Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol III,
IMECS 2010, March 17 - 19, 2010, Hong Kong

An Inexact Proximal Alternating Directions Method For Structured Variational Inequalities*

Zheng Peng[1,2†] Donghua Wu[2]

Abstract—**In this paper we propose an appealing inexact proximal alternating directions method (abbreviated as In-PADM) for solving a class of monotone variational inequalities with certain special structure, and this structure under consideration is common in practice. We prove convergence of In-PADM method while the inexact term is arbitrary but satisfied some suitable conditions. For solving the variational inequalities with the special structure under consideration, we prove the proximal point term in one of the variational inequalities can be taken out, while convergence of the proposed method is preserved. And to do in this way, some advantages are provided in the implementation of In-PADM method. Numerical tests on Compressed Sensor problem show applicability and availability of In-PADM method.**

Keywords: Structured variational inequalities; monotonicity; Inexact proximal alternating directions method; convergence; tractability

1 Introduction

In this paper, we consider the following problem:

$$\begin{array}{ll} \text{minimize} & \theta_1(x)+\theta_2(y), \\ \text{subject to} & Ax-y=0. \end{array} \tag{1}$$

Where $x \in \mathcal{X}, y \in \mathcal{Y}$, $\mathcal{X}$ and $\mathcal{Y}$ are closed convex subset of R^n and R^m, respectively. $A \in R^{m\times n}$ is a given constant matrix. The functions θ_1 and θ_2 are closed proper convex functions on $\mathcal{X}$ and $\mathcal{Y}$ respectively. We assume the subgradient information of θ_1 and θ_2 is available.

There is a wide range of applications of Problem (1). For example, the following problem which is referred to as Compressed Sensor problem, is one of the research hotspot in image processing (for examples, see [7, 12], ect.)

$$\min \|Dx-b\|_2^2+\mu\|x\|_1, \tag{2}$$

where $x \in R^n, D \in R^{m\times n}$, $b \in R^m$, $\mu > 0$, $m < n$. Problem (2) can be reformulated to the following separable form:

$$\begin{array}{ll} \min & \frac{1}{2}\|Dx-b\|_2^2+\mu\|y\|_1, \\ \text{s.t.} & x-y=0, \end{array} \tag{3}$$

where $x, y \in R^n$, $D \in R^{m\times n}$, $b \in R^m$, $\mu > 0$. It is obvious that Problem (3) is a special case of Problem (1), with $\theta_1(x) = \dfrac{1}{2}\|Dx-b\|_2^2$, $\theta_2(y) = \mu\|y\|_1$, $A = I$, and $\mathcal{X} = \mathcal{Y} = R^n$.

Let $f(x) = \partial\theta_1(x)$ and $g(y) = \partial\theta_2(y)$ be the sub-gradient of $\theta_1(x)$ and $\theta_2(x)$ respectively. By convexity of θ_1 and θ_2, we know that f and g are monotone with respect to $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then Problem (1) is equivalent to the following monotone structured variational inequality

$$\text{find } (x,y) \in \Omega, \quad \begin{cases} (x'-x)^T f(x) \ge 0, \\ (y'-y)^T g(y) \ge 0, \end{cases} \quad \forall (x',y') \in \Omega, \tag{4}$$

where

$$\Omega = \{(x,y) | x \in \mathcal{X}, y \in \mathcal{Y} : Ax-y=0\}. \tag{5}$$

Chen and Teboulle [6] investigated Problem (1) and proposed a proximal-based decomposition method for solving it. Tseng [10] interpreted Chen&Teboulles' approach as an alternating version of the proximal point method and the extragradient method.

There are many methods to deal with the equivalent version (4–5) of Problem (1). Proximal point method and alternating directions method are two of the power tools for this problem. For examples, see [1, 2, 3, 5, 9, 11].

By attaching a Lagrangian multiplier $\lambda \in R^m$ to the linear constraint $Ax - y = 0$, Problem (4–5) can be reformulated into the following equivalent form:

$$w \in \mathcal{W}, \quad \begin{cases} (x'-x)^T[f(x)-A^T\lambda] \ge 0, \\ (y'-y)^T[g(y)+\lambda] \ge 0, \\ (\lambda'-\lambda)^T(Ax-y) \ge 0. \end{cases} \quad \forall w' \in \mathcal{W} \tag{6}$$

where $w = (x, y, \lambda)$ and $\mathcal{W} = \mathcal{X} \times \mathcal{Y} \times R^m$.

We make some standard assumptions on Problem (6):
Assumption A. The functions $f(x)$ and $g(y)$ are monotone with respect to $\mathcal{X}$ and $\mathcal{Y}$ respectively.

*This work is supported by Jiangsu Postdoctoral Research Foundation (0901020), Shanghai Nature Science Foundation (09ZR1411100) and Key Disciplines of Shanghai Municipality (S30104).

†1. Department of Mathematics, NanJing University, NanJing China, 210093; 2. Department of Mathematics, Shanghai University, Shanghai China, 200444. Email: pzheng008@yahoo.com.

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Assumption B. The solution set of Problem (6), denoted by $\mathcal{W}^*$, is nonempty.

From a given triplet $w^k = (x^k, y^k, \lambda^k)$, the primary proximal alternating directions method uses the following procedure to carry out the new iterate triplet $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ at each iteration:
1). Find $\hat{x}^k$ by solving variational inequality:

$$(x' - \hat{x})^T \Big\{ f(\hat{x}) - A^T[\lambda^k - \beta_k(A\hat{x} - y^k)] + r_k(\hat{x} - x^k) \Big\} \geq 0. \quad (7)$$

2). Find $\hat{y}^k$ by solving variational inequality:

$$(y' - \hat{y})^T \{ g(\hat{y}) + [\lambda^k - \beta_k(A\hat{x}^k - \hat{y})] + s_k(\hat{y} - y^k) \} \geq 0. \quad (8)$$

3). Update $\hat{\lambda}^k$ via

$$\hat{\lambda}^k = \lambda^k - \beta_k(A\hat{x}^k - \hat{y}^k). \quad (9)$$

Where β_k is a given penalty parameter of the linear constraint $Ax - y = 0$. The coefficients $r_k > 0$ in formulas (7) and $s_k > 0$ in (8) are referred to as proximal parameters. The method is convergent by setting $w^{k+1} = \hat{w}^k$ (for a proof see [1]).

But to solve subproblems (7) and (8) exactly is not an easy task, since each of them requires an implicit projection. He and Liao et al [2] suggested a method for solving subproblems (7) and (8) inexactly. This method is referred to as alternating projection based prediction-correction methods (abbreviated as APBPCM).

Essentially, to solve the following structured variational inequality

$$w \in \mathcal{W}, \quad \left\{ \begin{array}{ll} (x' - x)^T[f(x) - A^T\lambda] \geq 0, & \\ (y' - y)^T[g(y) - B^T\lambda] \geq 0, & \forall w' \in \mathcal{W}, \\ Ax + By = b, & \end{array} \right. \quad (10)$$

from a given triplet $w^k = (x^k, y^k, \lambda^k)$, the APBPCM method carries out the new iterate triplet $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ via the following two phases.
Prediction phase: Let H be a given positive matrix. The predictor $\hat{w}^k = (\hat{x}^k, \hat{y}^k, \hat{\lambda}^k)$ is generated by the following procedure:
1). Set

$$\hat{x}^k = P_{\mathcal{X}}\Big\{ x^k - \frac{1}{r_k}\big(f(x^k) - A^T[\lambda^k - H(Ax^k + By^k - b)] \big) \Big\}, \quad (11)$$

where $r_k > 0$ is a chosen parameter such that

$$\|\xi_x^k\| \leq \nu r_k \|x^k - \hat{x}^k\|, \xi_x^k := f(x^k) - f(\hat{x}^k) + A^T HA(x^k - \hat{x}^k). \quad (12)$$

2). Set

$$\hat{y}^k = P_{\mathcal{Y}}\Big\{ y^k - \frac{1}{s_k}\big(g(y^k) - B^T[\lambda^k - H(A\hat{x}^k + By^k - b)] \big) \Big\}, \quad (13)$$

where $s_k > 0$ is a chosen parameter such that

$$\|\xi_y^k\| \leq \nu s_k \|y^k - \hat{y}^k\|, \xi_y^k := g(x^k) - g(\hat{x}^k) + B^T HB(y^k - \hat{y}^k). \quad (14)$$

3). Update $\hat{\lambda}^k$ via

$$\hat{\lambda}^k = \lambda^k - H(A\hat{x}^k + B\hat{y}^k - b). \quad (15)$$

Denote $R_k = r_k I$ and $S_k = s_k I$, set

$$M_k = S_k + B^T HB, \quad (16)$$

$$G_k = \begin{pmatrix} R_k & 0 & 0 \\ 0 & M_k & 0 \\ 0 & 0 & H^{-1} \end{pmatrix}, \quad \xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix}, \quad (17)$$

and let

$$d(w^k, \tilde{w}^k, \xi^k) := (w^k - \hat{w}^k) - G_k^{-1}\xi^k. \quad (18)$$

Correction phase: The new iterate triplet $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ is accepted by using one of the following two strategies:

$$\text{I}). \ w_I^{k+1} = w^k - \alpha_k d(w^k, \tilde{w}^k, \xi^k), \quad (19)$$

$$\text{II}). \ w_{II}^{k+1} = P_{\mathcal{W}, G_k}\{ w^k - \alpha_k G_k^{-1} q(w^k, \hat{w}^k) \}, \quad (20)$$

where

$$q(w^k, \hat{w}^k) = Q(\hat{w}^k) + (A, B, 0)^T HB(y^k - \hat{y}^k), \quad (21)$$

$$Q(w) = \begin{pmatrix} f(x) - A^T\lambda \\ g(y) - B^T\lambda \\ Ax + By - b \end{pmatrix}. \quad (22)$$

Both correction strategies have the same step-size:

$$\alpha_k = \gamma \alpha_k^*, \quad (23)$$

where $\gamma \in (0, 2)$, and

$$\alpha_k^* = \frac{(\lambda^k - \hat{\lambda}^k)^T B(y^k - \hat{y}^k) + (w^k - \hat{w}^k) G_k d(w^k, \tilde{w}^k, \xi^k)}{\|d(w^k, \tilde{w}^k, \xi^k)\|^2}. \quad (24)$$

The problem under consideration in this paper is a special case of Problem (10) obviously. We have $B = -I$ and $b = 0$ in Problem (1). And this case is common in practice. The condition $B = -I$ provides some advantages for solving this problem, which is the proximal parameter s_k in (8) can be equal to zero and convergence of the proposed method is preserved. But in the previous works, for examples, [1], [2], [12] and [10], the authors have not made use of this condition. We also prove theoretically that, for any inexact term ξ^k, under suitable conditions we have convergence of the proposed method. This is a converse result of that one in [2].

This paper is organized as follows: In Section 2, we describe the proposed method and give some useful notations. In Section 3, we provide two descent directions and prove some contractive properties of the proposed method. In Section 4, we prove convergence of the proposed method under suitable assumptions. In Section 5, we give the details of the implementation of the proposed method. Numerical tests on the compressed sensor problem show its effectiveness. Finally, some conclusions are made in Section 6.

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

2 The proposed method

We describe the proposed method, and give some useful notations in this section. To solve Problem (6), the inexact proximal alternating directions method is as the following:

The proposed method (In-PADM)

For a given triplet $w^k = (x^k, y^k, \lambda^k)$, the In-PADM carries the new triplet $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ via the following procedures:

1). Find $\hat{x}^k$ by solving variational inequality:

$$(x'-x)^T\Big\{f(x)-A^T[\lambda^k-\beta_k(Ax-y^k)]+r_k(x-x^k)+\xi_x^k\Big\} \geq 0. \tag{25}$$

2). Find $\hat{y}^k$ by solving variational inequality:

$$(y'-y)^T\Big\{g(y)+[\lambda^k-\beta_k(A\hat{x}^k-y)]+\xi_y^k\Big\} \geq 0. \tag{26}$$

Remark: Here the proximal parameter s_k vanishes (Comparing with (13)).

3). Update $\hat{\lambda}^k$ by the following formula:

$$\hat{\lambda}^k = \lambda^k - \beta_k(A\hat{x}^k - \hat{y}^k). \tag{27}$$

4). Compute step-size $\alpha_k > 0$ and update w^{k+1} via

$$w^{k+1} = \mathrm{Cor}\left(w^k, \alpha_k, d(w^k, \hat{w}^k, \xi^k)\right). \tag{28}$$

In this proposed method, the unknown direction vector $d(w^k, \tilde{w}^k, \xi^k)$ and step-size α_k, and the unknown function $\mathrm{Cor}\left(w^k, \alpha_k, d(w^k, \tilde{w}^k, \xi^k)\right)$ in (28) will be defined later.

Denote $\hat{w}^k = (\hat{x}^k, \hat{y}^k, \hat{\lambda}^k)$, we define the direction vectors as the following:

$$d_1(w^k, \hat{w}^k, \xi^k) = \begin{pmatrix} r_k(x^k-\hat{x}^k)-\xi_x^k \\ \beta_k(y^k-\hat{y}^k)-\xi_y^k \\ \frac{1}{\beta_k}(\lambda^k-\hat{\lambda}^k) \end{pmatrix}, \tag{29}$$

and

$$d_2(w^k, \hat{w}^k) = \begin{pmatrix} f(\hat{x}^k) - A^T\hat{\lambda}^k - \beta_k A^T(y^k-\hat{y}^k) \\ g(\hat{y}^k) + \hat{\lambda}^k + \beta_k(y^k-\hat{y}^k) \\ A\hat{x}^k - \hat{y}^k \end{pmatrix}. \tag{30}$$

Combine (25–27) and rewrite it in a compact form:

$$(w' - \hat{w}^k)^T[d_2(w^k, \hat{w}^k) - d_1(w^k, \hat{w}^k, \xi^k)] \geq 0. \tag{31}$$

Here we give some useful notations:

$$M = \begin{pmatrix} rI & 0 & 0 \\ 0 & \beta I & 0 \\ 0 & 0 & \frac{1}{\beta}I \end{pmatrix}, \quad \xi = \begin{pmatrix} \xi_x \\ \xi_y \\ 0 \end{pmatrix}$$

and

$$F(w) = \begin{pmatrix} f(x) - A^T\lambda \\ g(y) + \lambda \\ Ax - y \end{pmatrix}, \eta(\beta, w, \hat{w}) = \begin{pmatrix} \beta A^T(y-\hat{y}) \\ -\beta(y-\hat{y}) \\ 0 \end{pmatrix}.$$

Then we get

$$d_1(w^k, \hat{w}^k, \xi^k) = M_k(w^k - \hat{w}^k) - \xi^k. \tag{32}$$

and

$$d_2(w^k, \hat{w}^k) = F(\hat{w}^k) - \eta(\beta_k, w^k, \hat{w}^k). \tag{33}$$

We define a key function $\varphi(w^k, \hat{w}^k)$ in the following:

$$\varphi(w^k, \hat{w}^k) = (\hat{\lambda}^k-\lambda^k)^T(y^k-\hat{y}^k)+(w^k-\hat{w}^k)^T d_1(w^k, \hat{w}^k, \xi^k). \tag{34}$$

This function $\varphi(w^k, \hat{w}^k)$ plays a crucial role in proof of convergence of the proposed method.

3 The descent directions of unknown function $\|w - w^*\|^2$

For convenience of the analysis in this section, we ignore the index k of all of the matrices, vectors and scalars.

For any $w^* \in \mathcal{W}^*$, $G(w - w^*)$ is the gradient of the unknown distance function $\frac{1}{2}\|w - w^*\|_G^2$ at the point w with a proper positive definite matrix G. The vector d is called the descent direction of $\|w - w^*\|_G^2$ if and only if $[G(w - w^*)]^T d < 0$. In this section, under suitable conditions, we will show that both $-d_1(w, \hat{w}, \xi)$ and $-d_2(w, \hat{w})$ [see (32) and (33) respectively] are the descent directions of $\|w - w^*\|^2$, while $w \in \mathcal{W} \setminus \mathcal{W}^*$.

Lemma 3.1 Let $\varphi(w, \hat{w})$ be defined by (34), where $\hat{w} = (\hat{x}, \hat{y}, \hat{\lambda})$ is generated by (25–27) from a given $w = (x, y, \lambda)$. Assume that the following conditions hold,

$$\|\xi_x\| \leq \nu r\|x - \hat{x}\|, \tag{35}$$

$$\|\xi_y\| \leq \frac{\nu\beta}{2\sqrt{2}}\|y - \hat{y}\|, \tag{36}$$

where $\nu \in (0, 1)$. Then we have

$$\varphi(w, \hat{w}) > \frac{\beta}{2}\|A\hat{x} - y\|^2 + \frac{\tau}{2}\|d_1(w, \hat{w}, \xi)\|^2, \tag{37}$$

where $\tau = \min\{\frac{1}{r}, \frac{1}{2\beta}, \beta\} > 0$.

Proof: By a manipulation, we get

$$\begin{aligned}
&\varphi(w, \hat{w}) \\
&= (\hat{\lambda} - \lambda)^T(y - \hat{y}) + (w - \hat{w})^T[M(w - \hat{w}) - \xi] \\
&= (\hat{\lambda} - \lambda)^T(y - \hat{y}) + \|w - \hat{w}\|_M^2 - (w - \hat{w})^T\xi \\
&= (\hat{\lambda} - \lambda)^T(y - \hat{y}) + r\|x - \hat{x}\|^2 + \beta\|y - \hat{y}\|^2 \\
&+\frac{1}{\beta}\|\lambda - \hat{\lambda}\|^2 - (x - \hat{x})^T\xi_x - (y - \hat{y})^T\xi_y \\
&= \Big[(\hat{\lambda} - \lambda)^T(y - \hat{y}) + \frac{1}{2}\beta\|y - \hat{y}\|^2 + \frac{1}{2\beta}\|\lambda - \hat{\lambda}\|^2 \\
&\qquad +\frac{1}{2\beta}\|\lambda - \hat{\lambda}\|^2\Big] && (38) \\
&+r\|x - \hat{x}\|^2 - (x - \hat{x})^T\xi_x && (39) \\
&+\frac{1}{2}\beta\|y - \hat{y}\|^2 - (y - \hat{y})^T\xi_y. && (40)
\end{aligned}$$

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Using $\hat{\lambda} - \lambda = -\beta(A\hat{x} - \hat{y})$, we get

$$\begin{aligned}
&(\hat{\lambda} - \lambda)^T(y - \hat{y}) + \frac{1}{2}\beta\|y - \hat{y}\|^2 + \frac{1}{2\beta}\|\lambda - \hat{\lambda}\|^2 \\
&= \frac{1}{2}\beta\left[-2(A\hat{x} - \hat{y})^T(y - \hat{y}) + \|y - \hat{y}\|^2 + \|A\hat{x} - \hat{y}\|^2\right] \\
&= \frac{1}{2}\beta\|A\hat{x} - y\|^2 \qquad (41)
\end{aligned}$$

By (35) and note $\nu \in (0,1)$, we get

$$\begin{aligned}
&r\|x - \hat{x}\|^2 - (x - \hat{x})^T\xi_x \\
&= \frac{1}{2}r\|x - \hat{x}\|^2 - (x - \hat{x})^T\xi_x + \frac{1}{2}r\|x - \hat{x}\|^2 \\
&= \frac{1}{2r}\left[r^2\|x - \hat{x}\|^2 - 2r(x - \hat{x})^T\xi_x + r^2\|x - \hat{x}\|^2\right] \\
&> \frac{1}{2r}\left[r^2\|x - \hat{x}\|^2 - 2r(x - \hat{x})^T\xi_x + \|\xi_x\|^2\right] \\
&= \frac{1}{2r}\|r(x - \hat{x}) - \xi_x\|^2. \qquad (42)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\frac{1}{2}\beta\|y - \hat{y}\|^2 - (y - \hat{y})^T\xi_y \\
&= \frac{1}{4}\beta\|y - \hat{y}\|^2 - (y - \hat{y})^T\xi_y + \frac{1}{4}\beta\|y - \hat{y}\|^2. \\
&= \frac{1}{\beta}\left[\|\frac{1}{2}\beta(y - \hat{y})\|^2 - \frac{1}{2}\beta(y - \hat{y})^T\xi_y + \|\frac{1}{2}\xi_y\|^2\right. \\
&\left. + \frac{1}{4}\beta^2\|y - \hat{y}\|^2 - \|\frac{1}{2}\xi_y\|^2 - \frac{1}{2}\beta(y - \hat{y})^T\xi_y\right] \\
&= \frac{1}{4\beta}\|\beta(y - \hat{y}) - \xi_y\|^2 \\
&+ \frac{1}{\beta}\left[\frac{1}{4}\beta^2\|y - \hat{y}\|^2 - \|\frac{1}{2}\xi_y\|^2 - \frac{1}{2}\beta(y - \hat{y})^T\xi_y\right]
\end{aligned}$$

By Cauchy-Schwarz inequality and condition (36), we have

$$\frac{\beta}{2}(y - \hat{y})^T\xi_y \le \frac{\beta}{2}\|y - \hat{y}\| \cdot \|\xi_y\| \le \frac{\beta^2}{4\sqrt{2}}\|y - \hat{y}\|^2$$

Using condition (36) again, we get

$$\begin{aligned}
&\frac{1}{\beta}\left[\frac{1}{4}\beta^2\|y - \hat{y}\|^2 - \|\frac{1}{2}\xi_y\|^2 - \frac{1}{2}\beta(y - \hat{y})^T\xi_y\right] \\
&\ge \frac{\beta}{4}\frac{7\sqrt{2} - 8}{8\sqrt{2}}\|y - \hat{y}\|^2 \ge 0.
\end{aligned}$$

Thus

$$\frac{1}{2}\beta\|y - \hat{y}\|^2 - (y - \hat{y})^T\xi_y \ge \frac{1}{4\beta}\|\beta(y - \hat{y}) - \xi_y\|^2 \qquad (43)$$

Substituting (41–43) into (40–42), we have

$$\begin{aligned}
\varphi(w, \hat{w}) \quad &> \frac{1}{2}\beta\|A\hat{x} - y\|^2 + \frac{1}{2r}\|r(x - \hat{x}) - \xi_x\|^2 \\
&+ \frac{1}{4\beta}\|\beta(y - \hat{y}) - \xi_y\|^2 + \frac{1}{2\beta}\|\lambda - \hat{\lambda}\|^2. \qquad (44)
\end{aligned}$$

Let

$$\tau = \min\{\frac{1}{r}, \frac{1}{2\beta}, \beta\} \qquad (45)$$

we obtain

$$\varphi(w, \hat{w}) > \frac{1}{2}\beta\|A\hat{x} - y\|^2 + \frac{\tau}{2}\|d_1(w, \hat{w}, \xi)\|^2,$$

which is (37). This completes the proof.

Lemma 3.2 Let $\varphi(w, \hat{w})$ be defined by (34), and $\hat{w} = (\hat{x}, \hat{y}, \hat{\lambda})$ is generated by (25–27) from given $w = (x, y, \lambda)$, and the conditions (35) and (36) hold, then we have

$$\varphi(w, \hat{w}) \ge (1-\nu)r\|x-\hat{x}\|^2 + \frac{1}{2}(1-\frac{\sqrt{2}}{2}\nu)\beta\|y-\hat{y}\|^2 + \frac{1}{2\beta}\|\lambda-\hat{\lambda}\|^2. \qquad (46)$$

Proof: By Cauchy-Schwarz inequality, we get

$$(x - \hat{x})^T\xi_x \le \|x - \hat{x}\| \cdot \|\xi_x\|,$$

$$(y - \hat{y})^T\xi_y \le \|y - \hat{y}\| \cdot \|\xi_y\|,$$

and $\beta > 0$,

$$\begin{aligned}
&\frac{1}{2}\left(\beta\|y - \hat{y}\|^2 + \frac{1}{\beta}\|\lambda - \hat{\lambda}\|^2\right) \\
&\ge \sqrt{\frac{1}{\beta}}\|\lambda - \hat{\lambda}\| \cdot \sqrt{\beta}\|y - \hat{y}\| \\
&\ge (\lambda - \hat{\lambda})^T(y - \hat{y}).
\end{aligned}$$

Then by using the previous three inequalities and the conditions (35–36) and the definitions (29, 34), we get

$$\begin{aligned}
&\varphi(w, \hat{w}) \\
&= (\hat{\lambda} - \lambda)^T(y - \hat{y}) + (w - \hat{w})^T[M(w - \hat{w}) - \xi] \\
&= (\hat{\lambda} - \lambda)^T(y - \hat{y}) + (x - \hat{x})^T r(x - \hat{x}) - (x - \hat{x})^T\xi_x \\
&+ (y - \hat{y})^T\beta(y - \hat{y}) - (y - \hat{y})^T\xi_y + (\lambda - \hat{\lambda})^T\frac{1}{\beta}(\lambda - \hat{\lambda}) \\
&\ge (1 - \nu)r\|x - \hat{x}\|^2 + (1 - \frac{\nu}{2\sqrt{2}})\beta\|y - \hat{y}\|^2 \\
&+ \frac{1}{\beta}\|\lambda - \hat{\lambda}\|^2 - (\lambda - \hat{\lambda})^T(y - \hat{y}) \\
&\ge (1 - \nu)r\|x - \hat{x}\|^2 + \frac{1}{2}(1 - \frac{\sqrt{2}}{2}\nu)\beta\|y - \hat{y}\|^2 + \frac{1}{2\beta}\|\lambda - \hat{\lambda}\|^2.
\end{aligned}$$

This is (46).

Lemma 3.3 Let $\varphi(w, \hat{w})$ be defined by (34) and $d_1(w, \hat{w}, \xi)$, $d_2(w, \hat{w})$ be defined by (29-30)(or more compactly, (32–33)) respectively. Let $\hat{w} = (\hat{x}, \hat{y}, \hat{\lambda})$ be generated by (25–27) from given $w = (x, y, \lambda)$. Then we have

$$(\hat{w}-w^*)^T d_2(w, \hat{w}) \ge \varphi(w, \hat{w}) - (w-\hat{w})^T d_1(w, \hat{w}, \xi). \qquad (47)$$

Proof: Since $f(x)$ and $g(x)$ are monotone with respect to $\mathcal{X}$ and $\mathcal{Y}$, respectively, it follows that $F(w)$ is a monotone operator with respect to $\mathcal{W}$. Then we get

$$(\hat{w} - w^*)^T[F(\hat{w}) - F(w^*)] \ge 0,$$

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

and consequently,

$$(\hat{w} - w^*)^T F(\hat{w}) \geq (\hat{w} - w^*)^T F(w^*). \tag{48}$$

Note that w^* is a solution of (6), this provides that $(\hat{w} - w^*)^T F(w^*) \geq 0$ for $\hat{w} \in \mathcal{W}$. Thus

$$(\hat{w} - w^*)^T F(\hat{w}) \geq 0. \tag{49}$$

Therefore, we have

$$\begin{aligned} & (\hat{w} - w^*)^T d_2(w, \hat{w}) \\ = \; & (\hat{w} - w^*)^T F(\hat{w}) - (\hat{w} - w^*)^T \eta(\beta_k, w^k, \hat{w}^k) \\ \geq \; & -(\hat{w} - w^*)^T \eta(\beta_k, w^k, \hat{w}^k) \\ = \; & -\beta(A\hat{x} - Ax^*)^T (y - \hat{y}) + (\hat{y} - y^*)^T \beta(y - \hat{y}) \\ = \; & (y - \hat{y})^T \beta(\hat{y} - A\hat{x}) \\ = \; & (y - \hat{y})^T (\hat{\lambda} - \lambda). \end{aligned} \tag{50}$$

Recalling the definition of $\varphi(w, \hat{w})$ (see (34)), the inequality (47) follows from (50) directly.

Theorem 3.1 Let $\hat{w} = (\hat{x}, \hat{y}, \hat{\lambda})$ be generated by (25–27) from given $w = (x, y, \lambda)$ and $w^* \in \mathcal{W}^*$. Then we have

$$(w - w^*)^T d_1(w, \hat{w}, \xi) \geq \varphi(w, \hat{w}, \xi), \tag{51}$$

and

$$(w - w^*)^T d_2(w, \hat{w}) \geq \varphi(w, \hat{w}, \xi), \quad \forall w \in \mathcal{W}. \tag{52}$$

Proof: By using (31), we get (by letting $w' = w^*$),

$$(\hat{w} - w^*)^T d_1(w, \hat{w}, \xi) \geq (\hat{w} - w^*)^T d_2(w, \hat{w}). \tag{53}$$

Combining (53) with (47), we have

$$(\hat{w} - w^*)^T d_1(w, \hat{w}, \xi) \geq \varphi(w, \hat{w}) - (w - \hat{w})^T d_1(w, \hat{w}, \xi). \tag{54}$$

Following from (54), we obtain (51) directly. Adding (47) and (31), we get

$$\begin{aligned} & (w' - w^*)^T d_2(w, \hat{w}) \\ & \geq \varphi(w, \hat{w}, \xi) - (w - w')^T d_1(w, \hat{w}, \xi), \\ & \forall w' \in \mathcal{W}, \; w^* \in \mathcal{W}^*. \end{aligned} \tag{55}$$

Assertion (52) follows from (55) by letting $w' = w$.

4 Convergence of the proposed method

In this section, we define the "cor" function in (28) in the first of the section, and prove some contractive properties (with the given "cor" function) of the proposed method. Then we prove convergence of the proposed method. Be lighted by the papers [2, 3] and others, we give two "cor" functions and prove convergence of the proposed method provided with each of them.

The first strategy is setting the "cor" function as follows:

$$\begin{aligned} w_I^{k+1} &= \text{cor}_1\left(w^k, \alpha_k, d(w^k, \tilde{w}^k, \xi^k)\right) \\ &= w^k - \alpha_k d_1(w^k, \tilde{w}^k, \xi^k). \end{aligned} \tag{56}$$

Lemma 4.1 Let $w^{k+1} = w_{\mathrm{I}}^{k+1}$ be generated by (25–28), then we have

$$\begin{aligned} & \|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2 \\ & \geq 2\alpha_k \varphi(w^k, \tilde{w}^k, \xi^k) - \alpha_k^2 \|d_1(w^k, \tilde{w}^k, \xi^k)\|^2. \end{aligned} \tag{57}$$

Proof: By a straightforward computation, and using Theorem 3.1 (51), we get

$$\begin{aligned} & \|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2 \\ & = \|w^k - w^*\|^2 - \|w^k - \alpha_k d(w^k, \hat{w}^k, \xi^k) - w^*\|^2 \\ & = 2\alpha_k (w^k - w^*)^T d_1(w^k, \hat{w}^k, \xi^k) - \alpha_k^2 \|d_1(w^k, \hat{w}^k, \xi^k)\|^2 \\ & \geq 2\alpha_k \varphi(w^k, \hat{w}^k) - \alpha_k^2 \|d_1(w^k, \hat{w}^k, \xi^k)\|^2. \end{aligned}$$

We then deal with the second strategy by defining the "cor" function as follows:

$$\begin{aligned} w_{II}^{k+1} &= \text{cor}_2\left(w^k, \alpha_k, d(w^k, \hat{w}^k, \xi^k)\right) \\ &= P_{\mathcal{W}}\left\{w^k - \alpha_k d_2(w^k, \hat{w}^k)\right\}. \end{aligned} \tag{58}$$

Corresponding to the Lemma 4.1, by using the second strategy, we have the same contractive property.

Lemma 4.2 Let $w^{k+1} = w_{II}^{k+1}$ be generated by (25–28), then we have

$$\begin{aligned} & \|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2 \\ & \geq 2\alpha_k \varphi(w^k, \tilde{w}^k, \xi^k) - \alpha_k^2 \|d_1(w^k, \tilde{w}^k, \xi^k)\|^2. \end{aligned} \tag{59}$$

Proof: It's well known (by using the property of projection) that $\forall v \in R^n \times R^m \times R^m$,

$$\|u - P_{\mathcal{W}}(v)\|^2 \leq \|v - u\|^2 - \|v - P_{\mathcal{W}}(v)\|^2, \forall u \in \mathcal{W}.$$

Using the previous inequality, we get

$$\begin{aligned} & \|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2 \\ & \geq \|w^k - w^*\|^2 - \big(\|w^k - w^* - \alpha_k d_2(w^k, \hat{w}^k)\|^2 \\ & -\|w^k - w^{k+1} - \alpha_k d_2(w^k, \hat{w}^k)\|\big) \\ & = \|w^k - w^{k+1}\|^2 + 2\alpha_k (w^{k+1} - w^*)^T d_2(w^k, \hat{w}^k). \end{aligned} \tag{60}$$

By (50), we get

$$\begin{aligned} & (w^{k+1} - w^*)^T d_2(w^k, \hat{w}^k) \\ & = (w^{k+1} - \hat{w}^k)^T d_2(w^k, \hat{w}^k) + (\hat{w}^k - w^*)^T d_2(w^k, \hat{w}^k) \\ & \geq (w^{k+1} - \hat{w}^k)^T d_2(w^k, \hat{w}^k) + (y^k - \hat{y}^k)^T (\hat{\lambda}^k - \lambda^k) \end{aligned} \tag{61}$$

On the other hand, since $\hat{w}_k$ solves the variational inequality (31), and $w^{k+1} \in \mathcal{W}$, it follows that

$$(w^{k+1} - \hat{w}^k)^T d_2(w^k, \hat{w}^k) \geq (w^{k+1} - \hat{w}^k)^T d_1(w^k, \hat{w}^k, \xi^k). \tag{62}$$

Since $\alpha_k > 0$, by adding (61) and (62), we get

$$\begin{aligned} & 2\alpha_k (w^{k+1} - w^*)^T d_2(w^k, \hat{w}^k) \\ & \geq 2\alpha_k \big[(w^{k+1} - \hat{w}^k)^T d_1(w^k, \hat{w}^k, \xi^k) \\ & +(y^k - \hat{y}^k)^T (\hat{\lambda}^k - \lambda^k)\big]. \end{aligned} \tag{63}$$

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Substituting (63) into (60), we have

$$\begin{aligned}
&\|w^k - w^*\|^2 - \|w^{k+1} - w^*\|^2 \\
&\geq \|w^k - w^{k+1}\|^2 + 2\alpha_k[(w^{k+1} - \hat{w}^k)^T d_1(w^k, \hat{w}^k, \xi^k) \\
&+(y^k - \hat{y}^k)^T(\hat{\lambda}^k - \lambda^k)] \\
&= \|w^{k+1} - w^k - \alpha_k d_1(w^k, \hat{w}^k, \xi^k)\|^2 \\
&-\alpha_k^2 \|d_1(w^k, \hat{w}^k, \xi^k)\|^2 \\
&+2\alpha_k[(w^k - \hat{w}^k)^T d_1(w^k, \hat{w}^k, \xi^k) \\
&+(y^k - \hat{y}^k)^T(\hat{\lambda}^k - \lambda^k)].
\end{aligned} \tag{64}$$

Recall the definition of $\varphi(w^k, \hat{w}^k)$ (see (34)), we obtain (59) from (64) immediately.

From Lemma 4.1 and Lemma 4.2, we know that no matter which direction $d_1(w^k, \hat{w}^k, \xi^k)$ is used or the direction $d_2(w^k, \hat{w}^k)$, we have the same conclusion:

$$\begin{aligned}
\|w^{k+1} - w^*\|^2 \quad \leq \quad & \|w^k - w^*\|^2 + \alpha_k^2 \|d_1(w^k, \hat{w}^k, \xi^k)\|^2 \\
& -2\alpha_k \varphi(w^k, \hat{w}^k).
\end{aligned}$$

From the above inequality, we can get the maximal dropout value at each iteration by maximizing the following function

$$\Delta_k(\alpha) = -\alpha^2 \|d_1(w^k, \hat{w}^k, \xi^k)\|^2 + 2\alpha\varphi(w^k, \hat{w}^k). \tag{65}$$

This quadratic function $\Delta_k(\alpha)$ reaches its maximum at the point

$$\alpha_k^* = \frac{\varphi(w^k, \hat{w}^k, \xi^k)}{\|d_1(w^k, \hat{w}^k, \xi^k)\|^2}, \tag{66}$$

with its value

$$\Delta_k(\alpha_k^*) = \alpha_k^* \varphi(w^k, \tilde{w}^k, \xi^k). \tag{67}$$

Then by using Lemma 3.1, we have:

Corollary 4.1 If the constant sequences $\{r_k\}, \{\beta_k\}$ are nonnegative and bounded above for any k, then we have

$$\alpha_k^* > \frac{\tau}{2}. \tag{68}$$

We will show in the next section that the conditions of Corollary 4.1 is satisfied in practice.

The numerical experiments [3] suggest that for fast convergence one can use a relaxation factor $\gamma \in [1, 2)$ and let $\alpha_k = \gamma \alpha_k^*$ at each iteration. When we do so, we get

$$\Delta_k(\gamma \alpha_k^*) = \gamma(2 - \gamma)\Delta_k(\alpha_k^*). \tag{69}$$

By using (57) or (59), we then get

$$\|w^{k+1} - w^*\| \leq \|w^k - w^*\| - \gamma(2 - \gamma)\alpha_k^* \varphi(w^k, \hat{w}^k). \tag{70}$$

Furthermore, using Corollary 4.1, and letting

$$\tau = \min_k \min\{\frac{1}{r_k}, \frac{1}{2\beta_k}, \beta_k\}, \tag{71}$$

we have

$$\|w^{k+1} - w^*\| \leq \|w^k - w^*\| - \frac{\tau}{2}\gamma(2 - \gamma)\varphi(w^k, \hat{w}^k). \tag{72}$$

This inequality (72) plays an important role in proving convergence of the proposed method. Let

$$N = \begin{pmatrix} 2(1-\nu)rI & 0 & 0 \\ 0 & (1 - \frac{\sqrt{2}}{2}\nu)\beta I & 0 \\ 0 & 0 & \frac{1}{\beta}I \end{pmatrix}, \tag{73}$$

then by using Lemma 3.2 and this notation, we get a compact form of (46),

$$\varphi(w, \hat{w}) \geq \frac{1}{2}\|w - \hat{w}\|_N^2. \tag{74}$$

Theorem 4.1 Let the sequence $\{w^k\}$ be generated by the proposed method In-PADM (25–28). The conditions (35–36) hold and the conditions of Corollary 4.1 holds. Then the sequence $\{w^k\}$ converges to w^∞, which is a solution of SVI (6).

Proof: By using (72) and (74), we get

$$\|w^{k+1} - w^*\|^2 \leq \|w^k - w^*\|^2 - \frac{\tau\gamma(2-\gamma)}{2}\|w^k - \hat{w}^k\|_N^2. \tag{75}$$

Consequently,

$$\|w^{k+1} - w^*\|^2 \leq \|w^0 - w^*\|^2, \quad \forall k. \tag{76}$$

It follows that the sequence $\{w^k\}$ is bounded. Moreover, we get

$$\frac{\tau\gamma(2-\gamma)}{2}\sum_{k=0}^{\infty}\|w^k - \hat{w}^k\|_N^2 \leq \|w^0 - w^*\|, \tag{77}$$

which follows that

$$\lim_{k\to\infty}\|w^k - \hat{w}^k\|_N^2 = 0. \tag{78}$$

Thus $\{\hat{w}^k\}$ is also bounded and it has at least one cluster point. Let w^∞ be a cluster point of $\{\hat{w}^k\}$, let the subsequence $\{\hat{w}^{k_j}\}$ converges to w^∞. It follows from (78) that

$$\lim_{k\to\infty}\|x^k - \hat{x}^k\| = 0, \lim_{k\to\infty}\|y^k - \hat{y}^k\| = 0, \lim_{k\to\infty}\|\lambda^k - \hat{\lambda}^k\| = 0,$$

and consequently (by using conditions (35) and (36)),

$$\lim_{k\to\infty}\|\xi_x^k\| = 0, \lim_{k\to\infty}\|\xi_y^k\| = 0. \tag{79}$$

Then by (25–27), we get,

$$\hat{w}^k \in \mathcal{W}, \begin{cases} \lim_{k\to\infty}(x' - \hat{x}^k)^T\{f(\hat{x}^k) - A^T\hat{\lambda}^k\} \geq 0, \\ \lim_{k\to\infty}(y' - \hat{y}^k)^T\{g(\hat{y}^k) + \hat{\lambda}^k\} \geq 0, \\ \lim_{k\to\infty}(A\hat{x}^k - \hat{y}^k) = 0, \end{cases} \tag{80}$$

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

and

$$\hat{w}^{k_j} \in \mathcal{W}, \quad \begin{cases} \lim\limits_{k_j\to\infty} (x' - \hat{x}^{k_j})^T \{f(\hat{x}^{k_j}) - A^T\hat{\lambda}^{k_j}\} \geq 0, \\ \lim\limits_{k_j\to\infty} (y' - \hat{y}^{k_j})^T \{g(\hat{y}^{k_j}) + \hat{\lambda}^{k_j}\} \geq 0, \\ \lim\limits_{k_j\to\infty} (A\hat{x}^{k_j} - \hat{y}^{k_j}) = 0. \end{cases} \tag{81}$$

Ineq. (81) implies that

$$w^\infty \in \mathcal{W}, \quad \begin{cases} (x' - x^\infty)^T \{f(x^\infty) - A^T\lambda^\infty\} \geq 0, \\ (y' - y^\infty)^T \{g(y^\infty) + \lambda^\infty\} \geq 0, \\ (Ax^\infty - y^\infty) = 0. \end{cases} \tag{82}$$

Thus w^∞ solves the variational inequalities (6) (or equally (10)). Since $\{\hat{w}^{k_j}\} \to w^\infty$, by (78), for any given $\varepsilon > 0$, there exists an integer $l > 0$ such that

$$\|w^{k_l} - \hat{w}^{k_l}\|_N < \frac{\varepsilon}{2}, \ \|\hat{w}^{k_l} - w^\infty\|_N < \frac{\varepsilon}{2}. \tag{83}$$

Therefore, for any $k \geq k_l$, by (75) and (83), we have

$$\begin{aligned} \|w^k - w^\infty\|_N &\leq \|w^{k_l} - w^\infty\|_N \\ &\leq \|w^{k_l} - \hat{w}^{k_l}\|_N + \|\hat{w}^{k_l} - w^\infty\| < \varepsilon, \end{aligned}$$

which implies the sequence $\{w^k\}$ converges to w^∞— a solution of variational inequalities (6).

5 Practical implementation and numerical results

We have to give the proximal parameter r_k and penalty parameter β_k in the practical implementation of the proposed In-PADM method.

From (45) and (71), we know if the sequences $\{r_k\}, \{\beta_k\}$ are positive and bounded, then τ is well defined and $\tau > 0$, and the condition of Corollary 4.1 holds.

On the penalty parameter sequence $\{\beta_k\}$, He and Yang, et al [4] proposed a method that adjusts the penalty parameter per iteration based on the iterate message, which is referred to as self-adaptive penalty parameters method, and showed that the sequence $\{\beta_k\}$ generated by this method is both bounded above and bounded below and away from zero. In the proposed method of this paper, we use this self-adaptive penalty parameters method (Method 3, Strategy S3 in [4]) directly and set

$$0 < \beta_l \leq \beta_k \leq \beta_h < +\infty, \quad \forall k. \tag{84}$$

Next we consider the proximal parameters r_k. For any $\nu \in (0,1)$, the variational inequality (26) is equivalent to

$$y \in \mathcal{Y}, \ (y'-y)^T \frac{1}{\beta_k} \left\{g(y) + [\lambda^k - \beta_k(A\hat{x}^k - y)] + \xi_y^k\right\} \geq 0, \tag{85}$$

and furthermore it is equivalent to the following projection equation

$$y = P_{\mathcal{Y}}\Big\{y - \frac{1}{\beta_k}\{g(y) + [\lambda^k - \beta_k(A\hat{x}^k - y)] + \xi_y^k\}\Big\}. \tag{86}$$

Then due to the variational inequality (26), we have the following Lemma.

Lemma 5.1 Let g be monotone and one of the following conditions is satisfied:

$$\|g(y^k) - g(\hat{y}^k)\| \leq \left(\frac{\nu\beta_k}{2\sqrt{2}}\right) \|y^k - \hat{y}^k\|. \tag{87}$$

Let

$$\xi_y^k = g(y^k) - g(\hat{y}^k). \tag{88}$$

Then the condition (36) holds, and the subproblem (26) (correspondingly (85)) is monotone and tractable.

Proof: It's obvious and we omit it here.

We finally focus on the proximal parameter sequence $\{r_k\}$. Note that the variational inequality (25) is equivalent to the following projection equation

$$x = P_{\mathcal{X}}\Big\{x - \frac{1}{r_k}\big[f(x) - A^T(\lambda^k - \beta_k(Ax - y^k)) + r_k(x - x^k) + \xi_x^k\big]\Big\} \tag{89}$$

By manipulation, we get

$$x = P_{\mathcal{X}}\Big\{x^k - \frac{1}{r_k}\big[f(x) + \beta_k A^T Ax + \xi_x^k - (A^T\lambda^k + \beta_k A^T y^k)\big]\Big\} \tag{90}$$

Lemma 5.2 If f is monotone and Lipschitz continuous with constant L_f, i.e.,

$$\|f(x) - f(\hat{x})\| \leq L_f\|x - \hat{x}\|, \ \forall x, \ \hat{x} \in \mathcal{X}. \tag{91}$$

Let

$$\xi_x^k = f(x^k) - f(x) + \beta_k A^T A(x^k - x), \tag{92}$$

and

$$r_k \geq \frac{L_f + \beta_k\|A^T A\|}{\nu}. \tag{93}$$

Then the condition (35) is satisfied and the variational inequality (25) is monotone and tractable.

Proof: By computing directly and using (91), (93), we get

$$\begin{aligned} \|\xi_x^k\| &= \|f(x^k) - f(x) + \beta_k A^T A(x^k - x)\| \\ &\leq \|f(x^k) - f(x)\| + \|\beta_k A^T A(x^k - x)\| \\ &\leq L_f\|x^k - x\| + \beta_k\|A^T A\|\|x^k - x\| \\ &\leq \nu r_k\|x^k - x\|. \end{aligned} \tag{94}$$

This is (36). Substituting (92) into (90), we have

$$x = P_{\mathcal{X}}\Big\{x^k - \frac{1}{r_k}\big[f(x^k) + \beta_k A^T Ax^k - (A^T\lambda^k + \beta_k A^T y^k)\big]\Big\}. \tag{95}$$

The right-hand of (95) is an explicit projection and is tractable. The monotonicity of (25) is obvious under the conditions.

From the previous discussions, we know that the parameter sequences $\{r_k\}$, $\{\beta_k\}$ is positive and bounded such that τ in (71) is well defined which guarantees convergence of the proposed method.

Finally in this section, we present some numerical results on the described problems. We deal with Problem (3)

ISBN: 978-988-18210-5-8
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

by using the proposed method, on a portable computer with: 166GHz CPU, 2.5GB RAM, Matlab 6.5.

The data of Problem (3) is generated by the following way: $D \in R^{m\times n}$ is a random matrix with $D_{ij} \in (1, 2)$ according to uniform distribution, $b \in R^m$ is a given signal vector with a random noise. $\mu = 0.01 * \|A^T b\|_\infty$. The parameters of the implementation of the proposed method are given in the following: $\nu = 0.95$, $r_0 = 0.51 * \max(\sigma(AA^T))$ (where $\sigma(AA^T)$ denotes the eigenvalue of AA^T), $\beta_0 = 1.35$, $x^0 = \mathbf{0}_{n\times1}$, $y^0 = \mathbf{1}_{n\times1}$, $\lambda^0 = \mathbf{1}_{n\times1}$, $\gamma = 1.2$. We restrict the terminate error $\varepsilon < 10^{-3}$. The experimental results are stated in table 5.1. The notations in table 5.1 is as the following: m, n and `t` denotes cputime; *sp* denotes the required sparse degree, iter. is the number of iterations, feval. is the number of computation of product on matrix and vector (which is the main cost of the proposed method), r is the ultimate proximal parameter.

Table 5.1 Numerical results

m	n	sp	iter.	feval.	t(s)	r
1024	4096	160/4096	172	520	7.78	2.240
1600	8192	320/8192	218	658	25.34	2.092
2048	12000	400/12000	248	746	58.83	2.006
3052	16348	512/16348	276	830	121.62	2.051

According to table 5.1, we can see that the proposed method (i.e., In-PADM) is applicable for Compressed Sensor problems.

6 Conclusions

In this paper, we present an inexact proximal alternating directions method (In-PADM) for solving a class of monotone structured variational inequalities. This method make full use of the special structure of the described problem. The problem with this special structure is common in practice. The research of this paper is a continuation and development of previous works. In this paper, we prove theoretically that for any inexact term ξ^k under suitable conditions we have convergence of the proposed method. By using the special structure of the described problem, i.e., $B = -I$, we confirm that, without using the proximal point term in the second variational inequality, convergence of the proposed method is preserved. And to do in this way provides some advantages for tractability of the implementation. Numerical experiments show the proposed method is applicable and valid.

Acknowledgement: I am grateful to Professor He Bingsheng, for his enlightening of the idea of this paper, and much useful advices to improve its quality.

References

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

[2] B.S. He, L.-Z. Liao and M.J. Qian, Alternating projection based predition-correction methods for structure variational inequalities. Journal of Computational Mathematics, 24(6) (2006), 693-710.

[3] B.S. He, S.L. Wang, H. Yang, A modified variable-penalty alternating directions method for monotone variational inequalities. Journal of Computational Mathematics, 21(4) (2003), 495-504.

[4] B.S. He, H. Yang and S.L. Wang , Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities, Journal Optimization Theory and Applications, 106(2) (2000), 337-356.

[5] B.S. He, J. Zhou, A modified alternating direction method for convex minimization problems, Applied Mathematics Letters, 13 (2000), 123-130.

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

[7] M. Andrecut, R. A. Este, S. A. Kauffman, Competitive optimization of compressed sensing, Journal of Physics A: Mathematical and Theoretical, 40 (2007), 299-305.

[8] M. Fukushima, Application of the alternating direction method of multipliers to separable convex programming problems, Computational Optimization and Applications, 1 (1992), 93-111.

[9] L. A. Parente, P. A. Lotito, M. V. Solodov, A class of inexact variable metric proximal point algorihtms, SIAM Journal on Optimization, 19(1) (2008), 240-260.

[10] P. Tseng, Alternating projection-proximal methods for convex programming and variational inequalities, SIAM Journal on Optimization, 7 (1997), 951-965.

[11] P. Smoczynski, M. A. Tawhid, Two numerical schemes for general variational inequalities, Journal of Industrial and Management Optimization, 4(2) (2008), 393-406.

[12] W. Yin, S. Osher, D. Goldfarb, et al, Bregman Iterative Algorithms for l_1-Minimization with Applications to Compressed Sensing, SIAM Journal of Imaging Sciences, 1(1) (2008), 143-168.