An inexact parallel splitting augmented Lagrangian method for large system of linear equations

Zheng Peng* [1, 2] DongHua Wu [2]

[1]. Mathematics Department of Nanjing University, Nanjing PR. China, 210093
[2]. Mathematics Department of Shanghai University, Shanghai PR. China, 200444

Abstract: Parallel iterative methods are powerful tool for solving large system of linear equations (LEs). The existing parallel computing research results are focussed mainly on sparse system or others with particular structure. And most are based on parallel implementation of the classical relaxation methods such as Gauss-Seidel, SOR, and AOR methods carried efficiently on multiprcessor systems. In this paper, we proposed a novel parallel splitting operator method based on a new approach. In this method, we divide the coefficient matrix into two or three parts. Then we convert the original problem (LEs) into a monotone (linear) variational inequalities problem (VIs) with separable structure. Finally, we propose an inexact parallel splitting augmented Lagrangian method to solve this variational inequalities problem (VIs). We avoid the matrix inverse operator by introducing proper inexact terms in subproblems, such that complexity of each step of this method is $O(n^2)$. This is different to a general iterative method with complexity $O(n^3)$. In addition, this method does not depend on any special structure of the problem which is to be solved. Convergence of the proposed methods, in dealing with two and three separable operators respectively, is proved. Numerical experiments are provided to show its applicability, especially its robustness with respect to the scale (dimensions) and condition (measured by the condition number of coefficient matrix A) of this method.

Keywords: System of linear equations; synchronous parallel iteration; matrix decomposition; inexact terms; convergence; robustness.

Mathematics subject classification: 90-08, 90C25, 90C90, 65F10, 65F30

1 Introduction

Many problems in scientific and engineering computing require finding a solution (or solutions) of a large system of linear equations. Some of these systems may have millions of variables. The large scale leads to difficulties. Researchers have used self-absorption in parallel computing—-a class of methods for finding solutions of a problem by using multiprcessor systems simultaneously. This methods save time to obtain solution of the large scale problem.

*Corresponding author. Tel: +86-02551730149. e-mail: pzheng008@yahoo.com. This work is partly supported by Natural Sciences Foundation of China (No. 10671117 and No. 10571083).

In this paper, we consider the following well-determined system of linear equations

$$Ax = b \tag{1.1}$$

where $A \in R^{n\times n},\ x \in R^n$ and $b \in R^n$, and A matrix is positive definite.

There are many methods for solving Problem (1.1). Iterative methods are powerful ways to solve this problem. A general iterative method is defined as follows:

Definition 1.1 [1] *Let A be a nonsingular matrix. A pair of matrices (M, N) with M nonsingular (and easily invertible in practice) such that*

$$A = M - N$$

is called a splitting (or regular decomposition) of A. An iterative method based on the splitting (M, N) is defined by

$$\begin{cases} x_0 \text{ given in } R^n \\ Mx_{k+1} = Nx_k + b \quad \forall k \geq 1 \end{cases} \tag{1.2}$$

By using (1.2) we get

$$x_{k+1} = M^{-1}Nx_k + M^{-1}b \tag{1.3}$$

Denotes $J = M^{-1}N$ and refer it as to the iterative matrix. An iterative method (1.2–1.3) converges if and only if the spectral radius of J satisfies $\rho(J) < 1$. In a general way, (1.3) is the main cost in an iterative method. The complexity of the operator M^{-1} is $O(n^3)$ theoretically.

The main iterative methods [1, 28, 22] are Jacobi method, Gauss-Seidel method and Successive Over-Relaxation (SOR) method. Let $A = (a_{ij})_{1\leq i,j\leq n}$ and

$$D = \text{diag}(a_{11}, a_{22}, ..., a_{nn}).$$

The Jacobi method is defined by

$$M = D \ \ and \ \ N = D - A$$

in (1.2-1.3). Let

$$\begin{cases} e_{ij} = -a_{ij} \quad if\ i > j,\ and\ 0\ otherwise \\ f_{ij} = -a_{ij} \quad if\ i < j,\ and\ 0\ otherwise \end{cases}$$

and $E = (e_{ij})_{1\leq i,j\leq n}$, $F = (f_{ij})_{1\leq i,j\leq n}$, then $A = D - E - F$. The Gauss-Seidel Method defined by

$$M = D - E \ and \ N = F$$

in (1.2-1.3). The SOR method is defined by

$$M = \frac{D}{\omega} - E \ and \ N = \frac{1-\omega}{\omega}D + F$$

where $\omega \in R^+$.

From the practical point of view, the main advantage of Jacobi and Gauss-Seidel method is the avoidance of the inverse operator M^{-1} in (1.3) in some sense. In Jacobi method, iterative formula (1.3) is equivalent to

$$x_i^{k+1} = \frac{1}{a_{ii}}\left[-a_{i1}x_1^k - a_{i2}x_2^k - ... - a_{i,i-1}x_{i-1}^k - a_{i,i+1}x_{i+1}^k ... - a_{in}x_n^k + b_i\right],\ i = 1, 2, ..., n$$

In the Gauss-Seidel method, iterative formula (1.3) is equivalent to

$$x_i^{k+1} = \frac{1}{a_{ii}} \left[-a_{i1}x_1^{k+1} - a_{i2}x_2^{k+1} - ... - a_{i,i-1}x_{i-1}^{k+1} - a_{i,i+1}x_{i+1}^{k} ... - a_{in}x_n^k + b_i \right], \ i = 1, 2, ..., n.$$

Many of research is based on the above three iterative methods. For example, Bagnara [3] gave a unified proof for the convergence of Jacobi and Gauss-Seidel methods. Salkuyeh [23] generalized the Jacobi and Gauss-Seidel methods. Wang and Huang [25] investigated the convergence and the monotone convergence theories for the alternating method when the coefficient matrix is a H-matrix or a monotone matrix, and gave some sufficient conditions for the induced splitting by the alternating method to be a regular splitting.

At the same time, many authors studied structured systems of linear equations. For example, Gu and his co-authors [12, 11] studied and established some fast algorithms for structured systems of linear equations, such as Toeplitz and Hankel systems. Lin, Wei and Zhang [20] studied convergence and quotient convergence of iterative methods for solving singular linear equations with index one. Other authors studied sparse systems, block-diagonal systems, etc.. They gave many effective algorithms or theoretical results on their research problems.

Bai [24, 6] studied necessary and sufficient convergence conditions for splitting iteration methods for non-Hermitian system of linear equations when the coefficient matrix is non-singular, and the convergence conditions for the additive and the multiplicative splitting iteration methods [7], convergence of parallel nonstationary multisplitting iteration methods [8],etc.. Bai [4, 26, 5] studied parallel splitting methods for large system with peculiar structure. For example, Bai [5] proposed a class of parallel decomposition-type accelerated over-relaxation methods for solving the large sparse systems of linear equations, and gave a sufficient conditions ensuring its convergence when the coefficient matrices of the linear systems of equations are respectively L-matrices, H-matrices and positive definite matrices.

Hadjidimos [19] established an accelerated over-relaxation method which is essentially based upon reasonable decomposition of the linear system (1.1) as well as technical application of the accelerated over-relaxation idea, and hence has natural and nice parallelism as well as good convergence behaviour. Frommer [9], and Neumann [21] give convergence of relaxed parallel multisplitting methods and parallel multisplitting iterative methods for M-matrices.

Another novel way to solve Problem (1.1) is to convert it into an optimization problem in the following form:

$$x^* = \text{Arg} \min_{x \in R^n} \|Ax - b\|_2^2 \tag{1.4}$$

One can solve Problem (1.4) by using projection methods, see C.Brezinski [2].

It is well known that Problem (1.4) is equivalent to the following linear variational inequality problem:

$$\text{Find } x^* \in R^n, \text{ such that } (x - x^*)^T (A^T(Ax^* - b)) \geq 0, \ \forall x \in R^n \tag{1.5}$$

Recent research shows that projection-contraction methods are powerful tool for solving Problem (1.5). For example, He [14] developed a class of projection-contraction (PC) methods for monotone linear variational inequalities (LVIs). Problem (1.5) is in

this class, when A^TA matrix is positive definite. In his follow-up research, He and his co-authors [15] developed many of PC-like methods for solving monotone VIs, includes LVIs.

Furthermore, when A matrix is large scale, as previously mentioned, operator splitting and parallel computing will be necessary. On operator splitting and parallel computing for variational inequalities, Glowinski [13], Fukushima [10] and He [16] have many useful resluts.

Consider the variational inequalities with two separable operators in the following form:

$$\text{Find } \; u = \begin{pmatrix} x \\ y \end{pmatrix} \in \Omega, \text{ such that } \; \begin{cases} (x' - x)^T f(x) \geq 0 \\ (y' - y)^T g(y) \geq 0 \end{cases} \; \forall u' \in \Omega \tag{1.6}$$

where $f(x), g(y)$ are given monotone operator with respect to Ω. And,

$$\Omega = \{(x, y) | x \in \mathcal{X}, y \in \mathcal{Y} : Ax + By = d\}, \tag{1.7}$$

$\mathcal{X} \in R^{n_1}, \mathcal{Y} \in R^{n_2}$ and $A \in R^{m \times n_1}, B \in R^{m \times n_1}$ are given matrices, $d \in R^m$ is a given vector.

He [16] proposed a method to deal with Problem (1.6–1.7). This method is referred to as parallel splitting augment Lagrangian method (abbreviate to PSALM). By attaching a Lagrange multiplier vector $\lambda \in R^m$ to the linear constraints $Ax + By = d$, Problem (1.6-1.7) can be explained in the following form:

$$\text{Find } w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} \in \mathcal{W}, \;\; \text{such that } \; \begin{cases} (x' - x)^T [f(x) - A^T \lambda] \geq 0, \\ (y' - y)^T [g(y) - B^T \lambda] \geq 0, \\ Ax + By - d = 0, \end{cases} \; \forall w' \in \mathcal{W}, \tag{1.8}$$

where $\mathcal{W} = \mathcal{X} \times \mathcal{Y} \times R^m$.

For solving Problem (1.8), for a given quadruple (x^k, y^k, λ^k), PSALM method finds iteratively $\widetilde{x}^k, \widetilde{y}^k$ by solving in parallel the following variational inequalities:

$$x \in \mathcal{X}, \;\; (x' - x)^T \left\{ f(x) - A^T \left[\lambda^k - H(Ax + By^k - d) \right] \right\} \geq 0, \; \forall x' \in \mathcal{X} \tag{1.9}$$

$$y \in \mathcal{Y}, \;\; (y' - y)^T \left\{ g(y) - B^T \left[\lambda^k - H(Ax^k + By - d) \right] \right\} \geq 0, \; \forall y' \in \mathcal{Y} \tag{1.10}$$

Let

$$\widetilde{\lambda}^k = \lambda^k - H(A\widetilde{x}^k + B\widetilde{y}^k - d). \tag{1.11}$$

where H is a given positive matrix. The PSALM method produces the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ by letting

$$w^{k+1} = w^k - \alpha^k G^{-1} M (w^k - \widetilde{w}^k) \tag{1.12}$$

where α^k, G, and M are given proper parameters or matrices of the algorithm. He also extends this method to deal with the variational inequalities with three separable operators, for details, see Section 4 of [16].

In this paper, we shall propose a new iterative method for solving system of linear equations. This is an inexact parallel splitting augmented Lagrangian method (abbreviate to in-PSALM). The advantages of this method are in the following: it decomposes the cost of computational loads to each of the processors which participate in solving

the problem and at the same time it can avoid the inverse matrix operator such that the complexity of each iteration is $O(n^2)$ in theory and in practice. Numerical results show it is robust for very broad range of system of LEs.

The rest of this paper is organized as follows: In Section 2, we propose the inexact parallel splitting augmented Lagrangian method for solving Problem (1.1), by splitting A matrix into two parts. In Section 3, we prove convergence of the proposed method. In Section 4, we generalize our method to deal with the general monotone variational inequalities with three separable operators and give its convergence. In Section 5, some numerical results are provided and conclusions are made.

2 The Proposed Method

In this section, we consider a new algorithm for solving the well-determined system of linear equations (1.1), i.e.,

$$Ax = b \tag{2.1}$$

where $A \in R^{n\times n},\ x \in R^n$ and $b \in R^n$, A is positive definite.

By using the notations in Matlab, we part A matrix (in row) into two parts as the following:

$$M = A(1:l,:),\quad N = A(l+1:n,:),\quad p = b(1:l),\ q = b(l+1:n),\ (l < n).$$

The system of LEs (2.1) can be driven from the following system of LEs,

$$\begin{cases} Mx = p \\ Nx = q \end{cases} \tag{2.2}$$

Problem (2.2) can be solved in Least-square approach. We solve the following minimization problem:

$$x^* = \text{Arg}\min_{x\in R^n}\ \|Mx - p\|^2 + \|Nx - q\|^2 \tag{2.3}$$

where $\|\cdot\|$ denotes Euclidean norm.

Clearly, Problem (2.3) is equivalent to the following separable form:

$$\begin{array}{ll} \min\limits_{x,y\in R^n} & \|Mx - p\|^2 + \|Ny - q\|^2 \\ \text{subject to} & x - y = 0 \end{array} \tag{2.4}$$

By using KKT condition of Problem (2.4), we get the following separable linear variational inequalities (LVIs):

$$\text{Find } w = (x, y, \lambda) \in \mathcal{W} \text{ such that } \begin{cases} (x' - x)^T[M^T(Mx - p) - \lambda] \geq 0 \\ (y' - y)^T[N^T(Ny - q) + \lambda] \geq 0 \\ x - y = 0 \end{cases} \quad \forall w' \in \mathcal{W} \tag{2.5}$$

where $\mathcal{W} = R^n \times R^n \times R^n$.

For given $w^k = (x^k, y^k, \lambda^k)$, we can carry out the candidate $\widehat{w}^k = (\widehat{x}^k, \widehat{y}^k, \widehat{\lambda}^k)$ via the following procedure: First we get $(\widehat{x}^k, \widehat{y}^k)$ by solving parallel the following LVIs:

$$(x' - x)^T[M^T(Mx - p) - (\lambda^k - \beta_k(x - y^k))] \geq 0 \tag{2.6}$$

$$(y' - y)^T[N^T(Ny - q) + (\lambda^k - \beta_k(x^k - y))] \geq 0 \tag{2.7}$$

and then update λ by

$$\widehat{\lambda}^k = \lambda^k - \beta_k(\widehat{x}^k - \widehat{y}^k) \tag{2.8}$$

This method (2.6–2.8) is refered to the parallel method. Linear variational inequalities (2.6– 2.7) can be solved directly. Their solutions are:

$$\widehat{x}^k = (M^T M + \beta_k I)^{-1} * (p + \lambda^k + \beta_k y^k);\ \widehat{y}^k = (N^T N + \beta_k I)^{-1} * (q - \lambda^k + \beta_k x^k). \tag{2.9}$$

In this approach we use PSALM for solving (2.5) directly. He [16] has proved its convergence. However, when we use PSALM to solve Problem (2.6–2.8), we find that (2.9) involves inverse matrix operators, which is expensive computationally.

Inexact proximal point method (In-PPA) can be used to solved LVIs (2.6–2.7), and a good "inexact term" may be help us to bypass the inverse matrix operators.

For given $w^k = (x^k, y^k, \lambda^k)$, In-PPA carries out the candidate $\widehat{w}^k = (\widehat{x}^k, \widehat{y}^k, \widehat{\lambda}^k)$ via the following procedure: First we get $(\widehat{x}^k, \widehat{y}^k)$ by solving in parallel the following LVIs:

$$(x' - x)^T \left\{ M^T(Mx - p) - [\lambda^k - \beta_k(x - y^k)] + r_k(x - x^k) + \xi_x^k \right\} \geq 0 \tag{2.10}$$

$$(y' - y)^T \left\{ N^T(Ny - q) + [\lambda^k - \beta_k(x^k - y)] + s_k(y - y^k) + \xi_y^k \right\} \geq 0. \tag{2.11}$$

Then we update $\widehat{\lambda}^k$ by (2.8).

By a proper choice of parameters $r_k,\ s_k$ and with inexact terms $\xi_x^k,\ \xi_y^k$, we construct a new algorithm to generate a sequence $\{(x^k, y^k, \lambda^k)\}$ which converges to $\{(x^*, y^*, \lambda^*)\}$– a solution of LVIs (2.5). This is the following method:

The proposed method (in-PSALM-II)

`Step` 0. Given the parameters $\varepsilon > 0, \gamma \in (0, 2), 0 < \beta_L \leq \beta \leq \beta_U, \nu \in (0, 1)$, initial parameters r_0, s_0 and initial point $w^0 = (x^0, y^0, \lambda^0) \in R^n \times R^n \times R^n$. Set $k = 0$, let $\beta_k = \beta$.

`Step` 1. Produce $(\widehat{x}^k, \widehat{y}^k)$ via solving parallel LVIs (2.10 and 2.11). Here r_k and ξ_x^k satisfy the following inequality

$$\|\xi_x^k\| \leq \nu r_k \|x^k - \widehat{x}^k\|,\ \ \xi_x^k = M^T M(x^k - \widehat{x}^k) + \beta_k(x^k - \widehat{x}^k). \tag{2.12}$$

Furthermore, s_k and ξ_y^k satisfy the following inequality

$$\|\xi_y^k\| \leq \nu s_k \|y^k - \widehat{y}^k\|,\ \ \xi_y^k = N^T N(y^k - \widehat{y}^k) + \beta_k(y^k - \widehat{y}^k). \tag{2.13}$$

Then update $\widehat{\lambda}^k$ by using (2.8).

`Step` 2. Check terminating condition: let

$$\varepsilon_k = \max(\|\lambda^k - \widehat{\lambda}^k\|, \max(\|y^k - \widehat{y}^k\|, \|x^k - \widehat{x}^k\|)),$$

if $\varepsilon_k < \varepsilon$, stop. Else goto next step.

`Step` 3. Produce the new iterative point $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ by using the following formula:

$$w^{k+1} = w^k - \alpha_k d_1(w^k, \widehat{w}^k). \tag{2.14}$$

Let $k := k + 1$, go to Step 1.

Here

$$d_1(w^k, \widehat{w}^k) = \begin{bmatrix} (x^k - \widehat{x}^k) - \dfrac{1}{r_k + \beta_k}\xi_x^k \\ (y^k - \widehat{y}^k) - \dfrac{1}{s_k + \beta_k}\xi_y^k \\ (\lambda^k - \widehat{\lambda}^k) \end{bmatrix}$$

and $\alpha_k = \gamma\alpha_k^*$,

$$\alpha_k^* = \frac{\varphi(w^k, \widehat{w}^k)}{\|d_1(w^k, \widehat{w}^k)\|_G^2}, \tag{2.15}$$

$$\varphi(w^k, \widehat{w}^k) = (\lambda^k - \widehat{\lambda}^k)^T(x^k - \widehat{x}^k) - (\lambda^k - \widehat{\lambda}^k)^T(y^k - \widehat{y}^k) + (w^k - \widehat{w}^k)^T G d_1(w^k, \widehat{w}^k). \tag{2.16}$$

$$G = \begin{pmatrix} (r_k + \beta_k)I & & \\ & (s_k + \beta_k)I & \\ & & \dfrac{1}{\beta_k}I \end{pmatrix} \text{ is positive definite.}$$

Remark: 2.1 From (2.10, 2.12), we can get $\widehat{x}^k, \widehat{y}^k$ easily, via the following projection (on R^n):

$$\widehat{x}^k = x^k - \frac{1}{r_k}\left[M^T(Mx^k - p) - (\lambda^k - \beta_k(x^k - y^k))\right]. \tag{2.17}$$

In the same way, we get $\widehat{y}^k$ via the following projection (on R^n):

$$\widehat{y}^k = y^k - \frac{1}{s_k}\left[N^T(Ny^k - q) + (\lambda^k - \beta_k(x^k - y^k))\right]. \tag{2.18}$$

3 Convergence of the proposed method

Let $w = (x, y, \lambda)$ and $\widehat{w} = (\widehat{x}, \widehat{y}, \widehat{\lambda})$. For convenience, we denote

$$d_2(w, \widehat{w}) = \begin{bmatrix} M^T(M\widehat{x} - p) - [\widehat{\lambda} + \beta(y - \widehat{y}) - \beta(x - \widehat{x})] \\ N^T(N\widehat{y} - q) + [\widehat{\lambda} + \beta(y - \widehat{y}) - \beta(x - \widehat{x})] \\ \widehat{x} - \widehat{y} \end{bmatrix} \tag{3.1}$$

$$F(\widehat{w}) = \begin{pmatrix} M^T(M\widehat{x} - p) - \widehat{\lambda} \\ N^T(N\widehat{y} - q) + \widehat{\lambda} \\ \widehat{x} - \widehat{y} \end{pmatrix} \quad \text{and} \quad \eta(w, \widehat{w}) = \begin{pmatrix} -\beta(y - \widehat{y}) + \beta(x - \widehat{x}) \\ \beta(y - \widehat{y}) - \beta(x - \widehat{x}) \\ 0 \end{pmatrix}$$

$$\xi_w = \begin{pmatrix} \xi_x \\ \xi_y \\ 0 \end{pmatrix}.$$

Then by a simple transmutation of (2.10, 2.11) and (2.8), we get

$$(w - \widehat{w})^T[F(\widehat{w}) + \eta(w, \widehat{w}) + G(\widehat{w} - w) + \xi_w] \geq 0 \tag{3.2}$$

and

$$d_1(w, \widehat{w}) = (w - \widehat{w}) - G^{-1}\xi_w, \quad d_2(w, \widehat{w}) = F(\widehat{w}) + \eta(w, \widehat{w}).$$

The variational inequality (3.2) is equivalent to

$$\textbf{\textit{Cond. 1.}} \quad \widehat{w} = P_{\mathcal{W}}\left\{\widehat{w} - [d_2(w, \widehat{w}) - Gd_1(w, \widehat{w})]\right\} \tag{3.3}$$

where $\mathcal{W} = R^n \times R^n \times R^n$.

Lemma 3.1. *$F(\widehat{w})$ is a monotone mapping with respect to $\mathcal{W}$.*

Proof: For given arbitrarily $w_1,\ w_2 \in \mathcal{W}$, by computing directly, we get

$$\begin{aligned}(w_1 - w_2)^T[F(w_1) - F(w_2)] &= \begin{pmatrix} x_1 - x_2 \\ y_1 - y_2 \\ \lambda_1 - \lambda_2 \end{pmatrix} \begin{pmatrix} M^TM(x_1 - x_2) - (\lambda_1 - \lambda_2) \\ N^TN(y_1 - y_2) + (\lambda_1 - \lambda_2) \\ (x_1 - x_2) - (y_1 - y_2) \end{pmatrix} \\ &= (x_1 - x_2)^T M^T M(x_1 - x_2) + (y_1 - y_2)^T N^T N(y_1 - y_2) \\ &= \|M(x_1 - x_2)\|^2 + \|N(y_1 - y_2)\|^2 \\ &\geq 0\end{aligned}$$

This proves Lemma 3.1.

Theorem 3.1. *Let $\widehat{w} = (\widehat{x}, \widehat{y}, \widehat{\lambda})$ be generated by the **in-PSALM-II** method, from a given $w = (x, y, \lambda)$. And let $w^* = (x^*, y^*, \lambda^*)$ be a solution of variational inequalities (2.5). Then we have*

$$\textbf{\textit{Cond. 2.}} \quad (\widehat{w} - w^*)^T d_2(w, \widehat{w}) \geq \varphi(w, \widehat{w}) - (w - \widehat{w})^T G d_1(w, \widehat{w}). \tag{3.4}$$

Proof: By using Lemma 3.1 and notation $d_2(w, \widehat{w})$ in (3.1), and w^* is a solution of (2.5), we have

$$\begin{aligned}(\widehat{w} - w^*)^T d_2(w, \widehat{w}) &= (\widehat{w} - w^*)^T [F(w, \widehat{w}) + \eta(w, \widehat{w})] \\ &\geq (\widehat{w} - w^*)^T F(w, w^*) + (\widehat{w} - w^*)^T \eta(w, \widehat{w}) \\ &\geq (\widehat{w} - w^*)^T \eta(w, \widehat{w}) \\ &= [(y - \widehat{y}) - (x - \widehat{x})]^T [\beta(\widehat{y} - y^*) - \beta(\widehat{x} - x^*)] \\ &= [\beta(\widehat{y} - \widehat{x}) - \beta(y^* - x^*)]^T [(y - \widehat{y}) - (x - \widehat{x})]\end{aligned}$$

Since $x^* - y^* = 0$ and $\widehat{\lambda} - \lambda = -\beta(\widehat{x} - \widehat{y})$, we get

$$(\widehat{w} - w^*)^T d_2(w, \widehat{w}) \geq (\widehat{\lambda} - \lambda)^T [(y - \widehat{y}) - (x - \widehat{x})]. \tag{3.5}$$

Using the notation $\varphi(w, \widehat{w})$ in (2.16) and inequality (3.5), we obtain (3.4).

Theorem 3.2. *Let $\widehat{w} = (\widehat{x}, \widehat{y}, \widehat{\lambda})$ be generated by the **in-PSALM-II** method, from a given $w = (x, y, \lambda)$. We have*

$$\varphi(w, \widehat{w}) \geq \frac{2 - \sqrt{2}}{2} \|d_1(w, \widehat{w})\|_G^2 + (\sqrt{2} - 1)(1 - \nu)[r\|x - \widehat{x}\|^2 + s\|y - \widehat{y}\|^2]. \tag{3.6}$$

Proof: Using (2.16), and the notations G and ξ_w, we get

$$\begin{aligned}\varphi(w, \widehat{w}) &= (w - \widehat{w})^T G(w - \widehat{w}) - (w - \widehat{w})^T \xi_w + (\lambda - \widehat{\lambda})^T (x - \widehat{x}) - (\lambda - \widehat{\lambda})^T (y - \widehat{y}) \\ &= (r + \beta)\|x - \widehat{x}\|^2 - (x - \widehat{x})^T \xi_x + (s + \beta)\|y - \widehat{y}\|^2 - (y - \widehat{y})^T \xi_y + \frac{1}{\beta}\|\lambda - \widehat{\lambda}\|^2 \\ &\quad + (\lambda - \widehat{\lambda})^T (x - \widehat{x}) + (\lambda - \widehat{\lambda})^T (\widehat{y} - y).\end{aligned} \tag{3.7}$$

Use Cauchy-Schwarz inequality, we get

$$(\lambda - \widehat{\lambda})^T (x - \widehat{x}) \geq -\frac{1}{2} \left(\sqrt{2}\beta \|x - \widehat{x}\|^2 + \frac{1}{\sqrt{2}\beta} \|\lambda - \widehat{\lambda}\|^2 \right) \tag{3.8}$$

$$(\lambda - \widehat{\lambda})^T(\widehat{y} - y) \geq -\frac{1}{2}\left(\sqrt{2}\beta\|y - \widehat{y}\|^2 + \frac{1}{\sqrt{2}\beta}\|\lambda - \widehat{\lambda}\|^2\right). \tag{3.9}$$

Substitute (3.8-3.9) into (3.7), we get

$$\begin{aligned}
\varphi(w, \widehat{w}) \quad &\geq \quad (r + \frac{2-\sqrt{2}}{2}\beta)\|x - \widehat{x}\|^2 - (x - \widehat{x})^T\xi_x && (3.10)\\
&\quad +(s + \frac{2-\sqrt{2}}{2}\beta)\|y - \widehat{y}\|^2 - (y - \widehat{y})^T\xi_y && (3.11)\\
&\quad +\frac{2-\sqrt{2}}{2}\frac{1}{\beta}\|\lambda - \widehat{\lambda}\|^2. && (3.12)
\end{aligned}$$

From (3.10), we get

$$\begin{aligned}
&(r + \frac{2-\sqrt{2}}{2}\beta)\|x - \widehat{x}\|^2 - (x - \widehat{x})^T\xi_x\\
= \quad &\frac{2-\sqrt{2}}{2}\left[((2+\sqrt{2})r + \beta)\|x - \widehat{x}\|^2 - (2+\sqrt{2})(x - \widehat{x})^T\xi_x\right]\\
= \quad &\frac{2-\sqrt{2}}{2}\left[(r+\beta)\|x - \widehat{x}\|^2 - 2(x - \widehat{x})^T\xi_x + r\|x - \widehat{x}\|^2\right] && (3.13)\\
&+\frac{2-\sqrt{2}}{2}\left[\sqrt{2}r\|x - \widehat{x}\|^2 - \sqrt{2}(x - \widehat{x})^T\xi_x\right]. && (3.14)
\end{aligned}$$

Use (2.12) to (3.13) and (3.14), we get

$$\begin{aligned}
&(r+\beta)\|x - \widehat{x}\|^2 - 2(x - \widehat{x})^T\xi_x + r\|x - \widehat{x}\|^2\\
= \quad &\frac{1}{r+\beta}\left[(r+\beta)^2\|x - \widehat{x}\|^2 - 2(r+\beta)(x - \widehat{x})^T\xi_x + r(r+\beta)\|x - \widehat{x}\|^2\right]\\
\geq \quad &\frac{1}{r+\beta}\left[(r+\beta)^2\|x - \widehat{x}\|^2 - 2(r+\beta)(x - \widehat{x})^T\xi_x + r^2\|x - \widehat{x}\|^2\right]\\
\geq \quad &\frac{1}{r+\beta}\left[(r+\beta)^2\|x - \widehat{x}\|^2 - 2(r+\beta)(x - \widehat{x})^T\xi_x + \|\xi_x\|^2\right]\\
= \quad &\frac{1}{r+\beta}\|(r+\beta)(x - \widehat{x}) - \xi_x\|^2 && (3.15)
\end{aligned}$$

and

$$\sqrt{2}r\|x - \widehat{x}\|^2 - \sqrt{2}(x - \widehat{x})^T\xi_x \geq \sqrt{2}(1-\nu)r\|x - \widehat{x}\|^2. \tag{3.16}$$

Substitute (3.15-3.16) into (3.13-3.14), then we have

$$\begin{aligned}
&(r + \frac{2-\sqrt{2}}{2}\beta)\|x - \widehat{x}\|^2 - (x - \widehat{x})^T\xi_x\\
\geq \quad &\frac{2-\sqrt{2}}{2}\frac{1}{r+\beta}\|(r+\beta)(x - \widehat{x}) - \xi_x\|^2 + (\sqrt{2}-1)(1-\nu)r\|x - \widehat{x}\|^2. && (3.17)
\end{aligned}$$

By the same method on (3.11), we get

$$(s+\frac{2-\sqrt{2}}{2}\beta)\|y-\widehat{y}\|^2-(y-\widehat{y})^T\xi_y \geq \frac{2-\sqrt{2}}{2}\frac{1}{s+\beta}\|(s+\beta)(y-\widehat{y})-\xi_y\|^2+(\sqrt{2}-1)(1-\nu)s\|y-\widehat{y}\|^2. \tag{3.18}$$

Substitute (3.17-3.18) into (3.10-3.11), and using the notation G, and use $\frac{1}{\beta}\|\lambda-\widehat{\lambda}\|^2 = \beta\|\frac{1}{\beta}(\lambda-\widehat{\lambda})\|^2$, we get

$$\begin{aligned}\varphi(w,\widehat{w}) &\geq \frac{2-\sqrt{2}}{2}\|Gd_1(w,\widehat{w})\|^2_{G^{-1}} + (\sqrt{2}-1)(1-\nu)[r\|x-\widehat{x}\|^2 + s\|y-\widehat{y}\|^2] \\ &= \frac{2-\sqrt{2}}{2}\|d_1(w,\widehat{w})\|^2_{G} + (\sqrt{2}-1)(1-\nu)[r\|x-\widehat{x}\|^2 + s\|y-\widehat{y}\|^2]\end{aligned}$$

This is (3.6). That completes the proof.

By Theorem 3.2, we get

$$\alpha_k^* \geq \frac{2-\sqrt{2}}{2}, \forall k. \tag{3.19}$$

Lemma 3.2. *For given w^k, $\widehat{w}^k$ is generated by in-PSALM-II method. Let w^* be a solution of variational inequality (2.5). Then we have*

$$(w^k - w^*)^T G d_1(w^k, \widehat{w}^k) \geq \varphi(w^k, \widehat{w}^k). \tag{3.20}$$

Proof: Here we omit this proof. For a similar proof see Lemma 3.1 in [18].

We are now already to give convergence of in-PSALM-II method. We can obtain convergence of this method directly, by using **cond. 1** and **cond. 2** and the theorems of the unified framework of proximal-like contraction methods for monotone variational inequalities [17]. Here we provide a brief proof on convergence.

Theorem 3.3. *Let $\{w^k\}$ be the sequence generated by the proposed method in-PSALM-II, and let w^* be a solution of variational inequality (2.5). Then we have*

$$\|w^{k+1} - w^*\|^2_G \leq \|w^k - w^*\|^2_G - \frac{3-2\sqrt{2}}{2}\gamma(2-\gamma)\|w^k - \widehat{w}^k\|^2_G, \tag{3.21}$$

where $0 < \gamma < 2$.

Proof: By using the iterative formulation (2.14), we get

$$\begin{aligned}\|w^{k+1} - w^*\|^2_G &= \|w^k - \alpha_k d_1(w^k, \widehat{w}^k) - w^*\|^2_G \\ &= \|w^k - w^*\|^2_G - 2\alpha_k (w^k - w^*)^T G d_1(w^k, \widehat{w}^k) + \alpha_k^2 \|d_1(w^k, \widehat{w}^k)\|^2_G.\end{aligned}$$

By Lemma 3.2, $(w^k - w^*)^T G d_1(w^k, \widehat{w}^k) \geq \varphi(w^k, \widehat{w}^k)$, and by using (2.15), we get

$$\begin{aligned}\|w^{k+1} - w^*\|^2_G &\leq \|w^k - w^*\|^2_G - 2\gamma\frac{\varphi^2(w^k, \widehat{w}^k)}{\|d_1(w^k, \widehat{w}^k)\|^2_G} + \gamma^2\frac{\varphi^2(w^k, \widehat{w}^k)}{\|d_1(w^k, \widehat{w}^k)\|^2_G} \\ &= \|w^k - w^*\|^2_G - \gamma(2-\gamma)\frac{\varphi(w^k, \widehat{w}^k)}{\|d_1(w^k, \widehat{w}^k)\|^2_G}\varphi(w^k, \widehat{w}^k) \\ &= \|w^k - w^*\|^2_G - \gamma(2-\gamma)\alpha_k^*\varphi(w^k, \widehat{w}^k).\end{aligned} \tag{3.22}$$

Use Theorem 3.2, we get

$$\varphi(w^k, \widehat{w}^k) \geq \frac{2-\sqrt{2}}{2}\|w^k - \widehat{w}^k\|^2_G.$$

Apply this inequality and (3.19) to (3.22), we obtain (3.21).

Theorem 3.3 and bounded property of the sequence $\{w^k\}$ (see Theorem 2.1 of [17]) provide convergence of the in-PSALM-II method.

4 A generalization of the proposed method

The in-PSALM-II method can be generalized in two ways: one is to generalize it from linear operators to general monotone operators, the other is from two operators to three operators. The first generalization is trivial, so in this section we give a generalization of the proposed method in handling variational inequalities with three separable monotone operators which has the following form:

$$\text{Find } u=(x,y,z)\in\Omega, \text{such that } \begin{cases} (x'-x)^T f(x)\geq 0 \\ (y'-y)^T g(y)\geq 0 \\ (z'-z)^T h(z)\geq 0 \end{cases} \quad \forall u'\in\Omega \tag{4.1}$$

where $\Omega=\{(x,y,z)|x\in\mathcal{X},y\in\mathcal{Y},z\in\mathcal{Z}:Px+Qy+Rz=d\}$, and $\mathcal{X}\subset R^{n_1},\mathcal{Y}\subset R^{n_2},\mathcal{Z}\subset R^{n_3}$ are convex sets. $P\in R^{m\times n_1},Q\in R^{m\times n_2},R\in R^{m\times n_3}$ are given matrices, $d\in R^m$ is a given vector, $f:R^{n_1}\to R^{n_1},g:R^{n_2}\to R^{n_2},h:R^{n_3}\to R^{n_3}$ are given monotone operators.

Problem (1.1) can be converted into a special case of Problem (4.1) via the following approach: First part the coefficient matrix A into three parts:

$$M=A(1:m,:),\ N=A(m+1:l,:),\ L=A(l+1:n,:),\ m<l<n,$$

and correspondingly,

$$p=b(1:m),\ q=b(m+1:l),\ v=b(l+1:n).$$

Then the system of LEs (1.1) can be educed by the following system:

$$\begin{cases} Mx=p \\ Nx=q \\ Lx=v \end{cases} \tag{4.2}$$

Furthermore, the system (4.2) is equivalent to

$$x^*=\text{Arg}\min_{x\in R^n}\|Mx-p\|^2+\|Nx-q\|^2+\|Lz-v\|^2. \tag{4.3}$$

A separable equivalent of (4.3) is:

$$\begin{array}{ll} \text{Arg}\min\limits_{x\in R^n} & \|Mx-p\|^2+\|Ny-q\|^2+\|Lz-v\|^2 \\ \text{s.t.} & x-y=0 \\ & y-z=0 \\ & z-x=0. \end{array} \tag{4.4}$$

Let

$$P=\begin{pmatrix} I \\ 0 \\ -I \end{pmatrix},\ Q=\begin{pmatrix} -I \\ I \\ 0 \end{pmatrix},\ R=\begin{pmatrix} 0 \\ -I \\ I \end{pmatrix}.$$

Problem (4.4) is equivalent to the following variational inequality:

$$\text{Find } u=(x,y,z)\in\Omega, \text{such that } \begin{cases} (x'-x)^T[M^T(Mx-p)]\geq 0 \\ (y'-y)^T[N^T(Ny-q)]\geq 0 \\ (z'-z)^T[L^T(Lz-v)]\geq 0 \end{cases} \quad \forall u'\in\Omega, \tag{4.5}$$

where $\Omega = \{(x, y, z)|x, y, z \in R^n : Px + Qy + Rz = 0\}$. This is the linear variational inequalities whose structure is as same as Problem (4.1).

We come back to the general case, namely Problem (4.1). Our goal of this section is to generalize in-PSALM-II for solving Problem (4.1), and prove its convergence.

By attaching a Lagrange multiplier vector $\lambda \in R^m$ to the linear constraint $Px + Qy + Rz = d$, Problem (4.1) can be explained as the following form:

$$\text{Find } w = \begin{pmatrix} x \\ y \\ z \\ \lambda \end{pmatrix} \in \mathcal{W}, \text{ such that } \begin{cases} (x' - x)^T[f(x) - P^T\lambda] \geq 0, \\ (y' - y)^T[g(y) - Q^T\lambda] \geq 0, \\ (z' - z)^T[h(z) - R^T\lambda] \geq 0, \\ Px + Qy + Rz - d = 0, \end{cases} \forall w' \in \mathcal{W} \quad (4.6)$$

where $\mathcal{W} = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \times R^m$.

For the problems in this section, we denote

$$H = \begin{pmatrix} 2\beta P^TP & \beta P^TQ & \beta P^TR & \\ \beta Q^TP & 2\beta Q^TQ & \beta Q^TR & \\ \beta R^TP & \beta R^TQ & 2\beta R^TR & \\ & & & \frac{1}{\beta}I \end{pmatrix}$$

In general, H is positive semi-definite, and thus when H is not positive definite $\|\cdot\|_H$ is not a norm. For convenience, we use the notation $\|w-\widehat{w}\|_H^2$ to denote $(w-\widehat{w})^TH(w-\widehat{w})$. We know $\|w - \widehat{w}\|_H^2 \geq 0$, see equation (4.6) in [16].

We are now ready to propose the new method for solving Problem (4.6).

The extended method (in-PSALM-III):

`Step` 0. Given the parameters $\varepsilon > 0, \gamma \in (0, 2), 0 < \beta_L \leq \beta \leq \beta_U, \nu \in (0, 1)$, initial parameters r_0, s_0, t_0 and initial point $w^0 = (x^0, y^0, z^0, \lambda^0) \in R^n \times R^n \times R^n \times R^m$. Set $k = 0$, let $\beta_k = \beta$.

`Step` 1. Produce $(\widehat{x}^k, \widehat{y}^k, \widehat{z}^k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ via solving parallel the following variational inequalities:

$$x \in \mathcal{X}, (x'-x)^T \left\{ f(x) - P^T[\lambda^k - \beta_k(Px + Qy^k + Rz^k - d)] + r_k(x - x^k) + \xi_x^k \right\} \geq 0, \forall x' \in \mathcal{X} \quad (4.7)$$

where r_k and ξ_x^k satisfy the following inequality

$$\|\xi_x^k\| \leq \nu r_k \|x^k - \widehat{x}^k\|, \quad \xi_x^k = f(x^k) - f(\widehat{x}^k) + \beta_k P^TP(x^k - \widehat{x}^k); \quad (4.8)$$

and

$$y \in \mathcal{Y}, (y'-y)^T \left\{ g(y) - Q^T[\lambda^k - \beta_k(Px^k + Qy + Rz^k - d)] + s_k(y - y^k) + \xi_y^k \right\} \geq 0, \forall y' \in \mathcal{Y} \quad (4.9)$$

where s_k and ξ_y^k satisfy the following inequality

$$\|\xi_y^k\| \leq \nu s_k \|y^k - \widehat{y}^k\|, \quad \xi_y^k = g(y^k) - g(\widehat{y}^k) + \beta_k Q^TQ(y^k - \widehat{y}^k); \quad (4.10)$$

and

$$z \in \mathcal{Z}, (z'-z)^T \left\{ h(z) - R^T[\lambda^k - \beta_k(Px^k + Qy^k + Rz - d)] + t_k(z - z^k) + \xi_z^k \right\} \geq 0, \forall z' \in \mathcal{Z} \quad (4.11)$$

where t_k and ξ_z^k satisfy the following inequality

$$\|\xi_z^k\| \leq \nu t_k \|z^k - \widehat{z}^k\|, \quad \xi_z^k = h(z^k) - h(\widehat{z}^k) + \beta_k R^T R(z^k - \widehat{z}^k). \tag{4.12}$$

`Step` 3. Update $\widehat{\lambda}^k$ by using the following formula:

$$\widehat{\lambda}^k = \lambda^k - \beta_k(P\widehat{x}^k + Q\widehat{y}^k + R\widehat{z}^k - d). \tag{4.13}$$

`Step` 4. Check termination condition: let

$$\varepsilon_k = \max(\|x^k - \widehat{x}^k\|, \|y^k - \widehat{y}^k\|, \|z^k - \widehat{z}^k\|, \|\lambda^k - \widehat{\lambda}^k\|),$$

if $\varepsilon_k < \varepsilon$, stop. Else goto next step.
`Step` 5. Produce the new iterative point $w^{k+1} = (x^{k+1}, y^{k+1}, z^{k+1}, \lambda^{k+1})$ by using the following formula:

$$w^{k+1} = w^k - \alpha_k d_1(w^k, \widehat{w}^k). \tag{4.14}$$

Let $k := k + 1$, go to Step 1.
Here,

$$d_1(w^k, \widehat{w}^k) = \begin{pmatrix} r_k(x^k - \widehat{x}^k) - \xi_x^k \\ s_k(y^k - \widehat{y}^k) - \xi_y^k \\ t_k(z^k - \widehat{z}^k) - \xi_z^k \\ 0 \end{pmatrix} + H(w^k - \widehat{w}^k), \quad \beta = \beta_k \text{ in matrix } H$$

and $\alpha_k = \gamma\alpha_k^*$,

$$\alpha_k^* = \frac{\varphi(w^k, \widehat{w}^k)}{\|d_1(w^k, \widehat{w}^k)\|^2}, \tag{4.15}$$

$$\varphi(w^k, \widehat{w}^k) = 2(\lambda^k - \widehat{\lambda}^k)^T[P(x^k - \widehat{x}^k) + Q(y^k - \widehat{y}^k) + R(z^k - \widehat{z}^k)] + (w^k - \widehat{w}^k)^T d_1(w^k, \widehat{w}^k). \tag{4.16}$$

Conveniently, we ignore the index variable k and the following notations are useful in this section.

$$F(\widehat{w}) = \begin{pmatrix} f(\widehat{x}) - P^T\widehat{\lambda} \\ g(\widehat{y}) - Q^T\widehat{\lambda} \\ h(\widehat{z}) - R^T\widehat{\lambda} \\ P\widehat{x} + Q\widehat{y} + R\widehat{z} - d \end{pmatrix}, \xi_w = \begin{pmatrix} \xi_x \\ \xi_y \\ \xi_z \\ 0 \end{pmatrix}$$

$$K = \begin{pmatrix} \beta P^T P & \beta P^T Q & \beta P^T R \\ \beta Q^T P & \beta Q^T Q & \beta Q^T R \\ \beta R^T P & \beta R^T Q & \beta R^T R \end{pmatrix},$$

$$u = (x, y, z)^T, \ \eta(w, \widehat{w}) = \begin{pmatrix} K(u - \widehat{u}) \\ 0 \end{pmatrix}.$$

Let

$$d_2(w, \widehat{w}) = F(\widehat{w}) + 2\eta(w, \widehat{w}).$$

Then variational inequalities (4.7, 4.9, 4.11, 4.13) can be rewritten in a compact form:

$$(w' - \widehat{w})^T[d_2(w, \widehat{w}) - d_1(w, \widehat{w})] \geq 0, \forall w' \in \mathcal{W}. \tag{4.17}$$

Variational inequality (4.17) is equivalent to the following projection equation,

$$\textbf{\textit{cond. 1.}} \quad \widehat{w} = P_{\mathcal{W}}\left\{\widehat{w} - [d_2(w,\widehat{w}) - d_1(w,\widehat{w})]\right\}. \tag{4.18}$$

Theorem 4.1. *Let $w^* = (x^*, y^*, z^*, \lambda^*)$ be a solution of Problem (4.17), and $\widehat{w}$ be generated by in-PSALM-III method for a given w. Then we have*

$$\textbf{\textit{cond. 2.}} \quad (\widehat{w} - w^*)^T d_2(w,\widehat{w}) \geq \varphi(w,\widehat{w}) - (w - \widehat{w})^T d_1(w,\widehat{w}). \tag{4.19}$$

Proof: Since $F(w)$ is monotone (by using monotonicity of operators f, g, h) and w^* is a solution of Problem (4.17) we have $(\widehat{w} - w^*)^T F(w^*) \geq 0$, and

$$\begin{aligned}(\widehat{w} - w^*)^T d_2(w,\widehat{w}) &= (\widehat{w} - w^*)^T[F(\widehat{w}) + 2\eta(w,\widehat{w})] \\ &\geq (\widehat{w} - w^*)^T[F(w^*) + 2\eta(w,\widehat{w})] \\ &\geq 2(\widehat{w} - w^*)^T\eta(w,\widehat{w}).\end{aligned} \tag{4.20}$$

But

$$\begin{aligned}&(\widehat{w} - w^*)^T\eta(w,\widehat{w}) \\ =\ & [\beta P(\widehat{x} - x^*) + \beta Q(\widehat{y} - y^*) + \beta R(\widehat{z} - z^*)]^T[P(x - \widehat{x}) + Q(y - \widehat{y}) + R(z - \widehat{z})] \\ =\ & [\beta(P\widehat{x} + Q\widehat{y} + R\widehat{z} - d) - \beta(Px^* + Qy^* + Rz^* - d)]^T \\ & [P(x - \widehat{x}) + Q(y - \widehat{y}) + R(z - \widehat{z})] \\ =\ & [\lambda - \widehat{\lambda}]^T[P(x - \widehat{x}) + Q(y - \widehat{y}) + R(z - \widehat{z})]\end{aligned} \tag{4.21}$$

in the last equation we use (4.13) and $Px^* + Qy^* + Rz^* - d = 0$. Substitute (4.21) into (4.20) and using the notation $\varphi(w,\widehat{w})$, we get (4.19).

By the inequality (4.17) of Lemma 4.1 in [16], we have

$$2[\lambda - \widehat{\lambda}]^T[P(x - \widehat{x}) + Q(y - \widehat{y}) + R(z - \widehat{z})] \geq -\frac{\sqrt{3}}{2}\|w - \widehat{w}\|_H^2. \tag{4.22}$$

Then we have:

Theorem 4.2. *For given $w = (x, y, z, \lambda)$, let $\widehat{w} = (\widehat{x}, \widehat{y}, \widehat{z}, \widehat{\lambda})$ be generated by in-PSALM-III method. Then we have*

$$\varphi(w,\widehat{w}) \geq \frac{2 - \sqrt{3}}{2}\|w - \widehat{w}\|_H^2 + (1 - \nu)[r\|x - \widehat{x}\|^2 + s\|y - \widehat{y}\|^2 + t\|z - \widehat{z}\|^2]. \tag{4.23}$$

Proof: Use (4.16) and (4.22), we get

$$\begin{aligned}\varphi(w,\widehat{w}) &\geq (w - \widehat{w})^T d_1(w,\widehat{w}) - \frac{\sqrt{3}}{2}\|w - \widehat{w}\|_H^2 \\ &= \|w - \widehat{w}\|_H^2 - \frac{\sqrt{3}}{2}\|w - \widehat{w}\|_H^2 + (w - \widehat{w})^T \begin{pmatrix} r(x - \widehat{x}) - \xi_x \\ s(y - \widehat{y}) - \xi_y \\ t(z - \widehat{z}) - \xi_z \\ 0 \end{pmatrix} \\ &= \frac{2 - \sqrt{3}}{2}\|w - \widehat{w}\|_H^2 + r\|x - \widehat{x}\|^2 - (x - \widehat{x})^T\xi_x \\ &\quad + s\|y - \widehat{y}\|^2 - (y - \widehat{y})^T\xi_y + t\|z - \widehat{z}\|^2 - (z - \widehat{z})^T\xi_z.\end{aligned} \tag{4.24}$$

By Cauchy-Schwarz inequality,

$$-(x-\widehat{x})^T\xi_x \geq \|x-\widehat{x}\|\|\xi_x\|,$$

and by (4.8), we get

$$-(x-\widehat{x})^T\xi_x \geq -\nu r\|x-\widehat{x}\|^2. \tag{4.25}$$

In the same way, we have

$$-(y-\widehat{y})^T\xi_y \geq -\nu s\|y-\widehat{y}\|^2, \quad -(z-\widehat{z})^T\xi_z \geq -\nu t\|y-\widehat{y}\|^2. \tag{4.26}$$

Substitute (4.25, 4.26) into (4.24), we get (4.23).

Lemma 4.1. *Let w^* be a solution of Problem (4.17). For given w, let $\widehat{w}$ be generated by in-PSALM-III method. Then we have*

$$(w-w^*)^T d_1(w,\widehat{w}) \geq \varphi(w,\widehat{w}). \tag{4.27}$$

Proof: Use (4.17), let $w' = w^*$, we get

$$(\widehat{w}-w^*)^T d_1(w,\widehat{w}) \geq (\widehat{w}-w^*)^T d_2(w,\widehat{w}). \tag{4.28}$$

Add(4.28) and (4.19), we have

$$(\widehat{w}-w^*)^T d_1(w,\widehat{w}) \geq \varphi(w,\widehat{w}) - (w-\widehat{w})^T d_1(w,\widehat{w}).$$

Re-arranging the last inequality we get (4.27) directly.

We are ready to give the key theorem of convergence of in-PSALM-III method.

Theorem 4.3. *Let w^* be a solution of Problem (4.17). Let the sequence $\{w^k\}$ and $\{\widehat{w}^k\}$ be generated by in-PSALM-III method. Then we have*

$$\|w^{k+1}-w^*\|^2 \leq \|w^k-w^*\|^2. \tag{4.29}$$

The equation $\|w^{k+1}-w^*\|^2 = \|w^k-w^*\|^2$ holds if and only if $\|w^k-\widehat{w}^k\| = 0$.

Proof: By the iterative formula (4.14), we have

$$\begin{aligned}
\|w^{k+1}-w^*\|^2 &= \|w^k-w^*-\alpha_k d_1(w,\widehat{w})\|^2 \\
&= \|w^k-w^*\|^2 - 2\alpha_k(w^k-w^*)^T d_1(w,\widehat{w}) + \alpha_k^2\|d_1(w,\widehat{w})\|^2 \\
&\leq \|w^k-w^*\|^2 - 2\alpha_k\varphi(w,\widehat{w}) + \alpha_k^2\|d_1(w,\widehat{w})\|^2 \qquad (4.30)
\end{aligned}$$

In last inequality we use Lemma 4.1. Note $\alpha_k = \gamma\alpha_k^*$ ($\gamma \in (0,2)$) and (4.15), we get

$$\begin{aligned}
\|w^{k+1}-w^*\|^2 &\leq \|w^k-w^*\|^2 - 2\gamma\frac{\varphi^2(w,\widehat{w})}{\|d_1(w,\widehat{w})\|^2} + \gamma^2\frac{\varphi^2(w,\widehat{w})}{\|d_1(w,\widehat{w})\|^2} \\
&= \|w^k-w^*\|^2 + \gamma(2-\gamma)\frac{\varphi^2(w,\widehat{w})}{\|d_1(w,\widehat{w})\|^2} \qquad (4.31) \\
&\leq \|w^k-w^*\|^2. \qquad (4.32)
\end{aligned}$$

By using (4.23), $\varphi(w,\widehat{w}) = 0$ iff $\|w^k-\widehat{w}^k\| = 0$, so the second term of (4.31) is equal to zero iff $\|w^k-\widehat{w}^k\| = 0$. This implies the second part of Theorem 4.3.

5 Numerical results and conclusions remarks

In this section, we employ three examples to show applicability, availability and robustness of the proposed methods.

Example 5.1. For Problem (1.1), let A be $n \times n$ Hilbert matrix, $x^* = (1, ..., 1)^T$ and $b = Ax^*$. We solve this problem by using in-PSALM-II method with stopping criterion $\varepsilon = 1.0 \times 10^{-5}$ (correspondingly residual $\|Ax - b\|^2 < 1.0 \times 10^{-2}$). We observe the relationship of variables n and iterations k, and obtain the result described by Figure 1.

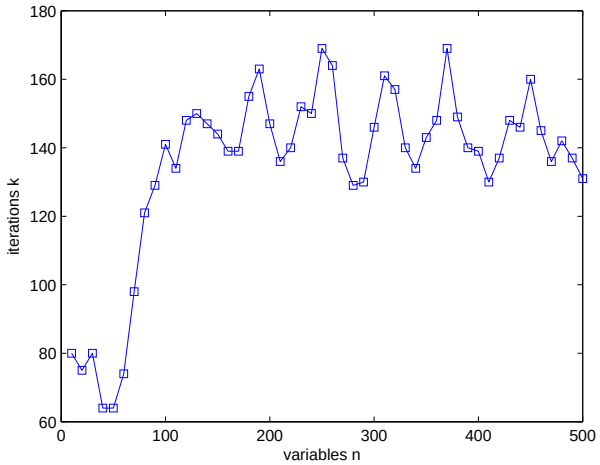

Figure 1: Iterations vs variables when A is Hilbert matrix

Example 5.2. For Problem (1.1), let $A = (a_{ij})$ be a 100×100 positive definite matrix with given condition number, which components a_{ij} are random variables, the right hand vector b and stopping criterion are as same as example 5.1. The Matlab Program for generating A is given in the following box (let $n = 100$):

```
Sigma=zeros(n,1); for j=1:n Sigma(j)=cos((pi*j)/(n+1)) +1; end;
Hcond=input('Input condition number of A, suggest 10^2 – 10^18, Hcond= ');
tau=(max(Sigma)- min(Sigma)*Hcond)/(Hcond-1);
Sigma=Sigma + ones(n,1)*tau; v=zeros(n,1); xop=zeros(n,1); xstart=zeros(n,1);
t=0; for j=1:n t=mod(t*31416+13846,46261); v(j)=t*(2/46261)-1; end;
v=v/norm(v); V = eye(n)- v*v'*2; H = V * diag(Sigma) * V; A=H;
```

We observe the relationship of condition number of A and iterations k, and obtain the result described by Figure 2.

Example 5.3. This example is as same as the second example of [5]. For completeness, we state it here. This system of linear equations is obtained from the discretization of a one dimensional two point boundary value problem of the Poisson equation and has the form

$$A = \text{tridiag}(-1, 2, -1), \quad b = (1, 0, ..., 0, 1)^T.$$

For the stopping criterion we set $\varepsilon < 1.0 \times 10^{-7}$ (correspondingly residual $\|Ax - b\|^2 < 1.0 \times 10^{-2}$). We observe the relationship of variables n and iterations k, and obtain the

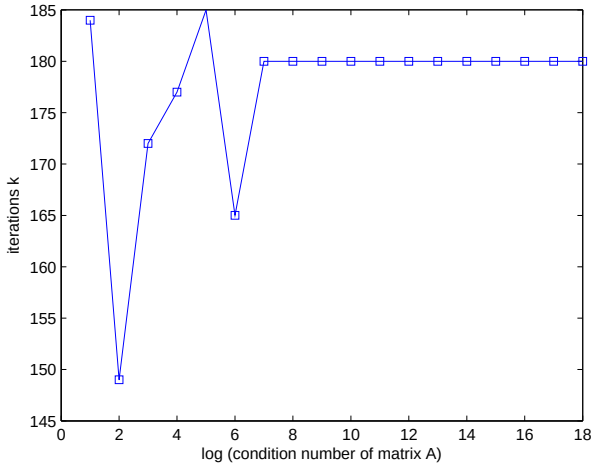

Figure 2: Iterations *vs* condition number when A is random positive definite matrix

result described by Figure 3. For a comparison, one can refer to table 1 of [5].

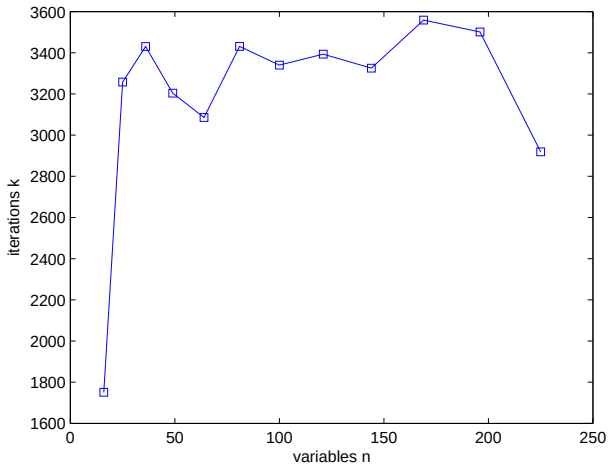

Figure 3: Iterations *vs* variables when A is a special tridiagonal matrix

We see from the numerical results that the superiority of the proposed method (in-PSALM-II) is robust. The numerical results suggest iterations to dimension is not too sensitive (see figure 1 and figure 3); and iterations to condition number is also not too sensitive (see figure 2). compare with table 1 of [5] and example 5.3, we see that the proposed method is efficient and more robust than the enumerative methods in table 1 of [5].

Conclusions: In this paper, we propose a novel method to solve a large system of linear equations, we have the inexact parallel splitting augmented Lagrangian method. The numerical results show the applicability, availability and robustness of the proposed method. We also give the proof of convergence of the proposed method and extend it to deal with general monotone variational inequalities problem with three separable

operators. Since we introduce into proper inexact terms, it steers clear of the inverse operator on the matrix when we employ this method to solve system of linear equations (LEs). Theoretically, to do in this way, the most expensive computational loads is the product of matrix and vector, see (2.17) and (2.18). Thus the complexity of this method is $O(n^2)$. This compare with the common iterative method with complexity $O(n^3)$. On the other hand, the total cost of memory in processors does not increase, since in each subproblem, we only need to store a partial matrix of coefficient matrix A, i.e., M or N. From the practical point of view, the most useful value of the proposed method is robustness with respect to the scale (dimension) and condition (measured by condition number of the coefficient matrix A).

However, as affirmed in [16], when $p > 3$ we can not extend this method to deal with the same problem with p separable operators directly. But on the parallel computational point of view, to handle problem with p ($p > 3$) separable operators is necessary. We guess one can handle this problem in this class by carving up these operators into several groups. Two or three operators are in one group, and two or three groups are in one super-group, and so on. Then we can use the proposed method in this paper to handle this problem effectively. This investigation is in working in our research group.

Acknowledgement: We are grateful to Professor BS He , for his help in this research. We also thank the anonymous reviewers for their helpful comments.

References

[1] G. Allaire, S. M. Kaber, *Numerical linear algorithm*, (Translated by K. Trabelsi), Springer Science+Business Media, LLC, 2008

[2] C. Brezinski, Projection methods for linear systems, *Journal of Computational and Applied Mathematics*, 77 (1997), 35-51

[3] R. Bagnara, A unified proof for the convergence of Jacobi and Gauss-Seidel methods, *SIAM Review*, 37(1)(1995), 93-97

[4] Z.Z. Bai, D.R. Wang, D.J. Evans, Models of asynchronous parallelmatrix multisplitting relaxed iterations, *Parallel Comput.* 21 (1995) 565-582.

[5] Z.Z. Bai, A class of parallel decomposition-type relaxation methods for large sparse systems of linear equations, *Linear Algebra and its Applications,* 282 (1998), 1-24

[6] Z.Z. Bai, G. H. Golub, M. K. Ng, On inexact Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems, *Linear Algebra and its Applications,* 428 (2008) 413–440

[7] Z.Z. Bai, On the convergence of additive and multiplicative splitting iterations for systems of linear equations, *Journal of Computational and Applied Mathematics,* 154 (2003) 195–214

[8] Z.Z. Bai, On the convergence of parallel nonstationary multisplitting iteration methods, *Journal of Computational and Applied Mathematics,* 159 (2003) 1–11

[9] A. Frommer, G. Mayer, Convergence of relaxed parallel multisplitting methods, *Linear Algebra and its Applications* 119 (1989) 141-152.

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

[11] S. Chandrasekaran, M. Gu, X. Sun, J. Xia, and J. Zhu, A superfast algorithm for Toeplitz systems of linear equations, *SIAM J. matrix Anal. Appl.*, 29(4)(2007), 1247–1266.

[12] M. Gu, Stable and efficient algorithms for structured systems of linear equations, *SIAM J. matrix Anal. Appl.*, 19(2)(1998), 279-306.

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

[14] B.S. He, A new method for a class of linear variational inequalities, *Mathematical Programming*, 66 (1994) 137-144.

[15] B.S. He, Solution and applications of a class of general linear variational inequalities, *Science in China, Series A*, 39 (1996), 395-404.

[16] B.S. He, Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities, *Computation optimization and applicactions*, 42 (2009), 195-212.

[17] B.S. He , Liao Lizhi and Wang Xiang, Proximal-like contraction methods for monotone variational inequalities in a unified framework, Optimization Online.

[18] B.S. He and M.H. Xu, A general framework of contraction methods for monotone variational inequalities, *Pacific journal of optimization*, 4(2) (2008) 195-212.

[19] A. Hadjidimos, Accelerated over-relaxation method, *Math. Comp.* 32 (1978) 149-157.

[20] L.J. Lin, Y.M. Wei, N.M. Zhang, Convergence and quotient convergence of iterative methods for solving singular linear equations with index one, *Linear Algebra and its Applications* 430 (2009) 1665-1674.

[21] M. Neumann, R.J. Plemmons, Convergence of parallel multisplitting iterative methods for M-matrices, *Linear Algebra and its Applications*, 88 (1987) 559-573.

[22] C.D. Meyer, *Matrix analysis and applied linear algebra*, SIAM Press, 2000.

[23] D.K. Salkuyeh, Generalized Jacobi and Gauss-Seidel Methods for Solving Linear ystem of Equations, *Numer. Math. J. Chinese Univ. (English Ser.),* 16(2) (2007), 164-170.

[24] L. Wang, Z.Z. Bai, Convergence conditions for splitting iteration methods for non-Hermitian linear systems, *Linear Algebra and its Applications,* 428 (2008), 453–468.

[25] C.L. Wang, T.Z. Huang, New convergence results for alternating methods, *Journal of Computational and Applied Mathematics,* 135 (2001), 325–333.

[26] D.R. Wang, Z.Z. Bai, D.J. Evans, A class of asynchronous parallel matrix multisplitting relaxation methods, *Parallel Algorithms Appl.* 2 (1994) 173–192.

[27] Y. Gao, Solving symmetric monotone linear variational inequalities by some modified Levitin-Polyak projection methods, *Applied Mathematics-A Journal of Chinese Universities*, 13(2)(1998), 150–158.

[28] X.D. Zhang, *Matrix analysis and applications*, Tsinghua University Press & Springer, 2004.