

Research Article

Sensitivity Analysis of the Proximal-Based Parallel Decomposition Methods

Feng Ma, Mingfang Ni, Lei Zhu, and Zhanke Yu

College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China

Correspondence should be addressed to Feng Ma; mafengnju@gmail.com

Received 30 June 2013; Revised 25 November 2013; Accepted 27 November 2013; Published 20 January 2014

Academic Editor: Dongdong Ge

Copyright © 2014 Feng Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The proximal-based parallel decomposition methods were recently proposed to solve structured convex optimization problems. These algorithms are eligible for parallel computation and can be used efficiently for solving large-scale separable problems. In this paper, compared with the previous theoretical results, we show that the range of the involved parameters can be enlarged while the convergence can be still established. Preliminary numerical tests on stable principal component pursuit problem testify to the advantages of the enlargement.

1. Introduction

Consider the constrained convex optimization problems with separable objective functions in the following form:

$$\begin{aligned} \min_{x \in \mathcal{X}, y \in \mathcal{Y}} \quad & f(x) + g(y) \\ \text{s.t.} \quad & Ax + By = b, \end{aligned} \quad (1)$$

where $A \in \mathfrak{R}^{m \times n}$, $B \in \mathfrak{R}^{m \times p}$, $b \in \mathfrak{R}^m$, $\mathcal{X} \subset \mathfrak{R}^n$, and $\mathcal{Y} \subset \mathfrak{R}^p$, are two nonempty, closed, and convex sets and $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$, $g : \mathfrak{R}^p \rightarrow \mathfrak{R}$ are convex functions. Problem of this type arises from a number of fields such as signal processing, compressed sensing, machine learning, and semidefinite programming (see, e.g., [1–7] and references cited therein).

To solve (1), the classical alternating direction method generates the new iterate via the following scheme:

$$\begin{aligned} x^{k+1} = \arg \min_{x \in \mathcal{X}} \quad & f(x) + g(y^k) \\ & + \frac{\beta}{2} \left\| Ax + By^k - b - \frac{1}{\beta} \lambda^k \right\|^2, \end{aligned}$$

$$\begin{aligned} y^{k+1} = \arg \min_{y \in \mathcal{Y}} \quad & f(x^{k+1}) + g(y) \\ & + \frac{\beta}{2} \left\| Ax^{k+1} + By - b - \frac{1}{\beta} \lambda^k \right\|^2, \\ \lambda^{k+1} = \lambda^k - \beta \quad & (Ax^{k+1} + By^{k+1} - b), \end{aligned} \quad (2)$$

where $\lambda \in \mathfrak{R}^m$ is the Lagrange multiplier associated with the linear constraint and $\beta > 0$ is a penalty parameter for the violation of the linear constraint.

At each iteration, ADM essentially splits the subproblem of the augmented Lagrangian method into two subproblems in Gauss-Seidel fashion. The subproblems can be solved in consecutive order, which makes ADM possible to exploit the individual structure of f and g . The decomposed subproblems in (2) are often easy when A and B in (1) are both identity matrices and the resolvent operators of f and g have closed-form solutions or can be efficiently solved up to a high precision. Here, the resolvent operator of a function (say, $\theta : \mathfrak{R}^n \rightarrow \mathfrak{R}$) is defined by

$$\left(I + \frac{1}{\beta} \partial \theta \right)^{-1} (v) = \arg \min_{z \in \mathcal{Z}} \theta(z) + \frac{\beta}{2} \|z - v\|^2, \quad (3)$$

where $v \in \mathcal{Z}$ and $\beta > 0$. However, in some cases, both A and B are not identity matrices; the two subproblems in ADM (2) are difficult to solve because the evaluation of the following minimization style

$$\left(M^T M + \frac{1}{\beta} \partial \theta\right)^{-1} (Mv) = \arg \min_{z \in \mathcal{Z}} \theta(z) + \frac{\beta}{2} \|Mz - v\|^2 \quad (4)$$

could be costly, where M is a given nonidentity matrix, for example, A or B .

For the purpose of parallel and easy computing, the first parallel decomposition method [8] (abbreviated as FPDM) generates the new iterative as follows:

$$\begin{aligned} x^{k+1} &= \arg \min_{x \in \mathcal{X}} f(x) \\ &+ \beta (x - x^k)^T \left(A^T \left(Ax^k + By^k - b - \frac{1}{\beta} \lambda^k \right) \right) \\ &+ \frac{r}{2} \|x - x^k\|^2, \\ y^{k+1} &= \arg \min_{y \in \mathcal{Y}} g(y) \\ &+ \beta (y - y^k)^T \left(B^T \left(Ax^k + By^k - b - \frac{1}{\beta} \lambda^k \right) \right) \\ &+ \frac{s}{2} \|y - y^k\|^2, \\ \lambda^{k+1} &= \lambda^k - \beta (Ax^{k+1} + By^{k+1} - b), \end{aligned} \quad (5)$$

where the parameters r, s are required to satisfy $r > 2\beta \|A^T A\|$ and $s > 2\beta \|B^T B\|$. Here, $\|C\|$ denotes the largest eigenvalue of matrix C . It is easy to verify that the proximal-based decomposition method proposed in [9] is a special case of the FPDM.

When (4) is easy to evaluate for f and g , the second parallel decomposition method [8] (abbreviated as SPDM) can be used, which generates the new iterative as follows:

$$\begin{aligned} x^{k+1} &= \arg \min_{x \in \mathcal{X}} f(x) + \frac{\beta}{2} \|Ax + By^k - b - \frac{1}{\beta} \lambda^k\|^2 \\ &+ \frac{r}{2} \|x - x^k\|^2, \\ y^{k+1} &= \arg \min_{y \in \mathcal{Y}} g(y) + \frac{\beta}{2} \|Ax^k + By - b - \frac{1}{\beta} \lambda^k\|^2 \\ &+ \frac{s}{2} \|y - y^k\|^2, \\ \lambda^{k+1} &= \lambda^k - \beta (Ax^{k+1} + By^{k+1} - b), \end{aligned} \quad (6)$$

where the parameters r, s are required to satisfy $r > \beta \|A^T A\|$ and $s > \beta \|B^T B\|$.

Note that the subproblems in FPDM and SPDM can be processed in a parallelized fashion because the first subproblem involving x is independent on the second subproblem

involving y . Thus, FPDM and SPDM are suitable for solving large-scale distributed machine learning and big-data-related optimization problems.

ADM was first described in [10] and is closely related to many other algorithms, such as augmented Lagrangian methods, proximal point algorithm [11], and split Bregman methods [12]. Recently, the convergence of ADM has been analyzed under certain assumptions (see e.g., [13–16]) and the direct extension of ADM for multiblock convex minimization problems has already been proved not necessarily convergent [17].

In this paper, we study the proximal-based parallel decomposition methods from the perspective of variational inequalities. We show that the requirement ranges of the parameters r, s , and β can be significantly enlarged. Our contributions are as follows.

- (i) For the FPDM, we show that the requirements of the step sizes r, s , and β can be uniformly relaxed by

$$\frac{\beta \|A^T A\|}{r} + \frac{\beta \|B^T B\|}{s} \in (0, 1). \quad (7)$$

- (ii) For the SPDM, we show that the requirements of the step sizes r, s , and β can be uniformly relaxed by

$$\frac{\beta \|A^T A\|}{r + \beta \|A^T A\|} + \frac{\beta \|B^T B\|}{s + \beta \|B^T B\|} \in (0, 1). \quad (8)$$

- (iii) We provide a new application example in machine learning, that is, stable principal component pursuit problem. Preliminary numerical experiments testify to the advantages of the enlargement.

The rest of this paper is organized as follows. In Section 2 we derive a variational reformulation of (1) and summarize some preliminaries of variational inequalities. In Section 3, we describe our main theoretical results and analyze their convergence. We report some numerical results in Section 4 and make some conclusions in Section 5.

2. Preliminaries

2.1. Variational Inequality Characterization. In this section, we derive a variational reformulation of (1) which will be used in subsequent analysis.

Since the functions f and g are all assumed to be convex, by invoking the first-order optimality condition for (1), we can easily verify that solving (1) amounts to finding a vector $\mu^* \in \Omega$ of the variational inequality (VI):

$$(\mu' - \mu^*)^T F(\mu^*) \geq 0, \quad \forall \mu' \in \Omega, \quad (9)$$

with

$$\mu := \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad F(\mu) := \begin{pmatrix} \partial f(x) - A^T \lambda \\ \partial g(y) - B^T \lambda \\ Ax + By - b \end{pmatrix}, \quad (10)$$

where

$$\Omega = \mathcal{X} \times \mathcal{Y} \times \mathfrak{R}^m. \quad (11)$$

The problem (9) is referred to as a structured variational inequality (SVI) and has been studied extensively both in the theoretical frameworks and applications. Recently, He et al. [18, 19] proposed a unified framework of proximal-like contraction methods for monotone VI. They also construct the $O(1/t)$ convergence rate of the projection and contraction methods for VI with Lipschitz continuous monotone operators [20]. Xu et al. [21] proposed two classes of correction methods for the SVI in which the mapping F does not have an explicit form. Yuan and Li [22] developed a logarithmic-quadratic proximal (LQP) based decomposition method by applying the LQP terms to regularize the ADM subproblems. Tao and Yuan [23] established the $O(1/t)$ convergence rate of ADM with LQP regularization. Bnouhachem et al. [24] studied a new inexact LQP alternating direction method by solving a series of related systems of nonlinear equations.

2.2. Some Properties of Variational Inequalities. In this section, we summarize some basic knowledge and related definitions of variational inequalities.

Let G be a symmetric positive definite matrix; the G -norm of the vector μ is denoted by $\|\mu\|_G := \sqrt{\langle \mu, G\mu \rangle}$. In particular, when $G = I$, $\|\mu\| := \sqrt{\langle \mu, \mu \rangle}$ is the Euclidean norm of μ . Let $P_{\Omega,G}(\cdot)$ be the projection operator onto Ω under the G -norm; that is,

$$P_{\Omega,G}(\nu) = \arg \min \{ \|\mu - \nu\|_G \mid \mu \in \Omega \}. \quad (12)$$

From the above definition, we have the following well-known properties:

$$\begin{aligned} (\mu' - P_{\Omega,G}(\mu'))^T G(\mu - P_{\Omega,G}(\mu')) &\leq 0 \\ \forall \mu' \in \mathfrak{R}^l, \quad \forall \mu \in \Omega, \\ \|P_{\Omega,G}(\mu) - P_{\Omega,G}(\mu')\|_G &\leq \|\mu - \mu'\|_G, \quad \forall \mu, \mu' \in \mathfrak{R}^l, \\ \|\mu - P_{\Omega,G}(\mu')\|_G^2 &\leq \|\mu - \mu'\|_G^2 - \|\mu' - P_{\Omega,G}(\mu')\|_G^2, \\ \forall \mu' \in \mathfrak{R}^l, \quad \forall \mu \in \Omega. \end{aligned} \quad (13)$$

The mapping F is said to be monotone with respect to Ω if

$$(\mu - \nu)^T (F(\mu) - F(\nu)) \geq 0, \quad \forall \mu, \nu \in \Omega. \quad (14)$$

The following lemma [25, page 267] states an important result which characterizes a VI by a projection equation.

Lemma 1. *Let Ω be a closed convex set in \mathfrak{R}^l and let G be any positive definite matrix; then μ^* is a solution of VI(Ω, F) if and only if it satisfies*

$$\mu^* = P_{\Omega,G}(\mu^* - \alpha G^{-1} F(\mu^*)), \quad \forall \alpha > 0. \quad (15)$$

3. Theoretical Results of the Relaxation

In this section, we show that the range of the parameters r, s , and β can be enlarged in FPDM and SPDM, which is broader than the previous theoretical results. We also establish the global convergence of FPDM and SPDM under the new conditions for the parameters.

3.1. The Parameters Relaxation of the FPDM. From (5), the subproblems of FPDM can be, respectively, characterized by the following VI form: find $\mu^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})^T \in \Omega$ such that

$$\begin{aligned} (\mu' - \mu^{k+1})^T \{ F(\mu^{k+1}) + \eta_I(\mu^{k+1}, \mu^k) - G(\mu^k - \mu^{k+1}) \} &\geq 0, \\ \forall \mu' \in \Omega, \end{aligned} \quad (16)$$

with

$$\begin{aligned} \eta_I(\mu^{k+1}, \mu^k) &= \beta \begin{pmatrix} A^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \\ B^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \\ 0 \end{pmatrix}, \\ G_I &= \begin{pmatrix} rI_n & & \\ & sI_p & \\ & & \beta^{-1}I_m \end{pmatrix}, \end{aligned} \quad (17)$$

where $\eta_I(\mu^{k+1}, \mu^k) \in \mathfrak{R}^{(n+p+m)}$ and $G_I \in \mathfrak{R}^{(n+p+m) \times (n+p+m)}$ is a positive definite matrix.

Lemma 2. *For a given $\mu^k = (x^k, y^k, \lambda^k)^T$, let $\mu^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})^T$ be generated by (16) and (17). If*

$$u_I = \sqrt{\frac{\beta \|A^T A\|}{r} + \frac{\beta \|B^T B\|}{s}} \in (0, 1), \quad (18)$$

then for any $\mu^* = (x^*, y^*, \lambda^*)^T$, one has

$$(\mu^{k+1} - \mu^*)^T \eta_I(\mu^{k+1}, \mu^k) \leq \frac{u_I}{2} \|\mu^k - \mu^{k+1}\|_{G_I}^2. \quad (19)$$

Proof. Since $Ax^* + By^* = b$ and $\beta(Ax^{k+1} + By^{k+1} - b) = \lambda^k - \lambda^{k+1}$ one has

$$\begin{aligned} &(\mu^{k+1} - \mu^*)^T \eta_I(\mu^{k+1}, \mu^k) \\ &= \beta \begin{pmatrix} x^{k+1} - x^* \\ y^{k+1} - y^* \\ \lambda^{k+1} - \lambda^* \end{pmatrix}^T \\ &\quad \times \begin{pmatrix} A^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \\ B^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \\ 0 \end{pmatrix} \\ &= (\lambda^k - \lambda^{k+1})^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})). \end{aligned} \quad (20)$$

Applying the Cauchy-Schwarz inequality to the right term, we get

$$\begin{aligned}
& (\mu^{k+1} - \mu^*)^T \eta_I(\mu^{k+1}, \mu^k) \\
& \leq \left| (\lambda^k - \lambda^{k+1})^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \right| \\
& \leq \frac{1}{2} \left\{ 2 \left| (\lambda^k - \lambda^{k+1})^T A(x^k - x^{k+1}) \right| \right\} \\
& \quad + \frac{1}{2} \left\{ 2 \left| (\lambda^k - \lambda^{k+1})^T B(y^k - y^{k+1}) \right| \right\} \\
& \leq \frac{1}{2} \left\{ \frac{\|A^T A\|}{u_I r} |\lambda^k - \lambda^{k+1}|^2 + \frac{u_I r}{\|A^T A\|} |A(x^k - x^{k+1})|^2 \right\} \\
& \quad + \frac{1}{2} \left\{ \frac{\|B^T B\|}{u_I s} |\lambda^k - \lambda^{k+1}|^2 + \frac{u_I s}{\|B^T B\|} |B(y^k - y^{k+1})|^2 \right\} \\
& = \frac{1}{2} \left\{ \left(\frac{\|A^T A\|}{u_I r} + \frac{\|B^T B\|}{u_I s} \right) |\lambda^k - \lambda^{k+1}|^2 \right. \\
& \quad \left. + \frac{u_I r}{\|A^T A\|} |A(x^k - x^{k+1})|^2 + \frac{u_I s}{\|B^T B\|} |B(y^k - y^{k+1})|^2 \right\} \\
& \leq \frac{1}{2} \left\{ \left(\frac{\|A^T A\|}{u_I r} + \frac{\|B^T B\|}{u_I s} \right) |\lambda^k - \lambda^{k+1}|^2 + u_I r |x^k - x^{k+1}|^2 \right. \\
& \quad \left. + u_I s |y^k - y^{k+1}|^2 \right\} \\
& = \frac{1}{2} \left\{ \frac{u_I}{\beta} |\lambda^k - \lambda^{k+1}|^2 + u_I r |x^k - x^{k+1}|^2 + u_I s |y^k - y^{k+1}|^2 \right\} \\
& = \frac{u_I}{2} \|\mu^k - \mu^{k+1}\|_{G_I}^2.
\end{aligned} \tag{21}$$

Theorem 3. Let the sequence $\{\mu^k\}$ be generated by FPD (16) and (17). If

$$u_I = \sqrt{\frac{\beta \|A^T A\|}{r} + \frac{\beta \|B^T B\|}{s}} \in (0, 1), \tag{22}$$

then one has

$$\begin{aligned}
& \|\mu^{k+1} - \mu^*\|_{G_I}^2 \\
& \leq \|\mu^k - \mu^*\|_{G_I}^2 - (1 - u_I) \|\mu^k - \mu^{k+1}\|_{G_I}^2, \quad \forall \mu^* \in \Omega^*.
\end{aligned} \tag{23}$$

Proof. Consider

$$\begin{aligned}
\|\mu^{k+1} - \mu^*\|_{G_I}^2 &= \|\mu^{k+1} - \mu^k + \mu^k - \mu^*\|_{G_I}^2 \\
&= \|\mu^k - \mu^*\|_{G_I}^2 - 2(\mu^k - \mu^*)^T G_I (\mu^k - \mu^{k+1}) \\
& \quad + \|\mu^{k+1} - \mu^k\|_{G_I}^2.
\end{aligned} \tag{24}$$

On the other hand, by setting $\mu' = \mu^*$ in (16), we have

$$\begin{aligned}
(\mu^{k+1} - \mu^*)^T G (\mu^k - \mu^{k+1}) &\geq (\mu^{k+1} - \mu^*)^T F (\mu^{k+1}) \\
& \quad + (\mu^{k+1} - \mu^*) \eta_I (\mu^{k+1}, \mu^k).
\end{aligned} \tag{25}$$

Using the fact that F is a monotone operator, we have

$$(\mu^{k+1} - \mu^*)^T F (\mu^{k+1}) \geq (\mu^{k+1} - \mu^*)^T F (\mu^*) \geq 0. \tag{26}$$

With rearrangement of the term (26) and using (27), we derive that

$$\begin{aligned}
(\mu^k - \mu^*)^T G (\mu^k - \mu^{k+1}) &\geq (\mu^k - \mu^{k+1})^T G (\mu^k - \mu^{k+1}) \\
& \quad + (\mu^{k+1} - \mu^*) \eta_I (\mu^{k+1}, \mu^k).
\end{aligned} \tag{27}$$

Substituting (28) into (24), we get

$$\begin{aligned}
\|\mu^{k+1} - \mu^*\|_{G_I}^2 &\leq \|\mu^k - \mu^*\|_{G_I}^2 - 2(\mu^{k+1} - \mu^*) \eta_I (\mu^{k+1}, \mu^k) \\
& \quad - \|\mu^{k+1} - \mu^k\|_{G_I}^2.
\end{aligned} \tag{28}$$

With Lemma 1, substituting (19) into (29), we get

$$\|\mu^{k+1} - \mu^*\|_{G_I}^2 \leq \|\mu^k - \mu^*\|_{G_I}^2 - (1 - u_I) \|\mu^k - \mu^{k+1}\|_{G_I}^2, \tag{29}$$

which completes the proof. \square

Remark 4. Compared to the requirement of the parameters r, s, β in [8], we now allow the step sizes r, s, β to be chosen according to rule (7). In fact, the restriction on r and s proposed in [8] is

$$r > 2\beta \|A^T A\|, \quad s > 2\beta \|B^T B\|, \tag{30}$$

which is a special case of the rule (18), since

$$\frac{\beta \|A^T A\|}{r} + \frac{\beta \|B^T B\|}{s} < \frac{\beta \|A^T A\|}{2\beta \|A^T A\|} + \frac{\beta \|B^T B\|}{2\beta \|B^T B\|} = 1; \tag{31}$$

that is, $u_I \in (0, 1)$. Hence, the requirement on the parameters is significantly relaxed.

3.2. *The Parameters Relaxation of the SPDM.* In this subsection, we extend our analysis to the SPDM. From (5), the subproblems of SPDM can be characterized by the following VI form: find $\mu^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})^T \in \Omega$ such that

$$(\mu' - \mu^{k+1})^T \{F(\mu^{k+1}) + \eta_{II}(\mu^{k+1}, \mu^k) - G(\mu^k - \mu^{k+1})\} \geq 0, \quad \forall \mu' \in \Omega, \quad (33)$$

with

$$\eta_{II}(\mu^{k+1}, \mu^k) = \beta \begin{pmatrix} A^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \\ B^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \\ 0 \end{pmatrix},$$

$$G_{II} = \begin{pmatrix} rI_n + \beta \|A^T A\| & & \\ & sI_p + \beta \|B^T B\| & \\ & & \beta^{-1} I_m \end{pmatrix}, \quad (34)$$

where $\eta_{II}(\mu^{k+1}, \mu^k) \in \mathfrak{R}^{(n+p+m)}$ and $G_{II} \in \mathfrak{R}^{(n+p+m) \times (n+p+m)}$ is a positive definite matrix.

Lemma 5. For a given $\mu^k = (x^k, y^k, \lambda^k)^T$, let $\mu^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})^T$ be generated by (33) and (34). If

$$u_{II} = \sqrt{\frac{\beta \|A^T A\|}{r + \beta \|A^T A\|} + \frac{\beta \|B^T B\|}{s + \beta \|B^T B\|}} \in (0, 1), \quad (35)$$

then for any $\mu^* = (x^*, y^*, \lambda^*)^T$ one has

$$(\mu^{k+1} - \mu^*)^T \eta_{II}(\mu^{k+1}, \mu^k) \leq \frac{u_{II}}{2} \|\mu^k - \mu^{k+1}\|_{G_{II}}^2. \quad (36)$$

Proof. Analogically, we have

$$\begin{aligned} & (\mu^{k+1} - \mu^*)^T \eta_{II}(\mu^{k+1}, \mu^k) \\ &= (\lambda^k - \lambda^{k+1})^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})). \end{aligned} \quad (37)$$

Applying the Cauchy-Schwarz inequality to the right term, we get

$$\begin{aligned} & (\mu^{k+1} - \mu^*)^T \eta_{II}(\mu^{k+1}, \mu^k) \\ & \leq \left| (\lambda^k - \lambda^{k+1})^T (A(x^k - x^{k+1}) + B(y^k - y^{k+1})) \right| \\ & \leq \frac{1}{2} \left\{ 2 \left| (\lambda^k - \lambda^{k+1})^T A(x^k - x^{k+1}) \right| \right\} \\ & \quad + \frac{1}{2} \left\{ 2 \left| (\lambda^k - \lambda^{k+1})^T B(y^k - y^{k+1}) \right| \right\} \\ & \leq \frac{1}{2} \left\{ \frac{\|A^T A\|}{u_{II}(r + \beta \|A^T A\|)} |\lambda^k - \lambda^{k+1}|^2 \right. \\ & \quad \left. + \frac{u_{II}(r + \beta \|A^T A\|)}{\|A^T A\|} |A(x^k - x^{k+1})|^2 \right\} \end{aligned}$$

$$\begin{aligned} & + \frac{1}{2} \left\{ \frac{\|B^T B\|}{u_{II}(s + \beta \|B^T B\|)} |\lambda^k - \lambda^{k+1}|^2 \right. \\ & \quad \left. + \frac{u_{II}(s + \beta \|B^T B\|)}{\|B^T B\|} |B(y^k - y^{k+1})|^2 \right\} \\ &= \frac{1}{2} \left\{ \left(\frac{\|A^T A\|}{u_{II}(r + \beta \|A^T A\|)} + \frac{\|B^T B\|}{u_{II}(s + \beta \|B^T B\|)} \right) \right. \\ & \quad \times |\lambda^k - \lambda^{k+1}|^2 + \frac{u_{II}(r + \beta \|A^T A\|)}{\|A^T A\|} |A(x^k - x^{k+1})|^2 \\ & \quad \left. + \frac{u_{II}(s + \beta \|B^T B\|)}{\|B^T B\|} |B(y^k - y^{k+1})|^2 \right\} \\ & \leq \frac{1}{2} \left\{ \left(\frac{\|A^T A\|}{u_{II}(r + \beta \|A^T A\|)} + \frac{\|B^T B\|}{u_{II}(s + \beta \|B^T B\|)} \right) \right. \\ & \quad \times |\lambda^k - \lambda^{k+1}|^2 + u_{II}(r + \beta \|A^T A\|) |x^k - x^{k+1}|^2 \\ & \quad \left. + u_{II}(s + \beta \|B^T B\|) |y^k - y^{k+1}|^2 \right\} \\ &= \frac{1}{2} \left\{ \frac{u_{II}}{\beta} |\lambda^k - \lambda^{k+1}|^2 + u_{II}(r + \beta \|A^T A\|) |x^k - x^{k+1}|^2 \right. \\ & \quad \left. + u_{II}(s + \beta \|B^T B\|) |y^k - y^{k+1}|^2 \right\} \\ &= \frac{u_{II}}{2} \|\mu^k - \mu^{k+1}\|_{G_{II}}^2. \end{aligned} \quad (38)$$

□

Theorem 6. Let the sequence $\{\mu^k\}$ be generated by SPDM (33) and (34). If

$$u_{II} = \sqrt{\frac{\beta \|A^T A\|}{r + \beta \|A^T A\|} + \frac{\beta \|B^T B\|}{s + \beta \|B^T B\|}} \in (0, 1), \quad (39)$$

then one has

$$\begin{aligned} \|\mu^{k+1} - \mu^*\|_{G_{II}}^2 & \leq \|\mu^k - \mu^*\|_{G_{II}}^2 - (1 - u_{II}) \|\mu^k - \mu^{k+1}\|_{G_{II}}^2, \\ & \quad \forall \mu^* \in \Omega^*. \end{aligned} \quad (40)$$

Proof. Consider

$$\begin{aligned} \|\mu^{k+1} - \mu^*\|_{G_{II}}^2 &= \|\mu^{k+1} - \mu^k + \mu^k - \mu^*\|_{G_{II}}^2 \\ &= \|\mu^k - \mu^*\|_{G_{II}}^2 - 2(\mu^k - \mu^*)^T G_{II}(\mu^k - \mu^{k+1}) \\ & \quad + \|\mu^{k+1} - \mu^k\|_{G_{II}}^2. \end{aligned} \quad (41)$$

$$\quad (42)$$

On the other hand, by setting $\mu' = \mu^*$ in (33), we have

$$\begin{aligned} (\mu^{k+1} - \mu^*)^T G(\mu^k - \mu^{k+1}) &\geq (\mu^{k+1} - \mu^*)^T F(\mu^{k+1}) \\ &\quad + (\mu^{k+1} - \mu^*) \eta_{II}(\mu^{k+1}, \mu^k). \end{aligned} \quad (43)$$

By using the monotonicity of F , we have

$$(\mu^{k+1} - \mu^*)^T F(\mu^{k+1}) \geq (\mu^{k+1} - \mu^*)^T F(\mu^*) \geq 0. \quad (44)$$

With rearrangement of the term (43), we derive that

$$\begin{aligned} (\mu^k - \mu^*)^T G(\mu^k - \mu^{k+1}) &\geq (\mu^k - \mu^{k+1})^T G(\mu^k - \mu^{k+1}) \\ &\quad + (\mu^{k+1} - \mu^*) \eta_{II}(\mu^{k+1}, \mu^k). \end{aligned} \quad (45)$$

Substituting (45) into (41), we get

$$\begin{aligned} \|\mu^{k+1} - \mu^*\|_{G_{II}}^2 &\leq \|\mu^k - \mu^*\|_{G_{II}}^2 \\ &\quad - 2(\mu^{k+1} - \mu^*) \eta_{II}(\mu^{k+1}, \mu^k) \\ &\quad - \|\mu^{k+1} - \mu^k\|_{G_{II}}^2. \end{aligned} \quad (46)$$

With Lemma 2, substituting (36) into (46), we get

$$\begin{aligned} \|\mu^{k+1} - \mu^*\|_{G_{II}}^2 &\leq \|\mu^k - \mu^*\|_{G_{II}}^2 \\ &\quad - (1 - u_{II}) \|\mu^k - \mu^{k+1}\|_{G_{II}}^2, \end{aligned} \quad (47)$$

which completes the proof. \square

Remark 7. Compared to the requirement of the parameters r , s , β in [8], we now allow the step sizes r , s , β to be chosen according to the rule (8). In fact, the restriction on r and s proposed in [8] is

$$r > \beta \|A^T A\|, \quad s > \beta \|B^T B\|, \quad (48)$$

which is a special case of the rule (35), since

$$\frac{\beta \|A^T A\|}{r + \beta \|A^T A\|} + \frac{\beta \|B^T B\|}{s + \beta \|A^T A\|} < \frac{\beta \|A^T A\|}{2\beta \|A^T A\|} + \frac{\beta \|B^T B\|}{2\beta \|B^T B\|} = 1; \quad (49)$$

that is, $u_{II} \in (0, 1)$. Hence, the requirement on the parameters is significantly relaxed.

3.3. The Convergence. In this subsection, we give the main convergence theorem of the FPDM and SPDM under the new required parameters conditions.

Theorem 8. *The sequence $\{\mu^k\}$ generated by the FPDM (resp., SPDM) under the conditions (18) (resp., (35)) converges to some solution μ^* , which is a solution of SVIs (9).*

Proof. Theorem 3 (resp., Theorem 6) means that the sequence $\{\mu^k\}$ generated is Fejér monotone with respect to the solution set Ω^* and the assertion follows immediately by using the property of Fejér monotonicity. \square

4. Numerical Experiments

In this section, we report the sensitivity of the involved parameters r , s , β of FPDM on the stable principal component pursuit problem (SPCP). Since SPDM is the extended version of FPDM and the sensitivity results of SPDM are similar to those of FPDM, we omit the numerical results of SPDM for the sake of succinctness. The problem tested is from Example 2 of [26]. Codes were all written in Matlab 2009b and all programs were run on HP notebook with Intel Core CPU 2.0 GHZ and 2 G memory.

SPCP arising from compressed sensing seeks to decompose a given observation matrix M into the sum of three matrices: $M := L + S + Z$, where L is a nonnegative and low-rank matrix, S is a sparse matrix, and Z is a noise matrix. The model of SPCP can be cast as

$$\begin{aligned} \min_{L,S,Z} \quad & \|L\|_* + \rho \|S\|_1 + \mathcal{J}(\|Z\|_F \leq \sigma) + \mathcal{J}(L \geq 0) \\ \text{s.t.} \quad & L + S + Z = M, \end{aligned} \quad (50)$$

where $\|\cdot\|_*$ is the so-called nuclear norm (the sum of all singular values), $\|\cdot\|_1$ is the l_1 norm, and $\mathcal{J}(\cdot)$ is an indicator function.

Following the procedure described in [26], by introducing an auxiliary variable K , grouping L and S as one big block $[L; S]$, and grouping Z and K as another big block $[Z; K]$, (50) can be reformulated as the standard form of (1) as follows:

$$\begin{aligned} \min_{L,S,Z,K} \quad & \|L\|_* + \rho \|S\|_1 + \mathcal{J}(\|Z\|_F \leq \sigma) + \mathcal{J}(K \geq 0) \\ \text{s.t.} \quad & \begin{pmatrix} I & I \\ I & 0 \end{pmatrix} \begin{pmatrix} L \\ S \end{pmatrix} + \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} Z \\ K \end{pmatrix} = \begin{pmatrix} M \\ 0 \end{pmatrix}. \end{aligned} \quad (51)$$

Then the largest singular value of the coefficient matrix of $[L; S]$ is $\sqrt{2.618}$. The largest singular value of the coefficient matrix of $[Z; K]$ is 1. For a better illustration, we denote $\tau_1 = 2.618$, $\tau_2 = 1$.

FPDM (5) applied to (51) yields the following iterative scheme:

$$\begin{aligned} L^{k+1} &= \operatorname{argmin} \|L\|_* \\ &\quad + \frac{r}{2} \|L - L^k \\ &\quad - \frac{1}{r} (\Lambda_1^k + \Lambda_2^k - \beta(2L^k + S^k + Z^k - K^k - M))\|_F^2, \end{aligned} \quad (52a)$$

$$S^{k+1} = \operatorname{argmin} \rho \|S\|_1 + \frac{r}{2} \left\| S - S^k - \frac{1}{r} \left(\Lambda_1^k - \beta (L^k + S^k + Z^k - M) \right) \right\|_F^2, \quad (52b)$$

$$Z^{k+1} = \operatorname{argmin} \mathcal{J} (\|Z\|_F \leq \sigma) + \frac{s}{2} \left\| Z - Z^k - \frac{1}{s} \left(\Lambda_1^k - \beta (L^k + S^k + Z^k - M) \right) \right\|_F^2, \quad (52c)$$

$$K^{k+1} = \operatorname{argmin} \mathcal{J} (K \geq 0) + \frac{s}{2} \left\| K - K^k + \frac{1}{s} \left(\Lambda_2^k - \beta (K^k - L^k) \right) \right\|_F^2, \quad (52d)$$

$$\Lambda_1^{k+1} = \Lambda_1^k - \beta (L^{k+1} + S^{k+1} + Z^{k+1} - M), \quad (52e)$$

$$\Lambda_2^{k+1} = \Lambda_2^k - \beta (L^{k+1} - K^{k+1}). \quad (52f)$$

There are two main advantages of FPDM applied to SPCP. First, all the generated minimizations in (52a)–(52d) have closed-form solutions. Second, the subproblems (52a)–(52d) are highly parallel, making FPDM appealing for parallel or distributed computing. Now, we elaborate on the strategy of solving the resulting subproblems at each iteration.

(i) The L -subproblem (52a) amounts to evaluate the proximal operator of the nuclear norm function and is given by the matrix shrinkage operation

$$L^{k+1} := \operatorname{MatShrink} \left(L^k + \frac{1}{r} \left(\Lambda_1^k + \Lambda_2^k - \beta (2L^k + S^k + Z^k - K^k - M) \right), \frac{1}{r} \right), \quad (53)$$

where the matrix shrinkage operator $\operatorname{MatShrink}(M, \xi)$ ($\xi > 0$) is defined as

$$\operatorname{MatShrink}(M, \xi) := U \operatorname{Diag}(\max\{\sigma - \xi, 0\}) V^T, \quad (54)$$

and $U \operatorname{Diag}(\sigma) V^T$ is the SVD of matrix M .

(ii) The closed-form solution of S -subproblem (52b) can be given by the l_1 shrinkage operation:

$$S^{k+1} := \operatorname{Shrink} \left(S^k + \frac{1}{r} \left(\Lambda_1^k - \beta (L^k + S^k + Z^k - M) \right), \frac{\rho}{r} \right), \quad (55)$$

where the l_1 shrinkage operator $\operatorname{Shrink}(M, \xi)$ is defined as

$$[\operatorname{Shrink}(M, \xi)]_{ij} := \begin{cases} M_{ij} - \xi, & \text{if } M_{ij} > \xi \\ M_{ij} + \xi, & \text{if } M_{ij} < -\xi \\ 0, & \text{if } |M_{ij}| \leq \xi. \end{cases} \quad (56)$$

(iii) The Z -subproblem (52c) amounts to projecting the matrix $W^k := M + (1/\beta)\Lambda_1^k - (L^k + S^k)$ onto the Euclidean ball $\|Z\|_F \leq \sigma$, whose closed-form solution is given by

$$Z^{k+1} := \frac{W^k}{\max\{1, \|W^k\|_F/\sigma\}}. \quad (57)$$

(iv) The K -subproblem (52d) amounts to projecting the matrix $L^k - (1/\beta)\Lambda_2^k$ onto the nonnegative orthant, whose closed-form solution is given by

$$K^{k+1} := \max \left\{ L^k - \frac{1}{\beta} \Lambda_2^k, 0 \right\}. \quad (58)$$

For detailed analytical methods of (52a)–(52d), the reader is referred to, for example, [26, 27].

In our experiment, we generate the data of (50) randomly in the same way as [26]. For given n , $r < n$, the $n \times n$ rank- r matrix L^* was generated by $L^* = R_1 R_2^T$, where R_1 and R_2 are $n \times r$ random matrices whose entries are independently and identically (i.i.d.) uniformly distributed in $[0, 1]$. Note that, in this experiment, L^* is a componentwise nonnegative and low-rank matrix we want to recover. The support of the sparse matrix S^* was chosen uniformly and randomly, and the nonzero entries of S^* were i.i.d. uniformly in the interval $[-500, 500]$. The entries of matrix Z^* for noise were generated as i.i.d. Gaussian with standard deviation 10^{-4} .

As in [26], we set $M := L^* + S^* + Z^*$; we chose $\rho := 1/\sqrt{n}$. The initial iterate for FPDM is $L^0 = K^0 = -M, S^0 = Z^0 = 0, \Lambda_1^0 = \Lambda_2^0 = 0$. The stopping criterion is set as

$$\operatorname{resid} := \frac{\|L + S + Z - M\|_F}{\|M\|_F} < \epsilon_r, \quad (59)$$

where ϵ_r is the tolerance set as $\epsilon_r = 10^{-4}$. We denoted $\operatorname{Rank}_r := r/n$ so that the rank of L^* is $n * \operatorname{Rank}_r$ and $\operatorname{Card}_r := \operatorname{cardinality}(S^*)/(n^2)$ so that the cardinality of S^* is $n^2 * \operatorname{Card}_r$. For some cases of dimension m , we report the iteration numbers (Iter.), relative error of the low-rank matrix $L(\operatorname{rel}_L)$, relative error of the sparse matrix $S(\operatorname{rel}_S)$, and CPU times in seconds (CPU(s)), where the relative errors are defined as

$$\operatorname{rel}_L := \frac{\|L - L^*\|_F}{\|L^*\|_F}, \quad \operatorname{rel}_S := \frac{\|S - S^*\|_F}{\|S^*\|_F}. \quad (60)$$

For each instance, we randomly created ten examples, so the results were averaged over ten runs.

The computational results are presented in Table 1. For a different instance, the value of β , r , and s chosen to satisfy the condition (7), respectively. It can be seen that if we choose r a little smaller than $2\beta\|A^T A\|$ and s a little larger than $2\beta\|B^T B\|$, the numerical performance of FPDM with the new selected parameters shows better than the case, where $r = 2\beta\|A^T A\|, s = 2\beta\|B^T B\|$.

5. Conclusions

In this paper, we show that the requirement ranges on the involved parameters of the proximal-based parallel decomposition methods can be significantly enlarged. We prove

TABLE 1: Numerical results for stable principal component pursuit problem.

n	r	s	Iter.	rel_L	rel_S	CPU (s)	r	s	Iter.	rel_L	rel_S	CPU (s)
$\beta = 0.01, \text{Rank}_r = 0.01, \text{Card}_r = 0.01$												
50	$2\beta\tau_1$	$2\beta\tau_2$	97	$9.01e-003$	$2.89e-005$	0.3	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	101	$9.18e-003$	$2.87e-005$	0.3
100	$2\beta\tau_1$	$2\beta\tau_2$	78	$5.71e-003$	$6.16e-005$	1.1	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	76	$6.57e-003$	$5.11e-005$	1.0
150	$2\beta\tau_1$	$2\beta\tau_2$	81	$3.52e-003$	$4.65e-005$	3.2	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	77	$3.75e-003$	$5.19e-005$	2.9
200	$2\beta\tau_1$	$2\beta\tau_2$	81	$3.60e-003$	$5.53e-005$	6.6	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	76	$3.61e-003$	$6.56e-005$	5.9
300	$2\beta\tau_1$	$2\beta\tau_2$	83	$2.33e-003$	$5.97e-005$	20.4	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	82	$2.57e-003$	$5.31e-005$	19.4
$\beta = 0.02, \text{Rank}_r = 0.01, \text{Card}_r = 0.01$												
50	$2\beta\tau_1$	$2\beta\tau_2$	181	$8.42e-003$	$2.37e-005$	0.5	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	180	$8.39e-003$	$2.36e-005$	0.5
100	$2\beta\tau_1$	$2\beta\tau_2$	105	$8.47e-003$	$2.03e-005$	1.4	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	106	$8.21e-003$	$2.34e-005$	1.4
150	$2\beta\tau_1$	$2\beta\tau_2$	127	$5.06e-003$	$2.09e-005$	4.7	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	126	$5.05e-003$	$2.14e-005$	4.5
200	$2\beta\tau_1$	$2\beta\tau_2$	93	$4.73e-003$	$1.87e-005$	7.2	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	91	$4.79e-003$	$2.67e-005$	7.0
300	$2\beta\tau_1$	$2\beta\tau_2$	84	$3.10e-003$	$3.37e-005$	20.3	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	85	$3.13e-003$	$2.63e-005$	20.2
$\beta = 0.01, \text{Rank}_r = 0.02, \text{Card}_r = 0.02$												
50	$2\beta\tau_1$	$2\beta\tau_2$	84	$1.14e-002$	$5.10e-005$	0.2	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	85	$1.24e-002$	$3.92e-005$	0.2
100	$2\beta\tau_1$	$2\beta\tau_2$	79	$5.76e-003$	$5.40e-005$	1.2	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	76	$5.71e-003$	$5.02e-005$	1.1
150	$2\beta\tau_1$	$2\beta\tau_2$	82	$3.55e-003$	$5.99e-005$	3.6	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	81	$3.80e-003$	$4.90e-005$	3.4
200	$2\beta\tau_1$	$2\beta\tau_2$	84	$2.90e-003$	$6.70e-005$	7.3	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	83	$2.92e-003$	$5.81e-005$	7.1
300	$2\beta\tau_1$	$2\beta\tau_2$	91	$2.18e-003$	$5.42e-005$	22.8	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	90	$2.33e-003$	$4.85e-005$	22.5
$\beta = 0.01, \text{Rank}_r = 0.03, \text{Card}_r = 0.03$												
50	$2\beta\tau_1$	$2\beta\tau_2$	99	$9.31e-003$	$3.09e-005$	0.3	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	101	$9.33e-003$	$2.99e-005$	0.3
100	$2\beta\tau_1$	$2\beta\tau_2$	83	$4.89e-003$	$5.03e-005$	1.4	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	81	$5.28e-003$	$4.53e-005$	1.3
150	$2\beta\tau_1$	$2\beta\tau_2$	86	$3.19e-003$	$5.84e-005$	3.8	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	84	$3.22e-003$	$5.68e-005$	3.7
200	$2\beta\tau_1$	$2\beta\tau_2$	91	$2.82e-003$	$5.67e-005$	8.5	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	90	$2.78e-003$	$4.92e-005$	8.4
300	$2\beta\tau_1$	$2\beta\tau_2$	115	$3.47e-003$	$1.88e-005$	30.8	$1.9\beta\tau_1$	$1.9\beta\tau_2/0.9$	116	$3.52e-003$	$1.85e-005$	29.0

the global convergence of the new scheme under the new conditions. Preliminary numerical experiments on the stable principal component pursuit problem testify to the advantages of the enlargement.

Conflict of Interests

We declare that there is no conflict of interests regarding the publication of this article.

Acknowledgment

The work of Mingfang Ni is supported in part by the Natural Science Foundation of China Grant no. NSFC-70971136.

References

- [1] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288–297, 2010.
- [2] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, vol. 49, pp. 185–212, Springer, New York, NY, USA, 2011.
- [3] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1 -problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.
- [4] S. Ma, L. Xue, and H. Zou, "Alternating direction methods for latent variable Gaussian graphical model selection," *Neural Computation*, vol. 25, no. 8, pp. 2172–2198, 2013.
- [5] 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*, vol. 3, no. 1, pp. 1–122, 2010.
- [6] B. He, M. Xu, and X. Yuan, "Solving large-scale least squares semidefinite programming by alternating direction methods," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 1, pp. 136–152, 2011.
- [7] Z. Wen, D. Goldfarb, and W. Yin, "Alternating direction augmented Lagrangian methods for semidefinite programming," *Mathematical Programming Computation*, vol. 2, no. 3-4, pp. 203–230, 2010.
- [8] B. He and X. Yuan, "A unified framework of some proximal-based decomposition methods for monotone variational inequalities with separable structures," *Pacific Journal of Optimization*, vol. 8, no. 4, pp. 817–844, 2012.
- [9] G. Chen and M. Teboulle, "A proximal-based decomposition method for convex minimization problems," *Mathematical Programming*, vol. 64, no. 1, pp. 81–101, 1994.
- [10] R. G. A. Marrocco, "Sur l'approximation par éléments finis d'ordre un et la résolution, par pénalisation dualité, d'une classe de problèmes de dirichlet non linéaires," *RAIRO*, vol. 9, pp. 41–76, 1975.
- [11] B. Martinet, "Régularisation d'inéquations variationnelles par approximations successives," vol. 4, no. 3, pp. 154–158, 1970.

- [12] T. Goldstein and S. Osher, "The split Bregman method for L_1 -regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [13] B. He and X. Yuan, "On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method," *SIAM Journal on Numerical Analysis*, vol. 50, no. 2, pp. 700–709, 2012.
- [14] M. Hong and Z. Q. Luo, "On the linear convergence of the alternating direction method of multipliers," In press, <http://arxiv.org/abs/1208.3922>.
- [15] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," Tech. Rep., 2012.
- [16] D. Han and X. Yuan, "A note on the alternating direction method of multipliers," *Journal of Optimization Theory and Applications*, vol. 155, no. 1, pp. 227–238, 2012.
- [17] C. Chen, B. He, Y. Ye, and X. Yuan, "The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent," In press, http://www.optimization-online.org/DB_HTML/2013/09/4059.html.
- [18] B. He, L.-Z. Liao, and X. Wang, "Proximal-like contraction methods for monotone variational inequalities in a unified framework I: effective quadruplet and primary methods," *Computational Optimization and Applications*, vol. 51, no. 2, pp. 649–679, 2012.
- [19] B. He, L.-Z. Liao, and X. Wang, "Proximal-like contraction methods for monotone variational inequalities in a unified framework II: general methods and numerical experiments," *Computational Optimization and Applications*, vol. 51, no. 2, pp. 681–708, 2012.
- [20] X. Cai, G. Gu, and B. He, "On the $O(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators," *Computational Optimization and Applications*, 2013.
- [21] M. H. Xu, J. L. Jiang, B. Li, and B. Xu, "An improved prediction-correction method for monotone variational inequalities with separable operators," *Computers & Mathematics with Applications*, vol. 59, no. 6, pp. 2074–2086, 2010.
- [22] X. Yuan and M. Li, "An LQP-based decomposition method for solving a class of variational inequalities," *SIAM Journal on Optimization*, vol. 21, no. 4, pp. 1309–1318, 2011.
- [23] M. Tao and X. Yuan, "On the $O(1/t)$ convergence rate of alternating direction method with logarithmic-quadratic proximal regularization," *SIAM Journal on Optimization*, vol. 22, no. 4, pp. 1431–1448, 2012.
- [24] A. Bnouhachem, H. Benazza, and M. Khalfaoui, "An inexact alternating direction method for solving a class of structured variational inequalities," *Applied Mathematics and Computation*, vol. 219, no. 14, pp. 7837–7846, 2013.
- [25] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation*, 1989.
- [26] S. Ma, "Alternating proximal gradient method for convex minimization," In press, http://www.optimization-online.org/DB_HTML/2012/09/3608.html
- [27] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 123–231, 2013.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

