + 2(w^{k+1} - w^*)^\top D_k(w^{k+1} - w^k).
$$

Then we have

$$
\|w^{k+1} - w^*\|^2_{D_k} + \beta\|By^{k+1} - By^*\|^2
= \|w^k - w^*\|^2_{D_k} + \beta\|By^k - By^*\|^2 - \left(\|w^{k+1} - w^k\|^2_{D_k} + \beta\|By^{k+1} - By^k\|^2\right)
+ 2(w^{k+1} - w^*)^\top D_k(w^{k+1} - w^k) + 2\beta(By^{k+1} - By^*)^\top (By^{k+1} - By^k).
$$

(2.28)

Since the term $2(w^{k+1} - w^*)^\top D_k(w^{k+1} - w^k)$ in equality (2.28) can be bounded by (2.26) in Lemma 2.7, we can rearrange (2.28) as

$$
\begin{align*}
\|w^{k+1} - w^*\|^2_{D_k} + \beta\|By^{k+1} - By^*\|^2
\leq & \|w^k - w^*\|^2_{D_k} + \beta\|By^k - By^*\|^2 - \|w^{k+1} - w^k\|^2_{D_k} - \beta\|By^{k+1} - By^k\|^2 \\
& + \frac{1}{2}\|y^{k-1} - y^{k}\|^2_{\Gamma_{k-1}} - \frac{1}{2}\|y^{k+1} - y^{k}\|^2_{\Gamma_k} - \|y^{k+1} - y^k\|^2_{\Delta_k} \\
& + \left(\frac{1}{4} + \frac{1}{2}c\right)\beta\|By^{k+1} - By^k\|^2 + (1 - c)\frac{1}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 - 2\|u^{k+1} - u^*\|^2_{\Sigma_y} \\
= & \|w^k - w^*\|^2_{D_k} + \beta\|By^k - By^*\|^2 \\
& - \|w^{k+1} - w^k\|^2_{\Gamma_k} - \left(\frac{3}{4} - \frac{1}{2}c\right)\beta\|By^{k+1} - By^k\|^2 - \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 \\
& + \frac{1}{2}\|y^{k-1} - y^{k}\|^2_{\Gamma_{k-1}} - \frac{1}{2}\|y^{k+1} - y^k\|^2_{\Gamma_k} - \|y^{k+1} - y^k\|^2_{\Delta_k} - 2\|u^{k+1} - u^*\|^2_{\Sigma_y},
\end{align*}
$$

(2.29)
where the last equality follows from the definitions of $P_k$ and $D_k$ in (2.6). Rearranging (2.29) further, we have

\[
\|w^{k+1} - w^*\|_{D_k}^2 + \beta \|By^{k+1} - By^*\|^2 + \frac{1}{2}\|y^{k+1} - y^k\|_{\Gamma_k}^2 \geq \|y^k - w^*\|_{D_k}^2 + \beta \|By^k - By^*\|^2 + \frac{1}{2}\|y^{k+1} - y^k\|_{\Gamma_k}^2 - 2\|u^{k+1} - u^*\|^2_{\Sigma} - \left(\|u^{k+1} - u^k\|_{P_k}^2 + \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 + \|\gamma^{k+1} - \gamma^k\|_{A_k+(\frac{3}{4} - \frac{1}{2}c)\beta B^\top B}^2\right),
\]

that is,

\[
\|w^k - w^*\|^2_{D_k} + \beta \|By^k - By^*\|^2 + \frac{1}{2}\|y^{k-1} - y^k\|^2_{\Gamma_k} + \|u^k - u^*\|^2_{\Sigma} - \left(\|w^{k+1} - w^*\|^2_{D_k} + \beta \|By^{k+1} - By^*\|^2 + \frac{1}{2}\|y^{k+1} - y^k\|^2_{\Gamma_k} + \|u^{k+1} - u^*\|^2_{\Sigma}\right)
\geq \|u^{k+1} - u^k\|^2_{P_k} + \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 + \|\gamma^{k+1} - \gamma^k\|^2_{A_k+(\frac{3}{4} - \frac{1}{2}c)\beta B^\top B} + \|u^k - u^*\|^2_{\Sigma} + \|u^{k+1} - u^*\|^2_{\Sigma}
+ \|u^k - u^*\|^2_{\Sigma} - \|u^{k+1} - u^k\|^2_{S_k} + 2\|u^{k+1} - u^k\|^2_{\Sigma}.
\]

(2.30)

From the definition of $G_k$ in (2.6), inequality (2.30) can be written as

\[
\|u^k - u^*\|^2_{\Sigma} + \frac{1}{2}\|y^{k-1} - y^k\|^2_{\Gamma_k} - \left(\|u^{k+1} - u^*\|^2_{G_k} + \frac{1}{2}\|y^{k+1} - y^k\|^2_{\Gamma_k}\right)
\geq \|u^{k+1} - u^k\|^2_{P_k} + \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 + \|\gamma^{k+1} - \gamma^k\|^2_{A_k+(\frac{3}{4} - \frac{1}{2}c)\beta B^\top B} + \|u^k - u^*\|^2_{\Sigma} + \|u^{k+1} - u^*\|^2_{\Sigma}
\geq \|u^{k+1} - u^k\|^2_{P_k} + \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2 + \|\gamma^{k+1} - \gamma^k\|^2_{A_k+(\frac{3}{4} - \frac{1}{2}c)\beta B^\top B} + \|u^k - u^*\|^2_{\Sigma} + \|u^{k+1} - u^k\|^2_{\Sigma}
\geq \|x^{k+1} - x^k\|^2_{S_k} + \|y^{k+1} - y^k\|^2_{S_k} + \frac{c}{\beta}\|\lambda^{k+1} - \lambda^k\|^2_{S_k} + \frac{c}{\beta}\|\Sigma\|_{\Sigma} + \frac{1}{2}\|u^{k+1} - u^k\|^2_{\Sigma}.
\]

where the second inequality follows from the well-known inequality $\|a\|_M^2 + \|b\|_M^2 \geq \frac{1}{2}\|a - b\|_M^2$ with $M = \Sigma$, $a = u^k - u^*$ and $b = u^{k+1} - u^*$.

From the definitions (2.7b) and (2.7c), we have that

\[
\Delta_k = T_k + \frac{1}{2}\Sigma + \Lambda_k + \left(\frac{3}{4} - \frac{1}{2}c\right)\beta B^\top B.
\]

Thus the proof is completed. \qed

Condition (2.1) (a) implies $\|x^{k+1} - x^k\|^2_{\Sigma + \frac{1}{2}\Sigma f} \geq 0$ for all $k$. Moreover, Condition (2.1) (g) implies $\|y^{k+1} - y^k\|^2_{A_k} \geq 0$ for all $k$. Therefore, Term 1 in (2.27) is always nonnegative, which indicates the contraction of the sequence $\{w_k\}$.

It follows from the definition of $\{G_k\}$ and Condition (2.1) (a), (c) and (e) that $0 \leq G_{k+1} \leq (1 + \gamma_k)G_k$ for all $k$. We define two constants $C_s$ and $C_p$ as follows:

\[
C_s := \sum_{k=0}^{\infty} \gamma_k \quad \text{and} \quad C_p := \prod_{k=0}^{\infty} (1 + \gamma_k).
\]
From the assumption \( \sum_0^\infty \gamma_k < \infty \) and \( \gamma_k \geq 0 \), we have \( 0 \leq C_s < \infty \) and \( 1 \leq C_p < \infty \). Moreover, we can easily get
\[
0 \leq G_k \leq C_p G_0, \quad \forall k \geq 0,
\]
which means that the sequences \( \{G_k\} \) is bounded.

Now we give the main convergent theorem of this subsection.

**Theorem 2.9.** Let \( w^* = (x^*, y^*, \lambda^*) \in \Omega^* \), and let \( \{w^k\} \) be a sequence generated by (1.6). Suppose that \( \{T_k\} \) is a sequence satisfying Condition 2.7. Then the sequence \( \{w^k\} \) converges to a point \( w^* \in \Omega^* \).

**Proof.** First we show that the sequence \( \{w^k\} \) is bounded. Since \( 0 \leq G_{k+1} \leq (1 + \gamma_k)G_k \), we have
\[
||w^{k+1} - w^*||_{G_{k+1}}^2 \leq (1 + \gamma_k)||w^{k+1} - w^*||_{G_k}^2.
\]
Combining the inequality (2.31) with (2.27) in Lemma 2.8 we have
\[
||w^{k+1} - w^*||_{G_{k+1}}^2 + \frac{1}{2}||y^{k+1} - y^k||_{\Gamma_k}^2 \leq (1 + \gamma_k)\left(||w^k - w^*||_{G_k}^2 + \frac{1}{2}||y^k - y^k||_{\Gamma_{k-1}}^2\right) + \text{Term1}.
\]
It then follows that for all \( k \),
\[
||w^{k+1} - w^*||_{G_{k+1}}^2 + \frac{1}{2}||y^{k+1} - y^k||_{\Gamma_k}^2 \leq \left(\prod_{i=0}^{k}(1 + \gamma_i)\right)\left(||w^0 - w^*||_{G_0}^2 + \frac{1}{2}||y^0 - y^1||_{\Gamma_0}^2\right)
\]
\[
\leq C_p \left(||w^0 - w^*||_{G_0}^2 + \frac{1}{2}||y^0 - y^1||_{\Gamma_0}^2\right).
\]
Note that
\[
||w^{k+1} - w^*||_{G_{k+1}}^2 = ||x^{k+1} - x^*||_{S + \frac{1}{2} \Sigma_f}^2 + ||y^{k+1} - y^*||_{T_k + \Sigma_y + \beta B^T B}^2 + \frac{1}{\beta}||\lambda^{k+1} - \lambda^*||^2,
\]
\( T_k + \Sigma_y + \beta B^T B \) is positive definite from Condition 2.1(d), and \( C_p \left(||w^0 - w^*||_{G_0}^2 + \frac{1}{2}||y^0 - y^1||_{\Gamma_0}^2\right) \) is a constant. It then follows from (2.33) that \( \{y^k\} \) and \( \{\lambda^k\} \) are bounded. We now show that \( \{x^k\} \) is also bounded.

From (2.32) and (2.33), we have
\[
\text{Term1} = ||x^{k+1} - x^*||_{S + \frac{1}{2} \Sigma_f}^2 + ||y^{k+1} - y^k||_{\Gamma_k}^2 \leq ||x^k - w^*||_{G_k}^2 + \frac{1}{\beta}||\lambda^{k+1} - \lambda^*||^2
\]
\[
+ \gamma_k \left(||w^k - w^*||_{G_k}^2 + \frac{1}{2}||y^{k-1} - y^k||_{\Gamma_{k-1}}^2\right)
\]
\[
\leq ||w^k - w^*||_{G_k}^2 + \frac{1}{2}||y^{k-1} - y^k||_{\Gamma_{k-1}}^2 + \frac{1}{2}||y^{k-1} - y^k||_{\Gamma_{k-1}}^2 - \frac{1}{2}||y^{k-1} - y^k||_{\Gamma_{k-1}}^2
\]
\[
+ C_p \left(||w^0 - w^*||_{G_0}^2 + \frac{1}{2}||y^0 - y^1||_{\Gamma_0}^2\right).
\]
Summing up the inequalities, we obtain
\[
\sum_{k=1}^{\infty} \left( \|x^{k+1} - x^k\|_{S + \frac{1}{2} \Sigma_f}^2 + \|y^{k+1} - y^k\|_{\Delta_k}^2 + \frac{c}{\beta} \|\lambda^{k+1} - \lambda^k\|^2 \right) 
\leq \|w^0 - w^*\|_{\bar{G}_0}^2 + \frac{1}{2} \|y^0 - y^1\|_{\bar{F}_0}^2 + \left( \sum_{k=0}^{\infty} \gamma_k \right) \left( \|w^0 - w^*\|_{\bar{G}_0}^2 + \frac{1}{2} \|y^0 - y^1\|_{\bar{F}_0}^2 \right) 
\leq (1 + C_s C_p) \left( \|w^0 - w^*\|_{\bar{G}_0}^2 + \frac{1}{2} \|y^0 - y^1\|_{\bar{F}_0}^2 \right). 
\]

Since \((1 + C_s C_p) \left( \|w^0 - w^*\|_{\bar{G}_0}^2 + \frac{1}{2} \|y^0 - y^1\|_{\bar{F}_0}^2 \right)\) is a finite constant, we have
\[
\lim_{k \to \infty} \|x^{k+1} - x^k\|_{S + \frac{1}{2} \Sigma_f}^2 + \|y^{k+1} - y^k\|_{\Delta_k}^2 + \frac{c}{\beta} \|\lambda^{k+1} - \lambda^k\|^2 = 0, 
\]
which indicates that
\[
\lim_{k \to \infty} \|\lambda^{k+1} - \lambda^k\| = \lim_{k \to \infty} \beta \|Ax^{k+1} + By^{k+1} - b\| = 0. \tag{2.35} \]

Note that \(Ax^* + By^* - b = 0\), and
\[
\|Ax^{k+1} - Ax^*\| = \|Ax^{k+1} + By^{k+1} - b - B(y^{k+1} - y^k)\| \leq \|Ax^{k+1} + By^{k+1} - b\| + \|B(y^{k+1} - y^k)\|. 
\]

It then follows from (2.35) that \(\|A(x^{k+1} - x^*)\|\) is bounded. Moreover, inequalities (2.33) and (2.34) imply \(\|x^{k+1} - x^*\|_{S + \frac{1}{2} \Sigma_f}^2\) is bounded. Therefore \(\|x^{k+1} - x^*\|_{S + \frac{1}{2} \Sigma_f + \beta A^T A}^2\) is abounded since
\[
\|x^{k+1} - x^*\|_{S + \frac{1}{2} \Sigma_f + \beta A^T A}^2 = \|x^{k+1} - x^*\|_{S + \frac{1}{2} \Sigma_f}^2 + \beta \|A(x^{k+1} - x^*)\|^2. 
\]

From the positive definiteness of \(S + \frac{1}{2} \Sigma_f + \beta A^T A\) in Condition 2.1(b), it shows that \(\{x^k\}\) is also bounded. Consequently, the sequence \(\{w^k\}\) is bounded.

Next we should show that any cluster point of the sequence \(\{w^k\}\) is an optimal solution of (1.1) and the sequence \(\{w^k\}\) has only one cluster point. This can be done in a way similar to the proof of that in [25].

\section{VMIP-ADMM with the BFGS update}

As shown in the recent researches [25][26], a special variable metric proximal term via the BFGS update can get a solution faster on the iteration and CPU time than the proximal ADMM [18][29] with a fixed proximal matrix \(T\). Moreover, in their experiments, a slightly indefinite variable also performs well without the theoretical analysis. Note that this choice should have an assumption that the \(y\)-subproblems (1.6b) should be unconstrained quadratic programming problem. Based on the analysis above and the previous studies, we propose indefinite proximal terms \(\{T_k\}\) updated by the BFGS update, and show that \(\{T_k\}\) satisfies Condition 2.1.
3.1 Construction of the indefinite proximal matrix $T_k$ via the BFGS update

Inspired by the semidefinite proximal ADMM with the BFGS update [25, 26], we construct the indefinite matrix $T_k$ by the BFGS update.

We first explain the pure BFGS update for the following unconstrained quadratic optimization:

$$\min \frac{1}{2} x^\top M x,$$

where $M \in \mathbb{R}^{n \times n}$ is a positive definite matrix. Let $s \in \mathbb{R}^n$ and $l = Ms$. Note that $s^\top l > 0$ when $s \neq 0$. The BFGS update generates a sequence of approximate matrices $\{B_k\}$ of $M$, and its inverse $H_k = B_k^{-1}$.

For a given matrix $B_k$, the BFGS update generates $B_{k+1}^{\text{BFGS}}$ and $H_{k+1}^{\text{BFGS}}$ with $s$ and $l$ as follows

$$B_{k+1}^{\text{BFGS}} = B_k + \frac{ll^\top}{l^\top s} - \frac{B_k ss^\top B_k^\top}{s^\top B_k s},$$  \hspace{1cm} (3.1)

$$H_{k+1}^{\text{BFGS}} = (I - \frac{sl^\top}{s^\top l}) H_k (I - \frac{ls^\top}{s^\top l}) + \frac{ss^\top}{s^\top l}.$$  \hspace{1cm} (3.2)

Note that $B_{k+1}^{\text{BFGS}}$ and $H_{k+1}^{\text{BFGS}}$ are positive definite whenever $B_k, H_k > 0$ since $s^\top l > 0$. Note also that $H_{k+1}^{\text{BFGS}} l = s = M^{-1} l$.

We now explain how to construct $T_k$ via the BFGS update. Throughout this section we suppose that $g$ in the objective function (1.1) is a convex quadratic function. Then $y$-subproblems (1.6b) are unconstrained quadratic programming problems, and the Hessian matrix of the augmented Lagrangian function (1.2) is a constant matrix given as

$$M : = \nabla_{yy}^2 L_\beta(x, y, \lambda) = \tilde{M} + \beta B^\top B,$$

where $\tilde{M} : = \nabla_{yy}^2 g(y)$. Note that $M$ is always positive semidefinite since $\tilde{M} \succeq 0$.

We consider a perturbed matrix $M^\delta : = M + \delta I \succ 0$ with a sufficiently small $\delta > 0$, and construct an approximate matrix $B_k$ of $M^\delta$ via the BFGS update (3.1). Let $s_k = x^{k+1} - x^k$, where $\{x^k\}$ is a sequence generated by (1.6). We propose that $\{B_k\}$ is generated as

$$B_{k+1} = B_k + c_k \left( \frac{l_k l_k^\top}{s_k^\top s_k} - \frac{B_k s_k s_k^\top B_k^\top}{s_k^\top B_k s_k} \right),$$  \hspace{1cm} (3.3)

where $\tilde{l}_k = Ms_k + \delta s_k = M^\delta s_k$, and $\{c_k\}$ is a sequence such that $c_k \in [0, 1]$, and $\sum_{k=0}^\infty c_k < \infty$. We can rewrite the update formula (3.3) as

$$B_{k+1} = B_k + c_k (B_{k+1}^{\text{BFGS}} - B_k),$$

where $B_{k+1}^{\text{BFGS}}$ is updated by the pure BFGS update (3.1) with respect to $M^\delta$ at every iteration. Note that $B_{k+1} = B_{k+1}^{\text{BFGS}}$ when $c_k = 1$.

We then propose the following construction of $T_k$ via the BFGS update.
Lemma 2.1, there exist $c, s$ and $H$ if a given matrix.

Let $s_k 
eq 0$, then set $\bar{t}_k = M^\delta s_k$ and update $B_{k+1}$ via

$$B_{k+1} = B_k + c_k \left( \frac{\bar{t}_k l_k^	op}{l_k s_k} - B_k s_k s_k^	op B_k^	op \right);$$

Otherwise

$$B_{k+1} = B_k;$$

Construct $T_{k+1}$ as

$$T_{k+1} = B_{k+1} - M.$$

3.2 Discussion on the Condition 2.1 for the indefinite matrix $T_k$

We now consider matrices $\{T^k_+\}$ and $T_-$ such that $T_k = T^k_+ - T_-$, $T^k_+ \succeq 0$, $T_- \succeq 0$ in Condition 2.1 (c). Let

$$T^k_+ = B_k - \tau M \quad \text{and} \quad T_- = (1 - \tau)M, \quad \text{with} \quad \tau \in [0, 1).$$

Note that $T_k = T^k_+ - T_- = B_k - M$ and $T_- \succeq 0$. Thus we only show that $T^k_+$ is positive semidefinite.

To this end, we give an extension result related to Theorem 2.2 in [25].

**Lemma 3.1.** Let $M \in \mathbb{R}^{n \times n}$ be a positive definite matrix. Let $s \in \mathbb{R}^n$ such that $s \neq 0$, and let $l = Ms$.

If a given matrix $H_k \in \mathbb{R}^{n \times n}$ satisfies $H_k \preceq \tau_1 M^{-1}$ with $\tau_1 \geq 1$, then $H_k^{BFGS}$ which is generated by the BFGS update (3.2) with respect to $M$ also satisfies $H_k^{BFGS} \preceq \tau_1 M^{-1}$.

**Proof.** Let $v$ be an arbitrary nonzero vector in $\mathbb{R}^n$, and $\Psi = \{ z \in \mathbb{R}^n \mid s^\top z = 0 \}$. As shown in [25, Lemma 2.1], there exist $c \in \mathbb{R}$ and $z \in \Psi$ such that $v = cl + z$. Together with $H_{k+1}^{BFGS}l = s = M^{-1}l$ and $s^\top z = 0$, we can obtain that for any $\tau_1 \geq 1$,

$$v^\top H_{k+1}^{BFGS} v = (cl + z)^\top H_{k+1}^{BFGS} (cl + z) = c^2 l^\top s + 2c s^\top z + z^\top H_{k+1}^{BFGS} z = c^2 l^\top M^{-1} l + z^\top H_{k+1}^{BFGS} z = c^2 l^\top M^{-1} l + z^\top H_k z - 2z^\top \left( \frac{sl^\top}{s^\top l} H_k \right) z + z^\top \left( \frac{sl^\top}{s^\top l} H_k \frac{ls^\top}{s^\top l} \right) z + \frac{z^\top s s^\top z}{s^\top l}$$

where the forth equality follows from (3.2), and the inequality follows from the positive definiteness of $M^{-1}$ and the assumption that $H_k \preceq \tau_1 M^{-1}$. Since $v$ is arbitrary, we have $H_{k+1}^{BFGS} \preceq \tau_1 M^{-1}$. □
Lemma 3.1 implies that $B_{k+1}^{\text{BFGS}} \succeq \tau M^\delta$ when $B_k \succeq \tau M^\delta$ with $\tau = \frac{1}{\tau_1} \leq 1$, and hence

$$B_{k+1} = (1 - c_k)B_k + c_kB_{k+1}^{\text{BFGS}} \succeq \tau M^\delta.$$ 

That is, if $B_0 \succeq \tau M^\delta$ and $\tau \leq 1$, we have $B_k \succeq \tau M^\delta$ for all $k$, and hence $T_+^k \succeq 0$ for all $k$. When $\tau = 1$, it is reduced to the variable metric semi-proximal ADMM in [25].

For instance, we can choose the initial matrix $B_0$ as

$$B_0 = \xi I, \text{ with } \xi = \tau \lambda_{\max}(M^\delta), \text{ } \tau \in (0, 1).$$

It is easy to see that $B_0 \succeq \tau M^\delta$.

Next we show that the $T_k$, $T_+^k$ and $T_-$ satisfy Condition 2.1 (d)-(g). We suppose that $B_0 \succeq \tau M^\delta$ and $\tau \in \left(\frac{3}{4}, 1\right)$.

First we show Condition 2.1 (e). Note that $s_k^\top B_k s_k \geq \tau s_k^\top M^\delta s_k \geq \tau \delta s_k^\top M^\delta s_k$ and $M$ is the constant matrix. Therefore, we can suppose that $\|B_{k+1}^{\text{BFGS}} - B_k\|$ is bounded above by some constant $Q > 0$, that is, $-QI \preceq B_{k+1}^{\text{BFGS}} - B_k \preceq QI$. Moreover, $T_+^k = B_k - \tau M \succeq \tau M^\delta - \tau M \succeq \tau \delta I$. Then we can obtain that

$$T_+^{k+1} = B_{k+1} - \tau M = B_k + c_k (B_{k+1}^{\text{BFGS}} - B_k) - \tau M$$

$$= T_+^k + c_k (B_{k+1}^{\text{BFGS}} - B_k)$$

$$\preceq T_+^k + \frac{c_k Q}{\tau \delta} \tau \delta I$$

$$= T_+^k + \frac{c_k Q}{\tau \delta} T_+^k$$

$$= \left(1 + \frac{c_k Q}{\tau \delta}\right) T_+^k.$$ 

On the other hand, we have

$$T_+^k = T_+^{k+1} - c_k (B_{k+1}^{\text{BFGS}} - B_k)$$

$$\preceq T_+^{k+1} + \frac{c_k Q}{\tau \delta} \tau \delta I$$

$$\preceq T_+^{k+1} + \frac{c_k Q}{\tau \delta} T_+^{k+1}$$

$$= \left(1 + \frac{c_k Q}{\tau \delta}\right) T_+^{k+1}.$$ 

Let $\gamma_k = \frac{Q}{\tau \delta} c_k$. Then we have

$$\frac{1}{1 + \gamma_k} T_+^k \preceq T_+^{k+1} \preceq (1 + \gamma_k) T_+^k \text{ for all } k. \quad (3.4)$$

Note that $\tilde{M} = \nabla^2_{yy} g(y) = \Sigma_g$. Then $T_k + \Sigma_g + \beta B^\top B = B_k - M + \Sigma_g + \beta B^\top B = B_k > 0$ which shows that Condition (d) holds.

Next we show Condition (f). Since (3.4) implies that $B_{k+1} - \tau M = T_+^{k+1} \preceq (1 + \gamma_k) T_+^k = (1 + \gamma_k) (B_k - \tau M)$ and $M$ is positive semidefinite, we have

$$B_{k+1} \preceq (1 + \gamma_k) B_k - \gamma_k \tau M \preceq (1 + \gamma_k) B_k.$$
Obviously,

\[ T_{k+1} + \Sigma_g + \beta B^\top B = B_{k+1} \preceq (1 + \gamma_k)B_k = (1 + \gamma_k)(T_k + \Sigma_g + \beta B^\top B). \]

Finally, we show Condition (g). From the definition of \( M \), we have

\[
T_k + \frac{3}{2} \Sigma_g - \frac{\gamma_k}{2} T_k + 2T_+ + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \\
= B_k - M + \frac{3}{2} M - \frac{\gamma_k}{2} (B_k - \tau M) - 2(1 - \tau) M + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \\
= \left( 1 - \frac{\gamma_k}{2} \right) B_k + \frac{3}{2} M - M + \frac{\gamma_k}{2} \tau M - 2(1 - \tau) M + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \\
\succeq \left( 1 - \frac{\gamma_k}{2} \right) \tau M + \frac{3}{2} M - M + \frac{\gamma_k}{2} \tau M - 2(1 - \tau) M + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \\
= (3\tau - 3)(\bar{M} + \beta B^\top B) + \frac{3}{2} \bar{M} + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \\
= (3\tau - 3)\bar{M} + (3\tau - \frac{9}{4} - \frac{1}{2} c) \beta B^\top B,
\]

where the matrix inequality follows from \( B_k \succeq \tau M^\delta = \tau M + \tau \delta I \succeq \tau M \). Note that there exist \( k \) such that \( \gamma_k \leq 1 \) for all \( k \geq k \). Without loss of generality, we assume \( k = 0 \) and thus \( \left( 1 - \frac{\gamma_k}{2} \right) \geq 0 \) for all \( k \).

Let \( c = 2(\tau - \frac{3}{4}) \). It is easy to see that \( c \in (0, \frac{1}{4}) \). Moreover, \( 3\tau - \frac{3}{4} > 0 \) and \( 3\tau - \frac{9}{4} - \frac{1}{2} c = 2\tau - \frac{3}{2} > 0 \).

As a conclusion of the above discussion, the indefinite proximal term \( T_k \) generated via the BFGS update can satisfy Condition [2.1]. Obviously, the VMIP-ADMM can cover the general indefinite proximal ADMM as the following remark.

**Remark 3.2.** When \( \{T_k\} \) be a constant sequence for all \( k \), that is, \( T_k = T \), then we can write \( T = T_+ - T_- \), where \( T_+, T_- \succeq 0 \). It is easy to check that the boundness Condition (e) and (f) immediately hold when \( \gamma_k \equiv 0 \). Let \( T_+ = \tau (r I - \beta B^\top B) \succeq 0 \) and \( T_- = (1 - \tau) \beta B^\top B \succeq 0 \), we choose

\[ T = T_+ - T_- = \tau r I - \beta B^\top B, \text{ with } r > \beta \|B^\top B\|. \]

**Condition (d) holds.** For \( \tau \in (0.75, 1) \), taking \( c = 2(\tau - \frac{3}{4}) \), then Condition (g) turns to be

\[
T + \frac{3}{2} \Sigma_g - 2T_+ + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \succeq \tau \beta B^\top B - \beta B^\top B - 2(1 - \tau) \beta B^\top B + \left( \frac{3}{4} - \frac{1}{2} c \right) \beta B^\top B \\
= (3\tau - \frac{9}{4} - \frac{1}{2} c) \beta B^\top B \\
\succeq 0.
\]

It is reduced to the indefinite proximal ADMM in [22].

### 4 Conclusions

In this paper, we proposed a variable metric indefinite proximal ADMM whose indefinite proximal term can be chosen differently at every iterative step. We proved the global convergence of the proposed method under some requirements by applying an analysis technique in [24]. Moreover, for a special
problem whose $y$-subproblems are unconstrained quadratic programming problem, we proposed to construct the indefinite term $T_k$ via the BFGS update. We showed that such construction can satisfy the general convergent conditions.

Note that a strictly contractive version of the original ADMM which is known as the Peaceman-Rachford splitting method (PRSM) sometimes performs better in numerical experiments with some penalty parameters [28]. An indefinite proximal version of the PRSM also has been studied by many researchers [21, 32]. A further extension is to consider the variable metric indefinite term for PRSM. We leave this topic as one of our future work.

On the other hand, how to choose an adjusted proximal term is important to design a more efficient algorithm. The BFGS update provides better performance for some special problems whose $y$-subproblem is quadratic problem. It is worth developing some efficient proximal term for a general nonlinear sub-problem.

References

[1] H. Attouch, L. M. Briceno-Arias, and P. L. Combettes, A parallel splitting method for coupled monotone inclusions, SIAM Journal on Control and Optimization, 48 (2010), pp. 3246–3270.

[2] S. Banert, R. I. Bot, and E. R. CsetneK, Fixing and extending some recent results on the ADMM algorithm, arXiv preprint arXiv:1612.05057, (2016).

[3] H. H. Bauschke, P. L. Combettes, et al., Convex analysis and monotone operator theory in Hilbert spaces, vol. 408, Springer, 2011.

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends® in Machine Learning, 3 (2011), pp. 1–122.

[5] A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of mathematical imaging and vision, 40 (2011), pp. 120–145.

[6] G. Chen and M. Teboulle, A proximal-based decomposition method for convex minimization problems, Mathematical Programming, 64 (1994), pp. 81–101.

[7] P. Chen, J. Huang, and X. Zhang, A primal–dual fixed point algorithm for convex separable minimization with applications to image restoration, Inverse Problems, 29 (2013), p. 025011.

[8] P. L. Combettes, Iterative construction of the resolvent of a sum of maximal monotone operators, J. Convex Anal, 16 (2009), pp. 727–748.

[9] W. Deng and W. Yin, On the global and linear convergence of the generalized alternating direction method of multipliers, tech. rep., DTIC Document, 2012.

[10] On the global and linear convergence of the generalized alternating direction method of multipliers, Journal of Scientific Computing, 66 (2016), pp. 889–916.
[11] J. Douglas and H. Rachford, *On the numerical solution of heat conduction problems in two and three space variables*, Transactions of the American Mathematical Society, (1956), pp. 421–439.

[12] J. Eckstein, *Some saddle-function splitting methods for convex programming*, Optimization Methods and Software, 4 (1994), pp. 75–83.

[13] J. Eckstein and D. P. Bertsekas, *On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, 55 (1992), pp. 293–318.

[14] J. Eckstein and W. Yao, *Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives*, Pac. J. Optim., 11 (2015), pp. 619–644.

[15] ———, *Approximate ADMM algorithms derived from Lagrangian splitting*, Computational Optimization and Applications, 68 (2017), pp. 363–405.

[16] ———, *Relative-error approximate versions of Douglas–Rachford splitting and special cases of the ADMM*, Mathematical Programming, 170 (2018), pp. 417–444.

[17] E. Esser, X. Zhang, and T. Chan, *A general framework for a class of first order primal-dual algorithms for tv minimization*, UCLA Cam Report, (2009), pp. 09–67.

[18] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, *Hankel matrix rank minimization with applications to system identification and realization*, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 946–977.

[19] M. Fortin and R. Glowinski, *Chapter iii on decomposition-coordination methods using an augmented lagrangian*, in *Studies in Mathematics and Its Applications*, vol. 15, Elsevier, 1983, pp. 97–146.

[20] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computers & Mathematics with Applications, 2 (1976), pp. 17–40.

[21] B. Gao and F. Ma, *Symmetric alternating direction method with indefinite proximal regularization for linearly constrained convex optimization*, Journal of Optimization Theory and Applications, 176 (2018), pp. 178–204.

[22] R. Glowinski and A. Marroco, *Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de dirichlet non linéaires*, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 9 (1975), pp. 41–76.

[23] M. L. N. Gonçalves, M. M. Alves, and J. G. Melo, *Pointwise and ergodic convergence rates of a variable metric proximal alternating direction method of multipliers*, Journal of Optimization Theory and Applications, 177 (2018), pp. 448–478.
[24] Y. Gu, B. Jiang, and D. Han, *An indefinite-proximal-based strictly contractive Peaceman-Rachford splitting method*, tech. rep., 2019.

[25] Y. Gu and N. Yamashita, *An alternating direction method of multipliers with the BFGS update for structured convex quadratic optimization*, arXiv e-prints arXiv:1903.02270, (2019).

[26] Y. Gu and N. Yamashita, *A proximal ADMM with the Broyden family for convex optimization problems*, Available on http://www. optimization-online. org, (2019).

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

[28] B. He, H. Liu, Z. Wang, and X. Yuan, *A strictly contractive Peaceman–Rachford splitting method for convex programming*, SIAM Journal on Optimization, 24 (2014), pp. 1011–1040.

[29] B. He, F. Ma, and X. Yuan, *Optimal linearized alternating direction method of multipliers for convex programming*, Available on http://www. optimization-online. org, (2017).

[30] B. He and X. Yuan, *On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 700–709.

[31] ———, *Block-wise alternating direction method of multipliers for multiple-block convex programming and beyond*, SMAI Journal of Computational Mathematics, 1 (2015), pp. 145–174.

[32] F. Jiang, Z. Wu, and X. Cai, *Generalized admm with optimal indefinite proximal term for linearly constrained convex optimization*, Journal of Industrial & Management Optimization, (2018), pp. 183–202.

[33] K. Koh, S.-J. Kim, and S. Boyd, *An interior-point method for large-scale $l_1$-regularized logistic regression*, Journal of Machine learning research, 8 (2007), pp. 1519–1555.

[34] M. Li, D. Sun, and K.-C. Toh, *A majorized admm with indefinite proximal terms for linearly constrained convex composite optimization*, SIAM Journal on Optimization, 26 (2016), pp. 922–950.

[35] P.-L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM Journal on Numerical Analysis, 16 (1979), pp. 964–979.

[36] P. A. Lotito, L. A. Parente, and M. Solodov, *A class of variable metric decomposition methods for monotone variational inclusions*, J. Convex Anal, 16 (2009), pp. 857–880.

[37] Y. Nesterov, *Gradient methods for minimizing composite functions*, Mathematical Programming, 140 (2013), pp. 125–161.

[38] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: nonlinear phenomena, 60 (1992), pp. 259–268.

[39] H. Trevor, T. Robert, and F. JH, *The elements of statistical learning: data mining, inference, and prediction*, 2009.
[40] P. Tseng, Approximation accuracy, gradient methods, and error bound for structured convex optimization, Mathematical Programming, 125 (2010), pp. 263–295.

[41] P. Tseng and S. Yun, A coordinate gradient descent method for nonsmooth separable minimization, Mathematical Programming, 117 (2009), pp. 387–423.

[42] M. Xu and T. Wu, A class of linearized proximal alternating direction methods, Journal of Optimization Theory and Applications, 151 (2011), pp. 321–337.

[43] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, Bregman iterative algorithms for $l_1$-minimization with applications to compressed sensing, SIAM Journal on Imaging sciences, 1 (2008), pp. 143–168.

[44] X. Yuan, The improvement with relative errors of he et al.'s inexact alternating direction method for monotone variational inequalities, Mathematical and computer modelling, 42 (2005), pp. 1225–1236.