Abstract and Applied Analysis

Some Recent Trends in Variational Inequalities and Optimization Problems with Applications

Guest Editors: Abdellah Bnouhachem, Abdelouahed Hamdi, and Xu Minghua

Some Recent Trends in Variational Inequalities and Optimization Problems with Applications

Abstract and Applied Analysis

Some Recent Trends in Variational Inequalities and Optimization Problems with Applications

Guest Editors: Abdellah Bnouhachem, Abdelouahed Hamdi, and Xu Minghua

Editorial Board

Contents

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 215098, 2 pages
http://dx.doi.org/10.1155/2014/215098

Editorial

Some Recent Trends in Variational Inequalities and Optimization Problems with Applications

Abdellah Bnouhachem,[1,2] Abdelouahed Hamdi,[3] and Xu Minghua[4]

[1] *School of Management Science and Engineering, Nanjing University, Nanjing 210093, China*
[2] *Ibn Zohr University, ENSA, BP 1136, Agadir, Morocco*
[3] *Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, PB 2713, Doha, Qatar*
[4] *School of Mathematics and Physics, Changzhou University, Changzhou, Jiangsu Province 213164, China*

Correspondence should be addressed to Abdellah Bnouhachem; babedallah@yahoo.com

Received 26 November 2013; Accepted 26 November 2013; Published 23 January 2014

Variational inequalities theory, which was introduced in the sixties, has emerged as an interesting and fascinating branch of applicable mathematics with a wide range of applications in industry, finance, economics, social, and pure and applied sciences. This field is dynamic and is experiencing an explosive growth in both theory and applications; as a consequence, research techniques and problems are drawn from various fields. The ideas and techniques of variational inequalities are being applied in a variety of diverse areas of sciences and prove to be productive and innovative. It has been shown that this theory provides the most natural, direct, simple, unified, and efficient framework for a general treatment of a wide class of unrelated linear and nonlinear problems. Variational inequalities have been extended and generalized in several directions using novel and new techniques. In parallel, optimization methods based on proximal point and proximal-like type methods have attracted a large number of researchers in the last three decades. In the same spirit, we can cite, for instance, the alternating direction multipliers method, which is based on the augmented lagrangian algorithm, which itself can be seen as a direct application of the proximal point algorithm to the dual problem of a constrained optimization problem.

The aim of this special issue is to present new approaches and theories for variational inequalities arising in mathematics and applied sciences. This special issue includes 14 high-quality peer-reviewed papers that deal with different aspects of variational inequalities. These papers contain some new, novel, and innovative techniques and ideas. We hope that all the papers published in this special issue can motivate and foster further scientific works and development of the research in the area of theory, algorithms, and applications of variational inequalities.

The summaries of the 14 papers in this issue are listed as follows.

The paper of C. Chen et al. considers a class of linearly constrained separable convex programming problems without coupled variables. They weaken some conditions to obtain convergence of the alternating direction method of multipliers and they propose also a relaxed ADMM involving an additional computation of optimal step size and establish its global convergence under mild conditions.

The paper of H. Sun and Y. Wang revisits the global error bound for the generalized nonlinear complementarity problem over a polyhedral cone (GNCP) and sharpens the global error bound for the GNCP under weaker conditions, which improves the existing error bound estimation for the problem.

The paper of M. Ma concerns the design and the convergence analysis of algorithms to split variational inequality and equilibrium problems.

The paper of Y. Wang and C. Wang gives a new modified Ishikawa type iteration algorithm for common fixed points of total asymptotically strict pseudocontractive semigroups.

Strong and weak convergence are proved under mild conditions. Furthermore, the main results presented in this work extend and improve some recent results.

The paper of X. Fu presents an implementable proximal step by a slight relaxation to the subproblem of proximal point algorithm (PPA) to solve linearly constrained convex programming. Self-adaptive strategies are proposed to improve the convergence in practice. The paper also discusses some applications and performs some numerical experiments to confirm the efficiency of the proposed method.

The paper of H. Xu establishes the strong convergence of prediction-correction and relaxed hybrid steepest-descent method (PRH method) for variational inequalities under some suitable conditions that simplify the proof. Further, the author shows the efficiency of the proposed algorithm through a well-designed set of practical numerical experiments.

The paper of M. Xu et al. considers the study of some matrix optimization problems using the proximal alternating direction method. The authors show that the restriction on the proximal parameters can be relaxed for solving these kinds of problems and give some numerical experiments to conclude that their modified method presents better performance than the classical proximal alternating direction method.

The paper of J.-L. Jiang et al. considers the locations of multiple facilities in the space $\mathbb{R}^p$, with the aim of minimizing the sum of weighted distances between facilities and regional customers, where the proximity between a facility and a regional customer is evaluated by the closest distance. And the authors propose a new location-allocation heuristic scheme to solve their problem. Convergence is proved under mild assumptions; and furthermore some preliminary numerical results are reported to show the effectiveness of the new algorithm.

The paper of A. Roldan et al. studies the existence and uniqueness of coincidence point for nonlinear mappings of any number of arguments under a weak (ψ, φ)-contractivity condition in partial metric spaces. The obtained results generalize, extend, and unify several classical and very recent related results in the literature in metric spaces and in partial metric spaces.

The paper of Z. Jia et al. extends the convergence analysis given by Han and Yuan for alternating direction method of multipliers (ADMM) from the strongly convex to a more general case. Further, the authors prove under the assumption that the individual functions are composites of strongly convex functions and linear functions that the classical ADMM for separable convex programming with two blocks can be extended to the case with more than three blocks.

The paper of M. Li and Y. You presents a simple proof for the same convergence rate of the relaxed proximal point algorithm (PPA) in both ergodic and nonergodic senses.

The paper of W.-S. Du et al. extends, generalizes, and improves several fundamental results on the existence (and uniqueness) of coincidence points and fixed points for well-known maps in the literature. Furthermore, some fixed coincidence point theorems for multivalued nonself maps in the context of complete metric spaces are given.

The paper of F. Ma et al. develops, studies, and implements a new prediction-correction method for monotone variational inequalities with separable structure. At each iteration, the proposed algorithm also allows the involved subvariational inequalities to be solved in parallel.

The paper of A. Barbagallo and P. Mauro concerns a dynamic oligopolistic market equilibrium problem in the realistic case in which the presence of capacity constraints and production excesses are allowed and, moreover, the production function depends not only on the time but also on the equilibrium distribution. The authors prove the equivalence between this equilibrium definition and a suitable evolutionary quasi-variational inequality, and they study the analysis of existence, regularity, and sensitivity of solutions.

Acknowledgments

We would like to express our sincere thanks to the authors for contributing to this issue, as well as to the anonymous reviewers for their generous and accurate refereeing process and their valuable comments and suggestions. Without their evident support this special issue would not have come out. We want to also acknowledge with gratitude the assistance and help provided by the editorial board members of this journal during the preparation of this special issue.

Abdellah Bnouhachem
Abdelouahed Hamdi
Xu Minghua

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 598563, 10 pages
http://dx.doi.org/10.1155/2014/598563

Research Article

Proximal Alternating Direction Method with Relaxed Proximal Parameters for the Least Squares Covariance Adjustment Problem

Minghua Xu,[1] Yong Zhang,[1] Qinglong Huang,[1] and Zhenhua Yang[2]

[1] *School of Mathematics and Physics, Changzhou University, Jiangsu 213164, China*
[2] *College of Science, Nanjing University of Posts and Telecommunications, Jiangsu 210003, China*

Correspondence should be addressed to Minghua Xu; xuminghua@cczu.edu.cn

Received 13 June 2013; Accepted 27 July 2013; Published 21 January 2014

Academic Editor: Abdellah Bnouhachem

We consider the problem of seeking a symmetric positive semidefinite matrix in a closed convex set to approximate a given matrix. This problem may arise in several areas of numerical linear algebra or come from finance industry or statistics and thus has many applications. For solving this class of matrix optimization problems, many methods have been proposed in the literature. The proximal alternating direction method is one of those methods which can be easily applied to solve these matrix optimization problems. Generally, the proximal parameters of the proximal alternating direction method are greater than zero. In this paper, we conclude that the restriction on the proximal parameters can be relaxed for solving this kind of matrix optimization problems. Numerical experiments also show that the proximal alternating direction method with the relaxed proximal parameters is convergent and generally has a better performance than the classical proximal alternating direction method.

1. Introduction

This paper concerns the following problem:

$$\min_X \left\{ \frac{1}{2}\|X - C\|_F^2 \mid X \in S_+^n \cap S_B \right\}, \tag{1}$$

where $C \in R^{n\times n}$ is a given symmetric matrix,

$$S_+^n = \left\{ X \in R^{n\times n} \mid X^T = X, X \succeq 0 \right\},$$

$$S_B = \left\{ X \in R^{n\times n} \mid \operatorname{Tr}(A_i X) = b_i, i = 1, 2, \ldots, p, \right. \tag{2}$$
$$\left. \operatorname{Tr}(G_j X) \le d_j, j = 1, 2, \ldots, m \right\},$$

matrices $A_i \in R^{n\times n}$ and $G_j \in R^{n\times n}$ are symmetric and scalars, b_i and d_j are the problem data, $X \succeq 0$ denotes that X is a positive semidefinite matrix, Tr denotes the trace of a matrix, and $\|\cdot\|_F$ denotes the Frobenius norm; that is,

$$\|X\|_F = \left(\operatorname{Tr}\left(X^T X\right)\right)^{1/2} = \left(\sum_{i,j=1}^{n} X_{ij}^2\right)^{1/2}, \tag{3}$$

and $S_+^n \cap S_B$ is nonempty. Throughout this paper, we assume that the Slater's constraint qualification condition holds so that there is no duality gap if we use Lagrangian techniques to find the optimal solution to problem (1).

Problem (1) is a type of matrix nearness problem, that is, the problem of finding a matrix that satisfies some properties and is nearest to a given one. Problem (1) can be called the *least squares covariance adjustment problem* or the *least squares semidefinite programming problem* and solved by many methods [1–4]. In a least squares covariance adjustment problem, we make adjustments to a symmetric matrix so that it is consistent with prior knowledge or assumptions and a valid covariance matrix [2, 5, 6]. The matrix nearness problem has many applications especially in several areas of numerical linear algebra, finance industry, and statistics in [6]. A recent survey of matrix nearness problems can be found in [7]. It is clear that the matrix nearness problem considered here is a convex optimization problem. It thus follows from the strict feasibility and coercivity of the objective function that the minimum of (1) is attainable and unique.

In the literature of interior point algorithms, S_+^n is called the semidefinite cone and the related problem (1) belongs to the class of semidefinite programming (SDP) and second-order cone programming (SOCP) [8]. In fact, it is possible to reformulate problem (1) into a mixed SDP and SOCP as in [3, 9]:

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \langle A_i, X\rangle = b_i, \quad i = 1, 2, \ldots, p, \\ & \langle G_j, X\rangle \le d_j, \quad j = 1, 2, \ldots, m, \\ & t \ge \|X - C\|_F, \\ & X \in S_+^n, \end{aligned} \tag{4}$$

where $\langle X, Y\rangle = \operatorname{Tr}(X^T Y)$.

Thus, problem (1) can be efficiently solved by standard interior-point methods such as SeDuMi [10] and SDPT3 [11] when the number of variables (i.e., entries in the matrix X) is modest, say under 1000 (corresponds to n around 32) and the number of equality and inequality constraints is not too large (say 5,000) [2, 3, 12].

Specially, let

$$S_B = \{X \in R^{n\times n} \mid \operatorname{Diag}(X) = e\}, \tag{5}$$

where $\operatorname{Diag}(X)$ is the vector of diagonal elements of X and e is the vector of 1s. Then problem (1) can be viewed as the nearest correlation matrix problem. For the nearest correlation matrix problem, a quadratically convergent Newton algorithm was presented recently by Qi and Sun [13], and improved by Borsdorf and Higham [1]. For problem (1) with equality and inequality constraints, one difficulty in finding an efficient method for solving this problem is the presence of the inequality constraints. In [3], Gao and Sun overcome this difficulty by reformulating the problem as a system of semismooth equations with two level metric projection operators and then design an inexact smoothing Newton method to solve the resulting semismooth system. For the problem (1) with large number of equality and inequality constraints, the numerical experiments in [14] show that the alternating direction method (hereafter alternating direction method is abbreviated as ADM) is more efficient in computing time than the inexact smoothing Newton method which additionally requires solving a large system of linear equations at each iteration. The ADM has many applications in solving optimization problems [15, 16]. Papers written by Zhang, Han, Li, Yuan, and Bauschke and Borwein show that the ADM can be applied to solve convex feasibility problems [17–19].

The proximal ADM is a class of ADM type methods which can also be easily applied to solve the matrix optimization problems. Generally, the proximal parameters (i.e., the parameters r and s in (14) and (15)) of the proximal ADM are greater than zero. In this paper, we will show that the restriction on the proximal parameters can be relaxed while the proximal ADM is used to solve problem (1). Numerical experiments also show that the proximal ADM with the relaxed proximal parameters generally has a better performance than the classical proximal ADM.

The paper is organized as follows. In Section 2, we give some preliminaries about the proximal alternating direction method. In Section 3, we convert the problem (1) to a structured variational inequality and apply the proximal ADM to solve it. The basic analysis and convergent results of the proximal ADM with relaxed proximal parameters are built in Section 4. Preliminary numerical results are reported in Section 5. Finally, we give some conclusions in Section 6.

2. Proximal Alternating Direction Method

In order to introduce the proximal ADM, we first consider the following structured variational inequality problem which includes two separable subvariational inequality problems: find $(x, y) \in \Omega$ such that

$$\begin{aligned} &(x' - x)^T f(x) \ge 0, \\ &(y' - y)^T g(y) \ge 0, \end{aligned} \quad \forall (x', y') \in \Omega, \tag{6}$$

where

$$\Omega = \{(x, y) \mid Ax + By = b, x \in \mathscr{X}, y \in \mathscr{Y}\}, \tag{7}$$

$f : R^{n_1} \to R^{n_1}$ and $g : R^{n_2} \to R^{n_2}$ are monotone; that is,

$$\begin{aligned} &(\tilde{x} - x)^T (f(\tilde{x}) - f(x)) \ge 0, \quad \forall \tilde{x}, x \in R^{n_1}, \\ &(\tilde{y} - y)^T (g(\tilde{y}) - g(y)) \ge 0, \quad \forall \tilde{y}, y \in R^{n_2}, \end{aligned} \tag{8}$$

$A \in R^{l\times n_1}$, $B \in R^{l\times n_2}$, and $b \in R^l$; $\mathscr{X} \subset R^{n_1}$ and $\mathscr{Y} \subset R^{n_2}$ are closed convex sets. Studies of such variational inequality can be found in Glowinski [20], Glowinski and Le Tallec [21], Eckstein and Fukushima [22–24], He and Yang [25], He et al. [26], and Xu [27].

By attaching a Lagrange multiplier vector $\lambda \in R^l$ to the linear constraint $Ax + By = b$, problem (6)-(7) can be explained as the following form (see [20, 21, 24]): find $w = (x, y, \lambda) \in \mathscr{W}$ such that

$$\begin{aligned} &(x' - x)^T [f(x) - A^T\lambda] \ge 0 \\ &(y' - y)^T [g(y) - B^T\lambda] \ge 0, \quad \forall w' = (x', y', \lambda') \in \mathscr{W}, \\ &Ax + By - b = 0, \end{aligned} \tag{9}$$

where

$$\mathscr{W} = \mathscr{X} \times \mathscr{Y} \times R^l. \tag{10}$$

For solving (9)-(10), Gabay [28] and Gabay and Mercier [29] proposed the ADM method. In the classical ADM method, the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) \in \mathscr{W}$ is generated from a given triple $w^k = (x^k, y^k, \lambda^k) \in \mathscr{W}$ via the following procedure.

First, x^{k+1} is found by solving the following problem:

$$(x' - x)^T \{f(x) - A^T [\lambda^k - \beta (Ax + By^k - b)]\} \ge 0, \quad \forall x' \in \mathscr{X}, \tag{11}$$

where $x \in \mathcal{X}$. Then, y^{k+1} is obtained by solving

$$\left(y' - y\right)^T \left\{g\left(y\right) - B^T\left[\lambda^k - \beta\left(Ax^{k+1} + By - b\right)\right]\right\} \geq 0, \quad \forall y' \in \mathcal{Y}, \tag{12}$$

where $y \in \mathcal{Y}$. Finally, the multiplier is updated by

$$\lambda^{k+1} = \lambda - \beta\left(Ax^{k+1} + By^{k+1} - b\right), \tag{13}$$

where $\beta > 0$ is a given penalty parameter for the linearly constraint $Ax + By - b = 0$. Most of the existing ADM methods require that the subvariational inequality problems (11)-(12) should be solved exactly at each iteration. Note that the involved subvariational inequality problem (11)-(12) may not be well-conditioned without strongly monotone assumptions on f and g. Hence, it is difficult to solve these subvariational inequality problems exactly in many cases. In order to improve the condition of solving the subproblem by the ADM, some proximal ADMs were proposed (see, e.g., [26, 27, 30–34]). The classical proximal ADM is one of the attractive ADMs. From a given triple $w^k = (x^k, y^k, \lambda^k) \in \mathcal{W}$, the classical proximal ADM produces the new iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) \in \mathcal{W}$ by the following procedure.

First, x^{k+1} is obtained by solving the following variational inequality problem:

$$\begin{aligned}\left(x' - x\right)^T \left\{f\left(x\right) - A^T\left[\lambda^k - \beta\left(Ax + By^k - b\right)\right]\right.\\ \left. + r\left(x - x^k\right)\right\} \geq 0, \quad \forall x' \in \mathcal{X},\end{aligned} \tag{14}$$

where $r > 0$ is the given proximal parameter and $x \in \mathcal{X}$. Then, y^{k+1} is found by solving

$$\begin{aligned}\left(y' - y\right)^T \left\{g\left(y\right) - B^T\left[\lambda^k - \beta\left(Ax^{k+1} + By - b\right)\right]\right.\\ \left. + s\left(y - y^k\right)\right\} \geq 0, \quad \forall y' \in \mathcal{Y},\end{aligned} \tag{15}$$

where $s > 0$ is the given proximal parameter and $y \in \mathcal{Y}$. Finally, the multiplier is updated by

$$\lambda^{k+1} = \lambda^k - \beta\left(Ax^{k+1} + By^{k+1} - b\right). \tag{16}$$

In this paper, we will conclude that problem (1) can be solved by the proximal ADM and the restriction on the proximal parameters $r > 0$, $s > 0$ can be relaxed as $r > -1/2$, $s > -1/2$ when the proximal ADM is applied to solve problem (1). Our numerical experiments later also show that the numerical performance of the proximal ADM with smaller value of proximal parameters is generally better than the proximal ADM with comparatively larger value of proximal parameters.

3. Converting Problem (1) to a Structured Variational Inequality

In order to solve the problem (1) with proximal ADM, we convert problem (1) to the following equivalent one:

$$\begin{aligned}\min_{X,Y} \quad & \frac{1}{2}\|X - C\|_F^2 + \frac{1}{2}\|Y - C\|_F^2 \\ \text{s.t.} \quad & X - Y = 0, \\ & X \in S_+^n, \quad Y \in S_B.\end{aligned} \tag{17}$$

Following the KKT condition of (17), the solution to (17) can be found by finding $w = (X, Y, \Lambda) \in \mathcal{W}$ such that

$$\begin{aligned}&\left\langle X' - X, (X - C) - \Lambda\right\rangle \geq 0, \\ &\left\langle Y' - Y, (Y - C) + \Lambda\right\rangle \geq 0, \qquad \forall w' = \left(X', Y', \Lambda'\right) \in \mathcal{W}, \\ &\qquad X - Y = 0,\end{aligned} \tag{18}$$

where

$$\mathcal{W} = S_+^n \times S_B \times R^{n\times n}. \tag{19}$$

It is easy to see that problem (18)-(19) is a special case of the structured variational inequality (9)-(10) and thus can be solved by proximal ADM. For given $w^k = (X^k, Y^k, \Lambda^k) \in \mathcal{W}$, it is fortunate that the $w^{k+1} = (X^{k+1}, Y^{k+1}, \Lambda^{k+1})$ can be exactly obtained by the proximal ADM in the following way:

$$X^{k+1} = P_{S_+^n}\left\{\frac{1}{1+\beta+r}\left(C + rX^k + \beta Y^k + \Lambda^k\right)\right\}, \tag{20}$$

$$Y^{k+1} = P_{S_B}\left\{\frac{1}{1+\beta+s}\left(C + \beta X^{k+1} + sY^k - \Lambda^k\right)\right\}, \tag{21}$$

$$\Lambda^{k+1} = \Lambda^k - \beta\left(X^{k+1} - Y^{k+1}\right), \tag{22}$$

where the projection of v on a nonempty closed convex set S of $R^{m\times n}$ under Frobenius norm, denoted by $P_S(v)$, is the unique solution to the following problem; that is,

$$P_S\left(v\right) = \arg\min_u \left\{\|u - v\|_F^2 \mid u \in S\right\}. \tag{23}$$

It follows that the solution to

$$\min\left\{\frac{1}{2}\|Z - X\|_F^2 \mid Z \in S_+^n\right\} \tag{24}$$

is called the projection of X on S_+^n and denoted by $P_{S_+^n}(X)$. Using the fact that matrix Frobenius norm is invariant under unitary transform, it is known (see [35]) that

$$P_{S_+^n}\left(X\right) = Q\widetilde{\Lambda}Q^T, \tag{25}$$

where

$$Q^T XQ = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_n\right) \tag{26}$$

is the symmetric Schur decomposition of X ($Q = (q_1, \ldots, q_n)$ is an orthogonal matrix whose column vector q_i, $i = 1, \ldots, n$,

is the eigenvector of X, and λ_i, $i = 1, \ldots, n$, is the related eigenvalue),

$$\widetilde{\Lambda} = \operatorname{diag}\left(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n\right), \quad \widetilde{\lambda}_i = \max\left(\lambda_i, 0\right). \tag{27}$$

In order to obtain the projection $P_{S_B}(X)$, we need to solve the following quadratic program:

$$\begin{aligned} \min_{Z} \quad & \frac{1}{2}\|Z - X\|_F^2 \\ \text{s.t.} \quad & \operatorname{Tr}\left(A_i Z\right) = b_i, \quad i = 1, 2, \ldots, p, \\ & \operatorname{Tr}\left(G_j Z\right) \le d_j, \quad j = 1, 2, \ldots, m. \end{aligned} \tag{28}$$

The dual problem of (28) can be written as

$$\begin{aligned} \min_{v} \quad & \frac{1}{2} v^T H v + q^T v \\ \text{s.t.} \quad & v \in R^p \times R_+^m, \end{aligned} \tag{29}$$

where H is positive semidefinite and H and q have the following form, respectively:

$$H = \begin{pmatrix} \operatorname{Tr}\left(A_1 A_1^T\right) & \cdots & \operatorname{Tr}\left(A_1 A_p^T\right) & \operatorname{Tr}\left(A_1 G_1^T\right) & \cdots & \operatorname{Tr}\left(A_1 G_m^T\right) \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \operatorname{Tr}\left(A_p A_1^T\right) & \cdots & \operatorname{Tr}\left(A_p A_p^T\right) & \operatorname{Tr}\left(A_p G_1^T\right) & \cdots & \operatorname{Tr}\left(A_p G_m^T\right) \\ \operatorname{Tr}\left(G_1 A_1^T\right) & \cdots & \operatorname{Tr}\left(G_1 A_p^T\right) & \operatorname{Tr}\left(G_1 G_1^T\right) & \cdots & \operatorname{Tr}\left(G_1 G_m^T\right) \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \operatorname{Tr}\left(G_m A_1^T\right) & \cdots & \operatorname{Tr}\left(G_m A_p^T\right) & \operatorname{Tr}\left(G_m G_1^T\right) & \cdots & \operatorname{Tr}\left(G_m G_m^T\right) \end{pmatrix}, \quad q = \begin{pmatrix} b_1 - \operatorname{Tr}\left(A_1 X\right) \\ \vdots \\ b_p - \operatorname{Tr}\left(A_p X\right) \\ d_1 - \operatorname{Tr}\left(G_1 X\right) \\ \vdots \\ d_m - \operatorname{Tr}\left(G_m X\right) \end{pmatrix}. \tag{30}$$

Problem (29) is often a medium-scale quadratic programming (QP) problem. A variety of methods for solving the QP are commonly used, including interior-point methods and active set algorithm (see [36, 37]).

Particularly, if S_B is the following special case:

$$S_B = \left\{X \in R^{n\times n} \mid X^T = X, H_L \le X \le H_U\right\}, \tag{31}$$

where $H \ge 0$ expresses that each element of H is nonnegative, H_L and H_U are given $n \times n$ symmetric matrices, and $X \le H_U$ means that $H_U - X \ge 0$; then $P_{S_B}(X)$ is easy to be carried out and is given by

$$P_{S_B}(X) = \min\left(\max\left(X, H_L\right), H_U\right), \tag{32}$$

where $\max(X, Y)$ and $\min(X, Y)$ compute the element-wise maximum and minimum of matrix X and Y, respectively.

4. Main Results

Let $\{w^k\}$ be the sequence generated by applying the procedure (14)–(16) to problem (18)-(19); then for any $w' = (X', Y', \Lambda') \in \mathscr{W}$, we have that

$$\begin{aligned} & \left\langle X' - X^{k+1}, X^{k+1} - C - \Lambda^{k+1} - \beta\left(Y^k - Y^{k+1}\right) \right. \\ & \quad \left. + r\left(X^{k+1} - X^k\right)\right\rangle \ge 0, \\ & \left\langle Y' - Y^{k+1}, Y^{k+1} - C + \Lambda^{k+1} + s\left(Y^{k+1} - Y^k\right)\right\rangle \ge 0, \\ & \Lambda^{k+1} = \Lambda^k - \beta\left(X^{k+1} - Y^{k+1}\right). \end{aligned} \tag{33}$$

Further, letting

$$F\left(w^{k+1}\right) = \begin{pmatrix} X^{k+1} - C - \Lambda^{k+1} \\ Y^{k+1} - C + \Lambda^{k+1} \\ X^{k+1} - Y^{k+1} \end{pmatrix},$$

$$d_1\left(w^k, w^{k+1}\right) = \begin{pmatrix} rI_n & 0 & 0 \\ 0 & (s+\beta) I_n & 0 \\ 0 & 0 & \frac{1}{\beta} I_n \end{pmatrix} \begin{pmatrix} X^k - X^{k+1} \\ Y^k - Y^{k+1} \\ \Lambda^k - \Lambda^{k+1} \end{pmatrix}, \tag{34}$$

where $I_n \in R^{n\times n}$ is the unit matrix, and

$$d_2\left(w^k, w^{k+1}\right) = F\left(w^{k+1}\right) - \beta \begin{pmatrix} I_n \\ -I_n \\ 0 \end{pmatrix} \left(Y^k - Y^{k+1}\right), \tag{35}$$

then we can get the following lemmas.

Lemma 1. *Let $\{w^k\}$ be the sequence generated by applying the proximal ADM to problem* (18)-(19) *and let $w^* \in \mathscr{W}^*$ be any solution to problem* (18)-(19)*; then one has*

$$\begin{aligned} & \left\langle w^{k+1} - w^*, d_2\left(w^k, w^{k+1}\right)\right\rangle \\ & \quad \ge -\left\langle \Lambda^k - \Lambda^{k+1}, Y^k - Y^{k+1}\right\rangle + \left\|X^{k+1} - X^*\right\|_F^2 \\ & \quad + \left\|Y^{k+1} - Y^*\right\|_F^2. \end{aligned} \tag{36}$$

Proof. From (22) and (35), we have

$$\left\langle w^{k+1}-w^{*}, d_2\left(w^{k}, w^{k+1}\right)\right\rangle = -\left\langle \Lambda^{k}-\Lambda^{k+1}, Y^{k}-Y^{k+1}\right\rangle + \left\langle w^{k+1}-w^{*}, F\left(w^{k+1}\right)\right\rangle. \tag{37}$$

Since (9) and w^* are a solution to problem (18)-(19) and $X^{k+1} \in S_+^n$, $Y^{k+1} \in S_B$, we have

$$\left\langle w^{k+1}-w^{*}, F\left(w^{*}\right)\right\rangle \geq 0. \tag{38}$$

From (38), it follows that

$$\left\langle w^{k+1}-w^{*}, F\left(w^{k+1}\right)-F\left(w^{k+1}\right)+F\left(w^{*}\right)\right\rangle \geq 0. \tag{39}$$

Thus, we have

$$\begin{aligned}
&\left\langle w^{k+1}-w^{*}, F\left(w^{k+1}\right)\right\rangle \\
&\quad\geq \left\langle w^{k+1}-w^{*}, F\left(w^{k+1}\right)-F\left(w^{*}\right)\right\rangle \\
&\quad= \left\langle X^{k+1}-X^{*}, X^{k+1}-X^{*}-\left(\Lambda^{k+1}-\Lambda^{*}\right)\right\rangle \\
&\qquad+\left\langle Y^{k+1}-Y^{*}, Y^{k+1}-Y^{*}+\left(\Lambda^{k+1}-\Lambda^{*}\right)\right\rangle \\
&\qquad+\left\langle \Lambda^{k+1}-\Lambda^{*}, X^{k+1}-X^{*}-\left(Y^{k+1}-Y^{*}\right)\right\rangle \\
&\quad= \left\langle X^{k+1}-X^{*}, X^{k+1}-X^{*}\right\rangle+\left\langle Y^{k+1}-Y^{*}, Y^{k+1}-Y^{*}\right\rangle \\
&\quad= \left\|X^{k+1}-X^{*}\right\|_F^2+\left\|Y^{k+1}-Y^{*}\right\|_F^2.
\end{aligned} \tag{40}$$

Substituting (40) into (37), we get the assertion of this lemma. □

Lemma 2. *Let $\{w^k\}$ be the sequence generated by applying the proximal ADM to problem* (18)-(19) *and let $w^* \in \mathscr{W}^*$ be any solution to problem* (18)-(19)*; then one has*

$$\begin{aligned}
&\left\langle w^{k}-w^{*}, G_0\left(w^{k}-w^{k+1}\right)\right\rangle \\
&\geq \left\langle w^{k}-w^{k+1}, G_0\left(w^{k}-w^{k+1}\right)\right\rangle-\left\langle \Lambda^{k}-\Lambda^{k+1}, Y^{k}-Y^{k+1}\right\rangle \\
&\quad+\left\|X^{k+1}-X^{*}\right\|_F^2+\left\|Y^{k+1}-Y^{*}\right\|_F^2,
\end{aligned} \tag{41}$$

where

$$G_0=\begin{pmatrix} rI_n & 0 & 0 \\ 0 & (s+\beta)I_n & 0 \\ 0 & 0 & \frac{1}{\beta}I_n \end{pmatrix}. \tag{42}$$

Proof. It follows from (33) that

$$\left\langle w'-w^{k+1}, d_2\left(w^{k}, w^{k+1}\right)-d_1\left(w^{k}, w^{k+1}\right)\right\rangle \geq 0, \quad \forall w' \in \mathscr{W}. \tag{43}$$

Thus, we have

$$\begin{aligned}
&\left\langle w^{k+1}-w^{*}, d_1\left(w^{k}, w^{k+1}\right)\right\rangle \\
&\quad\geq \left\langle w^{k+1}-w^{*}, d_2\left(w^{k}, w^{k+1}\right)\right\rangle \\
&\quad\geq -\left\langle \Lambda^{k}-\Lambda^{k+1}, Y^{k}-Y^{k+1}\right\rangle+\left\|X^{k+1}-X^{*}\right\|_F^2 \\
&\qquad+\left\|Y^{k+1}-Y^{*}\right\|_F^2.
\end{aligned} \tag{44}$$

From the above inequality, we get

$$\begin{aligned}
&\left\langle w^{k}-w^{*}, G_0\left(w^{k}-w^{k+1}\right)\right\rangle \\
&\quad\geq \left\langle w^{k}-w^{k+1}, G_0\left(w^{k}-w^{k+1}\right)\right\rangle \\
&\qquad-\left\langle \Lambda^{k}-\Lambda^{k+1}, Y^{k}-Y^{k+1}\right\rangle+\left\|X^{k+1}-X^{*}\right\|_F^2 \\
&\qquad+\left\|Y^{k+1}-Y^{*}\right\|_F^2.
\end{aligned} \tag{45}$$

Hence, (41) holds and the proof is completed. □

Theorem 3. *Let $\{w^k\}$ be the sequence generated by applying the proximal ADM to problem* (18)-(19) *and let $w^* \in \mathscr{W}^*$ be any solution to problem* (18)-(19)*; then one has*

$$\left\|w^{k+1}-w^{*}\right\|_G^2 \leq \left\|w^{k}-w^{*}\right\|_G^2 - \left\langle w^{k}-w^{k+1}, M\left(w^{k}-w^{k+1}\right)\right\rangle, \tag{46}$$

where

$$G=\begin{pmatrix} (r+1)I_n & 0 & 0 \\ 0 & (1+s+\beta)I_n & 0 \\ 0 & 0 & \frac{1}{\beta}I_n \end{pmatrix},$$
$$M=\begin{pmatrix} \left(\frac{1}{2}+r\right)I_n & 0 & 0 \\ 0 & \left(\frac{1}{2}+s+\beta\right)I_n & -I_n \\ 0 & -I_n & \frac{1}{\beta}I_n \end{pmatrix}, \tag{47}$$

and $\|w\|_G^2 = \langle w, Gw\rangle$.

Proof. From (41), we have

$$\begin{aligned}
&\left\|w^{k+1}-w^{*}\right\|_{G_0}^2 \\
&\quad= \left\|w^{k}-w^{*}-\left(w^{k}-w^{k+1}\right)\right\|_{G_0}^2 \\
&\quad\leq \left\|w^{k}-w^{*}\right\|_{G_0}^2-2\left\|w^{k}-w^{k+1}\right\|_{G_0}^2
\end{aligned}$$

$$
\begin{aligned}
&+2\left\langle \Lambda^k-\Lambda^{k+1}, Y^k-Y^{k+1}\right\rangle-2\left\|X^{k+1}-X^*\right\|_F^2\\
&-2\left\|Y^{k+1}-Y^*\right\|_F^2+\left\|w^k-w^{k+1}\right\|_{G_0}^2\\
&=\left\|w^k-w^*\right\|_{G_0}^2-\left\|w^k-w^{k+1}\right\|_{G_0}^2\\
&+2\left\langle \Lambda^k-\Lambda^{k+1}, Y^k-Y^{k+1}\right\rangle-2\left\|X^{k+1}-X^*\right\|_F^2\\
&-2\left\|Y^{k+1}-Y^*\right\|_F^2.
\end{aligned}
\tag{48}
$$

Rearranging the inequality above, we find that

$$
\begin{aligned}
\left\|w^{k+1}-w^*\right\|_G^2 \le \left\|w^k-w^*\right\|_G^2 &- \left\langle w^k-w^{k+1}, \begin{pmatrix} rI_n & 0 & 0\\ 0 & (s+\beta)I_n & -I_n\\ 0 & -I_n & \frac{1}{\beta}I_n\end{pmatrix}\left(w^k-w^{k+1}\right)\right\rangle - \left(\left\|X^{k+1}-X^*\right\|_F^2+\left\|X^k-X^*\right\|_F^2\right)\\
&-\left(\left\|Y^{k+1}-Y^*\right\|_F^2+\left\|Y^k-Y^*\right\|_F^2\right).
\end{aligned}
\tag{49}
$$

Using the Cauchy-Schwarz Inequality on the last term of the right-hand side of (49), we obtain

$$
\begin{aligned}
\left\|X^{k+1}-X^*\right\|_F^2+\left\|X^k-X^*\right\|_F^2 &\ge \frac{1}{2}\left\|X^{k+1}-X^k\right\|_F^2,\\
\left\|Y^{k+1}-Y^*\right\|_F^2+\left\|Y^k-Y^*\right\|_F^2 &\ge \frac{1}{2}\left\|Y^{k+1}-Y^k\right\|_F^2.
\end{aligned}
\tag{50}
$$

Substituting (50) into (49), we get

$$
\begin{aligned}
\left\|w^{k+1}-w^*\right\|_G^2 \le \left\|w^k-w^*\right\|_G^2 &\\
-\left\langle w^k-w^{k+1}, M\left(w^k-w^{k+1}\right)\right\rangle&.
\end{aligned}
\tag{51}
$$

Thus, the proof is completed. □

Based on the Theorem 3, we get the following lemma.

Lemma 4. *Let $\{w^k\}$ be the sequence generated by applying proximal ADM to problem* (18)-(19), *$w^* \in \mathscr{W}^*$ any solution to problem* (18)-(19), *$r > -1/2$, and $s > -1/2$; then one has the following.*

(1) *The sequence $\{\|w^k-w^*\|_G^2\}$ is nonincreasing;*

(2) *The sequence $\{w^k\}$ is bounded;*

(3) $\lim_{k\to\infty}\|w^{k+1}-w^k\|_F^2=0$;

(4) *G and M are both symmetric positive-definite matrices.*

Proof. Since

$$
\begin{vmatrix} \left(\frac{1}{2}+s+\beta\right)I_n & -I_n\\ -I_n & \frac{1}{\beta}I_n\end{vmatrix} = \frac{((1/2)+s)}{\beta},
\tag{52}
$$

it is easy to check that if $r > -1/2$, $s > -1/2$, then G and M are symmetric positive-definite matrices.

Let $\tau > 0$ be the smallest eigenvalue of matrix M. Then, from (46), we have

$$
\left\|w^{k+1}-w^*\right\|_G^2 \le \left\|w^k-w^*\right\|_G^2-\tau\left\|w^k-w^{k+1}\right\|_F^2.
\tag{53}
$$

Following (53), we immediately have that $\|w^k-w^*\|_G^2$ is nonincreasing and thus the sequence $\{w^k\}$ is bounded. Moreover, we have

$$
\left\|w^{k+1}-w^*\right\|_G^2 \le \left\|w^0-w^*\right\|_G^2-\tau\sum_{j=0}^{k}\left\|w^j-w^{j+1}\right\|_F^2.
\tag{54}
$$

So, we get

$$
\sum_{j=0}^{k}\left\|w^j-w^{j+1}\right\|_F^2<\infty, \quad \forall k>0,
\tag{55}
$$

then

$$
\lim_{k\to\infty}\left\|w^k-w^{k+1}\right\|_F^2=0.
\tag{56}
$$

Thus, the proof is completed. □

Following Lemma 4, now we are in the stage of giving the main convergence results of proximal ADM with $r > -1/2$ and $s > -1/2$ for problem (18)-(19).

Theorem 5. *Let $\{w^k\}$ be the sequence generated by applying proximal ADM to problem* (18)-(19), *$r > -1/2$, and $s > -1/2$; then $\{w^k\}$ converges to a solution point of* (18)-(19).

Proof. Since the sequence $\{w^k\}$ is bounded (see point (2) of Lemma 4), it has at least one cluster point. Let w^∞ be a cluster point of $\{w^k\}$ and the subsequence $\{w^{k_j}\}$ converges to w^∞. It follows from (33) that

$$
\begin{aligned}
\lim_{j\to\infty}\left\langle X'-X^{k_j+1}, X^{k_j+1}-C-\Lambda^{k_j+1}-\beta\left(Y^{k_j}-Y^{k_j+1}\right)\right.&\\
\left.+r\left(X^{k_j+1}-X^{k_j}\right)\right\rangle \ge 0,&
\end{aligned}
$$

Table 1: Numerical results of Example 6.

n	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	31	0.292	34	0.331	72	0.764
200	33	1.346	39	1.570	84	3.364
300	38	4.265	41	5.746	90	9.991
400	40	9.872	43	9.919	94	22.03
500	39	15.83	45	18.39	98	39.91

$$\begin{aligned}&\lim_{j\to\infty}\left\langle Y'-Y^{k_j+1}, Y^{k_j+1}-C+\Lambda^{k_j+1}+s\left(Y^{k_j+1}-Y^{k_j}\right)\right\rangle\\&\quad\geq 0,\quad \forall w'\in\mathscr{W},\\&\qquad\lim_{j\to\infty}\Lambda^{k_j+1}=\Lambda^{k_j}-\beta\left(X^{k_j+1}-Y^{k_j+1}\right).\end{aligned}\tag{57}$$

Following point (3) of Lemma 4, we have

$$\begin{aligned}&\left\langle X'-X^{\infty}, X^{\infty}-C-\Lambda^{\infty}\right\rangle\geq 0,\\&\left\langle Y'-Y^{\infty}, Y^{\infty}-C+\Lambda^{\infty}\right\rangle\geq 0,\quad \forall w'\in\mathscr{W},\\&\qquad X^{\infty}-Y^{\infty}=0.\end{aligned}\tag{58}$$

This means that w^{∞} is a solution point of (18)-(19). Since $\{w^{k_j}\}$ converges to w^{∞}, we have that, for any given $\varepsilon>0$, there exists an integer $N>0$ such that

$$\left\|w^{k_j}-w^{\infty}\right\|_G^2<\varepsilon,\quad \forall k_j\geq N.\tag{59}$$

Furthermore, using the inequality (53), we have

$$\left\|w^{k}-w^{\infty}\right\|_G^2<\left\|w^{k_j}-w^{\infty}\right\|_G^2,\quad \forall k\geq k_j.\tag{60}$$

Combining (59) and (60), we get that

$$\left\|w^{k}-w^{\infty}\right\|_G^2<\varepsilon,\quad \forall k>N.\tag{61}$$

This implies that the sequence $\{w^k\}$ converges to w^{∞}. So the proof is completed. □

5. Numerical Experiments

In this section, we implement the proximal ADM to solve the problem (1) and show the numerical performances of proximal ADM with different proximal parameters. Additionally, we compare the classical ADM (i.e., the proximal ADM with proximal parameters $r=0$ and $s=0$) with the alternating projections method proposed by Higham [6] numerically and show that the alternating projections method is not equivalent to proximal ADM with zero proximal parameters. All the codes were written in Matlab 7.1 and run on IBM notebook PC R400.

Example 6. In the first numerical experiment, we set the C_1 as an $n\times n$ matrix whose entries are generated randomly in $[-1,1]$. Let $C=(C_1+C_1^T)/2$ and further let the diagonal elements of C be 1 that is, $C_{ii}=1$, $i=1,2,\ldots,n$. In this test example, we simply let S_B be in the form of (31) and

$$\begin{aligned}&H_L=\left(l_{ij}\right)\in R^{n\times n},\\&l_{ij}=\begin{cases}-0.5, & i\neq j\\ 1, & i=j,\end{cases}\quad i,j=1,2,\ldots,n,\\&H_U=\left(u_{ij}\right)\in R^{n\times n},\\&u_{ij}=\begin{cases}0.5, & i\neq j\\ 1, & i=j,\end{cases}\quad i,j=1,2,\ldots,n.\end{aligned}\tag{62}$$

Moreover, let $X^0=\mathrm{eye}(n)$, $Y^0=\mathrm{eye}(n)$, $\Lambda^0=\mathrm{zeroes}(n)$, $\beta=4$, and $\varepsilon=10^{-6}$, where eye(n) and zeroes(n) are both the Matlab functions. For different problem size n and different proximal parameters r and s, Table 1 shows the computational results. There, we report the number of iterations (It.) and the computing time in seconds (CPU.) it takes to reach convergence. The stopping criterion of the proximal ADM is

$$\left\|w^{k+1}-w^{k}\right\|_{\max}<\varepsilon,\tag{63}$$

where $\|X\|_{\max}=\max(\max(\mathrm{abs}(X)))$ is the maximum absolute value of the elements of the matrix X.

Remark 7. Note that if the proximal parameters are equal to zero, that is, $r=0$ and $s=0$, then the proximal ADM is the classical ADM.

Example 8. All the data are the same as in Example 6 except that C_1 is an $n\times n$ matrix whose entries are generated randomly in $[-1000,1000]$,

$$\begin{aligned}&H_L=\left(l_{ij}\right)\in R^{n\times n},\\&l_{ij}=\begin{cases}-500, & i\neq j\\ 1000, & i=j,\end{cases}\quad i,j=1,2,\ldots,n,\\&H_U=\left(u_{ij}\right)\in R^{n\times n},\\&u_{ij}=\begin{cases}500, & i\neq j\\ 1000, & i=j,\end{cases}\quad i,j=1,2\ldots,n.\end{aligned}\tag{64}$$

The computational results are reported in Table 2.

Example 9. Let S_B be in the form of (31) and $l_{ij}=0$, $u_{ij}=+\infty$, $i,j=1,2,\ldots,n$. Assume that C, X_0, Y_0, Λ_0, β, ε, and the stopping criterion are the same as those in Example 6, but the diagonal elements of matrix C are replaced by

$$C_{ii}=\alpha+(1-\alpha)\times\mathrm{rand},\quad i=1,2,\ldots,n,\tag{65}$$

where $\alpha\in(0,1)$ is a given number, rand is the Matlab function generating a number randomly in $[0,1]$. In the following numerical experiments, we let $\alpha=0.2$. For different problem size n and different proximal parameters r and s, Table 3 shows the number of iterations and the computing time in seconds it takes to reach convergence.

Table 2: Numerical results of Example 8.

n	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	49	0.476	54	0.551	116	1.837
200	51	2.197	57	2.334	128	5.430
300	59	6.614	61	8.108	136	15.25
400	56	12.74	63	14.51	140	31.65
500	58	23.90	66	26.90	147	59.98

Table 3: Numerical results of Example 9.

n	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	32	0.282	35	0.288	70	0.566
200	33	1.295	36	1.397	72	4.006
300	34	3.745	37	4.156	73	8.285
400	34	7.885	37	8.571	73	16.73
500	34	14.07	37	15.42	74	29.87

Table 4: Numerical results of Example 10.

5	$r=-0.3, s=-0.3$		$r=0, s=0$		$r=3, s=3$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	32	0.259	35	0.300	70	0.557
200	33	1.306	36	1.424	72	2.880
300	33	3.750	37	4.087	72	7.958
400	34	7.799	37	8.546	74	16.98
500	34	13.96	37	16.10	74	30.77

Table 5: (a) Numerical results of Example 11 with $\hat{n}_r = 5$. (b) Numerical results of Example 11 with $\hat{n}_r = 10$.

(a)

n	$r=-0.3,\ s=-0.3$		$r=0,\ s=0$		$r=1,\ s=1$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	22	0.293	25	0.354	34	0.448
200	25	2.119	28	2.425	40	3.436
300	27	7.141	30	8.024	44	11.64
400	29	17.40	31	18.59	46	27.32
500	30	34.17	33	37.45	48	53.84

(b)

n	$r=-0.3,\ s=-0.3$		$r=0,\ s=0$		$r=1,\ s=1$	
	It.	CPU.	It.	CPU.	It.	CPU.
100	23	0.309	25	0.342	33	0.439
200	24	2.029	27	2.305	38	3.162
300	27	7.150	29	7.801	42	11.29
400	28	16.68	31	18.47	45	26.60
500	29	32.73	32	36.37	47	53.06

Example 10. All the data are the same as in Example 9 except that $\alpha = 0$. The computational results are reported in Table 4.

Example 11. Let C_1 be an $n \times n$ matrix whose entries are generated randomly in $[-0.5, 0.5]$, $C = (C_1 + C_1^T)/2$, and let the diagonal elements of C be 1. And let

$$S_B = \left\{X \in R^{n\times n} \mid X = X^T, X_{ij} = e_{ij}, (i,j) \in \mathscr{B}_e, \right. \\ X_{ij} \geq l_{ij}, (i,j) \in \mathscr{B}_l, \\ \left. X_{ij} \leq u_{ij}, (i,j) \in \mathscr{B}_u\right\}, \tag{66}$$

where $\mathscr{B}_e$, $\mathscr{B}_l$, $\mathscr{B}_u$ are subsets of $\{(i,j) \mid 1 \leq i,j \leq n\}$ denoting the indexes of such entries of X that are constrained by equality, lower bounds, and upper bounds, respectively. In this test example, we let the index sets $\mathscr{B}_e$, $\mathscr{B}_l$, and $\mathscr{B}_u$ be the same as in Example 5.4 of [3]; that is, $\mathscr{B}_e = \{(i,i) \mid i = 1,2,\ldots,n\}$ and $\mathscr{B}_l, \mathscr{B}_u \subset \{(i,j) \mid 1 \leq i < j \leq n\}$ consist of the indices of $\min(\hat{n}_r, n-i)$ randomly generated elements at the ith row of X, $i = 1,2,\ldots,n$ with $\hat{n}_r = 5$ and $\hat{n}_r = 10$, respectively. We take $e_{ii} = 1$ for $(i,i) \in \mathscr{B}_e$, $l_{ij} = -0.1$ for $(i,j) \in \mathscr{B}_l$, and $u_{ij} = 0.1$ for $(i,j) \in \mathscr{B}_u$.

Moreover, let X_0, Y_0, Λ_0, β, ε, and the stopping criterion be the same as those in Example 6. For different problem size n, different proximal parameters r and s, and different values of $\hat{n}_r$, Tables 5(a) and 5(b) show the number of iterations and the computing time in seconds it takes to reach convergence, respectively.

Numerical experiments show that the proximal ADM with relaxed parameters is convergent. Moreover, we draw the conclusion that the proximal ADM with smaller value of proximal parameters generally converges more quickly than the proximal ADM with comparatively larger value of proximal parameters to solve the problem (1).

Example 12. In this test example, we apply the proximal ADM with $r = 0$, $s = 0$ (i.e., the classical ADM) to solve the nearest correlation matrix problem, that is, problem (1) with S_B in the form of (5), and compare the classical ADM numerically with the alternating projections method (APM) [6]. The APM computes the nearest correlation matrix to a symmetric $C \in R^{n\times n}$ by the following process:

$\Delta S_0 = 0, Y_0 = C;$
for $k = 1,2,\ldots$
$R_k = Y_{k-1} - \Delta S_{k-1};$
$X_k = P_{S_+^n}(R_k);$
$\Delta S_k = X_k - R_k;$
$Y_k = P_{S_B}(X_k);$
end.

In this numerical experiment, the stopping criterion of the APM is

$$\max\left\{\|X_k - X_{k-1}\|_{\max}, \|Y_k - Y_{k-1}\|_{\max}, \|X_k - Y_k\|_{\max}\right\} < \varepsilon. \tag{67}$$

Let the matrix C and the initial parameters of classical ADM be the same as those in Example 6. Table 6(a) reports the numerical performance of proximal ADM and the APM for computing the nearest correlation matrix to C.

Table 6: (a) Numerical results of Example 12. (b) Numerical results of Example 12.

(a)

n	ADM		APM	
	It.	CPU.	It.	CPU.
100	28	0.381	47	0.743
200	33	2.878	59	5.443
300	36	9.462	70	20.68
400	38	22.50	81	54.38
500	39	43.32	89	114.7

(b)

n	ADM		APM	
	It.	CPU.	It.	CPU.
100	27	0.634	42	0.582
200	30	2.590	59	5.428
300	32	8.524	65	19.36
400	34	20.34	75	50.79
500	35	39.43	86	111.6

Further, let C_1 be an $n \times n$ matrix whose entries are generated randomly in $[0, 1]$ and $C = (C_1 + C_1^T)/2$. The other data are the same as above. Table 6(b) reports the numerical performance of the classical ADM and the APM for computing the nearest correlation matrix to the matrix C. Numerical experiments show that the classical ADM generally exhibits a better numerical performance than the APM for the test problems above.

6. Conclusions

In this paper, we apply the proximal ADM to a class of matrix optimization problems and find that the restriction of proximal parameters can be relaxed. Moreover, numerical experiments show that the proximal ADM with relaxed parameters generally has a better numerical performance in solving the matrix optimization problem than the classical proximal alternating direction method.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank the referees very sincerely for their valuable suggestions and careful reading of their paper. This research is financially supported by a research Grant from the Research Grant Council of China (Project no. 10971095).

References

[1] R. Borsdorf and N. J. Higham, "A preconditioned Newton algorithm for the nearest correlation matrix," *IMA Journal of Numerical Analysis*, vol. 30, no. 1, pp. 94–107, 2010.

[2] S. Boyd and L. Xiao, "Least-squares covariance matrix adjustment," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 2, pp. 532–546, 2005.

[3] Y. Gao and D. Sun, "Calibrating least squares semidefinite programming with equality and inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1432–1457, 2009.

[4] S. Gravel and V. Elser, "Divide and concur: a general approach to constraint satisfaction," *Physical Review E*, vol. 78, Article ID 036706, 2008.

[5] N. J. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra and its Applications*, vol. 103, pp. 103–118, 1988.

[6] N. J. Higham, "Computing the nearest correlation matrix—a problem from finance," *IMA Journal of Numerical Analysis*, vol. 22, no. 3, pp. 329–343, 2002.

[7] N. J. Higham, "Matrix nearness problems and applications," in *Applications of Matrix Theory*, M. Gover and S. Barnett, Eds., vol. 22, pp. 1–27, Oxford University Press, Oxford, UK, 1989.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[9] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[10] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11/12, no. 1–4, pp. 625–653, 1999.

[11] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, vol. 95, no. 2, pp. 189–217, 2003.

[12] J. Malick, "A dual approach to semidefinite least-squares problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 1, pp. 272–284, 2004.

[13] H. Qi and D. Sun, "A quadratically convergent Newton method for computing the nearest correlation matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 2, pp. 360–385, 2006.

[14] 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.

[15] P. M. Pardalos and M. G. C. Resende, *Handbook of Applied Optimization*, Oxford University Press, Oxford, UK, 2002.

[16] P. M. Pardalos, T. M. Rassias, and A. A. Khan, *Nonlinear Analysis and Variational Problems*, vol. 35 of *Springer Optimization and Its Applications*, Springer, New York, NY, USA, 2010, In honor of George Isac, Edited by Panos M. Pardalos, Themistocles M. Rassias and Akhtar A. Khan.

[17] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Review*, vol. 38, no. 3, pp. 367–426, 1996.

[18] W. Zhang, D. Han, and Z. Li, "A self-adaptive projection method for solving the multiple-sets split feasibility problem," *Inverse Problems*, vol. 25, no. 11, 2009.

[19] W. Zhang, D. Han, and X. Yuan, "An efficient simultaneous method for the constrained multiple-sets split feasibility problem," *Computational Optimization and Applications*, vol. 52, no. 3, pp. 825–843, 2012.

[20] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, NY, USA, 1984.

[21] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, vol. 9 of

SIAM Studies in Applied Mathematics, SIAM, Philadelphia, Pa, USA, 1989.

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

[23] J. Eckstein and M. Fukushima, "Some reformulations and applications of the alternating direction method of multipliers," in *Large Scale Optimization: State of the Art*, W. W. Hager, D. W. Hearn, and P. M. Pardalos, Eds., pp. 115–134, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.

[24] M. Fukushima, "Application of the alternating direction method of multipliers to separable convex programming problems," *Computational Optimization and Applications*, vol. 1, no. 1, pp. 93–111, 1992.

[25] B. He and H. Yang, "Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities," *Operations Research Letters*, vol. 23, no. 3-5, pp. 151–161, 1998.

[26] B. He, L.-Z. Liao, D. Han, and H. Yang, "A new inexact alternating directions method for monotone variational inequalities," *Mathematical Programming*, vol. 92, no. 1, pp. 103–118, 2002.

[27] M. H. Xu, "Proximal alternating directions method for structured variational inequalities," *Journal of Optimization Theory and Applications*, vol. 134, no. 1, pp. 107–117, 2007.

[28] D. Gabay, "Applications of the method of multipliers to variational inequalities," in *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, M. Fortin and R. Glowinski, Eds., pp. 299–331, North-Holland, Amsterdam, The Netherlands, 1983.

[29] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximations," *Computer and Mathematics with Applications*, vol. 2, pp. 17–40, 1976.

[30] O. Güler, "New proximal point algorithms for convex minimization," *SIAM Journal on Optimization*, vol. 2, no. 4, pp. 649–664, 1992.

[31] W. W. Hager and H. Zhang, "Asymptotic convergence analysis of a new class of proximal point methods," *SIAM Journal on Control and Optimization*, vol. 46, no. 5, pp. 1683–1704, 2007.

[32] W. W. Hager and H. Zhang, "Self-adaptive inexact proximal point methods," *Computational Optimization and Applications*, vol. 39, no. 2, pp. 161–181, 2008.

[33] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[34] M. Teboulle, "Convergence of proximal-like algorithms," *SIAM Journal on Optimization*, vol. 7, no. 4, pp. 1069–1083, 1997.

[35] W. K. Glunt, "An alternating projections method for certain linear problems in a Hilbert space," *IMA Journal of Numerical Analysis*, vol. 15, no. 2, pp. 291–305, 1995.

[36] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, NY, USA, 1999.

[37] N. Narendra, "A new polynomial time algorithm for linear programming," *Combinatorica*, vol. 4, pp. 373–395, 1987.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 952915, 12 pages
http://dx.doi.org/10.1155/2013/952915

Research Article

A Quasi-Variational Approach for the Dynamic Oligopolistic Market Equilibrium Problem

Annamaria Barbagallo[1] and Paolo Mauro[2]

[1] *Department of Mathematics and Applications "R. Caccioppoli", University of Naples "Federico II", Via Cintia, 80126 Naples, Italy*
[2] *Department of Mathematics and Computer Science, University of Catania, Viale A. Doria 6, 95125 Catania, Italy*

Correspondence should be addressed to Annamaria Barbagallo; annamaria.barbagallo@unina.it

Received 18 July 2013; Accepted 19 September 2013

Academic Editor: Abdellah Bnouhachem

The paper is concerned with the dynamic oligopolistic market equilibrium problem in the realistic case in which we allow the presence of capacity constraints and production excesses and, moreover, we assume that the production function depends not only on the time but also on the equilibrium distribution. As a consequence, we introduce the generalized dynamic Cournot-Nash principle in the elastic case and prove the equivalence between this equilibrium definition and a suitable evolutionary quasi-variational inequality. For completeness we make the analysis of existence, regularity, and sensitivity of the solution. In the end, a numerical example is provided.

1. Introduction

The aim of the paper is to improve the results obtained in [1] concerning the dynamic oligopolistic market equilibrium problem in presence of production excesses by introducing the dependence on the equilibrium commodity shipment in the production function (see $\mathbb{K}^*(x^*)$ in (4)) and, as a consequence, studying the so-called elastic model. This is a more realistic situation since it is reasonable to think that the production function is influenced not only by the time, but also by the evaluation of the amount of commodity shipment, namely, the forecasted equilibrium solution. The presence of production excesses may be well justified in periods of economic crisis, so it is possible that some of the amounts of the commodity available are sold out whereas for a part of the products, an excess of production can occur.

In the last decade a lot of problems considering a feasible set depending on equilibrium solutions have been studied (see, e.g., [2–4]). It is well known that the equilibrium models with fixed constraint sets may be expressed in terms of evolutionary variational inequalities, while models with elastic constraint sets are expressed by evolutionary quasi-variational inequalities. Moreover, the dependence on time leads to considering variational and quasi-variational inequalities in an infinite dimensional setting, for example, a Lebesgue space.

Let us remember that a dynamic oligopolistic market equilibrium problem is the problem of finding a trade equilibrium in a supply-demand market between a finite number of spatially separated firms producing homogeneous goods in a fixed time interval. Moreover, the firms act in a noncooperative behavior. This problem has its origin with Cournot [5]. He considered only two firms and for this reason it was called the duopoly problem. Later, Nash [6, 7] extended Cournot's duopoly problem to n agents. A more complete and efficient study was done by Nagurney et al. in [8–11], but the problem was still faced in a static case through a finite dimensional variational approach. Finally, in [12] the time dependence was considered in the model. It allows to explore the change of behavior of equilibrium states for the oligopolistic market models over a finite time interval of interest. As Beckmann and Wallace stressed, for the first time, in [13], "the time-dependent formulation of equilibrium problems allows one to explore the dynamics of adjustment processes in which a delay on time response is operating." Of course a delay on time response always happens because the processes do not have an infinite speed. Usually, such adjustment processes can be represented by means of a memory term which depends

on previous equilibrium solutions according to the Volterra operator (see, e.g., [14, 15]).

Furthermore, in [16] the authors describe the behavior of the market by using the Lagrange multipliers of the infinite dimensional duality theory developed in [17–21]. Such results make use of the notion of tangent cone, normal cone, and quasi-relative interior of sets (see [22, 23]), important tools to overcome the difficulty of the emptiness of the topological interior of the ordering cone which defines constraints of several infinite dimensional problems (see [24, 25]). Moreover, a sensitivity result has been obtained which states that, under additional assumptions, small changes of the solution happen in correspondence with small changes of the profit function.

Lately, in [1, 26], the model presented in [12] has been improved with the addition of production excesses and both production and demand excesses, respectively. Another important question is to find some regularity properties for the solution. In [1, 26], the continuity of solution is proved under suitable assumptions, and it results to be very helpful in order to introduce numerical schemes to compute equilibrium solutions (see [27, 28]).

In [29, 30], the authors abandon the study of the problem from a producer's point of view whose purpose is to maximize his own profit and focus their attention on the policy-maker's perspective whose aim is to control the commodity exportations by means of the imposition of taxes or incentives and formulate the resulting optimization problem as an inverse variational inequality.

This paper is structured as follows. In Section 2 we present the dynamic oligopolistic market equilibrium problem with the elastic production function and after that we give the definition of equilibrium according to the generalized Cournot-Nash principle. Moreover, we prove the equivalence with a suitable evolutionary quasi-variational inequality. Section 3 is devoted to prove a result of existence of the solution, while in Section 4 Kuratowski's set convergence will be a preliminary property in order to prove the continuity of the equilibrium solution. In Section 5 we establish a sensitivity result that shows how the equilibrium solution can change if the data have been perturbed. In Section 6 a numerical example is provided to make the theoretical model presented in the previous sections more clearer.

2. Quasi-Variational Inequalities in Dynamic Oligopolistic Markets

Let us consider m firms P_i, $i = 1, \ldots, m$, that produce a homogeneous commodity and n demand markets Q_j, $j = 1, \ldots, n$, that are generally spatially separated. Assume that the homogeneous commodity, produced by the m firms and consumed by the n markets, is involved during a time interval $[0, T]$, $T > 0$.

Let $x_{ij}(t)$, $i = 1, \ldots, m$ and $j = 1, \ldots, n$, denote the nonnegative commodity shipment between the supply market P_i and the demand market Q_j at the time $t \in [0, T]$. In particular, let us set the vector $x_i(t) = (x_{i1}(t), \ldots, x_{in}(t))$, $i = 1, \ldots, m$ and $t \in [0, T]$, as the strategy vector for the firm P_i.

Let us group the commodity shipments into a matrix function $x : [0, T] \to \mathbb{R}_+^{mn}$ and suppose that $x \in L^2([0, T], \mathbb{R}_+^{mn})$. Furthermore, we assume that the nonnegative commodity shipment x_{ij} between the producer P_i and the demand market Q_j has to satisfy time-dependent constraints, namely, there exist two nonnegative functions $\underline{x}, \overline{x} : [0, T] \to \mathbb{R}_+^{mn}$ such that

$$0 \le \underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \quad \forall i = 1, \ldots, m, \quad \forall j = 1, \ldots, n, \text{ a.e. in } [0, T], \tag{1}$$

and suppose that $\underline{x}, \overline{x} \in L^2([0, T], \mathbb{R}_+^{mn})$.

Let us denote

$$D = \left\{ x \in L^2([0, T], \mathbb{R}^{mn}) : \underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \quad \forall i = 1, \ldots, m, \quad \forall j = 1, \ldots, n, \text{ a.e. in } [0, T] \right\}. \tag{2}$$

It is easy to verify that D is a nonempty, compact, and convex subset of $L^2([0, T], \mathbb{R}^{mn})$. Let $p_i(t, x(t))$, $i = 1, \ldots, m$, denote the nonnegative commodity output produced by firm P_i at the time $t \in [0, T]$. Let us group the production output into a vector function $p : [0, T] \times D \to \mathbb{R}_+^m$ and let us suppose that $p \in L^1([0, T] \times D, \mathbb{R}_+^m)$.

Now, let us introduce the production excesses. Let $\epsilon_i(t)$, $i = 1, \ldots, m$, be the nonnegative production excess for the commodity of the firm P_i at the time $t \in [0, T]$. Let us group the production excess into a vector function $\epsilon : [0, T] \to \mathbb{R}_+^m$ and let us assume that $\epsilon \in L^2([0, T], \mathbb{R}^m)$.

We consider a formulation of equilibrium problems where the dependence of the production on the unknown solution x^* is in the average sense with respect to the time; namely, the following feasibility condition holds:

$$\sum_{j=1}^{n} x_{ij}(t) + \epsilon_i(t) = \frac{1}{T} \int_0^T p_i(t, x^*(\tau))\, d\tau, \quad i = 1, \ldots, m, \text{ a.e. in } [0, T]. \tag{3}$$

Hence, condition (3) states that the average of the quantity produced by each firm P_i, in the time interval $[0, T]$, must be equal to the commodity shipments from that firm to all the demand markets plus the production excess, at the time $t \in [0, T]$. In fact, the production is supposed to depend on the firms' evaluation of the commodity shipments. So one can expect the producers not to evaluate the market practicability instantly, but by an average with respect to the whole time interval.

The set of feasible vectors $(x,\epsilon) \in L^2([0,T],\mathbb{R}^{mn+m})$ is then given by the set-valued map $\mathbb{K} : D \to 2^{L^2([0,T],\mathbb{R}_+^{mn+m})}$ defined as:

$$\begin{aligned}
&\mathbb{K}^*(x^*) \\
&\quad = \Big\{ (x,\epsilon) \in L^2([0,T],\mathbb{R}^{mn+m}) : \\
&\qquad \underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \ \forall i = 1,\dots,m, \\
&\qquad \forall j = 1,\dots,n, \text{ a.e. in } [0,T], \\
&\qquad \sum_{j=1}^{n} x_{ij}(t) + \epsilon_i(t) = \frac{1}{T}\int_0^T p_i(t,x^*(\tau))\,d\tau, \\
&\qquad \forall i = 1,\dots,m, \text{ a.e. in } [0,T], \ \epsilon_i(t) \ge 0, \\
&\qquad \forall i = 1,\dots,m, \text{ a.e. in } [0,T] \Big\}.
\end{aligned} \tag{4}$$

Moreover, let us associate with each firm P_i a production cost f_i^*, $i = 1,\dots,m$, and assume that the production cost of a firm P_i may depend upon the entire production pattern; namely,

$$f_i^* = f_i^*(t,x(t),\varepsilon(t)). \tag{5}$$

Similarly, let us associate with each demand market Q_j a demand price for unity of the commodity d_j, $j = 1,\dots,n$, and assume that the demand price of a demand market Q_j may depend, in general, upon the entire consumption pattern; namely,

$$d_j = d_j(t,x(t)). \tag{6}$$

Let g_i^*, $i = 1,\dots,m$, denote the storage cost of the commodity produced by the firm P_i and assume that this cost may depend upon the entire production pattern; namely,

$$g_i^* = g_i^*(t,x(t),\varepsilon(t)). \tag{7}$$

Finally, let c_{ij}, $i = 1,\dots,m$ and $j = 1,\dots,n$, denote the transaction cost, which includes the transportation cost associated with trading the commodity between firm P_i and demand market Q_j. Here we permit the transaction cost to depend upon the entire shipment pattern; namely,

$$c_{ij} = c_{ij}(t,x(t)). \tag{8}$$

Hence, we have the following mappings:

$$\begin{aligned}
&f^* : [0,T] \times L^2([0,T],\mathbb{R}_+^{mn}) \\
&\qquad \times L^2([0,T],\mathbb{R}_+^{m}) \longrightarrow L^2([0,T],\mathbb{R}_+^{m}), \\
&d : [0,T] \times L^2([0,T],\mathbb{R}_+^{mn}) \longrightarrow L^2([0,T],\mathbb{R}_+^{n}), \\
&g^* : [0,T] \times L^2([0,T],\mathbb{R}_+^{mn}) \\
&\qquad \times L^2([0,T],\mathbb{R}_+^{m}) \longrightarrow L^2([0,T],\mathbb{R}_+^{m}), \\
&c : [0,T] \times L^2([0,T],\mathbb{R}_+^{mn}) \longrightarrow L^2([0,T],\mathbb{R}_+^{mn}).
\end{aligned} \tag{9}$$

The profit $v_i^*(t,x(t),\varepsilon(t))$, $i = 1,\dots,m$, of the firm P_i at the time $t \in [0,T]$ is, then,

$$\begin{aligned}
&v_i^*(t,x(t),\varepsilon(t)) \\
&\quad = \sum_{j=1}^{n} d_j(t,x(t))\,x_{ij}(t) \\
&\qquad - f_i^*(t,x(t),\varepsilon(t)) - g_i^*(t,x(t),\varepsilon(t)) \\
&\qquad - \sum_{j=1}^{n} c_{ij}(t,x(t))\,x_{ij}(t);
\end{aligned} \tag{10}$$

namely, it is equal to the price that the demand markets are disposed to pay minus the production costs, the storage costs, and the transportation costs.

By virtue of (3), we can express the nonnegative production excess $\epsilon_i(t)$ at the time $t \in [0,T]$ in terms of the integral average of the production function and the commodity shipment. As a consequence, we get

$$\begin{aligned}
&\sum_{j=1}^{n} x_{ij}(t) \le \frac{1}{T}\int_0^T p_i(t,x^*(\tau))\,d\tau, \\
&\qquad \forall i = 1,\dots,m, \text{ a.e. in } [0,T].
\end{aligned} \tag{11}$$

Then, the production costs and the storage costs, by virtue of (5) and (7), respectively, become

$$\begin{aligned}
f_i(t,x(t)) &= f_i^*(t,x(t),\epsilon(t)), \\
g_i(t,x(t)) &= g_i^*(t,x(t),\epsilon(t)),
\end{aligned} \tag{12}$$

and, analogously, the profit (10) becomes

$$\begin{aligned}
v_i(t,x(t)) &= v_i^*(t,x(t),\epsilon(t)) \\
&= \sum_{j=1}^{n} d_j(t,x(t))\,x_{ij}(t) - f_i(t,x(t)) \\
&\quad - g_i(t,x(t)) - \sum_{j=1}^{n} c_{ij}(t,x(t))\,x_{ij}(t).
\end{aligned} \tag{13}$$

As a consequence, the set of feasible vectors $x \in L^2([0,T],\mathbb{R}^{mn})$ becomes the set-valued map $\mathbb{K} : D \to 2^{L^2([0,T],\mathbb{R}_+^{mn})}$, defined as

$$\begin{aligned}
&\mathbb{K}(x^*) \\
&\quad = \Big\{ x \in L^2([0,T],\mathbb{R}_+^{mn}) : \\
&\qquad \underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \ \forall i = 1,\dots,m, \\
&\qquad \forall j = 1,\dots,n, \text{ a.e. in } [0,T], \\
&\qquad \sum_{j=1}^{n} x_{ij}(t) \le \frac{1}{T}\int_0^T p_i(t,x^*(\tau))\,d\tau, \\
&\qquad \forall i = 1,\dots,m, \text{ a.e. in } [0,T] \Big\}.
\end{aligned} \tag{14}$$

Let us denote $x_i = \{x_{ij}\}_{j=1,\dots,n}$, $i = 1,\dots,m$, and $\nabla_D v = (\partial v_i/\partial x_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$. Let us assume the following assumptions:

(i) $v_i(t, x(t))$ is continuously differentiable for each $i = 1,\dots,m$, a.e. in $[0,T]$,

(ii) $\nabla_D v = (\partial v_i/\partial x_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ is a Carathéodory function, such that

$$\exists \gamma \in L^2([0,T]) : \|\nabla_D v(t,x)\|_{mn} \le \gamma(t) + \|x\|_{mn}, \quad \forall x \in \mathbb{R}^{mn}, \text{ a.e. in } [0,T], \tag{15}$$

(iii) $v_i(t, x(t))$ is pseudoconcave with respect to the variables x_i, $i = 1,\dots,m$, a.e. in $[0,T]$.

For the reader's convenience, we recall that a function v, continuously differentiable, is called *pseudoconcave* with respect to x_i, $i = 1,\dots,m$, a.e. in $[0,T]$ (see [31]), if the following holds a.e. in $[0,T]$:

$$\begin{aligned} &\langle \nabla_D v_i(t, x_1,\dots,x_i,\dots,x_m), x_i - y_i\rangle \\ &\quad = \sum_{j=1}^{n} \frac{\partial v_i(t,x)}{\partial x_{ij}}(x_{ij} - y_{ij}) \ge 0 \\ &\quad \Longrightarrow v_i(t, x_1,\dots,x_i,\dots,x_m) \\ &\quad \ge v_i(t, x_1,\dots,y_i,\dots,x_m). \end{aligned} \tag{16}$$

Moreover, we recall that in the Hilbert space $L^2([0,T],\mathbb{R}^k)$, we define the canonical bilinear form on $L^2([0,T],\mathbb{R}^k)^* \times L^2([0,T],\mathbb{R}^k)$ by

$$\langle\langle \phi, w\rangle\rangle := \int_0^T \langle \phi(t), w(t)\rangle\, dt, \tag{17}$$

where $\phi \in (L^2([0,T],\mathbb{R}^k))^* = L^2([0,T],\mathbb{R}^k)$, $w \in L^2([0,T],\mathbb{R}^k)$, and

$$\langle \phi(t), w(t)\rangle = \sum_{l=1}^{k} \phi_l(t)\, w_l(t). \tag{18}$$

Now, let us consider the dynamic oligopolistic market, in which the m firms supply the commodity in a noncooperative fashion, each one trying to maximize its own profit at the time $t \in [0,T]$. We seek to determine a nonnegative commodity distribution matrix function x^* for which the m firms and the n demand markets will be in a state of equilibrium according to the dynamic Cournot-Nash principle.

Definition 1. $x^* \in \mathbb{K}(x^*)$ is a dynamic oligopolistic market equilibrium in presence of production excesses if and only if for each $i = 1,\dots,m$ and a.e. in $[0,T]$

$$v_i(t, x^*(t)) \ge v_i(t, x_i(t), \hat{x}_i^*(t)), \quad \text{a.e. in } [0,T], \tag{19}$$

where $x_i(t) = (x_{i1}(t),\dots,x_{in}(t))$ and $\hat{x}_i^*(t) = (x_1^*(t),\dots,x_{i-1}^*(t), x_{i+1}^*(t),\dots, x_m^*(t))$, for $i = 1,\dots,m$, a.e. in $[0,T]$.

Definition 1 states that each firm P_i maximizes its own profit, at the time $t \in [0,T]$, considering the given optimal strategy $\hat{x}_i^*(t)$ of the other firms.

Theorem 2. *Suppose that assumptions (i), (ii), and (iii) are satisfied. Then, $x^* \in \mathbb{K}(x^*)$ is a dynamic oligopolistic market equilibrium according to Definition 1 if and only if it satisfies the evolutionary quasi-variational inequality*

$$\int_0^T \sum_{i=1}^{m}\sum_{j=1}^{n} \left(-\frac{\partial v_i(t, x^*(t))}{\partial x_{ij}}\right)(x_{ij}(t) - x_{ij}^*(t))\, dt \ge 0 \quad \forall x \in \mathbb{K}(x^*). \tag{20}$$

Proof. First of all, let us prove that the evolutionary quasi-variational inequality (20), that we can write as follows:

$$\begin{aligned} &\int_0^T \sum_{i=1}^{m}\sum_{j=1}^{n} -\frac{\partial v_i(t, x^*(t))}{\partial x_{ij}}(x_{ij}(t) - x_{ij}^*(t))\, dt \\ &\quad = \langle\langle -\nabla_D v(x^*), x - x^*\rangle\rangle \\ &\quad = \int_0^T \langle -\nabla_D v(t, x^*(t)), x(t) - x^*(t)\rangle\, dt \ge 0 \\ &\qquad \forall x \in \mathbb{K}(x^*), \end{aligned} \tag{21}$$

is equivalent to the following point-to-point quasi-variational inequality:

$$\begin{aligned} &\langle -\nabla_D v(t, x^*(t)), x(t) - x^*(t)\rangle \\ &\quad = \sum_{i=1}^{m}\sum_{j=1}^{n} -\frac{\partial v_i(t, x^*(t))}{\partial x_{ij}}(x_{ij}(t) - x_{ij}^*(t)) \ge 0 \\ &\qquad \forall x(t) \in \mathbb{K}(t, x^*), \text{ a.e. in } [0,T], \end{aligned} \tag{22}$$

where

$$\begin{aligned} &\mathbb{K}(t, x^*) \\ &\quad = \Bigg\{ x(t) \in \mathbb{R}_+^{mn} : \underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \\ &\qquad \forall i = 1,\dots,m,\ \forall j = 1,\dots,n, \\ &\qquad \sum_{j=1}^{n} x_{ij}(t) \le \frac{1}{T}\int_0^T p_i(t, x^*(\tau))\, d\tau, \\ &\qquad \forall i = 1,\dots,m \Bigg\}. \end{aligned} \tag{23}$$

In fact, let us suppose by absurdum that (22) does not hold, namely, $\exists \overline{x}(t) \in \mathbb{K}(x^*)$, $\exists I \subseteq [0,T]$ with $m(I) > 0$ such that

$$\langle -\nabla_D v(t, x^*(t)), \overline{x}(t) - x^*(t)\rangle < 0 \quad \text{a.e. in } I. \tag{24}$$

Let us choose, now,

$$x(t) = \begin{cases} x^*(t), & \text{in } [0,T] \setminus I, \\ \overline{x}(t), & \text{in } I. \end{cases} \tag{25}$$

Hence, let us consider

$$\begin{aligned}&\langle\langle -\nabla_D v(t,x^*(t)), x-x^*\rangle\rangle\\&\quad=\int_{[0,T]\setminus I}\langle -\nabla_D v(t,x^*(t)), x(t)-x^*(t)\rangle\,dt\\&\qquad+\int_I\langle -\nabla_D v(t,x^*(t)), \overline{x}(t)-x^*(t)\rangle\,dt<0,\end{aligned}\tag{26}$$

that is a contradiction. The vice versa is immediate.

So the equivalence between the evolutionary quasi-variational inequalities (20) and (22) is proved.

Let us prove, now, the equivalence between the dynamic Cournot-Nash principle and the evolutionary quasi-variational inequality (20).

Let us suppose that $x^*\in\mathbb{K}(x^*)$ is an equilibrium point according to Definition 1; namely,

$$\begin{aligned}&v_i(t,x^*(t))\geq v_i(t,x(t),\widehat{x}^*(t))\quad\forall x(t)\in\mathbb{K}(t,x^*),\\&\qquad\text{a.e. in }[0,T],\ \forall i=1,\dots,m.\end{aligned}\tag{27}$$

For well known theorems of optimization, we have that the necessary and sufficient condition to get (27) is that for all $i=1,\dots,m$, for all $x(t)\in\mathbb{K}(t,x^*)$, a.e. in $[0,T]$

$$\begin{aligned}&\langle -\nabla_D v_i(t,x^*(t)), x_i(t)-x_i^*(t)\rangle\\&\quad=\sum_{j=1}^n-\frac{\partial v_i(t,x^*(t))}{\partial x_{ij}}\left(x_{ij}(t)-x_{ij}^*(t)\right)\geq 0.\end{aligned}\tag{28}$$

By assumption that $\nabla_D v_i$ is a Carathéodory function such that

$$\begin{aligned}&\exists\gamma\in L^2([0,T]):\|\nabla_D v_i(t,x)\|_{mn}\\&\qquad\leq\gamma(t)+\|x\|_{mn},\quad\forall x\in\mathbb{R}^{mn},\text{ a.e. in }[0,T],\end{aligned}\tag{29}$$

moreover, x and $x^*\in L^2([0,T],\mathbb{R}^{mn})$, so we have

$$t\longrightarrow\langle -\nabla_D v_i(t,x^*(t)), x_i(t)-x_i^*(t)\rangle\in L^2([0,T],\mathbb{R}).\tag{30}$$

Then, we get

$$\langle\langle -\nabla_D v_i(t,x^*(t)), x_i-x_i^*\rangle\rangle\geq 0\quad\forall x\in\mathbb{K}(x^*),\tag{31}$$

from which, by summing up each firm P_i, for $i=1,\dots,m$, we obtain

$$\begin{aligned}&\sum_{i=1}^m\langle\langle -\nabla_D v_i(t,x^*(t)), x-x^*\rangle\rangle\\&\quad=\langle\langle -\nabla_D v(t,x^*(t)), x_i-x_i^*\rangle\rangle\geq 0\quad\forall x\in\mathbb{K}(x^*).\end{aligned}\tag{32}$$

Vice versa, let us suppose that $x^*(t)$ is a solution to evolutionary quasi-variational inequality (20), but not an equilibrium solution according to the dynamic Cournot-Nash principle, namely, $\exists I\subseteq[0,T]$ with $m(I)>0$, $\exists\bar{i}\in\{1,\dots,m\}$ and $\exists\widetilde{x}_{\bar{i}}$ such that

$$v_{\bar{i}}(t,x^*(t))<v_{\bar{i}}(t,\widetilde{x}_i(t),\widehat{x}^*(t))\quad\text{in } I.\tag{33}$$

Since the profit function $v_{\bar{i}}(t,x(t))$ is pseudoconcave with respect to $x_{\bar{i}}$, we get

$$\left\langle -\nabla_D v_{\bar{i}}(t,x^*(t)), x_{\bar{i}}^*(t)-\widetilde{x}_{\bar{i}}(t)\right\rangle<0\quad\text{in } I.\tag{34}$$

If we choose $x\in\mathbb{K}(x^*)$ such that

$$x_i(t)=\begin{cases}x_i^*(t)\ \text{in}\ [0,T]\setminus I, & \forall i=1,\dots,m,\\ x_i^*(t)\ \text{in}\ I, & \text{if } i\neq\bar{i},\\ \widetilde{x}_i\ \text{in}\ I, & \text{if } i=\bar{i},\end{cases}\tag{35}$$

then

$$\begin{aligned}&\int_0^T\langle -\nabla_D v(t,x^*(t)), x(t)-x^*(t)\rangle\,dt\\&\quad=\int_I\left\langle -\nabla_D v_{\bar{i}}(t,x^*(t)), \widetilde{x}_{\bar{i}}(t)-x_{\bar{i}}^*(t)\right\rangle dt<0,\end{aligned}\tag{36}$$

so we get the contradiction. □

3. An Existence Theorem for Equilibrium Solutions

Now, we prove an existence result for the equilibrium solution to the dynamic elastic oligopolistic market equilibrium problem. To this aim, we recall a general existence result for solutions to quasi-variational inequalities in topological linear locally convex Hausdorff spaces due to Tan [32].

Theorem 3. *Let X be a topological linear locally convex Hausdorff space and let D be a convex compact nonempty subset of X. Let $C:D\to 2^{X^*}$ be an upper semicontinuous multivalued mapping with $C(x)$, $x\in D$, convex compact nonempty, let $\mathbb{K}:D\to 2^D$ be a closed lower semicontinuous multivalued mapping with $\mathbb{K}(x)$, $x\in D$, convex compact nonempty, and let $\varphi:D\to\mathbb{R}$ be a proper convex lower semicontinuous function. Then, there exists $x^*\in D$ such that:*

(i) $x^*\in\mathbb{K}(x^*)$,

(ii) *there exists $y^*\in\mathbb{K}(x^*)$ for which*

$$\langle x-x^*,y^*\rangle+\varphi(x)-\varphi(x^*)\geq 0,\quad\forall x\in\mathbb{K}(x^*).\tag{37}$$

Now, we are able to prove our main result.

Theorem 4. *Let $v:[0,T]\to\mathbb{R}^m$ and $p:[0,T]\to\mathbb{R}^m$ be two vector functions such that assumptions (i) and (iii) are satisfied and*

(I) *$\nabla_D v(t,x)$ is measurable in t, for all $x\in\mathbb{R}_+^{mn}$, continuous in x, a.e. in $[0,T]$, such that $\exists\gamma\in L^2([0,T])$: $\|\nabla_D v(t,x)\|\leq\gamma(t)+\|x\|$, for all $x\in\mathbb{R}_+^{mn}$, a.e. in $[0,T]$;*

(II) $p(t,x)$ *is measurable in* t, *for all* $x \in \mathbb{R}_+^{mn}$, *continuous in* x, *a.e. in* $[0,T]$, *such that* $\exists \phi \in L^1([0,T]) : \|p(t,x)\| \le \phi(t) + \|x\|^2$, *for all* $x \in \mathbb{R}_+^{mn}$, *a.e. in* $[0,T]$;

(III) $\exists \nu(t) \ge 0$, *a.e. in* $[0,T]$, $\eta \in L^\infty([0,T])$ *such that*

$$\|p(t,x_1) - p(t,x_2)\| \le \eta(t)\|x_1 - x_2\|, \quad \forall x_1, x_2 \in \mathbb{R}_+^{mn}, \text{ a.e. in } [0,T]. \tag{38}$$

Then, evolutionary quasi-variational inequality (20) *admits a solution.*

Proof. At first, observe that under the hypotheses (I) and (II) and if $x \in L^2([0,T],\mathbb{R}_+^{mn})$,

$$t \longmapsto \nabla_D \nu(t,x(t)) \in L^2([0,T],\mathbb{R}_+^{mn}), \quad t \longmapsto p(t,x(t)) \in L^1([0,T],\mathbb{R}_+^{m}). \tag{39}$$

Moreover, by (I) and (II) it follows that $\nabla_D \nu$ and p belong to the class of nemytskii operators. Therefore if $\{x^k\}_{k\in\mathbb{N}}$ is a sequence such that $x^k \to x$, in $L^2([0,T],\mathbb{R}^{mn})$, we have

$$\left\|\nabla_D \nu(t,x^k) - \nabla_D \nu(t,x)\right\|_{L^2} \longrightarrow 0, \quad \left\|p(t,x^k) - p(t,x)\right\|_{L^1} \longrightarrow 0, \tag{40}$$

where the functions $\nabla_D \nu$ and p are L^2- and L^1-continuous, respectively.

Now, in order to show that $\mathbb{K}(x^*)$ is a closed multifunction, we prove that the following condition holds. For every two arbitrary sequences $\{x^k\}_{k\in\mathbb{N}}$ and $\{y^k\}_{k\in\mathbb{N}}$ such that $x^k \to x$ and $y^k \to y$ in $L^2([0,T],\mathbb{R}^{mn})$, with $y^k \in \mathbb{K}(x^k)$, $\forall n \in \mathbb{N}$, then $y \in \mathbb{K}(x)$. To this aim, let us consider two arbitrary convergent sequences in $L^2([0,T],\mathbb{R}^{mn})$, $\{x^k\}_{k\in\mathbb{N}}$ and $\{y^k\}_{k\in\mathbb{N}}$ to x and y, respectively. Since $y^k \in \mathbb{K}(x^k)$, $\underline{x}_{ij}(t) \le y_{ij}^k(t) \le \overline{x}_{ij}(t)$, for $i = 1,\dots,m$, $j = 1,\dots,n$, and a.e. in $[0,T]$, and the convergence of the sequence $\{y^k\}_{k\in\mathbb{N}}$ in $L^2([0,T],\mathbb{R}^{mn})$ implies that also y satisfies the capacity constraints.

Moreover, the following relationship holds:

$$\sum_{j=1}^{m} y_{ij}^k(t) \le \frac{1}{T}\int_0^T p_i\left(t,x^k(\tau)\right)d\tau, \quad i = 1,\dots,m, \text{ a.e. in } [0,T]. \tag{41}$$

The left-hand side converges almost everywhere to $\sum_{j=1}^m y_{ij}(t)$; for the right-hand side, meanwhile, $i = 1,\dots,m$, we have

$$\begin{aligned} &\sup_{[0,T]}\left\|\int_0^T p\left(t,x^k(\tau)\right)d\tau - \int_0^T p(t,x(\tau))\,d\tau\right\| \\ &\le \sup_{[0,T]}\int_0^T \left\|p\left(t,x^k(\tau)\right) - p(t,x(\tau))\right\| d\tau \\ &\le \|\eta\|_{L^\infty([0,T])}\int_0^T \left\|x^k(\tau) - x(\tau)\right\| d\tau. \end{aligned} \tag{42}$$

By considering that the convergence of $\{x^k\}$ in L^2 implies the convergence even in L^1, hence the sequence $\{(1/T)\int_0^T p_i(t, x^k(\tau))d\tau\}_{k\in\mathbb{N}}$ converges uniformly to $(1/T)\int_0^T p_i(t,x(\tau))d\tau$ in $L^1([0,T],\mathbb{R}^m)$.

Now, let us show the lower semicontinuity of the multifunction $\mathbb{K}$. To this aim it suffices to prove that for every $\{x^k\}_{k\in\mathbb{N}}$ such that $x^k \to x$, in $L^2([0,T],\mathbb{R}^{mn})$, and for every $y \in \mathbb{K}(x)$, there exists a sequence $\{y^k\}_{k\in\mathbb{N}}$ such that $y^k \to y$, in $L^2([0,T],\mathbb{R}^{mn})$, with $y^k \in \mathbb{K}(x^k)$, for all $k \in \mathbb{N}$.

Let $\{x^k\}_{k\in\mathbb{N}}$ be an arbitrary sequence such that $x^k \to x$, in $L^2([0,T],\mathbb{R}^{mn})$, and let $y \in \mathbb{K}(x)$. Let us note that, for $i = 1,\dots,m$ and $j = 1,\dots,n$, and if

$$a_{ij}^k(t) = y_{ij}(t) - \underline{x}_{ij}(t) + \frac{1}{nT} \times \left[\int_0^T p_i(t,x(\tau))\,d\tau - \int_0^T p_i\left(t,x^k(\tau)\right)d\tau\right], \tag{43}$$

we obtain, by virtue of the uniform convergence of $\{(1/T)\int_0^T p_i(t,x^k(\tau))d\tau\}_{k\in\mathbb{N}}$ to $(1/T)\int_0^T p_i(t,x(\tau))d\tau$ in $L^1([0,T],\mathbb{R}^m)$, that

$$\lim_{k\to+\infty} a_{ij}^k(t) = y_{ij}(t) - \underline{x}_{ij}(t) \ge 0, \quad \text{a.e. in } [0,T]. \tag{44}$$

As a consequence, there exists an index ν such that for $k > \nu$ one has, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$,

$$a_{ij}^k(t) \ge 0, \quad \text{a.e. in } [0,T]. \tag{45}$$

Then, we consider the sequence $\{y^k\}_{k\in\mathbb{N}}$ such that

(i) for $k > \nu$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$,

$$y_{ij}^k(t) = \underline{x}_{ij}(t) + \min\left\{\overline{x}_{ij}(t) - \underline{x}_{ij}(t), a_{ij}^k(t)\right\}, \quad \text{a.e. in } [0,T], \tag{46}$$

(ii) for $k \le \nu$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$,

$$y_{ij}^k(t) = P_{\mathbb{K}(x^k)} y_{ij}(t), \quad \text{a.e. in } [0,T], \tag{47}$$

where $P_{\mathbb{K}(x^k)}(\cdot)$ denotes the Hilbertian projection on $\mathbb{K}(x^k)$.

It is easy to verify that if $k \le \nu$, for (47), $y^k \in \mathbb{K}(x^k)$. Instead, for $k > \nu$, since for (45), $\min\{\overline{x}_{ij}(t) - \underline{x}_{ij}(t), a_{ij}^k(t)\} \ge 0$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$, a.e. in $[0,T]$,

$$y_{ij}^k(t) \ge \underline{x}_{ij}(t), \quad \forall i = 1,\dots,m, \ \forall j = 1,\dots,n, \text{ a.e. in } [0,T]. \tag{48}$$

Moreover, since $\min\{\overline{x}_{ij}(t) - \underline{x}_{ij}(t), a_{ij}^k(t)\} \le \overline{x}_{ij}(t) - \underline{x}_{ij}(t)$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$, a.e. in $[0,T]$, we have

$$y_{ij}^k(t) \le \overline{x}_{ij}(t), \quad \forall i = 1,\dots,m, \ \forall j = 1,\dots,n, \text{ a.e. in } [0,T]. \tag{49}$$

Finally, we get

$$\begin{aligned}\sum_{j=1}^{n} y_{ij}^{k}(t) &\le \sum_{j=1}^{n} y_{ij}(t) + \frac{1}{T}\int_0^T p_i\left(t, x^k(\tau)\right) d\tau \\ &\quad - \frac{1}{T}\int_0^T p_i(t, x(\tau))\, d\tau \\ &\le \frac{1}{T}\int_0^T p_i(t, x(\tau))\, d\tau + \frac{1}{T}\int_0^T p_i\left(t, x^k(\tau)\right) d\tau \\ &\quad - \frac{1}{T}\int_0^T p_i(t, x(\tau))\, d\tau \\ &= \frac{1}{T}\int_0^T p_i\left(t, x^k(\tau)\right) d\tau, \quad \forall i = 1, \ldots, m, \text{ a.e. in } [0, T]. \end{aligned} \tag{50}$$

Hence, we can conclude that y^k belongs to $\mathbb{K}(x^k)$, for all $k \in \mathbb{N}$.

Let us prove now the convergence of $\{y^k\}_{k\in\mathbb{N}}$ to y in $L^2([0,T], \mathbb{R}^{mn})$. Let us observe that

$$\begin{aligned}&\lim_{k\to+\infty} \min\left\{\overline{x}_{ij}(t) - \underline{x}_{ij}(t), a_{ij}^{k}(t)\right\} \\ &\qquad = y_{ij}(t) - \underline{x}_{ij}(t), \quad \text{a.e. in } [0, T]. \end{aligned} \tag{51}$$

As a consequence, we have that the sequence $\{y^k\}_{k\in\mathbb{N}}$ converges to y. It is easy to show that $\mathbb{K}(x)$ is a closed, bounded, and convex subset of D and since the space D is compact, $\mathbb{K}(x)$, for all $x \in D$, is compact too. As a consequence, all the hypotheses of Theorem 3 are satisfied and the existence of at least one solution is guaranteed. □

4. Regularity Results for Equilibrium Solutions

In this section, we study the assumptions under which the continuity of solutions to evolutionary quasi-variational inequality, which expresses the equilibrium condition for the dynamic elastic oligopolistic market equilibrium problem in presence of production excesses, is ensured.

4.1. Set Convergence. First of all, we recall the notion of Kuratowski's set convergence that has an important role in order to establish regularity results. The classical notion of convergence for subsets of a given metric space (X, d) is introduced in the 1950s by Kuratowski (see [33]; see also [34, 35]).

Let $\{\mathbb{K}_n\}_{n\in\mathbb{N}}$ be a sequence of subsets of X. Recall that

$$\begin{aligned} d - \underline{\lim}_n\, \mathbb{K}_n &= \Big\{x \in X : \exists \{x_n\}_{n\in\mathbb{N}} \\ &\qquad \text{eventually in } \mathbb{K}_n \text{ such that } x_n \xrightarrow{d} x\Big\}, \\ d - \overline{\lim}_n\, \mathbb{K}_n &= \Big\{x \in X : \exists \{x_n\}_{n\in\mathbb{N}} \\ &\qquad \text{frequently in } \mathbb{K}_n \text{ such that } x_n \xrightarrow{d} x\Big\}, \end{aligned} \tag{52}$$

where *eventually* means that there exists $\delta \in \mathbb{N}$ such that $x_n \in \mathbb{K}_n$ for any $n \ge \delta$ and *frequently* means that there exists an infinite subset $N \subseteq \mathbb{N}$ such that $x_n \in \mathbb{K}_n$ for any $n \in N$ (in this last case, according to the notation given above, we also write that there exists a subsequence $\{x_{k_n}\}_{n\in\mathbb{N}} \subseteq \{x_n\}_{n\in\mathbb{N}}$ such that $x_{k_n} \in \mathbb{K}_{k_n}$).

In the following, we recall Kuratowski's set convergence.

Definition 5. We say that $\{\mathbb{K}_n\}$ converges to some subset $\mathbb{K} \subseteq X$ in Kuratowski's sense and we briefly write $\mathbb{K}_n \to \mathbb{K}$, if $d - \underline{\lim}_n \mathbb{K}_n = d - \overline{\lim}_n \mathbb{K}_n = \mathbb{K}$. Thus, in order to verify that $\mathbb{K}_n \to \mathbb{K}$, it suffices to check that

(i) $d - \overline{\lim}_n \mathbb{K}_n \subseteq \mathbb{K}$, that is, for any sequence $\{x_n\}_{n\in\mathbb{N}}$ frequently in $\mathbb{K}_n$ such that $x_n \xrightarrow{d} x$ for some $x \in S$, then $x \in \mathbb{K}$;

(ii) $\mathbb{K} \subset d - \underline{\lim}_n \mathbb{K}_n$, that is, for any $x \in \mathbb{K}$ there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ eventually in $\mathbb{K}_n$ such that $x_n \xrightarrow{d} x$.

The below lemma establishes that the feasible set $\mathbb{K}$ of the dynamic elastic oligopolistic market equilibrium problem in the presence of production excesses satisfies the property of Kuratowski's set convergence.

Lemma 6. *Let $\underline{x}, \overline{x} \in C^0([0,T], \mathbb{R}_+^{mn})$, let $p \in C^0([0,T] \times \mathbb{R}_+^{mn}, \mathbb{R}_+^{m})$ be such that*

$$\begin{aligned} \exists \phi \in C^0([0,T], \mathbb{R}_+) : \|p(t,x)\| \le \phi(t) + \|x\|^2, \\ \forall x \in \mathbb{R}_+^{mn}, \text{ in } [0,T], \end{aligned} \tag{53}$$

and let $\{t_k\}_{k\in\mathbb{N}}$ be a sequence such that $t_k \to t$, with $t \in [0,T]$, as $k \to +\infty$. Then, the sequence of sets

$$\begin{aligned} \mathbb{K}(t_k, x^*) = \Big\{&x(t_k) \in \mathbb{R}^{mn} : \underline{x}_{ij}(t_k) \le x_{ij}(t_k) \le \overline{x}_{ij}(t_k), \\ &\forall i = 1, \ldots, m, \ \forall j = 1, \ldots, n, \\ &\sum_{j=1}^{n} x_{ij}(t_k) \le \frac{1}{T}\int_0^T p_i(t_k, x^*(\tau))\, d\tau, \\ &\forall i = 1, \ldots, m\Big\}, \end{aligned} \tag{54}$$

for all $k \in \mathbb{N}$, converges to

$$\begin{aligned} \mathbb{K}(t, x^*) = \Big\{&x(t) \in \mathbb{R}^{mn} : \underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \\ &\forall i = 1, \ldots, m, \ \forall j = 1, \ldots, n, \\ &\sum_{j=1}^{n} x_{ij}(t) \le \frac{1}{T}\int_0^T p_i(t, x^*(\tau))\, d\tau, \\ &\forall i = 1, \ldots, m\Big\}, \end{aligned} \tag{55}$$

as $k \to +\infty$, in Kuratowski's sense.

Proof. Firstly, we prove condition (K1). Let $\{t_k\}_{k\in\mathbb{N}}$ be a sequence such that $t_k \to t$, with $t \in [0,T]$ as $k \to +\infty$. Making use of the continuity assumptions on $\underline{x}$, $\overline{x}$, and p, we get $\underline{x}(t_k) \to \underline{x}(t)$, $\overline{x}(t_k) \to \overline{x}(t)$, and $p(t_k, y) \to p(t, y)$ as $k \to +\infty$, respectively. Furthermore,

$$\exists \phi \in C^0([0,T], \mathbb{R}_+) : \|p(t, x^*(\tau))\| \le \phi(t) + \|x^*(\tau)\|^2, \tag{56}$$

for $t \in [0,T]$ and$\tau \in [0,T]$. Since $\phi \in C^0([0,T])$ and $x^* \in L^2([0,T], \mathbb{R}_+^{mn})$, then we have for $t \in [0,T]$ and $\tau \in [0,T]$

$$\|p(t, x^*(\tau))\| \le \phi(t) + \|x^*(\tau)\|^2 \in L^1([0,T]), \tag{57}$$

and by virtue of the continuity of p with respect to the first variable we also obtain

$$\lim_{n\to+\infty} p(t_n, x^*(\tau)) = p(t, x^*(\tau)), \tag{58}$$

for $\tau \in [0,T]$ and $x^* \in L^2([0,T], \mathbb{R}_+^{mn})$. Taking into account a well known generalization of Lebesgue's theorem,

$$\lim_{n\to+\infty} \int_0^T p(t_n, x^*(\tau))\, d\tau = \int_0^T p(t, x^*(\tau))\, d\tau, \tag{59}$$

for every $x^* \in L^2([0,T], \mathbb{R}_+^{mn})$.

Let $x(t) \in \mathbb{K}(t, x^*)$ be fixed and let us note that, for $i = 1,\dots,m$ and $j = 1,\dots,n$, and if

$$a_{ij}(t_k) = x_{ij}(t) - \underline{x}_{ij}(t_k) + \frac{1}{nT} \times \left[\int_0^T p_i(t_k, x^*(\tau))\, d\tau - \int_0^T p_i(t, x^*(\tau))\, d\tau\right], \tag{60}$$

we have

$$\lim_{k\to+\infty} a_{ij}(t_k) = x_{ij}(t) - \underline{x}_{ij}(t) \ge 0. \tag{61}$$

As a consequence, there exists an index ν such that for $k > \nu$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$,

$$a_{ij}(t_k) \ge 0. \tag{62}$$

As a consequence, we consider the sequence $\{x(t_k)\}_{k\in\mathbb{N}}$ such that

(i) for $k > \nu$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$,

$$x_{ij}(t_k) = \underline{x}_{ij}(t_k) + \min\left\{x_{ij}(t) - \underline{x}_{ij}(t), \overline{x}_{ij}(t_k) - \underline{x}_{ij}(t_k), a_{ij}(t_k)\right\}, \tag{63}$$

(ii) for $k \le \nu$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$,

$$x_{ij}(t_k) = P_{\mathbb{K}(t_k,x^*)} x_{ij}(t), \tag{64}$$

where $P_{\mathbb{K}(t_k,x^*)}(\cdot)$ denotes the Hilbertian projection on $\mathbb{K}(t_k, x^*)$.

Obviously if $k \le \nu$, for (64) we have $x(t_k) \in \mathbb{K}(t_k, x^*)$. Instead, for $k > \nu$, since for (62), $\min\{x_{ij}(t) - \underline{x}_{ij}(t), \overline{x}_{ij}(t_k) - \underline{x}_{ij}(t_k), a_{ij}(t_k)\} \ge 0$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$, we get

$$\underline{x}_{ij}(t_k) \le x_{ij}(t_k), \quad \forall i = 1,\dots,m,\ \forall j = 1,\dots,n. \tag{65}$$

Moreover, since $\min\{x_{ij}(t) - \underline{x}_{ij}(t), \overline{x}_{ij}(t_k) - \underline{x}_{ij}(t_k), a_{ij}(t_k)\} \le \overline{x}_{ij}(t_k) - \underline{x}_{ij}(t_k)$, for all $i = 1,\dots,m$, for all $j = 1,\dots,n$, we have

$$x_{ij}(t_k) \le \overline{x}_{ij}(t_k), \quad \forall i = 1,\dots,m,\ \forall j = 1,\dots,n. \tag{66}$$

Since

$$\begin{aligned}\min&\left\{x_{ij}(t) - \underline{x}_{ij}(t), \overline{x}_{ij}(t_k) - \underline{x}_{ij}(t_k), a_{ij}(t_k)\right\}\\ &\le a_{ij}(t_k) = x_{ij}(t) - \underline{x}_{ij}(t_k)\\ &\quad + \left[\frac{1}{nT}\int_0^T p_i(t_k, x^*(\tau))\, d\tau - \int_0^T p_i(t, x^*(\tau))\, d\tau\right]\\ &\qquad \forall i = 1,\dots,m,\ \forall j = 1,\dots,n,\end{aligned} \tag{67}$$

we have

$$\begin{aligned}x_{ij}(t_k) &\le x_{ij}(t) + \frac{1}{nT}\\ &\quad \times \left[\int_0^T p_i(t_k, x^*(\tau))\, d\tau - \int_0^T p_i(t, x^*(\tau))\, d\tau\right],\\ &\qquad \forall i = 1,\dots,m,\ \forall j = 1,\dots,n.\end{aligned} \tag{68}$$

Then, taking into account (68), we obtain

$$\begin{aligned}\sum_{j=1}^n x_{ij}(t_k) &\le \sum_{j=1}^n x_{ij}(t) + \frac{1}{T}\int_0^T p_i(t_k, x^*(\tau))\, d\tau\\ &\quad - \frac{1}{T}\int_0^T p_i(t, x^*(\tau))\, d\tau\\ &\le \frac{1}{T}\int_0^T p_i(t, x^*(\tau))\, d\tau\\ &\quad + \frac{1}{T}\int_0^T p_i(t_k, x^*(\tau))\, d\tau\\ &\quad - \frac{1}{T}\int_0^T p_i(t, x^*(\tau))\, d\tau\\ &= \frac{1}{T}\int_0^T p_i(t_k, x^*(\tau))\, d\tau, \quad \forall i = 1,\dots,m.\end{aligned} \tag{69}$$

Hence, $x(t_k) \in \mathbb{K}(t_k, x^*)$, for all $k \in \mathbb{N}$, and

$$\begin{aligned}\lim_{k\to+\infty} x_{ij}(t_k) &= \underline{x}_{ij}(t) + \min\{x_{ij}(t) - \underline{x}_{ij}(t), \\ &\qquad \overline{x}_{ij}(t) - \underline{x}_{ij}(t), x_{ij}(t) - \underline{x}_{ij}(t)\} \\ &= \underline{x}_{ij}(t) + x_{ij}(t) - \underline{x}_{ij}(t) = x_{ij}(t).\end{aligned} \tag{70}$$

Then, the proof of condition (K1) is completed.

Let us prove, now, condition (K2). Let $\{t_k\}_{k\in\mathbb{N}}$ be a sequence such that $t_k \to t$, with $t \in [0,T]$, as $k \to +\infty$. Let $\{x(t_k)\}_{k\in\mathbb{N}}$ be a sequence, such that $x(t_k) \in \mathbb{K}(t_k, x^*)$, for all $k \in \mathbb{N}$, and converging to $x(t)$, as $k \to +\infty$. We have to prove that $x(t) \in \mathbb{K}(t, x^*)$.

Since $x(t_k) \in \mathbb{K}(t_k, x^*)$, for all $k \in \mathbb{N}$,

$$\underline{x}_{ij}(t_k) \le x_{ij}(t_k) \le \overline{x}_{ij}(t_k), \quad \forall i = 1,\dots,m, \quad \forall j = 1,\dots,n, \ \forall k \in \mathbb{N}, \tag{71}$$

$$\sum_{j=1}^{n} x_{ij}(t_k) \le \frac{1}{T}\int_0^T p_i(t_k, x^*(\tau))\,d\tau, \quad \forall i = 1,\dots,m, \ \forall k \in \mathbb{N}. \tag{72}$$

Passing to the limit in (71) as $n \to +\infty$ and taking into account the continuity assumption on the functions $\underline{x}$, $\overline{x}$, and p, we have

$$\underline{x}_{ij}(t) \le x_{ij}(t) \le \overline{x}_{ij}(t), \quad \forall i = 1,\dots,m, \ \forall j = 1,\dots,n. \tag{73}$$

Now, passing to the limit for $n \to +\infty$ in the left-hand side of (72), we have

$$\lim_{n\to+\infty}\sum_{j=1}^{n} x_{ij}(t_k) = \sum_{j=1}^{n} x_{ij}(t), \quad \forall i = 1,\dots,m. \tag{74}$$

Then, from (74) and (59), we obtain

$$\sum_{j=1}^{n} x_{ij}(t) \le \frac{1}{T}\int_0^T p_i(t, x^*(\tau))\,d\tau, \quad \forall j = 1,\dots,n. \tag{75}$$

As a consequence,

$$x(t) \in \mathbb{K}(t, x^*), \tag{76}$$

and, hence, condition (K2) is achieved. □

4.2. Continuity of Solutions to Weighted Quasi-Variational Inequalities. In [2, 36–39] some continuity results for variational and quasi-variational inequalities in infinite dimensional spaces have been obtained. It is worth remarking that similar results have been proved for weighted variational and quasi-variational inequalities in nonpivot Hilbert spaces (see [4, 40]).

Now, we show a continuity result for equilibrium solutions to the dynamic elastic oligopolistic market equilibrium problem in presence of production excesses.

Theorem 7. *Let $\underline{x}, \overline{x} \in C^0([0,T], \mathbb{R}^{mn}_+)$, and let $p \in C^0([0,T] \times \mathbb{R}^{mn}_+, \mathbb{R}^m_+)$ be such that*

$$\begin{aligned}&\exists \phi \in C^0([0,T]) : \|p(t,x)\| \le \phi(t) + \|x\|^2, \\ &\qquad \forall x \in \mathbb{R}^{mn}, \text{ in } [0,T], \\ &\exists \nu \in C^0([0,T], \mathbb{R}_+) : \|p(t,x_1) - p(t,x_2)\| \le \nu\|x_1 - x_2\|, \\ &\qquad \forall x_1, x_2 \in \mathbb{R}^{mn}, \text{ in } [0,T].\end{aligned} \tag{77}$$

Moreover, let $v \in C^1([0,T] \times \mathbb{R}^{mn}_+, \mathbb{R}^m_+)$ be a vector function satisfying assumption (iii) and such that

$$\begin{aligned}&\exists \gamma \in C^0([0,T]) : \|\nabla_D v(t,x)\| \le \gamma(t) + \|x\|, \\ &\qquad \forall x \in \mathbb{R}^{mn}, \text{ in } [0,T], \\ &\exists \mu > 0 : \langle -\nabla_D v(t,x) + \nabla_D v(t,y), x - y\rangle \ge \mu\|x - y\|^2, \\ &\qquad \forall x, y \in \mathbb{R}^{mn}, \text{ in } [0,T].\end{aligned} \tag{78}$$

Then the dynamic elastic market equilibrium distribution in presence of production excesses $x^* \in \mathbb{K}(x^*)$ is continuous in $[0,T]$.

Proof. The existence of equilibrium solution is ensured by Theorem 4. Moreover, by applying Theorem 8 in [2] and taking into account Lemma 6, we obtain the continuity of $x^* \in \mathbb{K}(x^*)$ in $[0,T]$. □

5. A Sensitivity Result

In this section a theorem about the sensitivity of solution is presented. The following result establishes that a small change in profit function produces a small change in equilibrium distribution.

Theorem 8. *Assume that the profit function changes from $v(\cdot)$ to the perturbed function $\widetilde{v}(\cdot)$ and denote by x^* and $\widetilde{x}$ the correspondent solutions of the following quasi-variational inequalities:*

$$\langle\langle -\nabla_D v(x^*), x - x^*\rangle\rangle \ge 0, \quad \forall x \in \mathbb{K}(x^*), \tag{79}$$

$$\langle\langle -\nabla_D \widetilde{v}(\widetilde{x}), x - \widetilde{x}\rangle\rangle \ge 0, \quad \forall x \in \mathbb{K}(x^*). \tag{80}$$

Let $\nabla_D v(t,x)$ be a strongly monotone function of constant α, namely, for all $x, y \in \mathbb{K}(x^*)$, $\exists \alpha > 0$ such that

$$\langle\langle -\nabla_D v(x) + -\nabla_D v(y), x - y\rangle\rangle \ge \alpha\|x - y\|^2_{L^2([0,T],\mathbb{R}^{mn})}. \tag{81}$$

Moreover, let $\nabla_D v$ be a Carathéodory function such that

$$\begin{aligned}&\exists h \in L^2([0,T]) : \|\nabla_D v(t, x(t))\|_{mn} \\ &\qquad \le h(t) + \|x(t)\|_{mn}, \quad \text{a.e. in } [0,T].\end{aligned} \tag{82}$$

Then, it follows that

$$\|x^* - \tilde{x}\|_{L^2([0,T],\mathbb{R}^{mn})} \leq \frac{1}{\alpha}\|-\nabla_D \tilde{v}(\tilde{x}) + \nabla_D v(x^*)\|_{L^2([0,T],\mathbb{R}^{mn})}. \tag{83}$$

Proof. Choosing $x(t) = \tilde{x}(t)$ in (79) and $x(t) = x^*(t)$ in (80), by summing up the two new inequalities, we have

$$\langle\langle -\nabla_D \tilde{v}(\tilde{x}) + \nabla_D v(x^*), x^* - \tilde{x}\rangle\rangle \geq 0. \tag{84}$$

By adding and subtracting $-\nabla_D v(\tilde{x})$ in (84), we have

$$\begin{aligned}&\langle\langle -\nabla_D \tilde{v}(\tilde{x}) + \nabla_D v(\tilde{x}), x^* - \tilde{x}\rangle\rangle \\ &\quad \geq \langle\langle -\nabla_D v(x^*) + \nabla_D v(\tilde{x}), x^* - \tilde{x}\rangle\rangle.\end{aligned} \tag{85}$$

Moreover, by using the strong monotonicity, inequality (85), and the Cauchy-Schwartz inequality, we get

$$\begin{aligned}\alpha\|x^* - \tilde{x}\|^2_{L^2([0,T],\mathbb{R}^{mn})} &\leq \langle\langle -\nabla_D v(x^*) + \nabla_D v(\tilde{x}), x^* - \tilde{x}\rangle\rangle \\ &\leq \langle\langle -\nabla_D \tilde{v}(\tilde{x}) + \nabla_D v(\tilde{x}), x^* - \tilde{x}\rangle\rangle \\ &\leq \|-\nabla_D \tilde{v}(\tilde{x}) + \nabla_D v(\tilde{x})\|_{L^2([0,T],\mathbb{R}^{mn})} \\ &\quad \times \|x^* - \tilde{x}\|_{L^2([0,T],\mathbb{R}^{mn})},\end{aligned} \tag{86}$$

from which we get

$$\|x^* - \tilde{x}\|_{L^2([0,T],\mathbb{R}^{mn})} \leq \frac{1}{\alpha}\|-\nabla_D \tilde{v}(\tilde{x}) + \nabla_D v(x^*)\|_{L^2([0,T],\mathbb{R}^{mn})}. \tag{87}$$

□

6. A Numerical Example

This section is devoted to provide a numerical example of the theoretical achievements presented.

Let us consider two firms and two demand markets, as in Figure 1. Let $\underline{x}, \overline{x} \in L^2([0,1],\mathbb{R}^4)$ be the capacity constraints such that, a.e. in $[0,1]$,

$$\underline{x}(t) = \begin{pmatrix} 0 & \frac{2}{5}t \\ \frac{1}{2}t & 0 \end{pmatrix}, \qquad \overline{x}(t) = \begin{pmatrix} 10t & 5t \\ 12t & 10t \end{pmatrix}. \tag{88}$$

Let us denote

$$\begin{aligned}D = \Big\{ x \in L^2([0,1],\mathbb{R}^4) :\ & \underline{x}_{ij}(t) \leq x_{ij}(t) \leq \overline{x}_{ij}(t), \\ & \forall i = 1,2,\ \forall j = 1,2,\ \text{a.e. in } [0,1] \Big\}.\end{aligned} \tag{89}$$

Let $p \in L^1([0,1] \times D, \mathbb{R}^2)$ be the production function, such that, a.e. in $[0,1]$,

$$p(t) = \begin{pmatrix} 6t + 2x^*_{11}(t) \\ 3t + 2x^*_{11}(t) + x^*_{12}(t) \end{pmatrix}. \tag{90}$$

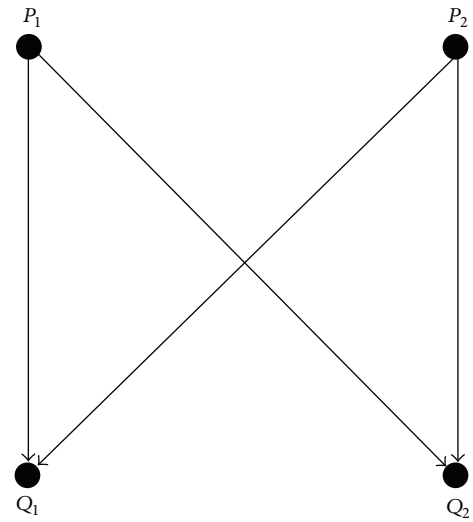

Figure 1: Network structure of the numerical dynamic spatial oligopoly problem.

As a consequence, the feasible set is the set value function $\mathbb{K} : D \to 2^{L^2([0,1],\mathbb{R}^4)}$ defined by

$$\begin{aligned}\mathbb{K}(x^*) = \Big\{ x \in L^2([0,1],\mathbb{R}^4) :\ & \underline{x}_{ij}(t) \leq x_{ij}(t) \leq \overline{x}_{ij}(t), \\ & \forall i = 1,2,\ \forall j = 1,2,\ \text{a.e. in } [0,1], \\ & \sum_{j=1}^{2} x_{ij}(t) \leq \int_0^1 p_i(t, x^*(\tau))\, d\tau, \\ & \forall i = 1,2,\ \text{a.e. in } [0,1] \Big\}.\end{aligned} \tag{91}$$

Let us consider the profit function $v \in C^1([0,1]\times D, \mathbb{R}^2)$ given by

$$\begin{aligned}v_1(t, x(t)) &= -4x^2_{11}(t) - 2x^2_{12}(t) - x_{11}(t)x_{12}(t) \\ &\quad + 6tx_{11}(t) + 3tx_{12}(t), \\ v_2(t, x(t)) &= -2x^2_{21}(t) - 5x^2_{22}(t) - 2x_{21}(t)x_{22}(t) \\ &\quad + 6tx_{21}(t) + 5tx_{22}(t).\end{aligned} \tag{92}$$

Then, the operator $\nabla_D v \in L^2([0,1] \times D, \mathbb{R}^4)$ is given by

$$\begin{aligned}&-\nabla_D v(t, x(t)) \\ &\quad = \begin{pmatrix} 8x_{11}(t) + x_{12}(t) - 6t & 4x_{12}(t) + x_{11}(t) - 3t \\ 4x_{21}(t) + 2x_{22}(t) - 6t & 10x_{22}(t) + 2x_{21}(t) - 5t \end{pmatrix}.\end{aligned} \tag{93}$$

The dynamic oligopolistic market equilibrium distribution in presence of the excesses is the solution to the evolutionary quasi-variational inequality:

$$\int_0^1 \sum_{i=1}^{2}\sum_{j=1}^{2} -\frac{\partial v_i\left(t, x^*(t)\right)}{\partial x_{ij}}\left(x_{ij}(t) - x_{ij}^*(t)\right) dt \geq 0, \quad \forall x \in \mathbb{K}\left(x^*\right). \tag{94}$$

Let us observe that all the hypotheses of Theorem 4 are satisfied; hence, evolutionary quasi-variational inequality (94) admits solutions.

In order to compute a solution to (94) we make use of the direct method (see [41]). We consider the following system:

$$\begin{gathered}
8x_{11}^*(t) + x_{12}^*(t) - 6t = 0, \qquad x_{11}^*(t) + 4x_{12}^*(t) - 3t = 0, \\
4x_{21}^*(t) + 2x_{22}^*(t) - 6t = 0, \qquad 2x_{21}^*(t) + 10x_{22}^*(t) - 5t = 0, \\
\underline{x}_{ij}(t) \leq x_{ij}^*(t) \leq \overline{x}_{ij}(t), \quad \forall i = 1,2, \forall j = 1,2, \text{ a.e. in } [0,1], \\
\sum_{j=1}^{2} x_{ij}^*(t) \leq \int_0^1 p_i\left(t, x^*(\tau)\right) d\tau, \quad \forall i = 1,2, \text{ a.e. in } [0,1],
\end{gathered} \tag{95}$$

and we get the following solution, a.e. in $[0,1]$,

$$x^*(t) = \begin{pmatrix} \frac{7}{10}t & \frac{3}{5}t \\ \frac{25}{18}t & \frac{2}{9}t \end{pmatrix}. \tag{96}$$

Let us observe that the solution satisfies all the constraints; in particular, if we compute

$$\begin{gathered}
\int_0^1 p_1\left(t, x^*(\tau)\right) d\tau = 6t + \frac{7}{10}, \\
\int_0^1 p_2\left(t, x^*(\tau)\right) d\tau = 3t + 1,
\end{gathered} \tag{97}$$

we are able to obtain the related production excesses:

$$\epsilon(t) = \begin{pmatrix} \frac{47t+7}{10} \\ \frac{25t+18}{18} \end{pmatrix}. \tag{98}$$

7. Conclusions

In [1] the dynamic oligopolistic market equilibrium problem was studied by introducing production excesses, and the dynamic Cournot-Nash equilibrium was characterized as a solution to a suitable evolutionary variational inequality. In this paper, in order to have a model closer to reality, it was supposed that the production function depends on the equilibrium commodity shipment. Hence, an elastic formulation was introduced that leads to an equivalent formulation by means of a suitable evolutionary quasi-variational inequality. By means of this mathematical formulation, results of existence and regularity of solutions were proved. Furthermore, a sensitivity analysis is provided. At last a numerical example was provided in order to clarify the theoretical results. In future work, it is possible to consider also demand excesses and elastic demand function, in order to have a more complete and realistic model.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work has been partially supported by F.A.R.O. 2012 "Metodi matematici per la modellizzazione di fenomeni naturali."

References

[1] A. Barbagallo and P. Mauro, "Evolutionary variational formulation for oligopolistic market equilibrium problems with production excesses," *Journal of Optimization Theory and Applications*, vol. 155, no. 1, pp. 288–314, 2012.

[2] A. Barbagallo, "Regularity results for evolutionary nonlinear variational and quasi-variational inequalities with applications to dynamic equilibrium problems," *Journal of Global Optimization*, vol. 40, no. 1-3, pp. 29–39, 2008.

[3] L. Scrimali, "Quasi-variational inequalities in transportation networks," *Mathematical Models & Methods in Applied Sciences*, vol. 14, no. 10, pp. 1541–1560, 2004.

[4] A. Barbagallo and S. Pia, "Weighted quasi-variational inequalities in non-pivot Hilbert spaces and applications," accepted on *Journal of Optimization Theory and Applications*.

[5] A. Cournot, "Researches into the mathematical principles of the theory of wealth," *Competition Policy International*, vol. 4, no. 1, pp. 283–305, 2008.

[6] J. F. Nash, Jr., "Equilibrium points in *n*-person games," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 36, pp. 48–49, 1950.

[7] J. Nash, "Non-cooperative games," *Annals of Mathematics. Second Series*, vol. 54, pp. 286–295, 1951.

[8] S. Dafermos and A. Nagurney, "Oligopolistic and competitive behavior of spatially separated markets," *Regional Science and Urban Economics*, vol. 17, no. 2, pp. 245–254, 1987.

[9] A. Nagurney, "Algorithms for oligopolistic market equilibrium problems," *Regional Science and Urban Economics*, vol. 18, no. 3, pp. 425–445, 1988.

[10] A. Nagurney, P. Dupuis, and D. Zhang, "A dynamical systems approach for network oligopolies and variational inequalities," *The Annals of Regional Science*, vol. 28, no. 3, pp. 263–283, 1994.

[11] A. Nagurney, *Network Economics: A Variational Inequality Approach*, vol. 1 of *Advances in Computational Economics*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[12] A. Barbagallo and M.-G. Cojocaru, "Dynamic equilibrium formulation of the oligopolistic market problem," *Mathematical and Computer Modelling*, vol. 49, no. 5-6, pp. 966–976, 2009.

[13] M. J. Beckmann and J. P. Wallace, "Continuous lags and the stability of market equilibrium," *Economica*, vol. 36, pp. 58–68, 1969.

[14] A. Barbagallo and A. Maugeri, "Memory term for dynamic oligopolistic market equilibrium problem," *Aplimat—Journal of Applied Mathematics*, vol. 3, no. 13, 23 pages, 2010.

[15] A. Barbagallo and R. Di Vincenzo, "Lipschitz continuity and duality for dynamic oligopolistic market equilibrium problem with memory term," *Journal of Mathematical Analysis and Applications*, vol. 382, no. 1, pp. 231–247, 2011.

[16] A. Barbagallo and A. Maugeri, "Duality theory for the dynamic oligopolistic market equilibrium problem," *Optimization*, vol. 60, no. 1-2, pp. 29–52, 2011.

[17] P. Daniele, S. Giuffrè, G. Idone, and A. Maugeri, "Infinite dimensional duality and applications," *Mathematische Annalen*, vol. 339, no. 1, pp. 221–239, 2007.

[18] P. Daniele and S. Giuffrè, "General infinite dimensional duality and applications to evolutionary network equilibrium problems," *Optimization Letters*, vol. 1, no. 3, pp. 227–243, 2007.

[19] P. Daniele, S. Giuffré, and A. Maugeri, "Remarks on general infinite dimensional duality with cone and equality constraints," *Communications in Applied Analysis for Theory and Applications*, vol. 13, no. 4, pp. 567–577, 2009.

[20] A. Maugeri and F. Raciti, "Remarks on infinite dimensional duality," *Journal of Global Optimization*, vol. 46, no. 4, pp. 581–588, 2010.

[21] M. B. Donato, "The infinite dimensional Lagrange multiplier rule for convex optimization problems," *Journal of Functional Analysis*, vol. 261, no. 8, pp. 2083–2093, 2011.

[22] J. M. Borwein and A. S. Lewis, "Partially finite convex programming. I. Quasi relative interiors and duality theory," *Mathematical Programming*, vol. 57, no. 1, pp. 15–48, 1992.

[23] J. Jahn, *Introduction to the Theory of Nonlinear Optimization*, Springer, Berlin, Germany, 2nd edition, 1996.

[24] M. B. Donato, A. Maugeri, M. Milasi, and C. Vitanza, "Duality theory for a dynamic Walrasian pure exchange economy," *Pacific Journal of Optimization*, vol. 4, no. 3, pp. 537–547, 2008.

[25] M. B. Donato and M. Milasi, "Lagrangean variables in infinite dimensional spaces for a dynamic economic equilibrium problem," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 15, pp. 5048–5056, 2011.

[26] A. Barbagallo and P. Mauro, "Time-dependent variational inequality for an oligopolistic market equilibrium problem with production and demand excesses," *Abstract and Applied Analysis*, vol. 2012, Article ID 651975, 35 pages, 2012.

[27] A. Barbagallo and P. Mauro, "On solving dynamic oligopolistic market equilibrium problems in presence of excesses," *Communications in Applied and Industrial Mathematics*, vol. 3, no. 1, pp. 1–20, 2012.

[28] A. Barbagallo, "Advanced results on variational inequality formulation in oligopolistic market equilirbrium problem," *Filomat*, vol. 5, pp. 935–947, 2012.

[29] A. Barbagallo and P. Mauro, "Inverse variational inequality approach and applications," submitted.

[30] A. Barbagallo and P. Mauro, "An inverse problem for the dynamic oligopolistic market equilibriumproblem in presence of excesses," accepted on *Procedia—Social and Behavioral Sciences*.

[31] O. L. Mangasarian, "Pseudo-convex functions," *Journal of the Society for Industrial and Applied Mathematics, Series A Control*, vol. 3, pp. 281–290, 1965.

[32] N. X. Tan, "Quasi-variational inequality in topological linear locally convex Hausdorff spaces," *Mathematische Nachrichten*, vol. 122, pp. 231–245, 1985.

[33] K. Kuratowski, *Topology. Vol. I*, Academic Press, New York, NY, USA, 1966.

[34] G. Salinetti and R. J.-B. Wets, "On the convergence of sequences of convex sets in finite dimensions," *SIAM Review*, vol. 21, no. 1, pp. 18–33, 1979.

[35] G. Salinetti and R. J.-B. Wets, "Addendum "On the convergence of convex sets infinite dimensions"," *SIAM Review*, vol. 22, p. 86, 1980.

[36] A. Barbagallo, "Regularity results for time-dependent variational and quasi-variational inequalities and application to the calculation of dynamic traffic network," *Mathematical Models & Methods in Applied Sciences*, vol. 17, no. 2, pp. 277–304, 2007.

[37] A. Barbagallo, "Existence and regularity of solutions to nonlinear degenerate evolutionary variational inequalities with applications to dynamic network equilibrium problems," *Applied Mathematics and Computation*, vol. 208, no. 1, pp. 1–13, 2009.

[38] A. Barbagallo, "On the regularity of retarded equilibrium in time-dependent traffic equilibrium problems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 12, pp. e2406–e2417, 2009.

[39] A. Barbagallo and M.-G. Cojocaru, "Continuity of solutions for parametric variational inequalities in Banach space," *Journal of Mathematical Analysis and Applications*, vol. 351, no. 2, pp. 707–720, 2009.

[40] A. Barbagallo and S. Pia, "Weighted variational inequalities in non-pivot Hilbert spaces with applications," *Computational Optimization and Applications*, vol. 48, no. 3, pp. 487–514, 2011.

[41] A. Maugeri, "Convex programming, variational inequalities, and applications to the traffic equilibrium problem," *Applied Mathematics and Optimization*, vol. 16, no. 2, pp. 169–185, 1987.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 209340, 13 pages
http://dx.doi.org/10.1155/2013/209340

Research Article

A Sharper Global Error Bound for the Generalized Nonlinear Complementarity Problem over a Polyhedral Cone

Hongchun Sun[1] and Yiju Wang[2]

[1] *School of Science, Linyi University, Linyi, Shandong 276005, China*
[2] *School of Management Science, Qufu Normal University, Rizhao, Shandong 276800, China*

Correspondence should be addressed to Yiju Wang; wyiju@hotmail.com

Received 12 July 2013; Revised 17 September 2013; Accepted 30 September 2013

Academic Editor: Abdellah Bnouhachem

We revisit the global error bound for the generalized nonlinear complementarity problem over a polyhedral cone (GNCP). By establishing a new equivalent formulation of the GNCP, we establish a sharper global error bound for the GNCP under weaker conditions, which improves the existing error bound estimation for the problem.

1. Introduction

Let $\mathscr{K} = \{v \in R^m \mid Av \geq 0, Bv = 0\}$ be a polyhedral cone in R^m for matrices $A \in R^{s\times m}$, $B \in R^{t\times m}$, and let $\mathscr{K}^\circ$ be its dual cone; that is,

$$\mathscr{K}^\circ = \left\{u \in R^m \mid u = A^\top \lambda_1 + B^\top \lambda_2,\ \lambda_1 \in R^s_+,\ \lambda_2 \in R^t\right\}. \tag{1}$$

For continuous mappings $F, G : R^n \to R^m$, the generalized nonlinear complementarity problem, abbreviated as GNCP, is to find vector $x^* \in R^n$ such that

$$F(x^*) \in \mathscr{K}, \qquad G(x^*) \in \mathscr{K}^\circ, \qquad F(x^*)^\top G(x^*) = 0. \tag{2}$$

Throughout this paper, the solution set of the GNCP, denoted by X^*, is assumed to be nonempty.

The GNCP is a direct generalization of the classical nonlinear complementarity problem and a special case of the general variational inequalities problem [1]. The GNCP was deeply discussed [2–5] after the work in [6]. The GNCP plays a significant role in economics, operation research, nonlinear analysis, and so forth (see [7, 8]). For example, the classical Walrasian law of competitive equilibria of exchange economies can be formulated as a generalized nonlinear complementarity problem in the price and excess demand variables (see [8]).

For the GNCP, the solution existence and the numerical solution methods for the GNCP were discussed [2, 3, 6]. As an important tool for a mathematical problem, the global error bound estimation for GNCP with the mapping being γ-strongly monotone and Hölder continuous was discussed in [5], and a global error bound for the GNCP for the linear and monotonic case was established in [4].

In this paper, we will establish a global error bound for the problem (2) without the Hölder continuity of the underlying mapping. To this end, we first develop some new equivalent reformulations of the GNCP under weaker conditions and then establish a sharper global error bound for the GNCP in terms of some easier computed residual functions. The results obtained in this paper can be taken as an improvement of the existing results for GNCP and variational inequalities problem [4, 5, 9–11].

To end this section, we give some notations used in this paper. Vectors considered in this paper are taken in the Euclidean space R^n equipped with the usual inner product, and the Euclidean 2-norm and 1-norm of vector in R^n are, respectively, denoted by $\|\cdot\|$ and $\|\cdot\|_1$. We use R^n_+ to denote the nonnegative orthant in R^n and use x_+ and x_- to denote the vectors composed by elements $(x_+)_i := \max\{x_i, 0\}$, $(x_-)_i := \max\{-x_i, 0\}$, $1 \leq i \leq n$, respectively. For simplicity, we use $(x; y)$ to denote vector $(x^\top, y^\top)^\top$, use I to denote the identity matrix with appropriate dimension, use $x \geq 0$ to denote

a nonnegative vector $x \in R^n$, and use $\operatorname{dist}(x, X^*)$ to denote the distance from point x to the solution set X^*.

2. Global Error Bound for the GNCP

First, we give some concepts used in the subsequent.

Definition 1. The mapping $F : R^n \to R^m$ is said to be

(i) monotone with respect to $G : R^n \to R^m$ if

$$\langle F(x) - F(y), G(x) - G(y)\rangle \geq 0, \quad \forall x, y \in R^n; \tag{3}$$

(ii) γ-strongly G-monotone with respect to $G : R^n \to R^m$ if there are constants $c_1 > 0$, $\gamma > 0$ such that

$$\langle F(x) - F(y), G(x) - G(y)\rangle \geq c_1 \|G(x) - G(y)\|^{1+\gamma}, \quad \forall x, y \in R^n. \tag{4}$$

Remark 2. Based on this definition, γ-strongly G-monotone implies monotonicity, and if $F(x) = Mx + p$, $G(x) = Nx + q$ with $M, N \in R^{m\times n}$, $p, q \in R^m$, then the above Definition 1(i) is equivalent to that the matrix $M^\top N$ is positive semidefinite.

Now, we give some assumptions for our analysis based on Definition 1.

Assumption 3. For mappings F, G and matrix A involved in the GNCP, we assume that

(A1) mapping F is monotone with respect to mapping G;

(A2) matrix $A^\top$ has full-column rank.

Remark 4. Under (A2) in the assumption, matrix $A^\top$ has left inverse $(AA^\top)^{-1}A$, that is, its pseudoinverse of $A^\top$. Certainly, the assumption on matrix $A^\top$ is weaker than that on matrix $(A^\top, B^\top)$ which has full-column rank [4]. In addition, when the mappings F, G are both linear, then Assumption 3(A1) coincides with Assumption (A1) in [4].

In the following, we will establish a new equivalent reformulation to the GNCP. First, we give the following conclusion established in [2].

Theorem 5. *A point $x^* \in R^n$ is a solution of the GNCP if and only if there exist $\lambda_1^* \in R^s$, $\lambda_2^* \in R^t$, such that*

$$\begin{aligned} AF(x^*) &\geq 0, \\ BF(x^*) &= 0, \\ \lambda_1^* &\geq 0, \\ (F(x^*))^\top G(x^*) &= 0, \\ G(x^*) &= A^\top\lambda_1^* + B^\top\lambda_2^*. \end{aligned} \tag{5}$$

From Theorem 5, under Assumption 3(A2), we can transform the system into a new system in which neither λ_1 nor λ_2 is involved. To this end, we need the following conclusion [12].

Lemma 6. *If the linear system $Hy = b$ is consistent, then $y = H^+b$ is the solution with the minimum 2-norm, where H^+ is the pesudo-inverse of H.*

Lemma 7. *Suppose that Assumption 3(A2) holds. Then, for any $x \in R^n$, the following statements are equivalent.*

(1) *There exist $\lambda_1 \in R_+^s$, $\lambda_2 \in R^t$ such that $G(x) = A^\top\lambda_1 + B^\top\lambda_2$.*

(2) *Consider*

$$\begin{aligned} &\left\{-A_L^{-1}B^\top\left[\left(A^\top A_L^{-1} - I\right)B^\top\right]^+\left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\}G(x) \geq 0, \\ &\left\{A^\top\left\{-A_L^{-1}B^\top\left[\left(A^\top A_L^{-1} - I\right)B^\top\right]^+\left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\}\right. \\ &\quad \left. + B^\top\left[\left(A^\top A_L^{-1} - I\right)B^\top\right]^+\left[A^\top A_L^{-1} - I\right] - I\right\}G(x) = 0, \end{aligned} \tag{6}$$

where $A_L^{-1} = (AA^\top)^{-1}A$.

Proof. The proof follows that of Lemma 2.1 in [4], and for completeness, we include it.

Set

$$\begin{aligned} X_1 &:= \left\{x \in R^n \mid G(x) = A^\top\lambda_1 + B^\top\lambda_2 \right. \\ &\qquad \left. \text{for some } \lambda_1 \in R_+^s, \lambda_2 \in R^t\right\}, \\ X_2 &:= \left\{x \in R^n \mid \left\{-A_L^{-1}B^\top\left[\left(A^\top A_L^{-1} - I\right)B^\top\right]^+\right.\right. \\ &\qquad \left. \times\left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\}G(x) \geq 0, \\ &\qquad \left\{A^\top\left\{-A_L^{-1}B^\top\left[\left(A^\top A_L^{-1} - I\right)B^\right]^+\right.\right. \\ &\qquad \left. \times\left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} \\ &\qquad + B^\top\left[\left(A^\top A_L^{-1} - I\right)B^\top\right]^+ \\ &\qquad \left.\left. \times\left[A^\top A_L^{-1} - I\right] - I\right\}G(x) = 0\right\}. \end{aligned} \tag{7}$$

Now, we show that these two sets are equal.

First, for any $x \in X_1$, there exist $\lambda_1 \in R_+^s$, $\lambda_2 \in R^t$ such that

$$G(x) = A^\top\lambda_1 + B^\top\lambda_2. \tag{8}$$

Premultiplying (8) by $A_L^{-1} := (AA^\top)^{-1}A$ gives

$$A_L^{-1}G(x) = \lambda_1 + A_L^{-1}B^\top\lambda_2. \tag{9}$$

Combining this with (8) yields that

$$\begin{aligned} G(x) &= A^\top\left(A_L^{-1}G(x) - A_L^{-1}B^\top\lambda_2\right) + B^\top\lambda_2 \\ &= A^\top A_L^{-1}G(x) - \left[A^\top A_L^{-1}B^\top - B^\top\right]\lambda_2; \end{aligned} \tag{10}$$

that is,

$$\left[A^\top A_L^{-1} B^\top - B^\top\right] \lambda_2 = \left[A^\top A_L^{-1} - I\right] G(x). \tag{11}$$

Recalling Lemma 6, we further have

$$\lambda_2 = \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] G(x). \tag{12}$$

Combining this with (9) yields that

$$\lambda_1 = \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} \times G(x). \tag{13}$$

Using (8), (12), and (13), we have

$$\left\{A^\top \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} + B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] - I\right\} G(x) = 0. \tag{14}$$

From the fact that $\lambda_1 \geq 0$, by (13), one has

$$\left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} G(x) \geq 0. \tag{15}$$

Combining this with (14) leads to that $x \in X_2$. This shows that $X_1 \subseteq X_2$.

Second, for any $x \in X_2$, let

$$\begin{aligned} \lambda_1 &= \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} \times G(x), \\ \lambda_2 &= \left\{\left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right]\right\} G(x). \end{aligned} \tag{16}$$

Then, $\lambda_1 \in R_+^s$, $\lambda_2 \in R^t$. From (14), one has

$$\begin{aligned} G(x) &= A^\top \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} \times G(x) \\ &\quad + B^\top \left\{\left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right]\right\} G(x) \\ &= A^\top \lambda_1 + B^\top \lambda_2; \end{aligned} \tag{17}$$

that is, $x \in X_1$. Hence, $X_2 \subseteq X_1$, and the desired result follows. □

Combining this conclusion with Theorem 5, we can establish the following equivalent formulation of the GNCP:

$$\begin{aligned} AF(x) &\geq 0, \\ BF(x) &= 0, \\ (F(x))^\top G(x) &= 0, \\ UG(x) &\geq 0, \\ VG(x) &= 0, \end{aligned} \tag{18}$$

where

$$\begin{aligned} U &= \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\}, \\ V &= \left\{A^\top \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \times \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} + B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right] - I\right\}. \end{aligned} \tag{19}$$

For the ease of description, we denote $\mu = F(x)$, $\nu = G(x)$. Thus, system (18) can be written as

$$\begin{aligned} A\mu &\geq 0, \\ B\mu &= 0, \\ \mu^\top \nu &= 0, \\ U\nu &\geq 0, \\ V\nu &= 0. \end{aligned} \tag{20}$$

For system (20), one has

$$\begin{aligned} \mu^\top \nu &= \mu^\top \left\{A^\top \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \times \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} + B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right]\right\} \nu \\ &= [A\mu]^\top \left\{-A_L^{-1} B^\top \left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \times \left[A^\top A_L^{-1} - I\right] + A_L^{-1}\right\} \nu \\ &\quad + [B\mu]^\top \left\{\left[\left(A^\top A_L^{-1} - I\right) B^\top\right]^+ \left[A^\top A_L^{-1} - I\right]\right\} \nu \\ &= [A\mu]^\top [U\nu], \end{aligned} \tag{21}$$

where the first equality follows from the last equality in (20), and the last equality uses the second equality in (20). Thus, system (20) can be further written as

$$\begin{aligned} A\mu &\geq 0, \qquad B\mu = 0, \\ (A\mu)^\top (U\nu) &= 0, \\ U\nu &\geq 0, \qquad V\nu = 0. \end{aligned} \tag{22}$$

Furthermore, for any $(\mu, \nu) \in R^m \times R^m$ with $A\mu \geq 0, U\nu \geq 0$, it holds from (21) that

$$\mu^\top \nu \geq 0. \tag{23}$$

Now, consider the following optimization problem:

$$\begin{aligned} \min \quad & f(\omega) = [(I,0)\,\omega]^\top [(0,I)\,\omega] \\ \text{s.t.} \quad & \omega \in \Omega, \end{aligned} \tag{24}$$

where $\omega = (\mu, \nu)$, $\Omega = \{\omega \in R^{2m} \mid A(I,0)\omega \geq 0,\ B(I,0)\omega = 0,\ U(0,I)\omega \geq 0,\ V(0,I)\omega = 0\}$. Denote the solution set of (24) by Ω^*.

Lemma 8. *Under Assumption 3(A1), $f(\omega)$ is a convex function.*

Proof. For any $\omega_1, \omega_2 \in R^{2m}$, $\tau \in [0,1]$, we have

$$\begin{aligned} & f(\tau\omega_1 + (1-\tau)\,\omega_2) - \tau f(\omega_1) - (1-\tau) f(\omega_2) \\ &= [(I,0)(\tau\omega_1 + (1-\tau)\,\omega_2)]^\top \\ &\quad \times [(0,I)(\tau\omega_1 + (1-\tau)\,\omega_2)] \\ &\quad - \tau[(I,0)\,\omega_1]^\top [(0,I)\,\omega_1] \\ &\quad - (1-\tau)[(I,0)\,\omega_2]^\top [(0,I)\,\omega_2] \\ &= \tau^2[(I,0)\,\omega_1]^\top [(0,I)\,\omega_1] \\ &\quad + (1-\tau)^2[(I,0)\,\omega_2]^\top [(0,I)\,\omega_2] \\ &\quad + \tau(1-\tau)[(I,0)\,\omega_1]^\top [(0,I)\,\omega_2] \\ &\quad + \tau(1-\tau)[(I,0)\,\omega_2]^\top [(0,I)\,\omega_1] \\ &\quad - \tau[(I,0)\,\omega_1]^\top [(0,I)\,\omega_1] \\ &\quad - (1-\tau)[(I,0)\,\omega_2]^\top [(0,I)\,\omega_2] \\ &= -\tau(1-\tau)[(I,0)\,\omega_1]^\top [(0,I)\,\omega_1] \\ &\quad - \tau(1-\tau)[(I,0)\,\omega_2]^\top [(0,I)\,\omega_2] \\ &\quad + \tau(1-\tau)[(I,0)\,\omega_1]^\top [(0,I)\,\omega_2] \\ &\quad + \tau(1-\tau)[(I,0)\,\omega_2]^\top [(0,I)\,\omega_1] \\ &= -\tau(1-\tau)([(I,0)\,\omega_1] - [(I,0)\,\omega_2])^\top \\ &\quad \times ((0,I)\,\omega_1 - (0,I)\,\omega_2) \leq 0, \end{aligned} \tag{25}$$

where the first inequality uses Assumption 3(A1). The desired result follows. □

Based on (20), combining (23) with Lemma 8, we can obtain the following conclusion.

Lemma 9. *A point $\omega^* = (\mu^*, \nu^*) \in R^{2m}$ is a solution of (20) if and only if ω^* is a global optimal solution with the objective vanishing of (24).*

In the following, we give the error bound for a polyhedral cone from [13] and error bound for a convex optimization from [14] to reach our aims.

Lemma 10. *For polyhedral cone $P = \{x \in R^n \mid D_1 x = d_1, B_1 x \leq b_1\}$ with $D_1 \in R^{l\times n}$, $B_1 \in R^{m\times n}$, $d_1 \in R^l$, and $b_1 \in R^m$, there exists a constant $c_2 > 0$ such that*

$$\operatorname{dist}(x, P) \leq c_2 \left[\|D_1 x - d_1\| + \|(B_1 x - b_1)_+\|\right], \quad \forall x \in R^n. \tag{26}$$

Lemma 11. *Let P be a convex polyhedron in R^n, and let θ be a convex quadratic function defined on R^n. Let S be the nonempty set of globally optimal solutions of the programming:*

$$\begin{aligned} \min \quad & \theta(x) \\ s.t. \quad & x \in P \end{aligned} \tag{27}$$

with θ_{opt} being the optimal value of θ on S. There exists a scalar $c_3 > 0$ such that

$$\begin{aligned} \operatorname{dist}(x, S) \leq c_3 \max \Big\{ & \operatorname{dist}(x, P), \left|\left[\theta(x) - \theta_{opt}\right]_+\right|, \\ & \left|\left[\theta(x) - \theta_{opt}\right]_+\right|^{1/2} \Big\}, \\ & \forall x \in R^n. \end{aligned} \tag{28}$$

Before proceeding, we present the following definition introduced in [15].

Definition 12. The mapping $G : R^n \to R^m$ is said to be strongly nonexpanding with a constant $\alpha > 0$ if $\|G(x) - G(y)\| \geq \alpha\|x - y\|$.

By Lemma 8, $f(\omega)$ is a convex function and the feasible set Ω is a polyhedral. Combining this with Lemmas 10 and 11, we immediately obtain the following conclusion.

Theorem 13. *Suppose that F is γ-strongly G-monotone with positive constants c_1, γ, respectively, and G is strongly nonexpanding with constant $\alpha > 0$. Then, there exists constant $\rho_1 > 0$ such that*

$$\begin{aligned} \operatorname{dist}(x, X^*) \leq \rho_1 \Big\{ & \|[AF(x)]_-\| + \|BF(x)\| \\ & + \|[UG(x)]_-\| + \|VG(x)\| \\ & + \left|[F(x)^\top G(x)]_+\right| \\ & + \left|[F(x)^\top G(x)]_+\right|^{1/2} \Big\}^{2/(1+\gamma)}, \\ & \forall x \in R^n. \end{aligned} \tag{29}$$

Proof. For any $x \in R^n$, let $\omega = (\mu, \nu) = (F(x), G(x)) \in R^{2m}$. Then, there exists $\omega^* = (\mu^*, \nu^*) = (F(x^*), G(x^*)) \in \Omega^*$ such that $\operatorname{dist}(\omega, \Omega^*) = \|\omega - \omega^*\|$. A direct computation yields that

$$\begin{aligned}
&\operatorname{dist}^{1+\gamma}(x, X^*) \\
&\quad \le \|x - x^*\|^{1+\gamma} \\
&\quad \le \frac{1}{\alpha^{1+\gamma}} \|G(x) - G(x^*)\|^{1+\gamma} \\
&\quad \le \frac{1}{c_1\alpha^{1+\gamma}} \left[(F(x) - F(x^*))^\top (G(x) - G(x^*))\right] \\
&\quad \le \frac{1}{c_1\alpha^{1+\gamma}} \|F(x) - F(x^*)\| \|G(x) - G(x^*)\| \\
&\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} \left\{\|F(x) - F(x^*)\|^2 + \|G(x) - G(x^*)\|^2\right\} \\
&\quad = \frac{1}{2c_1\alpha^{1+\gamma}} \|\omega - \omega^*\|^2 \\
&\quad = \frac{1}{2c_1\alpha^{1+\gamma}} \operatorname{dist}^2(\omega, \Omega^*) \\
&\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} c_3^2 \\
&\qquad \times \max\left\{\operatorname{dist}(\omega, \Omega), |[f(\omega)]_+|, |[f(\omega)]_+|^{1/2}\right\}^2 \\
&\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} c_3^2 \\
&\qquad \times \max\left\{c_2 \{\|[A(I,0)\omega]_-\| + \|B(I,0)\omega\| \right. \\
&\qquad\qquad + \|[U(0,I)\omega]_-\| + \|V(0,I)\omega\|\}, \\
&\qquad\qquad \left. |[f(\omega)]_+|, |[f(\omega)]_+|^{1/2}\right\}^2 \\
&\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} c_3^2 \\
&\qquad \times \max\{c_2, 1\}^2 \left\{\|[A\mu]_-\| + \|B\mu\| \right. \\
&\qquad\qquad + \|[U\nu]_-\| + \|V\nu\| \\
&\qquad\qquad \left. + \left|[\mu^\top \nu]_+\right| + \left|[\mu^\top \nu]_+\right|^{1/2}\right\}^2,
\end{aligned} \tag{30}$$

where the second inequality follows from Definition 12 with constant $\alpha > 0$, the third inequality follows from Definition 1(ii) with constants $c_1 > 0$, $\gamma > 0$, the fourth inequality follows from the Cauchy-Schwarz inequality, the fifth inequality follows from the fact that $(1/2)(a^2 + b^2) \ge ab$, for all $a, b \in R$, the sixth inequality follows from Lemma 11 with constant $c_3 > 0$ and Lemma 9, and the seventh inequality follows from Lemma 10 with constant $c_2 > 0$. By (30) and letting $\rho_1 = \{(1/2c_1\alpha^{1+\gamma})c_3^2 \max\{c_2, 1\}^2\}^{1/(1+\gamma)}$, then the desired result follows. □

Remark 14. It is clear that if F is γ-strongly G-monotone and G is strongly nonexpanding, then

$$\begin{aligned}
\langle F(x) - F(y), G(x) - G(y)\rangle &\ge c_1 \|G(x) - G(y)\|^{1+\gamma} \\
&\ge c_1\alpha^{1+\gamma} \|x - y\|^{1+\gamma}, \\
&\qquad \forall x, y \in R^n.
\end{aligned} \tag{31}$$

Moreover, the conditions which both F and G are Hölder continuous (or both F and G are Lipschitz continuous) in Theorem 13 are removed. Thus, Theorem 13 is stronger than Theorem 2.5 in [5]. Furthermore, by Theorem 2.1 in [5], the GNCP can be reformulated as general variational inequalities problem, and the conditions in Theorem 13 are also weaker than those in Theorem 3.1 in [15], Theorem 3.1 in [11], Theorem 3.1 in [10], and Theorem 2 in [9], respectively.

On the other hand, the condition that F is γ-strongly G-monotone and G is strongly nonexpanding in Theorem 13 is extended compared with the condition that F is strongly monotone with respect to G (i.e., $\gamma = 1$) in Theorems 3.4 and 3.6 in [15], and it is also extended than compared with the condition F is strongly monotone with respect to G (i.e., $\gamma = 1$) in Theorem 3.1 in [11], and compared with the condition that $F(x) = x$, $G(x)$ is strongly monotone (i.e., $\gamma = 1$) in Theorem 3.1 in [10].

Using the following Definition 15 developed from the complementarity conditions in (22), we can further detect the error bound of the GNCP.

Definition 15. A solution x_0 of the GNCP is said to be nondegenerate if it satisfies

$$AF(x_0) + UG(x_0) > 0. \tag{32}$$

Lemma 16. *Suppose that Assumptions 3(A1) and 3(A2) hold, and the GNCP has a nondegenerate solution, say x_0. Then,*

$$\begin{aligned}
\Omega^* = \{\omega \in \Omega \mid &[(I,0)\omega]^\top [(0,I)\omega_0] \\
&+ [(0,I)\omega]^\top [(I,0)\omega_0] = 0\},
\end{aligned} \tag{33}$$

where $\omega_0 = (\mu_0, \nu_0) = (F(x_0), G(x_0))$.

Proof. Since

$$[(I,0)\omega_0]^\top [(0,I)\omega_0] = 0, \tag{34}$$

by Assumption 3(A1), for any $\omega \in \Omega$, we have

$$\begin{aligned}
0 &\le (\mu - \mu_0)^\top (\nu - \nu_0) \\
&= [(I,0)\omega - (I,0)\omega_0]^\top [(0,I)\omega - (0,I)\omega_0] \\
&= [(I,0)\omega]^\top [(0,I)\omega] - [(I,0)\omega]^\top [(0,I)\omega_0] \\
&\quad - [(0,I)\omega]^\top [(I,0)\omega_0];
\end{aligned} \tag{35}$$

that is,

$$\begin{aligned}
&[(I,0)\omega]^\top [(0,I)\omega_0] + [(0,I)\omega]^\top [(I,0)\omega_0] \\
&\qquad \le [(I,0)\omega]^\top [(0,I)\omega].
\end{aligned} \tag{36}$$

To prove the assertion, we only need to show that the solution set Ω^* is equal to the set

$$W := \{\omega \in \Omega \mid [(I,0)\,\omega]^\top [(0,I)\,\omega_0] + [(0,I)\,\omega]^\top [(I,0)\,\omega_0] = 0\}. \tag{37}$$

For any $\tilde{\omega} \in \Omega^*$, combining Lemma 9 with (20) yields that

$$[(I,0)\,\tilde{\omega}]^\top [(0,I)\,\tilde{\omega}] = 0. \tag{38}$$

Letting $\omega = \tilde{\omega}$ in (36) yields that

$$[(I,0)\,\tilde{\omega}]^\top [(0,I)\,\omega_0] + [(0,I)\,\tilde{\omega}]^\top [(I,0)\,\omega_0] \le 0. \tag{39}$$

Since $\tilde{\omega}, \omega_0 \in \Omega$, using the similar technique to that of (21), we can obtain

$$\begin{aligned} &[(I,0)\,\tilde{\omega}]^\top [(0,I)\,\omega_0] + [(0,I)\,\tilde{\omega}]^\top [(I,0)\,\omega_0] \\ &\quad = \tilde{\mu}^\top \nu_0 + \tilde{\nu}^\top \mu_0 \\ &\quad = (A\tilde{\mu})^\top (U\nu_0) + (U\tilde{\nu})^\top (A\mu_0) \\ &\quad \ge 0, \end{aligned} \tag{40}$$

where $\tilde{\omega} = (\tilde{\mu}, \tilde{\nu})$. Combining (39) with (40), we have $\Omega^* \subseteq W$.

On the other hand, for any $\omega \in W$, one has

$$[(I,0)\,\omega]^\top [(0,I)\,\omega_0] + [(0,I)\,\omega]^\top [(I,0)\,\omega_0] = 0. \tag{41}$$

Since $\omega, \omega_0 \in \Omega$, using the similar arguments to that of (21), one has

$$\begin{aligned} [(I,0)\,\omega]^\top [(0,I)\,\omega_0] &= [A\mu]^\top [U\nu_0], \\ [(0,I)\,\omega]^\top [(I,0)\,\omega_0] &= [U\nu]^\top [A\mu_0]. \end{aligned} \tag{42}$$

Combining this with (41) yields that

$$[A\mu]^\top [U\nu_0] + [U\nu]^\top [A\mu_0] = 0. \tag{43}$$

From (32), we deduce that

$$[A\mu]^\top [U\nu] = 0. \tag{44}$$

Thus, using (21), one has

$$\mu^\top \nu = 0. \tag{45}$$

Hence, $\omega \in \Omega^*$. □

Based on Lemma 16, we obtain the following conclusion.

Corollary 17. *Suppose that the hypotheses of Lemma 16 hold. Then,*

$$\Omega^* = \{\omega \in \Omega \mid [(I,0)\,\omega]^\top [(0,I)\,\omega_0] + [(0,I)\,\omega]^\top [(I,0)\,\omega_0] \le 0\}. \tag{46}$$

Theorem 18. *Suppose that the hypotheses of Theorem 13 hold, and the GNCP has a nondegenerate solution. Then, there exists constant $\rho_2 > 0$ such that*

$$\begin{aligned} &\operatorname{dist}(x, X^*) \\ &\quad \le \rho_2 \left\{ \|[AF(x)]_-\| + \|BF(x)\| + \|[UG(x)]_-\| \right. \\ &\qquad \left. + \|VG(x)\| + \left|[F(x)^\top G(x)]_+\right| \right\}^{2/(1+\gamma)}, \quad \forall x \in R^n. \end{aligned} \tag{47}$$

Proof. For any $x \in R^n$, let $\omega = (\mu, \nu) = (F(x), G(x)) \in R^{2m}$. Then, there exists $\omega^* = (\mu^*, \nu^*) = (F(x^*), G(x^*)) \in \Omega^*$ such that $\operatorname{dist}(\omega, \Omega^*) = \|\omega - \omega^*\|$. Letting x_0 be a nondegenerate solution of GNCP and letting $\omega_0 = (F(x_0), G(x_0)) \in \Omega^*$, then

$$\begin{aligned} &\operatorname{dist}^{1+\gamma}(x, X^*) \\ &\quad \le \|x - x^*\|^{1+\gamma} \\ &\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} \operatorname{dist}^2(\omega, \Omega^*) \\ &\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} c_4^2 \\ &\qquad \times \left\{ \|[A(I,0)\,\omega]_-\| + \|[U(0,I)\,\omega]_-\| \right. \\ &\qquad\quad + \|B(I,0)\,\omega\| + \|V(0,I)\,\omega\| \\ &\qquad\quad + \left\| \{[(I,0)\,\omega]^\top [(0,I)\,\omega_0] \right. \\ &\qquad\qquad \left. \left. + [(0,I)\,\omega]^\top [(I,0)\,\omega_0]\}_+ \right\| \right\}^2 \\ &\quad \le \frac{1}{2c_1\alpha^{1+\gamma}} c_4^2 \\ &\qquad \times \left\{ \|[A(I,0)\,\omega]_-\| + \|[U(0,I)\,\omega]_-\| \right. \\ &\qquad\quad + \|B(I,0)\,\omega\| + \|V(0,I)\,\omega\| \\ &\qquad\quad \left. + \left\| \{[(I,0)\,\omega]^\top [(0,I)\,\omega]\}_+ \right\| \right\}^2 \\ &\quad = \frac{1}{2c_1\alpha^{1+\gamma}} c_4^2 \\ &\qquad \times \left\{ \|[A\mu]_-\| + \|[U\nu]_-\| + \|B\mu\| \right. \\ &\qquad\quad \left. + \|V\nu\| + \left\| \{\mu^\top \nu\}_+ \right\| \right\}^2, \end{aligned} \tag{48}$$

where the second equality uses the similar technique to that of (30), the third inequality follows from Corollary 17 and Lemma 10 with constant $c_4 > 0$, and the last inequality is based on (36). By (48) and letting $\rho_2 = \{(1/2c_1\alpha^{1+\gamma})c_4^2\}^{1/(1+\gamma)}$, the desired result follows. □

In the following, we give an error bound of the Hölderian type [14].

Lemma 19. *For $i = 1, 2, \ldots, m$, let $g_i(x)$ be a convex quadratic function. If the set $S := \{x \in R^n \mid g_1(x) \le 0,$*

$g_2(x) \le 0, \ldots, g_m(x) \le 0\}$ is nonempty, then there exist a positive integer $d \le n+1$ (called the degree of singularity of the inequality system) and a positive scalar c_5 such that

$$\operatorname{dist}(x, S) \le c_5 \max\left\{\left\|[g(x)]_+\right\|, \left\|[g(x)]_+\right\|^{1/2^d}\right\}, \quad \forall x \in R^n, \tag{49}$$

where $[g(x)]_+ = ([g_1(x)]_+, [g_2(x)]_+, \ldots, [g_m(x)]_+)$. Furthermore, if S contains an interior point, then $d = 0$.

Based on (18) and (21), the GNCP can also be written as

$$\begin{aligned} AF(x) &\ge 0, \\ BF(x) &= 0, \\ (F(x))^\top G(x) &\le 0, \\ UG(x) &\ge 0, \\ VG(x) &= 0. \end{aligned} \tag{50}$$

From Lemma 19, we can establish the following global error bound for GNCP.

Theorem 20. *Suppose that the hypotheses of Theorem 13 hold, and there exists point $\hat{x} \in R^n$ such that*

$$F(\hat{x})^\top G(\hat{x}) < 0. \tag{51}$$

Then, there exists constant $\rho_3 > 0$ such that

$$\begin{aligned} \operatorname{dist}(x, X^*) \le \rho_3 \Big\{ &\left\|[AF(x)]_-\right\| + \|BF(x)\| \\ &+ \left\|[UG(x)]_-\right\| + \|VG(x)\| \\ &+ \left|\left[F(x)^\top G(x)\right]_+\right| \Big\}^{2/(1+\gamma)}, \quad \forall x \in R^n. \end{aligned} \tag{52}$$

Proof. Let $S_1 := \{\omega \in R^{2m} \mid f(\omega) \le 0\}$, where $f(\omega) = [(I,0)\omega]^\top [(0,I)\omega]$. By Lemma 8, we have $f(\omega)$ is a convex quadratic function. Combining this with (51), using Lemma 19 with $d = 0$, this yields the following result

$$\operatorname{dist}(\omega, S_1) \le c_6 \left\|[f(\omega)]_+\right\|, \quad \forall \omega \in R^{2m}, \tag{53}$$

where c_6 is a positive constant.

Obviously, S_1 is a closed convex set. Thus, for any $\omega \in R^{2m}$, there exists a vector $\overline{\omega} \in S_1$ such that

$$\|\omega - \overline{\omega}\| = \operatorname{dist}(\omega, S_1). \tag{54}$$

For convenience, we also let

$$\begin{aligned} \Psi(\omega) = \ &(-A(I,0)\omega, -B(I,0)\omega, -U(0,I)\omega, \\ &-V(0,I)\omega, B(I,0)\omega, V(0,I)\omega)_+. \end{aligned} \tag{55}$$

From (50), we have $\Omega^* = \Omega \bigcap S_1$, where Ω is defined in (24), so for any $\omega \in S_1$, combining Lemma 10, one has

$$\begin{aligned} \operatorname{dist}(\omega, \Omega^*) &\le c_7 \left[\left\|(-A(I,0)\omega)_+\right\| + \left\|(-U(0,I)\omega)_+\right\| \right. \\ &\quad \left. + \|B(I,0)\omega\| + \|V(0,I)\omega\|\right] \\ &= c_7 \left[\left\|(-A(I,0)\omega)_+\right\| + \left\|(-U(0,I)\omega)_+\right\| \right. \\ &\quad + \left\|(B(I,0)\omega)_+\right\| + \left\|(-B(I,0)\omega)_+\right\| \\ &\quad \left. + \left\|(V(0,I)\omega)_+\right\| + \left\|(-V(0,I)\omega)_+\right\|\right] \\ &\le c_7 \{\left\|(-A(I,0)\omega)_+\right\|_1 + \left\|(-U(0,I)\omega)_+\right\|_1 \\ &\quad + \left\|(B(I,0)\omega)_+\right\|_1 + \left\|(-B(I,0)\omega)_+\right\|_1 \\ &\quad + \left\|(V(0,I)\omega)_+\right\|_1 + \left\|-(V(0,I)\omega)_+\right\|_1\} \\ &= c_7 \|\Psi(\omega)\|_1 \\ &\le c_7 \sqrt{2s + 2t + 2m}\, \|\Psi(\omega)\|, \end{aligned} \tag{56}$$

where c_7 is a positive constant, and the second and third inequalities follow from the fact that $\|x\| \le \|x\|_1 \le \sqrt{n}\|x\|$, for all $x \in R^n$.

Furthermore,

$$\begin{aligned} &\|\Psi(\omega) - \Psi(\overline{\omega})\| \\ &\quad = \|(-A(I,0)\omega, -B(I,0)\omega, -U(0,I)\omega, \\ &\qquad -V(0,I)\omega, B(I,0)\omega, V(0,I)\omega)_+ \\ &\qquad - (-A(I,0)\overline{\omega}, -B(I,0)\overline{\omega}, -U(0,I)\overline{\omega}, \\ &\qquad\quad -V(0,I)\overline{\omega}, B(I,0)\overline{\omega}, V(0,I)\overline{\omega})_+\| \\ &\quad = \Big\| P_{R_+^{2s+2t+2m}} \{(-A(I,0)\omega, -B(I,0)\omega, \\ &\qquad\qquad - U(0,I)\omega, -V(0,I)\omega, \\ &\qquad\qquad B(I,0)\omega, V(0,I)\omega)\} \\ &\qquad - P_{R_+^{2s+2t+2m}} \{(-A(I,0)\overline{\omega}, -B(I,0)\overline{\omega}, \\ &\qquad\qquad - U(0,I)\overline{\omega}, -V(0,I)\overline{\omega}, \\ &\qquad\qquad B(I,0)\overline{\omega}, V(0,I)\overline{\omega})\} \Big\| \\ &\quad \le \|\{(-A(I,0)\omega, -B(I,0)\omega, -U(0,I)\omega, \\ &\qquad -V(0,I)\omega, B(I,0)\omega, V(0,I)\omega)\} \\ &\qquad - \{(-A(I,0)\overline{\omega}, -B(I,0)\overline{\omega}, -U(0,I)\overline{\omega}, \\ &\qquad\quad -V(0,I)\overline{\omega}, B(I,0)\overline{\omega}, V(0,I)\overline{\omega}\}\| \\ &\quad \le \|A(I,0)\omega - A(I,0)\overline{\omega}\| \\ &\qquad + 2\|B(I,0)\omega - B(I,0)\overline{\omega}\| \\ &\qquad + \|U(0,I)\omega - U(0,I)\overline{\omega}\| \\ &\qquad + 2\|V(0,I)\omega - V(0,I)\overline{\omega}\| \end{aligned}$$

$$\begin{aligned}
&\leq (\|A(I,0)\| + 2\|B(I,0)\| + \|U(0,I)\| + 2\|V(0,I)\|) \\
&\quad \times \|\omega - \overline{\omega}\| \\
&= (\|A(I,0)\| + 2\|B(I,0)\| + \|U(0,I)\| + 2\|V(0,I)\|) \\
&\quad \times \operatorname{dist}(\omega, S_1),
\end{aligned} \tag{57}$$

where the second equality follows from the fact that

$$\min\{a,b\} = a - P_{R_+}(a-b), \quad \forall a,b \in R, \tag{58}$$

and the first inequality is by nonexpanding property of projection operator. Thus,

$$\begin{aligned}
\|\Psi(\overline{\omega})\| \leq\ & \|\Psi(\omega)\| \\
& + (\|A(I,0)\| + 2\|B(I,0)\| \\
& + \|U(0,I)\| + 2\|V(0,I)\|) \operatorname{dist}(\omega, S_1).
\end{aligned} \tag{59}$$

Combining (56) with (59), for any $\omega \in R^{2m}$, we have

$$\begin{aligned}
\operatorname{dist}(\omega, \Omega^*) \leq\ & \operatorname{dist}(\omega, S_1) + \operatorname{dist}(\overline{\omega}, \Omega^*) \\
\leq\ & \operatorname{dist}(\omega, S_1) + \sigma\|\Psi(\overline{\omega})\| \leq \operatorname{dist}(\omega, S_1) \\
& + \sigma(\|\Psi(\omega)\| \\
& + (\|A(I,0)\| + 2\|B(I,0)\| + \|U(0,I)\| \\
& + 2\|V(0,I)\|) \operatorname{dist}(\omega, S_1)) \\
\leq\ & \sigma\|\Psi(\omega)\| \\
& + [\sigma(\|A(I,0)\| + 2\|B(I,0)\| \\
& + \|U(0,I)\| + 2\|V(0,I)\|) + 1] \\
& \times \operatorname{dist}(\omega, S_1) \\
\leq\ & \sigma\|\Psi(\omega)\| \\
& + [\sigma(\|A(I,0)\| + 2\|B(I,0)\| \\
& + \|U(0,I)\| + 2\|V(0,I)\|) + 1]\, c_6 \\
& \times \|[f(\omega)]_+\| \\
\leq\ & \eta(\|\Psi(\omega)\| + \|[f(\omega)]_+\|) \\
\leq\ & \eta(\|\Psi(\omega)\|_1 + \|[f(\omega)]_+\|) \\
\leq\ & \eta(\|(-A(I,0)\,\omega)_+\|_1 + \|(-U(0,I)\,\omega)_+\|_1 \\
& + \|B(I,0)\omega\|_1 + \|V(0,I)\omega\|_1 + \|[f(\omega)]_+\|) \\
\leq\ & \eta(\sqrt{s}\|(-A(I,0)\,\omega)_+\| + \sqrt{s}\|(-U(0,I)\,\omega)_+\| \\
& + \sqrt{t}\|B(I,0)\,\omega\| + \sqrt{m}\|V(0,I)\,\omega\| \\
& + \|[f(\omega)]_+\|) \\
\leq\ & c_8(\|(-A(I,0)\,\omega)_+\| + \|(-U(0,I)\,\omega)_+\| \\
& + \|B(I,0)\,\omega\| \\
& + \|V(0,I)\,\omega\| + \|[f(\omega)]_+\|),
\end{aligned} \tag{60}$$

where the second inequality follows from (56) with constant $\sigma = c_7\sqrt{2s + 2t + 2m}$, the third inequality uses (59), the fifth inequality follows from (53), the sixth inequality follows from the fact that

$$\begin{aligned}
\eta = \max\{\sigma, [\sigma(\|A(I,0)\| + 2\|B(I,0)\| + \|U(0,I)\| \\
+ 2\|V(0,I)\|) + 1]\, c_6\},
\end{aligned} \tag{61}$$

the seventh and ninth inequalities follow from the fact that

$$\|x\| \leq \|x\|_1 \leq \sqrt{n}\|x\|, \quad \forall x \in R^n, \tag{62}$$

and the last inequality follows by letting $c_8 = \eta\max\{\sqrt{s}, \sqrt{t}, \sqrt{m}, 1\}$.

For any $x \in R^n$, letting $\omega := (\mu, \nu) = (F(x), G(x)) \in R^{2m}$, then there exists $\omega^* = (\mu^*, \nu^*) = (F(x^*), G(x^*)) \in \Omega^*$ such that $\operatorname{dist}(\omega, \Omega^*) = \|\omega - \omega^*\|$, and a direct computation yields that

$$\begin{aligned}
\operatorname{dist}^{1+\gamma}(x, X^*) \leq\ & \|x - x^*\|^{1+\gamma} \\
\leq\ & \frac{1}{2c_1\alpha^{1+\gamma}}\operatorname{dist}^2(\omega, \Omega^*) \\
\leq\ & \frac{1}{2c_1\alpha^{1+\gamma}}c_8^2 \\
& \times \{\|[A(I,0)\,\omega]_-\| + \|[U(0,I)\,\omega]_-\| \\
& + \|B(I,0)\,\omega\| + \|V(0,I)\,\omega\| \\
& + \|\{[(I,0)\,\omega]^\top [(0,I)\,\omega]\}_+\|\}^2 \\
=\ & \frac{1}{2c_1\alpha^{1+\gamma}}c_8^2\{\|[A\mu]_-\| + \|[U\nu]_-\| + \|B\mu\| \\
& + \|V\nu\| + \|\{\mu^\top\nu\}_+\|\}^2,
\end{aligned} \tag{63}$$

where the deduction of the second equality uses the similar technique to that of (30), and the third inequality is by (60). By (63) and letting $\rho_3 = \{(1/2c_1\alpha^{1+\gamma})c_8^2\}^{1/(1+\gamma)}$, then the desired result follows. □

Remark 21. When F is strongly monotone with respect to G, that is, $\gamma = 1$, without the requirement of nondegenerate solution, the square root term in the error bound estimation is removed as stated in Theorem 20. Hence, the error estimation becomes more practical than that in Theorem 4.1 in [4].

3. Global Error Bound for the GLCP

In this section, we consider the linear case of the GCP such that mappings F and G are both linear; that is, $F(x) = Mx+p$, $G(x) = Nx + q$ with $M, N \in R^{m\times n}$, $p, q \in R^m$:

$$\begin{aligned} \min \quad & H(x) = (Mx+p)^{\top}(Nx+q) \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{64}$$

where

$$X = \left\{ x \in R^n \mid \begin{array}{l} A(Mx+p) \geq 0,\ U(Nx+q) \geq 0, \\ B(Mx+p) = 0,\ V(Nx+q) = 0 \end{array} \right\}. \tag{65}$$

For problem (64), combining (18) with (23) and using a similar discussion in Lemmas 8 and 9, we also have the following conclusion.

Lemma 22. *Under Assumption 3(A1), $H(x)$ is a convex function.*

Lemma 23. *$x^* \in R^n$ is a solution of the GLCP if and only if x^* is global optimal solution with the objective vanishing of* (64).

Based on (64), using the argument similar to that of Theorem 13, we can obtain the following conclusion.

Theorem 24. *Under Assumptions 3(A1) and 3(A2), and that mappings F and G are both linear, there exists constant $\rho_4 > 0$ such that*

$$\begin{aligned} \operatorname{dist}(x, X^*) \leq \rho_4 \Big\{ & \|[AF(x)]_-\| + \|BF(x)\| + \|[UG(x)]_-\| \\ & + \|VG(x)\| + \left|[F(x)^{\top}G(x)]_+\right| \\ & + \left|[F(x)^{\top}G(x)]_+\right|^{1/2} \Big\}, \quad \forall x \in R^n. \end{aligned} \tag{66}$$

Proof. For any $x \in R^n$, a direct computation yields that

$$\begin{aligned} & \operatorname{dist}(x, X^*) \\ & \leq c_9 \max\left\{ \operatorname{dist}(x, X), \left|[H(x)]_+\right|, \left|[H(x)]_+\right|^{1/2} \right\} \\ & \leq c_9 \max\Big\{ c_{10}\{\|[AF(x)]_-\| + \|BF(x)\| \\ & \qquad + \|[UG(x)]_-\| + \|VG(x)\|\}, \\ & \qquad \left|[H(x)]_+\right|, \left|[H(x)]_+\right|^{1/2} \Big\} \\ & \leq c_9 \max\{c_{10}, 1\} \Big\{ \|[AF(x)]_-\| + \|BF(x)\| \\ & \qquad + \|[UG(x)]_-\| \\ & \qquad + \|VG(x)\| + \left|[(F(x))^{\top}G(x)]_+\right| \\ & \qquad + \left|[(F(x))^{\top}G(x)]_+\right|^{1/2} \Big\}, \end{aligned} \tag{67}$$

where the first inequality follows from Lemma 11 with constant $c_9 > 0$ and Lemma 23, and the second inequality uses Lemma 10 with constant $c_{10} > 0$. By (67) and letting $\rho_4 = c_9 \max\{c_{10}, 1\}$, the desired result follows. □

Remark 25. Obviously, Assumption 3(A2) in Theorem 24 is weaker than Assumption (A2) in Theorem 4.1 in [4], Assumption 3(A1) coincides with Assumption (A1) in [4]. In addition, Theorem 24 is sharper than Theorem 4.1 in [4].

The following result further estimates the error bound for the GLCP.

Theorem 26. *Suppose that the hypotheses of Theorem 24 hold, and the GLCP has a nondegenerate solution. Then, there exists constant $\rho_5 > 0$ such that*

$$\begin{aligned} \operatorname{dist}(x, X^*) \leq \rho_5 \Big\{ & \|[AF(x)]_-\| + \|BF(x)\| + \|[UG(x)]_-\| \\ & + \|VG(x)\| + \left|[F(x)^{\top}G(x)]_+\right| \Big\}, \\ & \forall x \in R^n. \end{aligned} \tag{68}$$

Proof. From Corollary 17, we have

$$\begin{aligned} X^* = \Big\{ & x \in X \mid (Mx+p)^{\top}(Nx_0+q) \\ & + (Nx+q)^{\top}(Mx_0+p) \leq 0 \Big\}, \end{aligned} \tag{69}$$

where x_0 is a nondegenerate solution of GLCP, and X is defined in (64). For any $x \in R^n$, a direct computation yields that

$$\begin{aligned} \operatorname{dist}(x, X^*) \leq c_{11} \Big\{ & \|[A(Mx+p)]_-\| + \|[U(Nx+q)]_-\| \\ & + \|B(Mx+p)\| + \|V(Nx+q)\| \\ & + \Big\|\big[(Mx+p)^{\top}(Nx_0+q) \\ & \quad + (Nx+q)^{\top}(Mx_0+p)\big]_+\Big\| \Big\} \\ \leq c_{11} \Big\{ & \|[A(Mx+p)]_-\| + \|[U(Nx+q)]_-\| \\ & + \|B(Mx+p)\| + \|V(Nx+q)\| \\ & + \Big\|\big[(Mx+p)^{\top}(Nx+q)\big]_+\Big\| \Big\}, \end{aligned} \tag{70}$$

where the first inequality follows from Lemma 10 with constant $c_{11} > 0$, and the second inequality uses (36). Letting $\rho_5 = c_{11}$, the desired result follows. □

Remark 27. The condition in Theorem 26 is weaker than that in Theorem 4.2 in [4].

Theorem 28. *Suppose that the hypotheses of Theorem 24 hold, and there exists point $\hat{x} \in R^n$ such that* (51) *holds. Then there exists constant $\rho_6 > 0$ such that*

$$\operatorname{dist}(x, X^*) \le \rho_6 \left\{ \left\|[AF(x)]_-\right\| + \|BF(x)\| + \left\|[UG(x)]_-\right\| + \|VG(x)\| + \left|\left[F(x)^\top G(x)\right]_+\right| \right\}, \quad \forall x \in R^n. \tag{71}$$

Proof. Let $S_2 := \{x \in R^n \mid H(x) \le 0\}$, where $H(x) = (Mx+p)^\top (Nx+q)$. By Lemma 22, $H(x)$ is a convex quadratic function, and S_2 is a closed convex set. For any $x \in R^n$, there exists a vector $\overline{x} \in S_2$ such that

$$\|x - \overline{x}\| = \operatorname{dist}(x, S_2). \tag{72}$$

Combining (51) and applying Lemma 19 yield the following result:

$$\operatorname{dist}(x, S_2) \le c_{12} \left\|[H(x)]_+\right\|, \quad \forall x \in R^n, \tag{73}$$

where c_{12} is a positive constant. For convenience, we let

$$\varphi(x) = (-AF(x), -BF(x), -UG(x) \; -VG(x), BF(x), VG(x))_+. \tag{74}$$

From (50), we have $X^* = X \bigcap S_2$, where X is defined in (64). So for any $x \in S_2$, combining Lemma 10 and using the similar technique to that of (56), one has

$$\begin{aligned} \operatorname{dist}(x, X^*) &\le c_{13} \left[\left\|(-AF(x))_+\right\| + \left\|(-UG(x))_+\right\| + \|BF(x)\| + \|VG(x)\| \right] \\ &\le c_{13} \sqrt{2s + 2t + 2m} \left\|\varphi(x)\right\|, \end{aligned} \tag{75}$$

where c_{13} is a positive constant.

Using the fact that

$$\min\{a, b\} = a - P_{R_+}(a - b), \quad \forall a, b \in R, \tag{76}$$

and using the similar technique to that of (57), one has

$$\begin{aligned} &\left\|\varphi(x) - \varphi(\overline{x})\right\| \\ &= \left\| (-AF(x), -BF(x), -UG(x), -VG(x), BF(x), VG(x))_+ - (-AF(\overline{x}), -BF(\overline{x}), -UG(\overline{x}), -VG(\overline{x}), BF(\overline{x}), VG(\overline{x}))_+ \right\| \\ &\le \|AF(x) - AF(\overline{x})\| + 2\|BF(x) - BF(\overline{x})\| + \|UG(x) - UG(\overline{x})\| + 2\|VG(x) - VG(\overline{x})\| \\ &\le (\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|)\|x - \overline{x}\| \\ &= (\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|) \operatorname{dist}(x, S_2), \end{aligned} \tag{77}$$

where the second inequality is by nonexpanding property of projection operator. Thus,

$$\left\|\varphi(\overline{x})\right\| \le \left\|\varphi(x)\right\| + (\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|) \operatorname{dist}(x, S_2). \tag{78}$$

Combining (75) with (78), we know that for any $x \in R^n$, it holds that

$$\begin{aligned} &\operatorname{dist}(x, X^*) \\ &\le \operatorname{dist}(x, S_2) + \operatorname{dist}(\overline{x}, X^*) \\ &\le \operatorname{dist}(x, S_2) + \sigma_1 \left\|\varphi(\overline{x})\right\| \\ &\le \operatorname{dist}(x, S_2) + \sigma_1 \left(\left\|\varphi(x)\right\| + (\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|) \operatorname{dist}(x, S_2) \right) \\ &\le \sigma_1 \left\|\varphi(x)\right\| + \left[\sigma_1 (\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|) + 1\right] \operatorname{dist}(x, S_2) \\ &\le \sigma_1 \left\|\varphi(x)\right\| + \left[\sigma_1 (\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|) + 1\right] c_{12} \left\|[H(x)]_+\right\| \\ &\le \eta_1 \left(\left\|\varphi(x)\right\| + \left\|[H(x)]_+\right\| \right) \\ &\le \eta_1 \left(\left\|\varphi(x)\right\|_1 + \left\|[H(x)]_+\right\| \right) \\ &\le \eta_1 \left(\left\|(-AF(x))_+\right\|_1 + \left\|(-UG(x))_+\right\|_1 + \|BF(x)\|_1 + \|VG(x)\|_1 + \left\|[H(x)]_+\right\| \right) \\ &\le \eta_1 \left(\sqrt{s} \left\|(-AF(x))_+\right\| + \sqrt{s} \left\|(-UG(x))_+\right\| + \sqrt{t}\|BF(x)\| + \sqrt{m}\|VG(x)\| + \left\|[H(x)]_+\right\| \right) \\ &\le \rho_6 \left(\left\|(-AF(x))_+\right\| + \left\|(-UG(x))_+\right\| + \|BF(x)\| + \|VG(x)\| + \left\|[H(x)]_+\right\| \right), \end{aligned} \tag{79}$$

where the second inequalities follows from (75) with constant $\sigma_1 = c_{13}\sqrt{2s + 2t + 2m}$, the third inequality follows from (78), the fifth inequality follows from (73), the sixth inequality follows by letting $\eta_1 = \max\{\sigma_1, [\sigma_1(\|AM\| + 2\|BM\| + \|UN\| + 2\|VN\|) + 1]c_{12}\}$, and the seventh and ninth inequality follow from the fact that

$$\|x\| \le \|x\|_1 \le \sqrt{n}\|x\|, \quad \forall x \in R^n. \tag{80}$$

By (79) and letting $\rho_6 = \eta_1 \max\{\sqrt{s}, \sqrt{t}, \sqrt{m}, 1\}$, the desired result follows. □

Remark 29. In Theorem 28, without the requirement of nondegenerate solution, the square root term in the error bound estimation is removed. Hence, the error estimation becomes more practical than that in Theorem 4.1 in [4].

4. Comparison with Existing Error Bound

In the end of this paper, we will present an example to compare Theorem 13 and Theorem 2.5 in [5]. Furthermore, we will present two examples to show the conclusion in Theorem 13 can provide a global error bound for the GNCP, while the conclusion in Theorem 2.5 in [5] cannot do.

Example 30. When $\mathscr{K} = R_+^m$, (2) reduces to the generalized nonlinear complementarity problem of finding vector $x^* \in R^n$ such that

$$F(x^*) \geq 0, \qquad G(x^*) \geq 0, \qquad F(x^*)^\top G(x^*) = 0. \quad (81)$$

For (81), using Theorem 13 with $\gamma = 1$, we have

$$\operatorname{dist}(x, X^*) \leq \tilde{\rho}\varphi(x), \quad \forall x \in R^n, \quad (82)$$

where $\varphi(x) =: \|[F(x)]_-\| + \|[G(x)]_-\| + |[F(x)^\top G(x)]_+| + |[F(x)^\top G(x)]_+|^{1/2}$.

Using Theorem 2.5 in [5] with $\gamma = 1$, $\nu_1 = \nu_2 = 1$, and $\beta = 1$, we have that there exists constant $\overline{\rho} > 0$ such that

$$\operatorname{dist}(x, X^*) \leq \overline{\rho} r(x), \quad (83)$$

where $r(x) =: \|\min\{F(x), G(x)\}\|$. In addition,

$$\varphi(x) \leq \left\{2\sqrt{m} + \|\max\{F(x), G(x)\}\|\right\}\left\{r(x) + r(x)^{1/2}\right\}. \quad (84)$$

In particular, when $\|x\| \leq c_{14}$ with constant $c_{14} > 0$, then there exists positive constant c_{15} such that

$$\varphi(x) \leq c_{15}\left\{r(x) + r(x)^{1/2}\right\}. \quad (85)$$

In fact, we have

$$\begin{aligned}
\varphi(x) &\leq \|(F(x))_-\|_1 + \|(G(x))_-\|_1 \\
&\quad + [(F(x))^\top G(x)]_+ + [(F(x))^\top G(x)]_+^{1/2} \\
&\leq 2\sqrt{m}\|\min\{F(x), G(x)\}\| \\
&\quad + [(F(x))^\top G(x)]_+ + [(F(x))^\top G(x)]_+^{1/2} \\
&\leq 2\sqrt{m}\|\min\{F(x), G(x)\}\| \\
&\quad + \sum_{i=1}^m [(F(x))_i (G(x))_i]_+ \\
&\quad + \left\{\sum_{i=1}^m [(F(x))_i (G(x))_i]_+\right\}^{1/2} \\
&\leq 2\sqrt{m}\|\min\{F(x), G(x)\}\| \\
&\quad + \sum_{i=1}^m |\min\{F(x), G(x)\}_i| \cdot |\max\{F(x), G(x)\}_i| \\
&\quad + \left\{\sum_{i=1}^m |\min\{F(x), G(x)\}_i| \cdot |\max\{F(x), G(x)\}_i|\right\}^{1/2} \\
&\leq 2\sqrt{m}\|\min\{F(x), G(x)\}\| \\
&\quad + \|\min\{F(x), G(x)\}\| \cdot \|\max\{F(x), G(x)\}\| \\
&\quad + \{\|\min\{F(x), G(x)\}\| \cdot \|\max\{F(x), G(x)\}\|\}^{1/2} \\
&= \{2\sqrt{m} + \|\max\{F(x), G(x)\}\|\} r(x) \\
&\quad + \|\max\{F(x), G(x)\}\|^{1/2} r(x)^{1/2} \\
&\leq \{2\sqrt{m} + \|\max\{F(x), G(x)\}\|\}\left\{r(x) + r(x)^{1/2}\right\},
\end{aligned} \quad (86)$$

where the first inequality follows from the fact that $\|x\| \leq \|x\|_1$, for all $x \in R^m$, the second inequality follows from the fact that $a_- \leq |\min\{a, b\}|$, for all $a, b \in R$, the third inequality follows from the fact that $(a+b)_+ \leq a_+ + b_+$, for all $a, b \in R$, the fourth inequality follows from the fact that $(ab)_+ \leq |\min\{a, b\}| \cdot |\max\{a, b\}|$, for all $a, b \in R$, and the fifth inequality follows from the Cauchy-Schwarz inequality.

Example 31. For mappings $F, G : R_+ \to R$ involved in problem (81), we set

$$F(x) = (x+1)^2, \qquad G(x) = \sqrt{x}. \quad (87)$$

It is easy to see that the solution set $X^* = \{0\}$, and one has

$$\begin{aligned}
\langle F(x) - F(y), G(x) - G(y)\rangle &= \frac{x+y+2}{\sqrt{x}+\sqrt{y}}(x-y)^2 \\
&\geq (x-y)^2,
\end{aligned} \quad (88)$$

where the first inequality follows from the fact that $x + y + 2 \geq \sqrt{x} + \sqrt{y}$.

In fact, we consider the following four cases.

Case 1 ($x \geq 1$ and $y \geq 1$). Then, $x \geq \sqrt{x}$ and $y \geq \sqrt{y}$, and the desired result follows.

Case 2 ($0 \leq x \leq 1$ and $0 \leq y \leq 1$). Then, $\sqrt{x} \leq 1$ and $\sqrt{y} \leq 1$, and the desired result follows.

Case 3 ($0 \leq x \leq 1$ and $y \geq 1$). Then, $\sqrt{x} \leq 1$ and $\sqrt{y} \leq y$, and the desired result follows.

Case 4 ($x \geq 1$ and $0 \leq y \leq 1$). Then, $\sqrt{x} \leq x$ and $\sqrt{y} \leq 1$, and the desired result follows.

For any $x(\epsilon) := \epsilon, \epsilon \geq 0$. By Theorem 13 with $\gamma = 1$, we can obtain

$$\frac{\|x(\epsilon) - 0\|}{\varphi(x(\epsilon))} = \frac{\epsilon}{(\epsilon+1)^2\sqrt{\epsilon} + \sqrt{(\epsilon+1)^2\sqrt{\epsilon}}} \longrightarrow 0 \tag{89}$$

as $\epsilon \to +\infty$. Thus, Theorem 13 provides a global error bound for the GNCP. Using Theorem 2.5 in [5], for $x(\epsilon)$, we have

$$\frac{\|x(\epsilon) - 0\|}{r(x(\epsilon))} = \frac{\epsilon}{\left\|\min\left\{(\epsilon+1)^2, \sqrt{\epsilon}\right\}\right\|} = \frac{\epsilon}{\sqrt{\epsilon}} = \sqrt{\epsilon} \longrightarrow +\infty \tag{90}$$

as $\epsilon \to +\infty$. Thus, Theorem 2.5 in [5] fails in providing an error bound for this GNCP.

Example 32. For mappings $F, G : R \to R$ involved in problem (81), we set

$$F(x) = \frac{1}{3}x^3 + x, \qquad G(x) = x. \tag{91}$$

It is easy to see that the solution set $X^* = \{0\}$. Without loss of generality, we let $x > y$, and one has

$$F(x) - F(y) \geq (x-y)^2, \tag{92}$$

where the inequality follows from the fact that

$$\begin{aligned}\left(\frac{1}{3}x^3 + x\right) - \left(\frac{1}{3}y^3 + y\right) - (x-y)^2 \\ = (x-y)\left[\frac{1}{3}x^2 + \frac{1}{3}y^2 \right. \\ \left. + \frac{1}{3}xy + 1 + (x-y)\right] \geq 0.\end{aligned} \tag{93}$$

In fact, we consider the following four cases.

Case 1 ($x > y \geq 0$). Then, $(1/3)x^2 + (1/3)y^2 + (1/3)xy + 1 + (x-y) \geq 0$, and the desired result follows.

Case 2 ($0 \geq x > y$). Then, $(1/3)x^2 + (1/3)y^2 + (1/3)xy + 1 + (x-y) \geq 0$, and the desired result follows.

Case 3 ($x \geq 0$, $y < 0$ and $x + y \geq 0$). Then,

$$\begin{aligned}\frac{1}{3}x^2 + \frac{1}{3}y^2 + \frac{1}{3}xy + 1 + (x-y) \\ = \frac{1}{3}x(x+y) + \frac{1}{3}y^2 + 1 + (x-y) \geq 0,\end{aligned} \tag{94}$$

and the desired result follows.

Case 4 ($x \geq 0$, $y < 0$, and $x + y \leq 0$). Then,

$$\begin{aligned}\frac{1}{3}x^2 + \frac{1}{3}y^2 + \frac{1}{3}xy + 1 + (x-y) \\ = \frac{1}{3}x^2 + \frac{1}{3}y(x+y) + 1 + (x-y) \geq 0,\end{aligned} \tag{95}$$

and the desired result follows.

Thus, we obtain

$$\langle F(x) - F(y), G(x) - G(y)\rangle \geq (x-y)^3. \tag{96}$$

For any $x(\epsilon) := \epsilon$, $\epsilon \geq 0$. By Theorem 13 with $\gamma = 2$, we can obtain

$$\frac{\|x(\epsilon) - 0\|}{\varphi(x(\epsilon))} = \frac{\epsilon}{\left[\left((1/3)\epsilon^3 + \epsilon\right)\epsilon + \sqrt{\left((1/3)\epsilon^3 + \epsilon\right)\epsilon}\right]^{2/3}} \longrightarrow 0 \tag{97}$$

as $\epsilon \to +\infty$. Thus, Theorem 13 provides a global error bound for the GNCP.

On the other hand, using Theorem 2.5 in [5], for $x(\epsilon)$, we have

$$\frac{\|x(\epsilon) - 0\|}{r(x(\epsilon))^\delta} = \frac{\epsilon}{\left\|\min\left\{(1/3)\epsilon^3 + \epsilon, \epsilon\right\}\right\|^\delta} = \frac{\epsilon}{\epsilon^\delta} \longrightarrow +\infty \tag{98}$$

as $\epsilon \to +\infty$, where δ is a constant with $1/3 < \delta \leq 1/2$. Thus, Theorem 2.5 in [5] fails in providing an error bound for this GNCP.

5. Conclusion

In this paper, we established some global error bounds on the generalized nonlinear complementarity problems over a polyhedral cone, which improves the result obtained for variational inequalities and the GNCP [4, 5, 9–11] by weakening the assumptions. Surely, under milder conditions, we may establish global error bounds for GNCP and use the error bounds estimation to establish quick convergence rate of the methods for the GNCP. This is a topic for future research.

Acknowledgments

The authors wish to give their sincere thanks to the associated editor and two anonymous referees for their valuable suggestions and helpful comments which improve the presentation of the paper. This work was supported by the Natural Science Foundation of China (nos. 11171180 and 11101303), the Specialized Research Fund for the Doctoral Program of Chinese Higher Education (20113705110002), the Shandong Provincial Natural Science Foundation (ZR2010AL005), the Shandong Province Science and Technology Development Projects (2013GGA13034), and the Applied Mathematics Enhancement Program of Linyi University.

References

[1] M. A. Noor, "General variational inequalities," *Applied Mathematics Letters*, vol. 1, no. 2, pp. 119–122, 1988.

[2] Y. J. Wang, F. M. Ma, and J. Z. Zhang, "A nonsmooth L-M method for solving the generalized nonlinear complementarity problem over a polyhedral cone," *Applied Mathematics and Optimization*, vol. 52, no. 1, pp. 73–92, 2005.

[3] X. Z. Zhang, F. M. Ma, and Y. J. Wang, "A Newton-type algorithm for generalized linear complementarity problem over a polyhedral cone," *Applied Mathematics and Computation*, vol. 169, no. 1, pp. 388–401, 2005.

[4] H. C. Sun, Y. J. Wang, and L. Q. Qi, "Global error bound for the generalized linear complementarity problem over a polyhedral cone," *Journal of Optimization Theory and Applications*, vol. 142, no. 2, pp. 417–429, 2009.

[5] H. C. Sun and Y. J. Wang, "Global error bound estimation for the generalized nonlinear complementarity problem over a closed convex cone," *Journal of Applied Mathematics*, vol. 2012, Article ID 245458, 11 pages, 2012.

[6] R. Andreani, A. Friedlander, and S. A. Santos, "On the resolution of the generalized nonlinear complementarity problem," *SIAM Journal on Optimization*, vol. 12, no. 2, pp. 303–321, 2002.

[7] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequality and Complementarity Problems*, Springer, New York, NY, USA, 2003.

[8] L. Walras, *Elements of Pure Economics*, Allen and Unwin, London, UK, 1954.

[9] M. V. Solodov, "Convergence rate analysis of iteractive algorithms for solving variational inquality problems," *Mathematical Programming*, vol. 96, no. 3, pp. 513–528, 2003.

[10] J.-S. Pang, "A posteriori error bounds for the linearly-constrained variational inequality problem," *Mathematics of Operations Research*, vol. 12, no. 3, pp. 474–484, 1987.

[11] N. H. Xiu and J. Z. Zhang, "Global projection-type error bounds for general variational inequalities," *Journal of Optimization Theory and Applications*, vol. 112, no. 1, pp. 213–228, 2002.

[12] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[13] A. J. Hoffman, "On approximate solutions of systems of linear inequalities," *Journal of Research of the National Bureau of Standards*, vol. 49, pp. 263–265, 1952.

[14] T. Wang and J.-S. Pang, "Global error bounds for convex quadratic inequality systems," *Optimization*, vol. 31, no. 1, pp. 1–12, 1994.

[15] M. A. Noor, "Merit functions for general variational inequalities," *Journal of Mathematical Analysis and Applications*, vol. 316, no. 2, pp. 736–752, 2006.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 624398, 15 pages
http://dx.doi.org/10.1155/2013/624398

Research Article

A Heuristic Algorithm for Constrained Multi-Source Location Problem with Closest Distance under Gauge: The Variational Inequality Approach

Jian-Lin Jiang, Saeed Assani, Kun Cheng, and Xiao-Xing Zhu

Department of Mathematics, College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Correspondence should be addressed to Jian-Lin Jiang; jiangjianlin_nju@yahoo.com

Received 1 August 2013; Accepted 27 August 2013

Academic Editor: Abdellah Bnouhachem

This paper considers the locations of multiple facilities in the space R^p, with the aim of minimizing the sum of weighted distances between facilities and regional customers, where the proximity between a facility and a regional customer is evaluated by the closest distance. Due to the fact that facilities are usually allowed to be sited in certain restricted areas, some locational constraints are imposed to the facilities of our problem. In addition, since the symmetry of distances is sometimes violated in practical situations, the gauge is employed in this paper instead of the frequently used norms for measuring both the symmetric and asymmetric distances. In the spirit of the Cooper algorithm (Cooper, 1964), a new location-allocation heuristic algorithm is proposed to solve this problem. In the location phase, the single-source subproblem with regional demands is reformulated into an equivalent linear variational inequality (LVI), and then, a projection-contraction (PC) method is adopted to find the optimal locations of facilities, whereas in the allocation phase, the regional customers are allocated to facilities according to the nearest center reclassification (NCR). The convergence of the proposed algorithm is proved under mild assumptions. Some preliminary numerical results are reported to show the effectiveness of the new algorithm.

1. Introduction

Due to the large number of practical applications in various fields such as marketing, urban planning, supply chain management, and transportation, the continuous facility location problem has aroused the attention of many researchers ever since the pioneering work [1, 2]. For a comprehensive review on this topic, see, for example, [3, 4]. More specifically, the continuous facility location problem in a space is to seek the optimal locations for facilities serving customers (also called demand points), with certain objectives such as minimizing the sum of distances between facilities and customers.

The vast majority of literature treats the locations of facilities and customers as points in a space. Hence, the distances between facilities and customers are just the point-to-point distance without any ambiguity. In practical applications, however, regional demand arises frequently in such scenarios as uncertain demand, mobile demand, or cumbersome discrete situation whose number of demand points is extremely large. For such scenarios, many authors (e.g., [5–11]) promote that the regional customer, that is, the locations of customers are geometrically connected regions rather than points, should be considered. Therefore, in this paper, we consider the case of regional customer.

The question to be emphasized here, however, is how to measure the proximity from a regional customer to a facility. In the literature, different kinds of distances have been used to measure the required proximity. For example, the average distance evaluated by the proximity between the facility and some mean point in the interior of a regional customer (e.g., [10–13]) and the farthest distance measured by the proximity between the facility and the farthest point on the boundary of a regional customer (e.g., [8, 9, 14, 15]). Definitely, the regional customers are treated by the average fashion when the average distance is considered, while the farthest distance realizes the worst-off nature in the sense that the regional customers are represented by their respective farthest points. In some real-life applications, however, the best-off nature is

of importance in facility location; see, for example, [6, 16, 17]. To realize precisely the desired best-off nature, we need to consider the closest distance; that is, the proximity from a regional customer to a facility is evaluated by the closest distance to this facility. Thus, in this paper, we consider the closest distance to measure such a proximity. Note that the three kinds of distances have been well justified in [14, 15] and the reader can be referred to them for more details.

Focusing on the real-life applications with vast eyes, the regional customers and the closest distances are highly essential to be considered. An example is about the communication network where several central servers are required to be sited. The demand points on the network are partially connected forming groups, each containing a large number of demand points. The points in the same group are connected to one another, and thus, each group becomes a regional customer in the plane. The servers need to be connected to the closest points in each regional customer, and the rest of the regional customer will be connected to the servers through that connection. Hence, the regional customers and the closest distances need to be used in the locations of these central servers. For more applications, we refer to, for example, [6, 15, 16].

As a matter of fact, in the literature of facility location, the distance is usually measured by norms such as l_p-norms and block norms. References [3, 18] discuss the approximations about the weighted l_p-norms based on statistical evidence. References [19, 20] investigate the use of block norms to obtain good approximation of actual distances. As it is known, the symmetry property of the norm assures that the distance from one point to another is always equal to the distance back. Nevertheless, as one of the first in-depth studies of mathematical location problems, [21] highlights the fact that in numerous real situations the symmetry of the distance is violated, for example, transportation in rush-hour traffic, flight in the presence of wind, navigation in the presence of currents, transportation on an inclined terrain, and so forth. For about two decades after the work of [21], however, no further research on this topic seems to be published, and only in the recent twenty years have the asymmetric distance problems started to attract the researchers' interest, for example, [22–28].

In this paper, we are interested in the locations of multiple facilities in the space R^p with the aim of minimizing the sum of weighted distance between these facilities and regional customers, where the distance between a facility and a regional customer is evaluated by the closest distance. In addition, we formulate this problem in a quite general way with the aim of enhancing its practical applicability. First, we recognize that, usually, not all the points in the space R^p are possible locations; that is, new facilities are often allowed to be sited within the confines of the restricted areas. Therefore, we introduce locational constraints so that both the unconstrained and constrained problems are taken into consideration in this paper. Second, since the distance in many practical situations is not necessarily symmetric, the *gauge* is used to measure the distance instead of the widely used norms. With the more general distance measuring function used, both the symmetric and asymmetric distances can be considered in our problem.

The rest of this paper is organized as follows. Section 2 states the formulation of our problem, which is shown to be nonconvex and NP-hard. The spirit of the well-known location-allocation heuristic algorithm, which consists of a location phase and an allocation phase in each iteration, is also discussed in this section. In Section 3, the subproblems arising in location phase and allocation phase are solved. More specifically, for the subproblem arising in allocation phase, the regional customers are allocated to facilities according to the nearest reclassification (NCR) heuristic, whereas for the single-source subproblem arising in location phase, the relationship between the subproblem and a monotone linear variational inequality (LVI) is firstly built, and then, a projection-contraction (PC) method is adopted to find the optimal location of the facility. A new location-allocation heuristic algorithm is proposed in Section 4 for solving our targeted problem, and its convergence is proved in Section 5. Preliminary numerical results are reported in Section 6 to verify the efficiency of the proposed algorithm. Finally, some conclusions are drawn in Section 7.

2. Model Description

This paper focuses on finding locations in the space R^p for m facilities, with the objective to minimize the sum of weighted closest distances between these facilities and n regional customers. Note that the minisum single-source models with closest distance have been well justified in [6, 16] where the distances are particularly measured by l_1-norm and l_2-norm, respectively. In our multi-source model, however, the gauge is used to measure the distances between facilities and regional customers, and thus, both the symmetric distance (including the l_1-norm used in [16] and the l_2-norm used in [6]) and the asymmetric distance are considered.

Let B be a compact convex set in R^p, and let the interior of B contain the origin; then, the gauge of B is defined by

$$\gamma(x) = \inf\left\{\lambda > 0 \mid \frac{x}{\lambda} \in B\right\}, \quad \forall x \in R^p. \tag{1}$$

B is called the unit ball of $\gamma(\cdot)$ due to

$$B = \left\{x \in R^p \mid \gamma(x) \le 1\right\}. \tag{2}$$

This way to define a gauge from a compact convex set was first introduced by Minkowski [29]. The gauge $\gamma(\cdot)$ satisfies the following properties:

(1) $\gamma(x) \ge 0$ and $\gamma(x) = 0 \Leftrightarrow x = 0$;

(2) $\gamma(tx) = t\gamma(x)$ for any $t \ge 0$;

(3) $\gamma(x + y) \le \gamma(x) + \gamma(y)$ for any x and $y \in R^p$.

It follows from (2) and (3) that any gauge $\gamma(x)$ is a convex function of x. The distance measuring function can be derived from a gauge by

$$D(x, y) = \gamma(y - x). \tag{3}$$

When B is symmetric around the origin, according to the definition of (1), we have

$$\gamma(-x) = \gamma(x), \quad \forall x \in R^p. \tag{4}$$

Combining (3) and (4), it follows that

$$D(x, y) = D(y, x), \tag{5}$$

which means that the distance measured by $\gamma(\cdot)$ is symmetric. On the contrary, when B is not symmetric around the origin, (4) does not hold any more, and thus, the distance measured by the gauge $\gamma(\cdot)$ is asymmetric. Thus, when different compact convex sets are used as unit balls, different gauges (symmetric or asymmetric) can be generated and employed to measure distances in location problems, which depends on the requirements of practical applications.

Let $\Lambda = \{A_j \subset R^p \mid j = 1, \ldots, n\}$ denote the set of regional customers, and let $x_i (i = 1, \ldots, m)$ be the location of the ith facility. Each regional customer A_j is simply assumed to be closed and convex. We denote the closest point in A_j to the facility x_i by

$$\begin{aligned} a_j(x_i) &:= \operatorname{argmin}\left\{D(q, x_i) \mid q \in A_j\right\} \\ &= \operatorname{argmin}\left\{\gamma(x_i - q) \mid q \in A_j\right\}. \end{aligned} \tag{6}$$

Then, the closest distance between the facility x_i and the regional customer A_j, denoted by $d_j(x_i)$, can be represented by

$$d_j(x_i) := \min_{q \in A_j} D(q, x_i) = \gamma\left(x_i - a_j(x_i)\right). \tag{7}$$

When the gauge is used to measure distances, we have the following proposition for $a_j(x)$ and $d_j(x)$ which is similar to that in [16].

Proposition 1. *Let x be the location of a facility; then, the closest point $a_j(x)$ in (6) is well defined, and the closest distance $d_j(x)$ in (7) is a convex function of x.*

Proof. Since A_j is a convex set and $\gamma(\cdot)$ is a convex function, (6) is a convex problem, and thus, $a_j(x)$ is well defined.

Now, we prove that $d_j(x)$ is a convex function of x as follows. Let x and y be two points in R^p and $\lambda \in [0, 1]$; then, due to $a_j(x)$ and $a_j(y)$ in A_j and the convexity of A_j, it follows that $\lambda a_j(x) + (1 - \lambda) a_j(y) \in A_j$, and thus, we have

$$\begin{aligned} & d_j(\lambda x + (1 - \lambda) y) \\ &\quad = \min_{a \in A_j} D(a, \lambda x + (1 - \lambda) y) \\ &\quad \le D\left(\lambda a_j(x) + (1 - \lambda) a_j(y), \lambda x + (1 - \lambda) y\right) \\ &\quad = \gamma\left(\lambda x + (1 - \lambda) y - \left(\lambda a_j(x) + (1 - \lambda) a_j(y)\right)\right) \\ &\quad = \gamma\left(\lambda\left(x - a_j(x)\right) + (1 - \lambda)\left(y - a_j(y)\right)\right) \\ &\quad \le \gamma\left(\lambda\left(x - a_j(x)\right)\right) + \gamma\left((1 - \lambda)\left(y - a_j(y)\right)\right) \\ &\quad = \lambda \gamma\left(x - a_j(x)\right) + (1 - \lambda)\gamma\left(y - a_j(y)\right) \\ &\quad = \lambda d_j(x) + (1 - \lambda) d_j(y). \end{aligned} \tag{8}$$

Therefore, $d_j(x)$ is convex with respect to x, and the proof is complete. □

Based on the notations introduced above, now the constrained multi-source location problem (abbreviated as CMLP) we consider in this paper takes the following formulation:

$$\begin{aligned} \text{CMLP:} \quad & \min_{X,W} \sum_{j=1}^{n}\sum_{i=1}^{m} w_{ij} d_j(x_i) = \sum_{j=1}^{n}\sum_{i=1}^{m} w_{ij}\gamma\left(x_i - a_j(x_i)\right) \\ & \text{s.t.} \quad \sum_{i=1}^{m} w_{ij} = s_j, \quad j = 1, 2, \ldots, n, \\ & \qquad x_i \in \Pi_i, \quad i = 1, 2, \ldots, m, \end{aligned} \tag{9}$$

where $s_j \ge 0$ is the given demand required by the jth customer, $X = (x_1^T, \ldots, x_m^T)^T$ is the variable of the locations of facilities to be determined, $W = (w_{ij})_{m \times n}$ is the undetermined variable of w_{ij} which denotes the unknown allocation from the ith facility to the jth customer, and Π_i is the locational constraint for the ith facility which is assumed to be a convex and closed set in R^p.

More explanations are required for our model (9). First, the locational constraint Π_i in (9) can also be chosen as R^p, and if all $\Pi_i (i = 1, \ldots, m)$ are R^p, the CMLP (9) is a unconstrained problem, and thus, both the constrained and unconstrained problems are considered in our model. Second, mark that the minisum models analyzed in [6, 16] are two special cases of CMLP (9), where $m = 1$, $\Pi_i = R^p$, and $\gamma(x)$ is particularly l_1-norm in [16] and Euclidean norm in [6].

It is noted that, with the presupposition that each facility is capable of providing sufficient services for the targeted customers, each customer is ultimately served only by the nearest facility in order to minimize the total sum of weighted distances. Therefore, the mathematical model CMLP also has the following form:

$$\begin{aligned} \text{CMLP}': \quad \min_{X \in \Pi} C(X) &= \sum_{j=1}^{n} s_j \min_{1 \le i \le m} d_j(x_i) \\ &= \sum_{j=1}^{n} s_j \min_{1 \le i \le m} \gamma\left(x_i - a_j(x_i)\right), \end{aligned} \tag{10}$$

where Π is the cartesian product of locational constraints; that is, $\Pi = \Pi_1 \times \cdots \times \Pi_m$.

When all A_j are points not regions and all Π_i are R^p, CMLP (9) reduces to the well-known multi-source Weber problem (MWP) which has wide applications in operations research, marketing, urban planning, and so forth; see, for example, [3, 30, 31]. Recall the fact that the multi-source Weber problem is nonconvex [32] and NP-hard [33], and therefore, heuristics algorithms are extremely popular and highly appreciated for overcoming the difficulty caused by its nonconvexity and NP-hardness; see, for example, [3, 30, 34–37]. In particular, the classical location-allocation heuristic, also called the Cooper algorithm, has received much attention ever since it was presented originally by Cooper in [34] for MWP, whose attractive characteristic is that each iteration consists of a location phase and an allocation phase. Now, as a more general problem of MWP, CMLP is also nonconvex and NP-hard. Hence, in this paper, we are interested in applying the location-allocation heuristic to solve the CMLP (9). Accordingly, some location subproblems and allocation subproblems occur. To clarify it, let $\mathcal{M} = \{1, 2, \ldots, m\}$, $\mathcal{N} = \{1, 2, \ldots, n\}$, and then $\Lambda = \{A_j : j \in \mathcal{N}\}$. At the $(k-1)$th iteration, let $\{\Lambda_1^{k-1}, \Lambda_2^{k-1}, \ldots, \Lambda_m^{k-1}\}$ be the disjoint partition of Λ in the sense that $\cup_{i=1}^m \Lambda_i^{k-1} = \Lambda$ and $\Lambda_i^{k-1} \cap \Lambda_j^{k-1} = \emptyset$ (for $i \neq j$), and each Λ_i^{k-1} $(i = 1, \ldots, m)$ in the partition is called one cluster. Then, at the kth iteration, the location phase finds the candidates of locations of facilities (denoted by $x_1^k, x_2^k, \ldots, x_m^k$) by solving the following m constrained single-source location problems (CSLP) with the closest distance under gauge for each cluster Λ_i^{k-1}, $i = 1, \ldots, m$:

CSLP:

$$x_i^k = \operatorname*{argmin}_{x \in \Pi_i} \left\{ C_i^k(x) := \sum_{\{j \in \mathcal{N} : A_j \in \Lambda_i^{k-1}\}} s_j \gamma\left(x - a_j(x)\right) = \sum_{\{j \in \mathcal{N} : A_j \in \Lambda_i^{k-1}\}} s_j \min_{q_j \in A_j} \gamma\left(x - q_j\right) \right\}. \tag{11}$$

After the location phase, the allocation phase then revises the current partition of Λ to generate a new disjoint partition of $\Lambda = \{\Lambda_1^k, \Lambda_2^k, \ldots, \Lambda_m^k\}$ by the following nearest center reclassification (NCR) heuristic (see [38]): for some customer $A_j \in \Lambda_h^{k-1}$ ($j \in \{1, 2, \ldots, n\}$ and $h \in \{1, 2, \ldots, m\}$), if x_l^k $(l \neq h)$ is the nearest point for A_j among all x_i^k computed by (11), then $\Lambda_h^k = \Lambda_h^{k-1} \setminus \{A_j\}$ and $\Lambda_l^k = \Lambda_l^{k-1} \cup \{A_j\}$. If x_i^k solved by (11) is the nearest facility for each regional customer in Λ_i^{k-1} for any $i \in \{1, 2, \ldots, m\}$, then x_i^k $(i = 1, 2, \ldots, m)$ are the desirable locations of facilities and stop. Otherwise, we set $k = k + 1$ and repeat the iterations.

3. The Subproblems in Location and Allocation Phases

In this section, we will discuss the subproblems arising in the location phase and allocation phase. The allocation phase will partition the customers to m clusters by the nearest center reclassification (NCR) heuristic, and the location phase will find the optimal location for each cluster by solving m CSLPs (11).

3.1. Nearest Center Reclassification for Allocation of Customers. The implementation of NCR heuristic to allocate regional customers can be executed by the following framework; see [38] for more details about this heuristic.

Algorithm 2 (the implementation of NCR). Given an initial partition $\Lambda^0 = \{\Lambda_1^0, \Lambda_2^0, \ldots, \Lambda_m^0\}$.
For $k = 1, 2, \ldots$, *do.*
Step 1. Set $t = 0$ (t stores the number of reassignments);
Step 2. Compute the facility x_i^k of Λ_i^{k-1} by solving CSLP (11), for $i = 1, 2, \ldots, m$;
Step 3. For $j = 1, 2, \ldots, n$ do:

$d_{ij} := \gamma(x_i^k - a_j(x_i^k))$ for $i = 1, \ldots, m$;
if $A_j \in \Lambda_h^{k-1}$ and $d_{lj} = \min_{i=1,\ldots,m;\, i \neq h}\{d_{ij}\} < d_{hj}$,
then $\Lambda_h^k = \Lambda_h^{k-1} \setminus \{A_j\}$, $\Lambda_l^k = \Lambda_l^{k-1} \cup \{A_j\}$;
$t = t + 1$.

Step 4. If $t = 0$, then the iteration terminates with $\{x_1^k, \ldots, x_m^k\}$ being the desirable locations for facilities and the customers in Λ_i^{k-1} being served by x_i^k $(i = 1, \ldots, m)$.

3.2. The Variational Inequality Approach for CSLP (11). According to the spirit of location-allocation heuristic algorithm, our central task for the CMLP is to solve CSLP (11) in location phase by an efficient means. Recall that the CSLP is a generalized problem of the minisum models discussed in [6, 16], where $\Pi_i = R^p$ and the gauge $\gamma(\cdot)$ are the particular l_1-norm in [16] and l_2-norm in [6]. For the model under l_1-norm in [16], by taking advantage of the piecewise linearity of the objective function, this model can be reduced to a standard minisum problem which can be easily solved by obtaining a median point for each coordinate separately. For the minisum model under l_2-norm in [6], an efficient Weiszfeld-type method is proposed, and the convergence of this method is analyzed. Similar to Weiszfeld procedure [2], one problem of the proposed method is that the singular case, that is, the current iterate happens to be within some location of regional customers, may occur during its implementation. Due to the use of the gradient of objective function in the iteration, this method will terminate unexpectedly once the singular case occurs. In order to tackle the undesirable singular case and make the Weiszfeld-type method computational effective, the authors suggest to ignore the gradient of $\|x - a_j(x)\|$ if $x \in A_j$ and then add an extra descent check and a boundary check to the iteration. As pointed out by Theorem 1 in [6], however, the sequence generated by the proposed Weiszfeld-type method is possible to be convergent to a nonoptimal point which is on the boundary of the regional customer.

In this section, a variational inequality approach is proposed to solve the general CSLP (11), where the locational constraints are imposed to the facility and the gauge is

used as distance measuring function. Note that the study of variational inequality has received much attention due to its various applications arising in engineering, operations research, economics, transportation, and so forth; see, for example, [39–45]. Specifically, the CSLP (11) considered in this paper is first reformulated into an equivalent linear variational inequality (LVI), and then, a projection-contraction (PC) method is adopted to solve the LVI. Consequently, a sequence will be generated by the variational inequality approach, which is shown to be convergent to the optimal location of the facility x_i^k of (11) even in the singular case. In addition, the closest points to the facility and the dual vectors with respect to the gauge can also be obtained from the generated sequence.

For convenience and succinctness, with the assumption that Λ_i^{k-1} contains d customers, throughout this section, we ignore some superscripts and subscripts in (11) and consider the simplified model of (11) without confusion:

$$\text{MCSLP:} \quad x = \underset{x\in\Pi}{\operatorname{argmin}} \left\{ C(x) := \sum_{j=1}^{d} s_j \gamma\left(x - a_j(x)\right) \right\}. \tag{12}$$

According to Proposition 1, it follows that CSLP (11), or equivalently MCSLP (12), is convex problem of x.

3.2.1. LVI Reformulation of MCSLP. For the gauge $\gamma(\cdot)$ in (12) which is defined by (1), there exists a dual gauge, $\gamma^d(\cdot)$, defined by

$$\gamma^d(z) = \max\left\{z^T x \mid \gamma(x) \le 1\right\}. \tag{13}$$

Let B^d be the unit ball of the dual gauge $\gamma^d(\cdot)$, which is also convex and compact and exactly the polar set of B. The dual gauge of $\gamma^d(\cdot)$ is again $\gamma(\cdot)$, that is,

$$\gamma(x) = \max\left\{z^T x \mid \gamma^d(z) \le 1\right\}, \tag{14}$$

which can also be rewritten as

$$\gamma(x) = \max_{z\in B^d} z^T x. \tag{15}$$

For more details about gauge and dual gauge, as well as their unit balls, the readers can be referred to [27].

According to (15), MCSLP (12) is equivalent to the following min-max problem:

$$\min_{x\in\Pi} \max_{z_j\in B_{s_j}^d} \sum_{j=1}^{d} z_j^T \left(x - a_j(x)\right), \tag{16}$$

where each z_j is a vector in $B_{s_j}^d = \{\xi \in R^p \mid \gamma^d(\xi) \le s_j\}$. Since $a_j(x)$ is the closest point to x in A_j, we can introduce $y_j \in A_j$ to replace $a_j(x)$. Hence, (16) is equivalent to

$$\min_{x\in\Pi, y_j\in A_j} \max_{z_j\in B_{s_j}^d} \sum_{j=1}^{d} z_j^T \left(x - y_j\right). \tag{17}$$

Denote

$$\begin{gathered} y = \left(y_1^T, \ldots, y_d^T\right)^T, \qquad z = \left(z_1^T, \ldots, z_d^T\right)^T, \\ A = A_1 \times \cdots \times A_d, \qquad \mathscr{B}^d = B_{s_1}^d \times \cdots \times B_{s_d}^d, \end{gathered} \tag{18}$$

and let $(x^*, y^*, z^*) \in \Pi \times A \times \mathscr{B}^d$ be the solution of (17); then, it follows that (x^*, y^*, z^*) is the saddle point of the objective function $\sum_{j=1}^{d} z_j^T (x - y_j)$; that is,

$$\sum_{j=1}^{d} z_j^T \left(x^* - y_j^*\right) \le \sum_{j=1}^{d} z_j^{*T} \left(x^* - y_j^*\right) \le \sum_{j=1}^{d} z_j^{*T} \left(x - y_j\right), \quad \forall (x, y, z) \in \Pi \times A \times \mathscr{B}^d. \tag{19}$$

Thus, (x^*, y^*, z^*) is the solution of the following linear variational inequality:

$$\begin{gathered} x^* \in \Pi, \quad y^* \in A, \quad z^* \in \mathscr{B}^d, \\ \left(x - x^*\right)^T \left(\sum_{j=1}^{d} z_j^*\right) \ge 0, \quad \forall x \in \Pi, \\ \left(y_j - y_j^*\right)^T \left(-z_j^*\right) \ge 0, \quad \forall y_j \in A_j, \\ \left(z_j - z_j^*\right)^T \left(-\left(x^* - y_j^*\right)\right) \ge 0, \quad \forall z_j \in B_{s_j}^d. \end{gathered} \tag{20}$$

A compact form of (20) is

$$\text{LVI}(\Omega, M, q): \quad u^* \in \Omega, \quad \left(u - u^*\right)^T \left(Mu^* + q\right) \ge 0, \quad \forall u \in \Omega, \tag{21}$$

where $u = (x^T, y^T, z^T)^T$, $\Omega = \Pi \times A \times \mathscr{B}^d$,

$$M = \begin{pmatrix} 0 & N \\ -N^T & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} I_2 & \cdots & \cdots & I_2 \\ -I_2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -I_2 \end{pmatrix}, \qquad q = 0. \tag{22}$$

Note that M in (22) is a skew-symmetric matrix, then it is positive semidefinite, and thus, the linear variational inequality (21)-(22) is monotone.

Based on the deduction above, we know that if (x^*, y^*, z^*) is the solution of (17), that is, x^* is the solution of the MCSLP (12), then (x^*, y^*, z^*) will be the solution of the LVI (21)-(22). Further, we can prove that the MCSLP (12) and the LVI (21)-(22) are equivalent in the following theorem.

Theorem 3. *The MCSLP* (12) *and the LVI* (21)-(22) *are equivalent in the sense that they have the same solution of* $x \in \Pi$.

Proof. Since the MCSLP (12) is equivalent to (17), we need to prove that (17) and the LVI (21)-(22) are equivalent. In the following, we will prove that (x^*, y^*, z^*) is a solution of (17) if and only if (x^*, y^*, z^*) is a solution of LVI (21)-(22).

Let (x^*, y^*, z^*) be the solution of (17); then, according to the deduction above, we know that (x^*, y^*, z^*) is the solution of LVI (21)-(22).

On the other hand, let (x^*, y^*, z^*) be the solution of LVI (21)-(22) and $\phi(x, y, z) = \sum_{j=1}^{d} z_j^T (x - y_j)$; then, the inequality (19) is true, which means that (x^*, y^*, z^*) is the saddle point of $\phi(x, y, z)$.

Note that (x^*, y^*, z^*) is the saddle point of $\phi(x, y, z)$ if and only if $(x^*, y^*, z^*) \in \Omega$ and

$$\begin{aligned} \max_{z \in \mathscr{B}^d} \phi(x^*, y^*, z) &= \phi(x^*, y^*, z^*) \\ &= \min_{x \in \Pi, y \in Y} \phi(x, y, z^*), \end{aligned} \tag{23}$$

which implies that

$$\begin{aligned} \min_{x \in \Pi, y \in Y} \max_{z \in \mathscr{B}^d} \phi(x, y, z) &\le \max_{z \in \mathscr{B}^d} \phi(x^*, y^*, z) \\ &= \phi(x^*, y^*, z^*) \\ &= \min_{x \in \Pi, y \in Y} \phi(x, y, z^*) \\ &\le \max_{z \in \mathscr{B}^d} \min_{x \in \Pi, y \in Y} \phi(x, y, z). \end{aligned} \tag{24}$$

On the other hand, let z' be any vector in $\mathscr{B}^d$; then, we have

$$\min_{x \in \Pi, y \in Y} \phi(x, y, z') \le \min_{x \in \Pi, y \in Y} \max_{z \in \mathscr{B}^d} \phi(x, y, z). \tag{25}$$

We choose z' in (25) as the maximum point of the left term over $z' \in \mathscr{B}^d$; then,

$$\max_{z \in \mathscr{B}^d} \min_{x \in \Pi, y \in Y} \phi(x, y, z) \le \min_{x \in \Pi, y \in Y} \max_{z \in \mathscr{B}^d} \phi(x, y, z). \tag{26}$$

Combining (24) and (26), it follows that all terms in (24) are equal, and therefore,

$$\phi(x^*, y^*, z^*) = \min_{x \in \Pi, y \in Y} \max_{z \in \mathscr{B}^d} \phi(x, y, z), \tag{27}$$

which implies that (x^*, y^*, z^*) is the solution of (17). □

Remark 4. It is worth pointing out that the equivalence between the MCSLP (12) and the linear variational inequality (21)-(22) can also be obtained by the duality theory and the variable z_j and the set $B_{s_j}^d$ $(j = 1, \ldots, d)$ in (16) are, respectively, the dual vector and dual ball in the space R^p which satisfy $z_j \in B_{s_j}^d$.

The norms especially l_1, l_2, and l_∞ are frequently used to measure distances in the literature; see, for example, [18, 19]. It should be noted that the gauge used in this paper is an extension of norms which include l_1, l_2, and l_∞. When the gauge $\gamma(\cdot)$ is chosen as the l_1, l_2, and l_∞-norm, the dual gauge $\gamma^d(\cdot)$ will be the l_∞, l_2, and l_1-norm, respectively. Let

$$B_{s_j,2} = \left\{\xi \in R^p \mid \|\xi\|_2 \le s_j\right\}, \qquad \mathscr{B}_2 = B_{s_1,2} \times \cdots \times B_{s_d,2}, \tag{28}$$

where $\|\cdot\|_2$ is the Euclidean norm; then, as a particular case of our single-source location problem (11), the problem under l_2-norm analyzed in [6] can be reformulated into the LVI (21)-(22) in which $\mathscr{B}^d$ is equal to $\mathscr{B}_2$ and $\Pi = R^p$.

Further let

$$\begin{aligned} B_{s_j,1} &= \left\{\xi \in R^p \mid \|\xi\|_1 \le s_j\right\}, \\ B_{s_j,\infty} &= \left\{\xi \in R^p \mid \|\xi\|_\infty \le s_j\right\}, \end{aligned} \tag{29}$$

where $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are the l_1, l_∞-norm, respectively, and

$$\mathscr{B}_1 = B_{s_1,1} \times \cdots \times B_{s_d,1}, \qquad \mathscr{B}_\infty = B_{s_1,\infty} \times \cdots \times B_{s_d,\infty}. \tag{30}$$

Then, the CSLP (11) under l_1-norm (the minisum model discussed in [16]) and the CSLP under l_∞-norm are equivalent to the LVI (21)-(22), where $\mathscr{B}^d$ is, respectively, equal to $\mathscr{B}_\infty$ and $\mathscr{B}_1$ and the locational constraints Π are both R^p.

3.2.2. A Projection-Contraction Method for LVI (21)-(22). Among numerous effective numerical algorithms for solving VI, especially LVI, one famous one is the projection-contraction (PC) method which was originally proposed by Uzawa [46]. The attractive characteristics of the PC method, for example, simpleness and effectiveness, have motivated further development on VI especially in computational aspects; see, for example, [39, 47–49]. In this section, we will summarize some concepts and results about linear variational inequalities and then adopt the projection-contraction method in [48] for solving LVI (21)-(22). More details about the proposed PC method can be referred to [48].

Let W be a nonempty closed convex set of $R^{\mathcal{Q}}$. For a given $v \in R^{\mathcal{Q}}$, the projection of v onto W denoted by $P_W(v)$ is the unique solution of the following problem:

$$P_W(v) = \operatorname{argmin}\left\{\|u - v\|_2 \mid u \in W\right\}. \tag{31}$$

A basic proposition of the projection mapping on a closed convex set is

$$\left(v - P_W(v)\right)^T \left(u - P_W(v)\right) \le 0, \quad \forall v \in R^{\mathcal{Q}}, \forall u \in W. \tag{32}$$

It is well known (see, e.g., [50]) that for any $\beta > 0$, u^* is the solution of LVI(Ω, M, q) if and only if

$$e(u^*, \beta) := u^* - P_\Omega\left[u^* - \beta(Mu^* + q)\right] = 0. \tag{33}$$

In the literature of variational inequalities, $e(u, \beta)$ is usually called the error bound of LVI, and it quantitatively measures how much u fails to be the solution of LVI(Ω, M, q). Therefore, $e(u, \beta)$ can serve as the stopping criterion for solving LVI(Ω, M, q) iteratively.

Let

$$e(u) = e(u,1), \qquad g(u) = M^T e(u) + (Mu + q), \qquad \varphi(u) = e(u)^T (Mu + q), \tag{34}$$

and Ω^* be the set of solutions of LVI(Ω, M, q); then, for the positive semidefinite (not necessarily symmetric) matrix M, the following theorem can be obtained.

Theorem 5 (Lemma 1 and Theorem 2 in [48]). *Let $u \in \Omega$, $u^* \in \Omega^*$, $g(u)$, and $\varphi(u)$ be defined as* (34). *Then, it holds that*

$$(u - u^*)^T g(u) \geq \varphi(u) \geq \|e(u)\|_2^2. \tag{35}$$

For $u \in \Omega \setminus \Omega^*$, it follows from Theorem 5 that $-g(u)$ is a descent direction of the unknown function $\|u - u^*\|_2^2$. We state the projection-contraction method in [48] as follows which is used to solve the LVI (21)-(22).

Algorithm 6 (the projection-contraction method for LVI (21)-(22)).
Step 0. Let $\varepsilon > 0$, $\beta = 1$, α_1, α_2 ($\alpha_1 > \alpha_2$) and $u^0 \in \Omega$. Set $k = 0$.
Step 1. Calculate $e(u^k)$. If $\|e(u^k)\| < \varepsilon$, stop.
Step 2. Calculate $g(u^k)$ and set $\alpha(u^k)$ as

$$\alpha(u^k) = \frac{\|e(u^k)\|^2}{\|e(u^k) + M^T e(u^k)\|^2}. \tag{36}$$

Step 3. Calculate u^{k+1} as

$$u^{k+1} = P_\Omega \left[u^k - \rho\alpha(u^k) g(w^k)\right], \quad \rho \in (0,2). \tag{37}$$

Step 4. Adjust β as follows

$$\beta = \begin{cases} \frac{3}{2}\beta & \sqrt{\alpha(u^k)} \geq \alpha_1, \\ \frac{2}{3}\beta & \sqrt{\alpha(u^k)} \leq \alpha_2, \\ \beta & \text{otherwise}, \end{cases} \tag{38}$$

and set

$$M = \beta M, \qquad q = \beta q. \tag{39}$$

Let $k = k + 1$ and go to Step 1.

Remark 7. In Step 4 of Algorithm 6, the parameter β is self-adaptive during the iterations according to the value of $\alpha(u^k)$. Note that M is skew-symmetric, and thus, $\alpha(u^k)$ can also be rewritten as

$$\alpha(u^k) = \frac{\|e(u^k)\|^2}{\left(\|e(u^k)\|^2 + \|M^T e(u^k)\|^2\right)}, \tag{40}$$

which is shown to be in $[0,1]$. It follows from (39)-(40) that the two terms $\|e(u^k)\|^2$ and $\|M^T e(u^k)\|^2$ in the denominator of $\alpha(u^k)$ are balanced by the self-adaptive parameter β.

Theorem 8 (Theorem 3 in [48]). *Let u^* be a solution of LVI* (21)-(22)*; then, the sequence $\{u^k\}$ generated by Algorithm 6 satisfies*

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \frac{\rho(2-\rho)}{\|I + M^T\|^2}\|e(u^k)\|^2. \tag{41}$$

As a result, $\{e(u^k)\}$ converges to zero, and thus, all accumulation points of $\{u^k\}$ are the solutions of LVI (21)-(22). However, it follows from (41) that $\|u^{k+1} - u^*\| \leq \|u^k - u^*\|$, which implies that $\{u^k\}$ has only one accumulation point. Thus, the sequence $\{u^k\}$ generated by Algorithm 6 will converge to the optimal solution of LVI (21)-(22).

4. A Location-Allocation Heuristic for CMLP (9)

Recall the fact that for the well-known multi-source Weber problem (MWP), heuristics algorithms are extremely popular and frequently used for overcoming its nonconvexity and NP-hardness. In particular, the location-allocation heuristic algorithm has drawn much attention ever since its presentation by Cooper [34]. Note that the targeted CMLP (9) is an extension of the MWP and it is harder than MWP, and thus, in this paper, we also focus on applying the location-allocation heuristic algorithm for solving the CMLP in the spirit of Cooper's work.

Our previous analysis indicates that each iteration of the location-allocation heuristic algorithm to be presented consists of an allocation phase and a location phase. The allocation task generates a new disjoint partition of all the regional customers according to the principle of NCR as in the Cooper algorithm, and the location phase identifies the optimal locations for the current partition of customers via implementing the variational inequality approach for solving m CSLPs.

Mark that the CMLP (9) differs from MWP mainly in that the customers are represented by regions rather than points. Consequently, the CSLPs involved in the location phase are constrained location problems with regional demand and closest distances under gauge. No doubt that the numerical implementation of the heuristic algorithm to be presented is expected to be more complicated than the location-allocation algorithms for MWP. Therefore, how to accelerate the convergence of the proposed heuristic deserves further consideration. To achieve this objective, we here consider a particular strategy for the initial partition of regional customers or the initial locations of facilities. In practical implementation, we suggest to choose the solution of the following constrained multi-source Weber problem (CMWP) as the initial locations of facilities for CMLP:

$$\text{CMWP:} \quad \min_{(x_1,\ldots,x_m)\in\Pi_1\times\cdots\times\Pi_m} C'(X) = \sum_{j=1}^{n} s_j \min_{1\leq i\leq m} \gamma(x_i - g_j), \tag{42}$$

where g_j's are geometric centers of the regional customers. Then, we apply the NCR to determine an initial partition of regional customers according to the solution of (42). For solving the constrained multi-source Weber problem (42), we

employ the location-allocation heuristic algorithm in [35]. As we will show by numerical experiments, this initialization strategy can accelerate the convergence of the proposed algorithm greatly.

In the spirit of Cooper's work, the new heuristic algorithm is ready to be presented for solving the targeted CMLP (9), and its iterative framework can be elaborated as follows.

Algorithm 9 (a location-allocation heuristic algorithm for CMLP). *Initialization:* Solve (42) by the location-allocation heuristic in [35] and use its heuristic solutions as the initial locations of facilities $(x_1^0, x_2^0, \ldots, x_m^0)$. Then, the initial partition of regional customers, which is denoted by $\Lambda^0 = \{\Lambda_1^0, \Lambda_2^0, \ldots, \Lambda_m^0\}$, is generated by the spirit of NCR heuristic (Step 3 in Algorithm 2). Set $k = 1$.

Step 1 (location phase). Solve the involved CSLP (11) and find the location of facility x_i^k for Λ_i^{k-1} by the variational inequality approach. Denote $X^k = ({x_1^k}^T, \ldots, {x_m^k}^T)^T$.

Step 2 (allocation phase). Update the partition of regional customers Λ from Λ_i^{k-1} to Λ_i^k based on the spirit of NCR heuristic.

Step 3 If $\|X^k - X^{k-1}\| < \varepsilon$, the current locations and partition are heuristic locations of facilities and heuristic partition of customers. Otherwise, set $k = k + 1$ and go to Step 1.

Remark 10. At the $(k+1)$th iteration, it is recommended to use x_i^k (and the corresponding y and z) in Step 1 as the initial iterate in the variational inequality approach for solving x_i^{k+1} $(i = 1, \ldots, m)$, considering the fact that Λ_i^k usually differs from Λ_i^{k-1} slightly in practical implementation.

Remark 11. Compared to the main body of the proposed location-allocation heuristic algorithm, the workload of the initialization is relatively less. However, this initialization strategy can reduce its number of iterations and computing time, which will be verified by the numerical experiments to be reported in Section 6.2. Hence, the convergence of the proposed algorithm is accelerated greatly by this initialization strategy.

5. Convergence of the Proposed Heuristic Algorithm

In this section, we analyze the convergence of the proposed location-allocation heuristic (Algorithm 9). For simplification of our discussion, some notations are introduced as follows. Let $A = A_1 \times \cdots \times A_n$, and recall $\Pi = \Pi_1 \times \cdots \times \Pi_m$. For any $X = (x_1^T, \ldots, x_m^T)^T \in \Pi$ and $C = (c_1^T, \ldots, c_n^T)^T \in A$, we can define an ordered pair (X, C) and we can also define the function $\omega(X, C)$, in the current partition of customers as follows:

$$\begin{aligned}\omega(X, C) &= \sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij}\gamma\left(x_i - c_j\right) \\ &= \sum_{i=1}^{m}\sum_{\{j\in\mathcal{N}:A_j\in\Lambda_i\}} s_j\gamma\left(x_i - c_j\right),\end{aligned} \tag{43}$$

which represents the objective functional value of CMLP (9) at (X, C).

During the implementation of Algorithm 9, we denote the map $\mathscr{L} : \Pi \times A \to \Pi \times A$ as the location operation in Step 1 and the map $\mathscr{A} : \Pi \times A \to \Pi \times A$ as the allocation operation in the Step 2. It follows that

$$\begin{aligned}\mathscr{L}\left(X^k, C^k\right) &= \left(X^{k+1}, \widetilde{C}^k\right), \\ \mathscr{A}\left(X^{k+1}, \widetilde{C}^k\right) &= \left(X^{k+1}, C^{k+1}\right),\end{aligned} \tag{44}$$

where X^k and X^{k+1} are, respectively, the locations of facilities in the kth and $(k+1)$th iteration, C^k is the variable of the closest points to X^k in the partition of Λ^k, $\widetilde{C}^k$ is the closest points to the new X^{k+1} in the partition of Λ^k, and C^{k+1} is the closest points to the new X^{k+1} in the new partition of Λ^{k+1}. Then, the iterate scheme of the location-allocation heuristic is

$$\left(X^{k+1}, C^{k+1}\right) = \mathscr{A}\mathscr{L}\left(X^k, C^k\right). \tag{45}$$

Let $S(X^0, C^0)$ denote the iterative sequence generated by the location-allocation heuristic for CMLP with the initial iterate (X^0, C^0). During the implementation of the proposed heuristic algorithm, we choose the initial iterate in location phase for solving CSLP as Remark 10 indicates. We first give the following proposition which reveals the monotonicity of the generated sequence $S(X^0, C^0)$.

Proposition 12. *$S(X^0, C^0)$ is strictly monotone in the sense that $\omega(X^{k+1}, C^{k+1}) < \omega(X^k, C^k)$ if $X^{k+1} \neq X^k$.*

Proof. Since $X^{k+1} \neq X^k$, there exists at least one $i \in \mathscr{M}$ such that $x_i^{k+1} \neq x_i^k$. For such i's according to the following convex optimization problem

$$\begin{aligned}&x_i^{k+1} \\ &= \underset{x\in\Pi_i}{\operatorname{argmin}} \left\{ C_i^{k+1}(x) = \sum_{\{j\in\mathcal{N}:A_j\in\Lambda_i^k\}} s_j\gamma\left(x - a_j(x)\right)\right\},\end{aligned} \tag{46}$$

we have

$$\omega\left(X^{k+1}, \widetilde{C}^k\right) \le \omega\left(X^k, C^k\right). \tag{47}$$

Based on Remark 10, we know that if x_i^k is the solution of (46), then x_i^{k+1} will be equal to x_i^k. Therefore, $x_i^{k+1} \neq x_i^k$ implies that x_i^k is not the solution of (46), and thus, $C_i^{k+1}(x^{k+1}) < C_i^k(x^k)$. It follows that

$$\omega\left(X^{k+1}, \widetilde{C}^k\right) < \omega\left(X^k, C^k\right). \tag{48}$$

On the other hand, based on the principle of NCR in the allocation phase of Algorithm 9, we also have

$$\omega\left(X^{k+1}, C^{k+1}\right) \le \omega\left(X^{k+1}, \widetilde{C}^k\right). \tag{49}$$

By Combining (48) and (49), it follows that

$$\omega\left(X^{k+1}, C^{k+1}\right) < \omega\left(X^{k}, C^{k}\right). \tag{50}$$

The proof is complete. □

Based on the monotonicity of the generated sequence $S(X^0, C^0)$, the following theorem can be proved.

Theorem 13. *Let $\{S_k\} := S(X^0, C^0)$. Then, the generated sequence $\{S_k\}$ satisfies that*

(1) *$\omega(S_k) \to \omega(S)$ for some $S \in \Pi \times A$,*

(2) *all accumulation points of $\{S_k\}$ have the same objective functional values.*

Proof. After a finite number of iterations, if $X^{J+1} = X^J$, then the iterates after S_J will be constant, and thus, $\{S_k\}$ is convergent to $(X^J, C^J) \in \Pi \times A$, and the two assertions are both true.

Below, we will discuss the case that $X^{k+1} \neq X^k$ for any $k \in N$. First, we prove the first assertion. Since $S_k \in \Pi \times A$ and $\Pi \times A$ is a compact space, it follows from the Bolzano-Weierstrass theorem that there exists a subsequence of $\{S_k\}$ which, say $\{S_k\}_K$, converges to an element $S \in \Pi \times A$; that is,

$$\lim_{K\to\infty} \{S_k\}_K \longrightarrow S, \quad S \in \Pi \times A. \tag{51}$$

Note that $\omega(X, C)$ is a continuous function according to (43); then,

$$\lim_{K\to\infty} \omega\left(\{S_k\}_K\right) \longrightarrow \omega(S), \quad S \in \Pi \times A. \tag{52}$$

Due to that $\{\omega(S_k)\}$ is a monotone sequence (Proposition 12) and has lower bound, then $\{\omega(S_k)\}$ is convergent. Thus, any subsequence of $\{\omega(S_k)\}$ will be convergent to the same value. Note that $\{\omega(\{S_k\}_K)\}$ is a subsequence of $\{\omega(S_k)\}$ and it is convergent to $\omega(S)$; then, it follows that

$$\lim_{k\to\infty} \omega(S_k) \longrightarrow \omega(S), \quad S \in \Pi \times A. \tag{53}$$

The second assertion can easily be proved. Let P be an accumulation point of $\{S_k\}$; then, there exists one subsequence $\{S_k\}_{K'}$ which converges to P, and due to the continuity of $\omega(X, C)$, we have

$$\omega\left(\{S_k\}_{K'}\right) \longrightarrow \omega(P). \tag{54}$$

The first assertion has shown that $\{\omega(S_k)\}$ is convergent, and note that $\{\omega(\{S_k\}_{K'})\}$ is a subsequence of $\{\omega(S_k)\}$; then, it follows that

$$\omega(P) = \lim_{K'\to\infty} \omega\left(\{S_k\}_{K'}\right) = \lim_{k\to\infty} \omega(S_k). \tag{55}$$

Thus, all accumulation points of $\{S_k\}$ have the same objective functional values equal to $\lim_{k\to\infty}\omega(S_k)$. □

Lemma 14. *$\mathscr{L} : \Pi \times A \to \Pi \times A$ which is defined in (44) is a closed map over $\Pi \times A$.*

Proof. Note that the CSLP (11) is a convex problem, then

$$\operatorname*{argmin}_{x\in\Pi_i} \left\{ \sum_{\{j\in\mathcal{N}: A_j \in \Lambda_i^{k-1}\}} s_j \min_{q_j\in A_j} \gamma\left(x_i - q_j\right) \right\}, \quad i = 1, 2, \ldots, m, \tag{56}$$

are continuous. Since $\Pi \times A$ is a compact space and also a Hausdorff space and every continuous map from a compact space to a Hausdorff space is closed, it follows that $\mathscr{L}$ is closed over $\Pi \times A$. □

Lemma 15. *Let v_0 be a given vector in $\Pi \times A$ and $\Delta := \{v \in \Pi \times A \mid \omega(v) \le \omega(v_0)\}$. Then, Δ is a compact set.*

Proof. It is known that every closed subset of a compact space is also compact, and therefore, it is enough to prove that Δ is a closed set.

For any sequence $\{v_k\}$ with $v_k \in \Delta$, since $\Pi \times A$ is compact, according to Bolzano-Weierstrass, there exists a convergent subsequence $\{v_k\}_K$ of $\{v_k\}$ such that

$$\lim_{K\to\infty} \{v_k\}_K \longrightarrow v. \tag{57}$$

Due to the continuity of ω, it follows that

$$\lim_{K\to\infty} \omega\left(\{v_k\}_K\right) = \omega(v). \tag{58}$$

On the other hand, $\{v_k\}_K \in \Delta$ implies

$$\omega\left(\{v_k\}_K\right) \le \omega(v_0). \tag{59}$$

By combining (58), (59), and the continuity of ω, the following inequality is obtained:

$$\omega(v) \le \omega(v_0), \tag{60}$$

and accordingly, $v \in \Delta$. This means that Δ is closed, and the proof is complete. □

Now, we are ready to prove the convergence of the proposed location-allocation heuristic (Algorithm 9). Let $\Xi \subseteq \Pi \times A$ be the nonempty local solution set of CMLP (9). Recall that in the location-allocation Cooper algorithm for MWP, if $X^{J+1} = X^J$ occurs after a finite number of iterations, the iterates after X^J will be constant. Then, no further improvement is possible for MWP, and it follows from [32, 34] that the X^J is a local solution of MWP. Similarly, in the proposed location-allocation heuristic algorithm for CMLP, if $X^{J+1} = X^J$ occurs, the iterates after $S_J = (X^J, C^J)$ will also be constant. Then, exactly as in the location-allocation Cooper algorithm for MWP, no further improvement is possible for CMWP, and a local solution, namely, S_J, is obtained. Hence in this case, $\{S_k\}$ is convergent to the $S_J \in \Xi$. However, it is not assured that $X^{J+1} = X^J$ always occurs during the implementation of Algorithm 9, and therefore, we assume that $X^{k+1} \neq X^k$ for any $k \in N$ and prove the convergence in this case.

Theorem 16. *Assume that $X^{k+1} \neq X^k$ for any $k \in N$; then, all the accumulation points of the sequence $\{S_k\}$ belong to Ξ.*

Proof. Let S be an accumulation point of $\{S_k\}$. Due to $S_k \in \Pi \times A$ and the compactness of $\Pi \times A$, we know that (1) $S \in \Pi \times A$ and (2) there exists a subsequence $\{S_k\}_K$ which is convergent to S. According to Theorem 13, we know that

$$\lim_{k \to \infty} \omega(S_k) = \omega(S), \quad S \in \Pi \times A. \tag{61}$$

So, it is enough to prove $S \in \Xi$. We prove this by contradiction. Assume that $S \notin \Xi$; that is, S is not a solution, and we consider the subsequence $\{S_{k+1}\}_K$. Denote $\Delta = \{v \in \Pi \times A \mid \omega(v) \leq \omega(S_0)\}$. Due to Proposition 12, it follows that for all $k \in N$ we have $(X^{k+1}, \widetilde{C}^k) \in \Delta$ and $(X^{k+1}, C^{k+1}) \in \Delta$. According to Lemma 15, Δ is compact, and thus, there exists $K' \subset K$ such that

$$\lim_{K' \to \infty} \left(X^{k+1}, \widetilde{C}^k\right)_{K'} = v_1, \qquad \lim_{K' \to \infty} \left(X^{k+1}, C^{k+1}\right)_{K'} = v_2. \tag{62}$$

According to Lemma 14, the map $\mathscr{L}$ is closed at $S \in \Pi \times A$; then, it follows that $v_1 = \mathscr{L}(S)$. Further, due to $S \notin \Xi$, v_1 will be not equal to S. Otherwise, we can choose S as the initial iterate of the location-allocation heuristic algorithm, then, the sequence generated by the algorithm will be constant. It follows from the first case (i.e., $X^{J+1} = X^J$) that S will be a local solution, which contradicts with $S \notin \Xi$. Therefore, we can obtain $v_1 \neq S$. Together with the monotone proposition of $\mathscr{L}$ (48), we have the inequality

$$\omega(v_1) < \omega(S). \tag{63}$$

On the other hand, note that $\mathscr{A}(X^{k+1}, \widetilde{C}^k) = (X^{k+1}, C^{k+1})$; then, by the monotonicity of $\mathscr{A}$ (49), it follows that

$$\omega\left(X^{k+1}, C^{k+1}\right)_{K'} \leq \omega\left(X^{k+1}, \widetilde{C}^k\right)_{K'}, \tag{64}$$

and thus, by taking the limit for (64) and by the continuity of ω, we know $\omega(v_2) \leq \omega(v_1)$. Combining this with (63), we obtain

$$\omega(v_2) < \omega(S). \tag{65}$$

However, note that v_2 and S are two accumulation points of $\{S_k\}$, and according to the second assertion of Theorem 13, $\omega(v_2) = \omega(S)$, which will contradict with (65). Therefore, our assumption is wrong, and thus, $S \in \Xi$. □

As a result, we have the following convergence theorem for the sequence generated by the proposed location-allocation algorithm.

Theorem 17. *The sequence $S(X^0, C^0)$ generated by the proposed location-allocation heuristic algorithm either converges to a point in Ξ or all accumulation points of $S(X^0, C^0)$ belong to Ξ.*

6. Numerical Results

This section reports some preliminary numerical results to verify the theoretical assertions proved in previous sections. Section 3.1, reports some numerical results of the proposed variational inequality approach for the CSLP (11) (or equivalently (12)) which includes (1) the results of the comparison between our approach and the Weiszfeld-type method by solving the example in [6] and some randomly generated unconstrained examples under Euclidean distances and (2) the results of our approach for solving some randomly generated constrained examples under a gauge. These numerical results demonstrate the efficiency of the proposed variational inequality approach for CSLP. In the second subsection, we apply the proposed location-allocation heuristic algorithm to solve some randomly generated examples of the CMLP (9). In particular, the effectiveness of the initialization strategy adopted in this heuristic for accelerating convergence will be justified. All the programming codes are written by MATLAB 2012b and were run on an ASUS notebook (Intel Core2 Duo T6670 2.20 GHz).

6.1. Numerical Results of Variational Inequality Approach for CSLP. When applying the variational inequality approach for solving CSLP and MCSLP (12), theoretically, the initial iteration u^0 in Algorithm 6 can be chosen arbitrarily in Ω. In practical implementation, however, we choose u^0 judiciously similar to the initialization strategy in the location-allocation heuristic: let g_j $(j = 1, \ldots, d)$ be the centers of regional customers, solve the following single-source Weber problem (SWP):

$$x^* = \operatorname*{argmin}_{x \in \Pi} \left\{ \sum_{j=1}^{d} \gamma\left(x - g_j\right) \right\} \tag{66}$$

by the projection-contraction method in [35], and then use its solution as the initial iterate for Algorithm 6. We call this the initialization strategy of variational inequality approach. In addition, throughout our experiments of VI approach, the α_1 and α_2 in Algorithm 6 are chosen as 1 and 0, respectively.

We first solve the example given in [16] by the proposed variational inequality approach and the Weiszfeld-type method in [16].

Example 18. Here, $d = 5$; that is, there are five regional customers, and all customers are unit squares whose sides are parallel to the axes. The geometric centers of the five customers are $(0.5, 0.5)$, $(4.5, 0.5)$, $(0.5, 2.5)$, $(2.5, 2.5)$, and $(4.5, 2.5)$, and $s_j = 1$, $j = 1, \ldots, 5$.

In order to clarify the comparison of two methods, we choose the same stopping criterion as $\|x^{k+1} - x^k\| \leq 10^{-4}$ (throughout this section, $\|\cdot\|$ is the l_∞-norm). We test this example for 100 times with the same initial iterate for the two methods which is randomly generated in $[0, 5] \times [0, 3]$, and the numerical results including the location of new facility, the closest points to the facility, number of iterations, and computing time in units of second are reported in Table 1.

Table 1: Numerical results for Example 1 given in [16].

Main results	VI approach	Weiszfeld-type method
x	(2.5000, 1.9484)	(2.5000, 1.9484)
$a_1(x)$	(1.0000, 1.0000)	(1.0000, 1.0000)
$a_2(x)$	(4.0000, 1.0000)	(4.0000, 1.0000)
$a_3(x)$	(4.0000, 2.0000)	(4.0000, 2.0000)
$a_4(x)$	(2.5000, 2.0000)	(2.5000, 2.0000)
$a_5(x)$	(1.0000, 2.0000)	(1.0000, 2.0000)
Objective value	6.6027	6.6027
Number of iterations	32.81/23.77	237.60
CPU	0.0124	0.0610

"32.81/23.77" means 32.81 Iter. for initialization and 23.77 Iter. for VI approach.

According to Table 1, it follows that both methods can get the optimal location of the new facility. Though our variational inequality approach needs smaller iteration numbers and less computing time, both methods are efficient for solving Example 1 given in [6].

Recall that the sequence generated by the Weiszfeld-type method in [16] is possible to be convergent to a nonoptimal point on the boundary of the regional customer, as indicated in [16]. The variational inequality approach, however, can obtain the optimal location of the new facility, which is guaranteed by the theoretical analysis in Section 3.2. To illustrate this, we compare two methods by solving the following particular example.

Example 19. Similar to Example 18, the number of customers $d = 5$, and all customers are unit squares whose edges are parallel to the axes. Let d_1 be the distance between any two neighboring customers, and we set d_1 equal to 1, 0.1, 0.01, and 0.001, respectively. The weights s_j $(j = 1, \ldots, 5)$ are randomly generated in the area of regional customers.

Note that in Example 19 the parameter d_1 reflects how close the customers are away from its neighborhoods. We test this example for a large number of times with the stopping criterion $\|x^{k+1} - x^k\| \leq 10^{-4}$, and the average numerical results are reported in Table 2. In this table, each row reports the average results by testing Example 19 for one hundred times. The column of "No. of Iter.$_0$" gives the average iteration times of initialization in the variational inequality approach, and the column of "No. of Iter." reports the average iteration times of both methods. The two columns of "CPU" give the average computing time in units of second for variational inequality approach (including the computing time for initialization) and Weiszfeld-type method, respectively. The columns of "Obj." give the average objective functional value obtained by the two methods, and the column of "Impro. Percent" gives the improvement percentage in objective functional values of the VI approach to Weiszfeld-type method. Remark that the convergence of VI approach to the optimal location of new facility is guaranteed, and then the column of "Freq. Num." reports the frequency among one hundred times that the Weiszfeld-type method can get the same solution as VI approach; that is, it does not converge to the nonoptimal solution on the boundary of the customer.

According to Table 2, it follows that both the VI approach and the Weiszfeld-type method are efficient for solving this particular example, and both of them need a small number of iterations and little computing time. In comparison of two methods, we can find that VI approach needs more iteration times and computing time than Weiszfeld-type method. From the column of "Impro. Percent," however, we can conclude that VI approach can obtain a better solution (in fact, the solution obtained by VI approach is the optimal location of new facility) than Weiszfeld-type method. In addition, according to the last column, we find that when d_1 decreases, that is, the customers become closer and closer, the frequency that Weiszfeld-type method obtains the optimal solution gets smaller and smaller. When d_1 is 0.001, which implies that the customers are quite close to the neighborhoods, this frequency is totally less than 20. In other words, for this particular example with $d_1 = 0.001$, the sequence generated by Weiszfeld-type method has a great possibility (more than 80%) to be convergent to a nonoptimal solution which is on the boundary of the customer. On the contrary, when d_1 is 1, which means that the customers are enough far away from one another, Weiszfeld-type method can obtain the optimal location of facility in most cases, and the frequency even exceeds 90.

Since our main effort in this paper is to solve the general location problem under gauge and locational constraint, it is necessary to apply the variational inequality approach to solve some CSLPs. In particular, we test a large number of randomly generated CSLPs with the number of customers d from 10 to 2000. In the experiments, all regional customers are assumed to be square units, and their edges are parallel to the coordinate axes. The geometric centers of all regional customers are randomly generated in $[-100, 100]^2$; the weights of the regional customer are all randomly chosen in (1,5); the locational constraint is $\|x - O\| \leq r$, where the center O is randomly generated in $[-100, 100]^2$, and the radius r is randomly generated in (1,5); the stopping criterion of VI approach is chosen as

$$\begin{gathered} \left\|x^{k+1} - x^k\right\| \leq 10^{-4}, \\ \left\|e\left(u^{k+1}\right)\right\| \leq 10^{-4}, \end{gathered} \tag{67}$$

and the initial iterate is randomly generated in $[-100, 100]^2$; the gauge $\gamma(\cdot)$ is generated with the unit ball set as

$$9\left(x + \frac{2}{3}\right)^2 + 12y^2 = 16. \tag{68}$$

For each d, we test one hundred randomly generated CSLPs, and the average numerical results are reported in Table 3. To illustrate the effect of the initialization strategy of VI approach, we also report the results of VI approach without the initialization strategy. The columns of "VI approach with Initial." and "VI approach without Initial.," respectively, report the average number of iterations and average computing time of variational inequality approach with and without the initialization strategy.

Table 2: Numerical results of VI approach and Weiszfeld-type method for Example 19.

d_1	VI approach				Weiszfeld-type method			Impro.	Freq.
	No. of Iter.$_0$	No. of Iter.	CPU	Obj.	No. of Iter.	CPU	Obj.	percent	num.
1	25.11	59.96	0.015	15.4465	52.99	0.0117	15.5664	0.78	90
	25.04	58.97	0.015	15.1803	27.28	0.0055	15.2926	0.74	92
	25.35	60.89	0.015	15.3693	31.11	0.0070	15.4834	0.74	92
	25.24	59.80	0.015	15.0924	35.45	0.0080	15.1961	0.69	90
	23.61	55.13	0.014	14.6321	34.27	0.0069	14.7055	0.50	93
0.1	32.95	278.26	0.060	4.2157	28.59	0.0082	4.6791	10.99	63
	32.11	270.81	0.059	4.0864	60.38	0.0135	4.4318	8.45	66
	34.07	301.67	0.067	4.2248	36.63	0.0081	4.6001	8.88	63
	34.16	278.24	0.065	4.1189	34.03	0.0067	4.4854	8.89	63
	32.63	280.00	0.062	4.0842	38.20	0.0069	4.3619	6.80	64
0.01	33.11	427.34	0.089	3.0633	25.85	0.0063	3.8005	24.07	28
	35.28	407.61	0.086	3.0615	39.92	0.0087	3.7669	23.04	25
	39.26	382.82	0.080	3.1259	26.53	0.0064	3.6954	18.22	29
	35.46	425.33	0.089	2.9975	35.06	0.0077	3.6500	21.77	35
	32.96	437.24	0.091	3.0691	57.90	0.0130	3.6925	20.31	33
0.001	35.44	151.52	0.035	3.0477	18.88	0.0045	3.8310	25.70	15
	36.87	140.58	0.032	3.0224	17.20	0.0048	3.7128	22.84	14
	33.43	91.72	0.022	2.9797	17.53	0.0042	3.9617	32.96	17
	33.79	115.79	0.027	2.9600	15.20	0.0044	3.7221	25.75	19
	37.19	96.08	0.022	3.1936	16.47	0.0047	4.1362	29.52	18

Table 3: Numerical results of VI approach for CSLP (12).

d	VI approach with Initial.			VI approach without Initial.	
	No. of Iter.$_0$	No. of Iter.	CPU	No. of Iter.	CPU
10	26.59	20.30	0.0108	25.26	0.0080
20	24.51	32.80	0.0239	38.20	0.0193
50	27.25	57.69	0.0847	64.89	0.0762
100	25.09	92.53	0.2664	106.84	0.2442
200	29.83	145.97	1.2196	170.66	1.3389
500	27.97	224.74	13.7783	252.99	15.9382
1000	26.71	259.03	56.8598	284.54	62.0384
2000	23.12	281.59	245.5680	346.52	295.2305

According to Table 3, it is easy to conclude that the variational inequality approach is effective for solving CSLP under gauge considering the difficulty of this problem. In addition, the number of iterations of "VI approach with Initial." is less than that of "VI approach without Initial.," which shows that the initialization strategy can accelerate the convergence of the variational inequality approach. This strategy, however, does not necessarily reduce the computing time of VI approach, especially when d is small, for example, $d = 10, 20, 50,$ and 100, which can be explained as follows. When the number of customers d is small, the variational inequality problem is small scale, and thus, it can be solved in a short time. In this case, the computational workload of initialization plays an important role in the total workload, and therefore, the computing time of "VI approach with Initial." is greater than that of "VI approach without Initial." due to the computational iterations for initialization. With d increasing, the scale of VI problem as well as the number of iterations becomes larger. Then, in the comparison of the workload of VI approach, the workload of initialization can almost be ignored, and therefore, the computing time of "VI approach with Initial." will be smaller than that of "VI approach without Initial." As a matter of fact, Table 3 reveals the computational necessity of the initialization strategy of variational inequality approach for large-scale CSLP; for example, the iteration number and computing time are reduced about 1/5 by the initialization strategy when $d = 2000$.

6.2. Numerical Results of Heuristic Algorithm for CMLP. This subsection applies the proposed location-allocation heuristic algorithm (Algorithm 9) to solve a large number of CMLPs

Table 4: Numerical results of location-allocation heuristic for CMLP.

n	m	Algorithm 9 with Initial.					Algorithm 9 without Initial.		
		Iter.$_0$	Iter.$_0$-PC	Iter.	Iter.-PC	CPU	Iter.	Iter.-PC	CPU
100	2	2.83	73.33	1.20	289.00	0.4868	3.04	176.70	0.7018
	4	3.61	99.88	1.29	176.80	0.2652	2.65	218.47	0.4992
	6	4.55	99.05	1.53	121.75	0.1710	3.50	130.96	0.2580
	8	3.60	126.23	1.82	149.37	0.2250	3.02	170.32	0.2836
	10	3.84	138.61	2.26	257.51	0.3212	4.65	240.48	0.5118
200	2	3.24	81.22	1.24	1620.80	5.4664	3.49	1192.33	11.4596
	4	4.01	112.96	1.71	447.08	1.2932	4.37	433.10	2.5694
	6	4.65	141.10	2.23	262.97	0.7958	3.88	305.75	1.4538
	8	4.26	130.25	2.65	246.80	0.6268	4.83	287.12	1.1734
	10	4.67	171.65	2.68	281.40	0.5866	4.46	346.18	1.0232
500	2	4.03	82.35	1.63	1998.20	64.5468	4.26	1815.90	139.2036
	4	4.66	104.28	2.27	1304.48	12.3492	5.25	1445.82	26.5510
	6	5.27	133.55	2.62	727.13	5.3322	7.60	875.26	18.2956
	8	7.60	169.97	2.84	549.60	2.8050	5.24	429.83	3.8626
	10	5.42	187.95	3.47	508.13	2.3742	5.02	507.70	3.3946
1000	2	4.49	56.69	1.91	718.17	93.2321	5.63	1587.31	365.6085
	4	4.81	80.60	2.10	1376.13	46.2636	5.49	1576.85	113.8830
	6	6.84	96.75	3.43	683.35	16.1398	7.65	1344.53	57.7762
	8	7.20	161.50	3.28	563.69	9.3570	6.06	760.93	20.0896
	10	6.63	171.89	4.05	615.07	7.5880	6.47	700.53	12.9448

(9) and also verifies the necessity of the initialization strategy in Algorithm 9. In the experiments, we again generate a large number of CMLPs with unit square customers and assume that the edges of these regions are parallel to the coordinate axes. The geometric centers of all the demand regions are randomly generated in $[-100 \quad 100]^2$, and all the demands, s_j $(j = 1, 2, \ldots, n)$, are randomly generated in $[\,1 \quad 10\,]$. The gauge $\gamma(\cdot)$ is also defined with the unit ball set as (68). We test the scenario with $n = 100, 200, 500,$ and 1000 and $m = 2, 4, 6, 8,$ and 10; the locational constraints are $\|x - O_j\| \leq r_j$ $(j = 1, \ldots, m)$, where the radius r_j is randomly generated in $[1, 10]$, and the center O_j is given in advance as follows:

$$\begin{aligned}
m = 2: &\quad O_1 = (-50, 0)^T, \quad O_2 = (50, 0)^T;\\
m = 4: &\quad O_1 = (-50, -50)^T, \quad O_2 = (50, -50)^T,\\
&\quad O_3 = (-50, 50)^T, \quad O_4 = (50, 50)^T;\\
m = 6: &\quad O_1 = (-50, -50)^T, \quad O_2 = (0, -50)^T,\\
&\quad O_3 = (50, -50)^T, \quad O_4 = (-50, 50)^T,\\
&\quad O_5 = (0, 50)^T, \quad O_6 = (50, 50)^T;\\
m = 8: &\quad O_1 = (-50, -50)^T, \quad O_2 = (0, -50)^T,\\
&\quad O_3 = (50, -50)^T, \quad O_4 = (-50, 50)^T,\\
&\quad O_5 = (0, 50)^T, \quad O_6 = (50, 50)^T,\\
&\quad O_7 = (-25, 0)^T, \quad O_8 = (25, 0)^T;\\
m = 10: &\quad O_1 = (-50, -50)^T, \quad O_2 = (0, -50)^T,\\
&\quad O_3 = (50, -50)^T, \quad O_4 = (-50, 50)^T,\\
&\quad O_5 = (0, 50)^T, \quad O_6 = (50, 50)^T,\\
&\quad O_7 = (-60, 0)^T, \quad O_8 = (-20, 0)^T,\\
&\quad O_9 = (20, 0)^T, \quad O_{10} = (60, 0)^T.
\end{aligned} \tag{69}$$

The initial locations of facilities are randomly generated in $[-100 \quad 100]^2$, and the stopping criterion used in Algorithms 6 and 9 is chosen as

$$\left\|x^{k+1} - x^k\right\| < 10^{-4}. \tag{70}$$

To show the significance of the initialization strategy, we compare the numerical performance of the location-allocation heuristic algorithm with initialization strategy (denoted by "Algorithm 9 with Initial.") and without this initialization strategy (denoted by "Algorithm 9 without Initial."). In the initialization step of Algorithm 9, the location-allocation algorithm in [35] is adopted to solve the corresponding CMWP (42), where a PC method is proposed to solve the subproblems in location phase, and the numbers of iterations of the algorithm in [35] (denoted by "Iter$_0$") and the average iteration numbers of PC method in one iteration of the algorithm (denoted by "Iter$_0$-PC") are reported. The columns of "Iter." and "CPU," respectively, report the number of iterations and computing time of Algorithm 9 with and without initialization strategy. Since the

efficiency of Algorithm 9 is mainly determined by the number of iterations of the variational inequality approach, we also report the average number of iterations of Algorithm 6 in one iteration of Algorithm 9 (denoted by "Iter.-PC"). For each given pair (n, m), we test the CMLP for 100 times, and the computational performance is reported in Table 4.

It follows from Table 4 that the proposed location-allocation heuristic, with or without the initialization strategy, is capable of tackling the CMLP (9) efficiently, even for large-scale cases. Also, the necessity of the initialization strategy is evident. In fact, this strategy reduces both the number of iterations and the computing time by about 50%.

Another interesting fact obtained from Table 4 deserves further illustration. Recall that with the number of new facilities (m) increasing, the number of subproblems (CSLPs) in location phase increases too. According to Table 4, however, we find that for fixed number of customers (n), with m increasing, the number of iterations for solving m CSLPs in one iteration of Algorithm 9 does not increase but almost decreases with m, especially for large-scale CMLP. This can be illustrated roughly as follows. For fixed n, when m increases, the average number of customers in each Λ_i^k becomes smaller, which implies that the scale of the involved CSLP (11) in location phase is smaller. According to Table 3, it follows that we need smaller number of iterations for small-scale CSLP, and thus, the total number of iterations for solving m CSLPs decreases. Similarly, due to the same reason, the computing time for solving CMLP also decreases with m increasing, as reported in the column of "CPU" in Table 4.

7. Conclusion

In this paper, we are interested in the locations of multiple facilities in the space R^p with regional demands, where the closest distance is used to measure the proximities between facilities and customers. With locational constraints introduced for the locations of new facilities and with the gauge used as the distance measuring function, the problem considered in this paper has much more applications in practice. Due to its nonconvexity and NP-hardness, a new location-allocation heuristic algorithm is proposed to solve this problem, and its convergence is proved under mild assumptions. Some preliminary numerical experiments are reported to verify the computational efficiency of the proposed algorithm.

Acknowledgments

Jian-lin Jiang is supported by NSFC no. 11101211, JSNSF no. BK2011719, the Fundamental Research Funds for the Central Universities no. NZ2012306, and the project sponsored by SRF for ROCS, SEM.

References

[1] A. Weber, *UBer Den Standort Der Industrien, 1. Teil: Reine Theorie Des Standortes*, Mohr Siebeck, Tübingen, Germany, 1909.

[2] E. Weiszfeld, "Sur le point pour lequel la somme des distances de n points donnes est minimum," *Tohoku Mathematical Journal*, vol. 43, pp. 355–386, 1937.

[3] R. F. Love, J. G. Morris, and G. O. Wesolowsky, *Facilities Location: Models and Methods,*, vol. 7, North-Holland Publishing, Amsterdam, Netherlands, 1988.

[4] F. Plastria, "Continuous location problems: research, results and questions," in *Facility Location: A Survey of Applications and Methods*, Z. Drezner, Ed., pp. 225–260, Springer, New York, NY, USA, 1995.

[5] C. D. Bennett and A. Mirakhor, "Optimal facility location with respect to several regions," *Journal of Regional Science*, vol. 14, no. 1, pp. 131–136, 1974.

[6] J. Brimberg and G. O. Wesolowsky, "Minisum location with closest Euclidean distances," *Annals of Operations Research*, vol. 111, pp. 151–165, 2002.

[7] Z. Drezner and G. O. Weslowsky, "Optimal location of a facility relative to area demands," *Naval Research Logistics Quarterly*, vol. 27, no. 2, pp. 199–206, 1980.

[8] J. Jiang and X. Yuan, "A Barzilai-Borwein-based heuristic algorithm for locating multiple facilities with regional demand," *Computational Optimization and Applications*, vol. 51, no. 3, pp. 1275–1295, 2012.

[9] A. Suzuki and Z. Drezner, "The p-center location problem in an area," *Location Science*, vol. 4, no. 1-2, pp. 69–82, 1996.

[10] G. O. Wesolowsky and R. F. Love, "Location of facilities with rectangular distances among point and area destinations," *Naval Research Logistics Quarterly*, vol. 18, pp. 83–90, 1971.

[11] E. Carrizosa, M. Muñoz-Márquez, and J. Puerto, "The weber problem with regional demand," *European Journal of Operational Research*, vol. 104, no. 2, pp. 358–365, 1998.

[12] R. E. Stone, "Some average distance results," *Transportation Science*, vol. 25, no. 1, pp. 83–91, 1991.

[13] J. Puerto and A. M. Rodríguez-Chía, "On the structure of the solution set for the single facility location problem with average distances," *Mathematical Programming*, vol. 128, no. 1-2, pp. 373–401, 2011.

[14] Z. Drezner and G. O. Wesolowsky, "Location models with groups of demand points," *INFOR Journal*, vol. 38, no. 4, pp. 359–372, 2000.

[15] O. Berman, Z. Drezner, and G. O. Wesolowsky, "Location of facilities on a network with groups of demand points," *IIE Transactions*, vol. 33, no. 8, pp. 637–648, 2001.

[16] J. Brimberg and G. O. Wesolowsky, "Note: facility location with closest rectangular distances," *Naval Research Logistics*, vol. 47, no. 1, pp. 77–84, 2000.

[17] S. Nickel, J. Puerto, and A. M. Rodriguez-Chia, "An approach to location models involving sets as existing facilities," *Mathematics of Operations Research*, vol. 28, no. 4, pp. 693–715, 2003.

[18] R. F. Love and J. G. Morris, "Mathematical models of road travel distances," *Management Science*, vol. 25, no. 2, pp. 130–139, 1979.

[19] J. E. Ward and R. E. Wendell, "A new norm for measuring distance which yields linear location problems," *Operations Research*, vol. 28, pp. 836–844, 1980.

[20] J. E. Ward and R. E. Wendell, "Using block norms for location modeling," *Operations Research*, vol. 33, no. 5, pp. 1074–1090, 1985.

[21] C. Witzgall, "Optimal location of a central facility, mathematical models and concepts," Report 8388, National Bureau of Standards, Washington, DC, USA, 1964.

[22] C. Michelot and O. Lefebvre, "A primal-dual algorithm for the Fermat-Weber problem involving mixed gauges," *Mathematical Programming*, vol. 39, no. 3, pp. 319–335, 1987.

[23] R. Durier, "On Pareto optima, the Fermat-Weber problem, and polyhedral gauges," *Mathematical Programming*, vol. 47, no. 1, pp. 65–79, 1990.

[24] E. Carrizosa, E. Conde, M. Munoz-Márquez, and J. Puerto, "Simpson points in planar problems with locational constraints. The polyhedral-gauge case," *Mathematics of Operations Research*, vol. 22, no. 2, pp. 291–300, 1997.

[25] S. Nickel, "Restricted center problems under polyhedral gauges," *European Journal of Operational Research*, vol. 104, no. 2, pp. 343–357, 1998.

[26] M. Cera, J. A. Mesa, F. A. Ortega, and F. Plastria, "Locating a central hunter on the plane," *Journal of Optimization Theory and Applications*, vol. 136, no. 2, pp. 155–166, 2008.

[27] F. Plastria, "Asymmetric distances, semidirected networks and majority in Fermat-Weber problems," *Annals of Operations Research*, vol. 167, pp. 121–155, 2009.

[28] I. Norman Katz and S. R. Vogl, "A Weiszfeld algorithm for the solution of an asymmetric extension of the generalized Fermat location problem," *Computers & Mathematics with Applications*, vol. 59, no. 1, pp. 399–410, 2010.

[29] H. Minkowski, *Theorie der Konvexen Körper*, Gesammelte Abhandlungen, Teubner, Berlin, 1911.

[30] Z. Drezner, *Facility Location: A Survey of Applications and Methods*, Springer, New York, NY, USA, 1995.

[31] A. Ghosh and G. Rushton, *Spatial Analysis and Location-Allocation Models*, Van Nostrand Reinhold, New York, NY, USA, 1987.

[32] L. Cooper, "Solutions of generalized location equilibrium models," *Journal of Regional Science*, vol. 7, pp. 1–18, 1967.

[33] N. Megiddo and K. J. Supowit, "On the complexity of some common geometric location problems," *SIAM Journal on Computing*, vol. 13, no. 1, pp. 182–196, 1984.

[34] L. Cooper, "Heuristic methods for location-allocation problems," *SIAM Review*, vol. 6, pp. 37–53, 1964.

[35] J.-L. Jiang and X.-M. Yuan, "A heuristic algorithm for constrained multi-source Weber problem: the variational inequality approach," *European Journal of Operational Research*, vol. 187, no. 2, pp. 357–370, 2008.

[36] J.-l. Jiang, K. Cheng, C.-C. Wang, and L.-p. Wang, "Accelerating the convergence in the single-source and multi-source Weber problems," *Applied Mathematics and Computation*, vol. 218, no. 12, pp. 6814–6824, 2012.

[37] Y. Levin and A. Ben-Israel, "A heuristic method for large-scale multi-facility location problems," *Computers & Operations Research*, vol. 31, no. 2, pp. 257–272, 2004.

[38] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, Mass, USA, 1990.

[39] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, NY, USA, 2004.

[40] M. C. Ferris and C. Kanzow, "Engineering and economic applications of complementarity problems," *SIAM Review*, vol. 39, no. 4, pp. 669–713, 1997.

[41] K. Addi, B. Brogliato, and D. Goeleven, "A qualitative mathematical analysis of a class of linear variational inequalities via semi-complementarity problems: applications in electronics," *Mathematical Programming*, vol. 126, no. 1, pp. 31–67, 2011.

[42] 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.

[43] A. Barbagallo and P. Mauro, "Time-dependent variational inequality for an oligopolistic market equilibrium problem with production and demand excesses," *Abstract and Applied Analysis*, vol. 2012, Article ID 651975, 35 pages, 2012.

[44] J. Gwinner, "Three-field modelling of nonlinear nonsmooth boundary value problems and stability of differential mixed variational inequalities," *Abstract and Applied Analysis*, vol. 2013, Article ID 108043, 10 pages, 2013.

[45] Y.-B. Zhao and J.-Y. Yuan, "An alternative theorem for generalized variational inequalities and solvability of nonlinear quasi-$P_x^M 2a$;-complementarity problems," *Applied Mathematics and Computation*, vol. 109, no. 2-3, pp. 167–182, 2000.

[46] H. Uzawa, "Iterative methods for concave programming," in *Studies in Linear and Nonlinear Programming*, K. J. Arrow, L. Hurwicz, and H. Uzawa, Eds., pp. 154–165, Stanford University Press, Stanford, Calif, USA, 1958.

[47] B. S. He, "A new method for a class of linear variational inequalities," *Mathematical Programming*, vol. 66, no. 2, pp. 137–144, 1994.

[48] B. S. He, "A modified projection and contraction method for a class of linear complementarity problems," *Journal of Computational Mathematics*, vol. 14, no. 1, pp. 54–63, 1996.

[49] N. Xiu, C. Wang, and J. Zhang, "Convergence properties of projection and contraction methods for variational inequality problems," *Applied Mathematics and Optimization*, vol. 43, no. 2, pp. 147–168, 2001.

[50] B. C. Eaves, "On the basic theorem of complementarity," *Mathematical Programming*, vol. 1, no. 1, pp. 68–75, 1971.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 183961, 7 pages
http://dx.doi.org/10.1155/2013/183961

Research Article

On the Convergence Analysis of the Alternating Direction Method of Multipliers with Three Blocks

Caihua Chen,[1] Yuan Shen,[2] and Yanfei You[3]

[1] *International Center of Management Science and Engineering, School of Management and Engineering, Nanjing University, Nanjing 210093, China*
[2] *School of Applied Mathematics, Nanjing University of Finance & Economics, Nanjing 210023, China*
[3] *Department of Mathematics, Nanjing University, Nanjing 210093, China*

Correspondence should be addressed to Caihua Chen; cchenhuayx@gmail.com

Received 4 July 2013; Accepted 5 September 2013

Academic Editor: Xu Minghua

We consider a class of linearly constrained separable convex programming problems whose objective functions are the sum of three convex functions without coupled variables. For those problems, Han and Yuan (2012) have shown that the sequence generated by the alternating direction method of multipliers (ADMM) with three blocks converges globally to their KKT points under some technical conditions. In this paper, a new proof of this result is found under new conditions which are much weaker than Han and Yuan's assumptions. Moreover, in order to accelerate the ADMM with three blocks, we also propose a relaxed ADMM involving an additional computation of optimal step size and establish its global convergence under mild conditions.

1. Introduction

In various fields of applied mathematics and engineering, many problems can be equivalently formulated as a separable convex optimization problem with two blocks; that is, given two closed convex functions $f_i : \Re^{n_i} \to \Re \cup \{+\infty\}$, $i = 1, 2$, to find a solution pair (x_1^*, x_2^*) of the following problem:

$$\begin{aligned} \min \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & A_1 x_1 + A_2 x_2 = b, \end{aligned} \tag{1}$$

where A_i is a matrix in $\Re^{p\times n_i}$, $i = 1, 2$, and b is a vector in $\Re^p$. The classical alternating direction method of multipliers (ADMM) [1, 2] applied to problem (1) yields the following scheme:

$$\begin{aligned} x_1^{k+1} &= \arg\min_{x_1\in\Re^{n_1}} f_1(x_1) - \left\langle A_1^T\lambda^k, x_1\right\rangle \\ &\quad + \frac{\beta}{2}\left\|A_1x_1 + A_2x_2^k - b\right\|^2, \\ x_2^{k+1} &= \arg\min_{x_2\in\Re^{n_2}} f_2(x_2) - \left\langle A_2^T\lambda^k, x_2\right\rangle \\ &\quad + \frac{\beta}{2}\left\|A_1x_1^{k+1} + A_2x_2 - b\right\|^2, \\ \lambda^{k+1} &= \lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right), \end{aligned} \tag{2}$$

where λ^k is a Lagrangian multiplier and $\beta > 0$ is a penalty parameter. Possibly due to its simplicity and effectiveness, the ADMM with two blocks has received continuous attention both in theoretical and application domains. We refer to [3–8] for theoretical results on ADMM with two blocks and [9–13] for its efficient applications in high-dimensional statistics, compressive sensing, finance, image processing, and engineering, to name just a few.

In this paper, we concentrate on the linearly constrained convex programming problem with three blocks:

$$\begin{aligned} \min \quad & f_1(x_1) + f_2(x_2) + f_3(x_3) \\ \text{s.t.} \quad & A_1x_1 + A_2x_2 + A_3x_3 = b, \end{aligned} \tag{3}$$

where $f_3 : \Re^{n_3} \to \Re \cup \{+\infty\}$ is a closed convex function and A_3 is a matrix in $\Re^{p\times n_3}$. For solving (3), a nature idea is to extend the ADMM with two blocks to the ADMM with three blocks in which the next iteration $(x_2^{k+1}, x_3^{k+1}, \lambda^{k+1})$ is updated by

$$\left(x_2^{k+1}, x_3^{k+1}, \lambda^{k+1}\right) := \left(\tilde{x}_2^k, \tilde{x}_3^k, \tilde{\lambda}^k\right), \tag{4}$$

where

$$\begin{aligned}
\tilde{x}_1^k &= \arg\min_{x_1\in\Re^{n_1}} f_1(x_1) - \langle A_1^T\lambda^k, x_1\rangle \\
&\quad + \frac{\beta}{2}\left\|A_1x_1 + A_2x_2^k + A_3x_3^k - b\right\|^2, \\
\tilde{x}_2^k &= \arg\min_{x_2\in\Re^{n_2}} f_2(x_2) - \left\langle A_2^T\lambda^k, x_2\right\rangle \\
&\quad + \frac{\beta}{2}\left\|A_1\tilde{x}_1^k + A_2x_2 + A_3x^k - b\right\|^2, \\
\tilde{x}_3^k &= \arg\min_{x_3\in\Re^{n_3}} f_3(x_3) - \left\langle A_3^T\lambda^k, x_3\right\rangle \\
&\quad + \frac{\beta}{2}\left\|A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3x_3 - b\right\|^2, \\
\tilde{\lambda}^k &= \lambda^k - \beta\left(A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3\tilde{x}_3^k - b\right).
\end{aligned} \tag{5}$$

Similar to the ADMM with two blocks, the ADMM with three blocks has found numerous applications in a broad spectrum of areas, such as doubly nonnegative cone programming [14], high-dimensional statistics [15, 16], imaging science [17], and engineering [18]. Even though its numerical efficiency is clearly seen from those applications, the theoretical treatment of ADMM with three blocks is challenging and the convergence of the ADMM is still open given only the convex assumptions of the objective function. To alleviate this difficulty, the authors of [19, 20] proposed prediction-correction type methods to solve the general separable convex programming; however, numerical results show that the direct ADMM outperforms its variants substantially. Therefore, it is of great significance to investigate the theoretical performance of the ADMM with three blocks even only to provide sufficient conditions to guarantee the convergence. To the best of our knowledge, there exist only two works aiming to attack the convergence problem of the direct ADMM with three blocks. By using an error bound analysis method, Hong and Luo [21] proved the linear convergence of the ADMM with m blocks for sufficiently small β subject to some technical conditions. However, the sufficiently small requirement on β makes the algorithm difficult to implement. In [22], Han and Yuan employed a contractive analysis method to establish the convergence of ADMM under the strongly convex assumptions of f_i and the parameter β less than a threshold depending on all the strongly convex moduli. In this paper, we firstly prove the convergence of ADMM with three blocks under two conditions weaker than those of [22]. In our conditions, the threshold on the parameter β only relies on the strongly convex moduli of f_2 and f_3, and furthermore f_1 is not necessarily strongly convex in one of our conditions. Also, the restricted range of β in this paper is shown to be at least three times as big as that of [22].

In order to accelerate the ADMM with three blocks, we also propose a relaxed ADMM with three blocks which involves an additional computation of optimal step size. Specifically, with the triple $(x_2^k, x_3^k, \lambda^k)$, we first generate a predictor $(\tilde{x}_2^k, \tilde{x}_3^k, \tilde{\lambda}^k)$ according to (5) and then obtain $(x_2^{k+1}, x_3^{k+1}, \lambda^{k+1})$ in the next iteration by

$$\begin{aligned}
x_2^{k+1} &= x_2^k - \gamma\alpha_k^*\left(x_2^k - \tilde{x}_2^k\right), \\
x_3^{k+1} &= x_3^k - \gamma\alpha_k^*\left(x_3^k - \tilde{x}_3^k\right), \\
\lambda^{k+1} &= \lambda^k - \gamma\alpha_k^*\left(\lambda^k - \tilde{\lambda}^k\right),
\end{aligned} \tag{6}$$

where $\gamma \in (0, 2)$ and α_k^* is special step size defined in (43). The convergence of the relaxed ADMM is also established under mild conditions. We should mention that it is possible to modify the analyses given in this paper to be problems with more than three blocks of separability. But this is not the focus of this paper.

The remaining parts of this paper are organized as follows. In Section 2, we list some preliminaries on the strongly convex function, subdifferential, and the ADMM with three blocks. In Section 3, we first show the contractive property of the distance between the sequence generated by ADMM with three blocks and the solution set and then prove the convergence of ADMM under certain conditions. In Section 4, we extend the direct ADMM with three blocks to the relaxed ADMM with an optimal step size and establish its convergence under suitable conditions. We conclude our paper in Section 5.

Notation. For any positive integer m, let I_m be the $m \times m$ identity matrix. We use $\|\cdot\|$ and $\|\cdot\|_2$ to denote the vector Euclidean norm and the spectral norm of matrices (defined as the maximum singular value of matrices). For any symmetric matrix $S \in \Re^{n\times n}$, we write $\|x\|_S^2 = x^TSx$ for any $x \in \Re^n$. G and M are two $(n_2 + n_3 + p) \times (n_2 + n_3 + p)$ matrices defined by

$$\begin{aligned}
G &:= \begin{pmatrix} \beta A_2^TA_2 & 0 & 0 \\ 0 & \beta A_3^TA_3 & 0 \\ 0 & 0 & \dfrac{I}{\beta} \end{pmatrix}, \\
M &:= \begin{pmatrix} 2\beta A_2^TA_2 & 0 & 0 \\ 0 & \beta A_3^TA_3 & 0 \\ 0 & 0 & \dfrac{I}{\beta} \end{pmatrix},
\end{aligned} \tag{7}$$

respectively. For given $x_1 \in \Re^{n_1}$, $x_2 \in \Re^{n_2}$, $x_3 \in \Re^{n_3}$, and $\lambda \in \Re^p$, we frequently use u and v to denote the joint vectors of x_2, x_3, λ and x_1, x_2, x_3, λ, respectively; that is,

$$u = \left[x_2^T, x_3^T, \lambda^T\right]^T, \qquad v = \left[x_1^T, x_2^T, x_3^T, \lambda^T\right]^T, \tag{8}$$

while $\tilde{u}$ and $\tilde{v}$ are the joint vectors corresponding to $\tilde{x}_2$, $\tilde{x}_3$, $\tilde{\lambda}$ and $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$, $\tilde{\lambda}$.

2. Preliminaries

Throughout this paper, we assume f_i, $i = 1, 2, 3$, are strongly convex functions with modulus $\mu_i \geq 0$; that is

$$\begin{aligned} & f_i\left((1-\alpha) z + \alpha z'\right) \\ & \quad \leq (1-\alpha) f_i(z) + \alpha f_i\left(z'\right) \\ & \quad - \frac{1}{2}\mu_i \alpha (1-\alpha) \left\|z - z'\right\|^2, \quad \forall z,\ z' \in \Re^{n_i}, \end{aligned} \tag{9}$$

for each i. Note that f_i is a strongly convex function with modulus 0 being equivalent to the convexity of f_i. Let x be a point of $\operatorname{dom}(f_i)$; the subdifferential of f_i at x is defined by

$$\partial f_i(x) := \left\{x^* \mid f(z) \geq f(x) + \left\langle x^*, z - x\right\rangle, \forall z\right\}. \tag{10}$$

From Proposition 6 in [23], we know that, for each i, ∂f_i is strongly monotone with modulus μ_i which means

$$\begin{aligned} & \left\langle z_1 - z_2, x_1 - x_2\right\rangle \geq \mu_i \left\|z_1 - z_2\right\|^2 \geq 0, \\ & \forall x_1, x_2, z_1 \in \partial f_i(x_1),\ z_2 \in \partial f_i(x_2). \end{aligned} \tag{11}$$

The next lemma introduced in [22] plays a key role in the convergence analysis of the ADMM and the relaxed ADMM with three blocks.

Lemma 1. *Let $(x_1^*, x_2^*, x_3^*, \lambda^*)$ be any KKT point of problem (3). Let $\tilde{v}^k$ be generated by (5) from given u^k. Then, one has*

$$\begin{aligned} & \left\langle \tilde{u}^k - u^*, G\left(u^k - \tilde{u}^k\right)\right\rangle \\ & \quad \geq \sum_{i=1}^{3} \mu_i \left\|\tilde{x}_i^k - x_i^*\right\|^2 + \left\langle \lambda^k - \tilde{\lambda}^k, \sum_{i=2}^{3} A_i\left(x_i^k - \tilde{x}_i^k\right)\right\rangle \\ & \quad + \beta \left\langle A_3\left(\tilde{x}_3^k - x_3^*\right), A_2\left(\tilde{x}_2^k - x_2^k\right)\right\rangle. \end{aligned} \tag{12}$$

3. The ADMM with Three Blocks

In this section, we first investigate the contractive property of the distance between the sequence generated by ADMM with three blocks and the solution set under the condition that $0 < \beta \leq \min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$.

Lemma 2. *Let $v^* = (x_1^*, x_2^*, x_3^*, \lambda^*)$ be a KKT point of problem (3) and let the sequence $\{v^k = (x_1^k, x_2^k, x_3^k, \lambda^k)\}$ be generated by the ADMM with three blocks. Then, it holds that*

$$\begin{aligned} \left\|u^{k+1} - u^*\right\|_M^2 \leq & \left\|u^k - u^*\right\|_M^2 - \beta\left\|A_3(x_3^{k+1} - x_3^k)\right\|^2 \\ & - \beta\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2 \\ & - 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 \\ & - 2\left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - \beta A_2^T A_2}^2 \\ & - 2\left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2. \end{aligned} \tag{13}$$

Proof. Since x_3^j minimizes $f_3(\cdot) - \langle A_3^T \lambda^j, \cdot\rangle$, we deduce from the first order optimality condition that

$$A_3^T \lambda^j \in \partial f_3\left(x_3^j\right), \quad j = 0, 1, \ldots, k. \tag{14}$$

By (14) and the monotonicity of $\partial f_3(\cdot)$ (11), it is easily seen that

$$\left\langle x_3^k - x_3^{k+1}, A_3^T \lambda^k - A_3^T \lambda^{k+1}\right\rangle \geq 0. \tag{15}$$

Then for each k,

$$\begin{aligned} & \left\langle u^{k+1} - u^*, G\left(u^k - u^{k+1}\right)\right\rangle \\ & \quad \geq \sum_{i=1}^{3} \mu_i \left\|x_i^{k+1} - x_i^*\right\|^2 + \left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_i^k - x_2^{k+1}\right)\right\rangle \\ & \quad + \beta\left\langle A_3\left(x_3^{k+1} - x_3^*\right), A_2\left(x_2^{k+1} - x_2^k\right)\right\rangle \\ & \quad \geq \sum_{i=1}^{2} \mu_i \left\|x_i^{k+1} - x_i^*\right\|^2 + \left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 \\ & \quad + \left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle \\ & \quad - \frac{\beta}{4}\left\|A_2\left(x_2^{k+1} - x_2^k\right)\right\|^2, \end{aligned} \tag{16}$$

where the last "$\geq$" follows from the elementary inequality

$$\langle x, y\rangle \geq -\|x\|^2 - \frac{1}{4}\|y\|^2. \tag{17}$$

Since

$$\begin{aligned} \left\|A_3\left(x_3^{k+1} - x_3^k\right)\right\|^2 \leq & 2\left\|A_3(x_3^{k+1} - x_3^*)\right\|^2 \\ & + 2\left\|A_3\left(x_3^k - x_3^*\right)\right\|^2, \end{aligned} \tag{18}$$

by direct computations, we further obtain that

$$\begin{aligned} \left\|u^k - u^*\right\|_G^2 \geq & \left\|u^{k+1} - u^*\right\|_G^2 \\ & + \left\|u^{k+1} - u^k\right\|_G^2 + 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 \\ & + 2\left\|x_2^{k+1} - x_2^*\right\|_{\mu_2 I_{n_2} - (\beta/2) A_2^T A_2}^2 \\ & + 2\left\|x_3^{k+1} - x_3^*\right\|_{\mu_3 I_{n_3} - \beta A_3^T A_3}^2 \\ & + 2\left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle \\ & - \beta\left\|A_2\left(x_2^k - x_2^*\right)\right\|^2, \end{aligned} \tag{19}$$

which, together with $G = M - \begin{pmatrix} \beta A_2^T A_2 & & \\ & 0 & \\ & & 0 \end{pmatrix}$, implies

$$\begin{aligned}\left\|u^k - u^*\right\|_M^2 \geq{}& \left\|u^{k+1} - u^k\right\|_G^2 \\ &+ \left\|u^{k+1} - u^*\right\|_M^2 + 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 \\ &+ 2\left\|x_2^{k+1} - x_2^*\right\|^2_{\mu_2 I_{n_2} - \beta A_2^T A_2} \\ &+ 2\left\|x_3^{k+1} - x_3^*\right\|^2_{\mu_3 I_{n_3} - \beta A_3^T A_3} \\ &+ 2\left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle.\end{aligned} \tag{20}$$

Note that

$$\begin{aligned}&\left\|x_2^k - x_2^{k+1}\right\|^2_{\beta A_2^T A_2} + 2\left\langle \lambda^k - \lambda^{k+1}, A_2\left(x_2^k - x_2^{k+1}\right)\right\rangle \\ &\quad + \frac{1}{\beta}\left\|\lambda^k - \lambda^{k+1}\right\|^2 \\ &\quad = \beta\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2.\end{aligned} \tag{21}$$

We complete the proof of this lemma. □

With the above preparation, we are ready to prove the convergence of the ADMM with three blocks for solving (3) given the following conditions.

Theorem 3. *Let $\{v^k = (x_1^k, x_2^k, x_3^k, \lambda^k)\}$ be the sequence generated by the ADMM with three blocks. Then $\{v^k\}$ converges to a KKT point of problem (3) if either of the following conditions holds:*

(i) *$\mu_1 > 0$ and $0 < \beta \leq \min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$;*

(ii) *A_1 is of full column rank, $0 < \beta < \mu_2\|A_2\|_2^2$, and $\beta \leq \mu_3\|A_3\|_2^2$.*

Proof. By the inequality (13), it follows that the sequence $\{A_2 x_2^k, A_3 x_3^k, \lambda^k\}$ is bounded. Recall that

$$A_1 x_1^{k+1} = \frac{\lambda^k - \lambda^{k+1}}{\beta} - A_2 x_2^{k+1} - A_3 x_3^{k+1} + b. \tag{22}$$

Hence $\{A_1 x_1^k\}$ is also bounded. Moreover, from (13) we see immediately that

$$\begin{aligned}+\infty >{}& \sum_{k=1}^{\infty}\beta\left\|A_3(x_3^{k+1} - x_3^k)\right\|^2 \\ &+ \beta\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2 \\ &+ \sum_{k=1}^{\infty} 2\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 \\ &+ 2\left\|x_2^{k+1} - x_2^*\right\|^2_{\mu_2 I_{n_2} - \beta A_2^T A_2} \\ &+ 2\left\|x_3^{k+1} - x_3^*\right\|^2_{\mu_3 I_{n_3} - \beta A_3^T A_3}.\end{aligned} \tag{23}$$

According to the condition that $0 < \beta \leq \min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$, we know

$$\begin{aligned}&\sum_{k=1}^{\infty}\| A_3\left(x_3^{k+1} - x_3^k\right)\|^2 < \infty, \\ &\sum_{k=1}^{\infty}\| A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\|^2 < +\infty, \\ &\sum_{k=1}^{\infty}\mu_1\| x_1^{k+1} - x_1^*\|^2 < +\infty, \\ &\sum_{k=1}^{\infty}\left\|x_2^{k+1} - x_2^*\right\|^2_{\mu_2 I_{n_2} - \beta A_2^T A_2} < +\infty, \\ &\sum_{k=1}^{\infty}\left\|x_3^{k+1} - x_3^*\right\|^2_{\mu_3 I_{n_3} - \beta A_3^T A_3} < +\infty.\end{aligned} \tag{24}$$

It therefore holds that

$$\begin{aligned}&\lim_{k\to\infty}\left\|A_3\left(x_3^{k+1} - x_3^k\right)\right\|^2 = 0, \\ &\lim_{k\to\infty}\left\|A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^{k+1} - b\right\|^2 = 0,\end{aligned} \tag{25}$$

$$\begin{aligned}&\lim_{k\to\infty}\mu_1\left\|x_1^{k+1} - x_1^*\right\|^2 = 0, \\ &\lim_{k\to\infty}\left\|x_2^{k+1} - x_2^*\right\|^2_{\mu_2 I_{n_2} - \beta A_2^T A_2} = 0, \\ &\lim_{k\to\infty}\left\|x_3^{k+1} - x_3^*\right\|^2_{\mu_3 I_{n_3} - \beta A_3^T A_3} = 0.\end{aligned} \tag{26}$$

Therefore, the sequence $\{\mu_1\|x_1^k\|^2, \|x_2^k\|^2_{\mu_2 I_{n_2} - \beta A_2^T A_2}, \|x_3^k\|^2_{\mu_3 I_{n_3} - \beta A_2^T A_2}\}$ is bounded, which, together with the boundedness of $\{A_1 x_1^k, A_2 x_2^k, A_3 x_3^k, \lambda^k\}$, implies that $\{x_2^k, x_3^k, \lambda^k\}$ is bounded, and $\{x_1^k\}$ is bounded given the condition $\mu_1 > 0$ or A_1 is of full column rank. Moreover, since

$$\begin{aligned}\left\|x_3^{k+1} - x_3^k\right\|^2 ={}& \left\|A_3 x_3^{k+1} - A_3 x_3^k\right\|^2 \\ &+ \left\|x_3^{k+1} - x_3^k\right\|^2_{\mu_3 I_{n_3} - A_3^T A_3},\end{aligned} \tag{27}$$

by the first equality in (25) and the third equality in (26), it holds that

$$\lim_{k\to\infty}\left\|x_3^{k+1} - x_3^k\right\| = 0. \tag{28}$$

We proceed to prove the convergence of ADMM by considering the following two cases.

Case 1 ($\mu_1 > 0$ and $\beta \le \min(\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2)$). In this case, the sequence $\{x_1^k\}$ converges to x_1^* and then

$$\lim_{k\to\infty}\left\|A_2x_2^{k+1} - A_2x_2^k\right\| = 0, \qquad \lim_{k\to\infty}\left\|\lambda^{k+1} - \lambda^k\right\| = 0. \tag{29}$$

By the second equality in (26), we deduce from (29) that

$$\lim_{k\to\infty}\left\|x_2^{k+1} - x_2^k\right\| = 0. \tag{30}$$

Since $\{x_2^k, x_3^k, \lambda^k\}$ is bounded, there exist a triple $(x_2^\infty, x_3^\infty, \lambda^\infty)$ and a subsequence $\{n_k\}$ such that

$$\lim_{k\to\infty} x_2^{n_k} = x_2^\infty, \qquad \lim_{k\to\infty} x_2^{n_k} = x_2^\infty, \qquad \lim_{k\to\infty} \lambda^{n_k} = \lambda^\infty, \tag{31}$$

which by combining (25), (29) with given conditions, implies

$$\lim_{k\to\infty} x_2^{n_k+1} = x_2^\infty, \qquad \lim_{k\to\infty} x_2^{n_k+1} = x_2^\infty, \qquad \lim_{k\to\infty} \lambda^{n_k+1} = \lambda^\infty. \tag{32}$$

Note that

$$\begin{aligned} 0 &\in \partial f_1\left(x_1^{k+1}\right) - A_1^T\lambda^{k+1} + A_1^TA_2\left(x_2^k - x_2^{k+1}\right) \\ &\quad + A_1^TA_3\left(x_3^k - x_3^{k+1}\right), \\ 0 &\in \partial f_2\left(x_2^{k+1}\right) - A_2^T\lambda^{k+1} + A_2^TA_3\left(x_3^k - x_3^{k+1}\right), \\ 0 &\in \partial f_3\left(x_3^{k+1}\right) - A_3^T\lambda^{k+1}, \\ \lambda^{k+1} &= \lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^{k+1}\right). \end{aligned} \tag{33}$$

Then, by taking the limits on both sides of (33), using (25) and (29), and invoking the upper semicontinuous of $\partial f_1(\cdot)$, $\partial f_2(\cdot)$, and $\partial f_3(\cdot)$ [24], one can immediately write

$$\begin{aligned} 0 &\in \partial f_1\left(x^*\right) - A_1^T\lambda^\infty, \\ 0 &\in \partial f_2\left(x_2^\infty\right) - A_2^T\lambda^\infty, \\ 0 &\in \partial f_3\left(x_3^\infty\right) - A_3^T\lambda^\infty, \\ A_1x^* &+ A_2x_2^\infty + A_3x_3^\infty = b, \end{aligned} \tag{34}$$

which indicates $(x_1^*, x_2^\infty, x_3^\infty, \lambda^\infty)$ is a KKT point of problem (3). Hence, the inequality (13) is also valid if $(x_1^*, x_2^*, x_3^*, \lambda^*)$ is replaced by $(x_1^*, x_2^\infty, x_3^\infty, \lambda^\infty)$. Then it holds that

$$\begin{aligned} &2\beta\left\|A_2x_2^{k+1} - A_2x_2^\infty\right\|^2 + \beta\left\|A_3x_3^{k+1} - A_3x_3^\infty\right\|^2 \\ &\quad + \frac{1}{\beta}\left\|\lambda^{k+1} - \lambda^\infty\right\|^2 \le 2\beta\left\|A_2x_2^k - A_2x_2^\infty\right\|^2 \\ &\quad + \beta\left\|A_3x_3^k - A_3x_3^\infty\right\|^2 + \frac{1}{\beta}\left\|\lambda^k - \lambda^\infty\right\|^2, \end{aligned} \tag{35}$$

which yields

$$\begin{aligned} \lim_{k\to\infty}\left\|x_2^k - x_2^\infty\right\|^2_{A_2^TA_2} &= 0, \\ \lim_{k\to\infty}\left\|x_3^k - x_3^\infty\right\|^2_{A_3^TA_3} &= 0, \end{aligned} \tag{36}$$

$$\lim_{k\to\infty}\lambda^k = \lambda^\infty. \tag{37}$$

By adding the last two equalities in (26) to (36), we know

$$\lim_{k\to\infty} x_2^k = x_2^\infty, \qquad \lim_{k\to\infty} x_3^k = x_3^\infty. \tag{38}$$

Therefore, we have shown that the whole sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ converges to $(x_1^*, x_2^\infty, x_3^\infty, \lambda^\infty)$ under condition (i) in Theorem 3.

Case 2 (A_1 is of full column rank, $0 < \beta < \mu_2/\|A_2\|_2^2$, and $\beta \le \mu_3/\|A_3\|_2^2$). In this case, the sequence $\{x_2^k\}$ converges to x_2^* and $\{x_1^k\}$ is bounded. From the second equality in (25) and (28), we have

$$\begin{aligned} \lim_{k\to\infty}\|A_1x_1^{k+1} - A_1x_1^k\| &= 0, \\ \lim_{k\to\infty}\|\lambda^k - \lambda^{k+1}\| &= 0. \end{aligned} \tag{39}$$

Since A_1 is of full column rank, it therefore holds that

$$\lim_{k\to\infty}\|x_1^{k+1} - x_1^k\| = 0. \tag{40}$$

Let $(x_1^\infty, x_3^\infty, \lambda^\infty)$ be a cluster point of the sequence $\{x_1^k, x_3^k, \lambda^k\}$. Following a similar proof in Case 1, we are able to show $(x_1^\infty, x_2^*, x_3^\infty, \lambda^\infty)$ is a KKT point of problem (3) and the whole sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ converges to this point. □

Remark 4 (see [22]). the authors proved the convergence of the ADMM under the conditions that f_1, f_2, and f_3 are strongly convex and $0 < \beta < \min_{1\le i\le 3}\{\mu_i/3\|A_i\|_2^2\}$. Our result improves the upper bound $\min_{1\le i\le 3}\{\mu_i/3\|A_i\|_2^2\}$ by $\min\{\mu_2/\|A_2\|_2^2, \mu_3/\|A_3\|_2^2\}$. Moreover, in our condition (ii), the strongly convexity assumption is only imposed on f_2 and f_3 while f_1 is not necessarily strongly convex with positive modulus.

4. The Relaxed ADMM with Three Blocks

For the ADMM with two blocks, Ye and Yuan [25] developed a variant of alternating direction method with an optimal step size. Numerical results demonstrated that an additional computation on the optimal step size would improve the efficiency of the new variant of ADMM. In this section, by adopting the essential idea of Ye and Yuan [25], we propose

a relaxed ADMM with three blocks to accelerate the ADMM via an optimal step size. For notational simplicity, we write

$$\begin{aligned}\Phi\left(u^k,\tilde{u}^k\right) := {} & \frac{3\beta}{4}\left\|A_2(x_2^k-\tilde{x}_2^k)\right\|^2 \\ & + \beta\left\|A_3(x_3-\tilde{x}_3^k)\right\|^2 + \frac{1}{\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2 \\ & + \left\langle \lambda^k-\tilde{\lambda}^k, A_2\left(x_2^k-\tilde{x}_2^k\right)+A_3\left(x_3^k-\tilde{x}_3^k\right)\right\rangle .\end{aligned} \tag{41}$$

With $u^k=(x_2^k,x_3^k,\lambda^k)$, the new iterate of extended ADMM is produced by

$$u^{k+1}=u^k-\gamma\alpha^*\left(u^k-\tilde{u}^k\right), \quad \gamma\in(0,2), \tag{42}$$

where $\tilde{u}^k$ is the solution of (5) and α^* is defined by

$$\alpha^* := \frac{\Phi\left(u^k,\tilde{u}^k\right)}{\|u^k-\tilde{u}^k\|_G^2}. \tag{43}$$

Lemma 5. *Let the sequence $\{u^k\}$ be generated by the relaxed ADMM with three blocks. Then, if $0<\beta\le\mu_3/\|A_3\|_2^2$, the following statements are valid:*

(i) $\Phi(u^k,\tilde{u}^k)\ge(1/6)\|u^k-u^{k+1}\|_G^2$ *and thus* $\alpha^*\ge 1/6$;

(ii) $\|u^{k+1}-u^*\|_G^2 \le \|u^k-u^*\|_G^2-(1/36)\gamma(2-\gamma)\|u^k-\tilde{u}^k\|_G^2-(1/3)\gamma\mu_1\|\tilde{x}_1^k-x_1^*\|^2-(1/3)\gamma\mu_2\|\tilde{x}_2^k-x_2^*\|^2-(1/3)\gamma\|\tilde{x}_3^k-x_3^*\|^2_{\mu_3 I-\beta A_3^TA_3}$.

Proof. By direct computations to $\Phi(u^k,\tilde{u}^k)$, we know that

$$\begin{aligned}\Phi\left(u^k,\tilde{u}^k\right) & = \frac{3\beta}{4}\left\|A_2\left(x_2^k-\tilde{x}_2^k\right)\right\|^2+\beta\left\|A_3(x_3-\tilde{x}_3^k)\right\|^2 \\ & \quad + \frac{1}{\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2 \\ & \quad + \left\langle \lambda^k-\tilde{\lambda}^k, A_2\left(x_2^k-\tilde{x}_2^k\right)+A_3\left(x_3^k-\tilde{x}_3^k\right)\right\rangle \\ & \ge \frac{3\beta}{4}\left\|A_2(x_2^k-\tilde{x}_2^k)\right\|^2+\beta\left\|A_3(x_3-\tilde{x}_3^k)\right\|^2 \\ & \quad + \frac{1}{\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2-\frac{\beta}{2}\left\|A_2\left(x_2^k-\tilde{x}_2^k\right)\right\|^2 \\ & \quad - \frac{1}{2\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2-\frac{3\beta}{4}\left\|A_3(x_3^k-\tilde{x}_3^k)\right\|^2 \\ & \quad - \frac{1}{3\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2=\frac{\beta}{4}\left\|A_2(x_2^k-\tilde{x}_2^k)\right\|^2 \\ & \quad + \frac{\beta}{4}\left\|A_3(x_3^k-\tilde{x}_3^k)\right\|^2+\frac{1}{6\beta}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2,\end{aligned} \tag{44}$$

where the second inequality follows Cauchy inequality. It therefore holds that

$$\Phi\left(u^k,\tilde{u}^k\right)\ge\frac{1}{6}\left\|u^k-\tilde{u}^k\right\|_G^2, \tag{45}$$

which completes the proof of the first part. By Lemma 1 and the elementary inequality (17), it can be easily verified that

$$\begin{aligned}&\left\langle u^k-u^*, G\left(u^k-\tilde{u}^k\right)\right\rangle \\ &\quad \ge \Phi\left(u^k,\tilde{u}^k\right)+\mu_1\left\|\tilde{x}_1^k-x_1^*\right\|^2+\mu_2\left\|\tilde{x}_2^k-x_2^*\right\|^2 \\ &\qquad + \left\|\tilde{x}_3^k-x_3^*\right\|^2_{\mu_3 I_{n_3}-\beta A_3^TA_3}\end{aligned} \tag{46}$$

and then

$$\begin{aligned}\left\|u^{k+1}-u^*\right\|_G^2 & = \left\|u^k-u^*-\gamma\alpha^*\left(u^k-\tilde{u}^k\right)\right\|_G^2 \\ & \le \left\|u^k-u^*\right\|_G^2-\gamma(2-\gamma)(\alpha^*)^2 \\ & \quad \times\left\|u^k-\tilde{u}^k\right\|_G^2-2\gamma\alpha^*\mu_1\left\|\tilde{x}_1^k-x_1^*\right\|^2 \\ & \quad -2\mu_2\gamma\alpha^*\left\|\tilde{x}_2^k-x_2^*\right\|^2 \\ & \quad -2\gamma\alpha^*\left\|\tilde{x}_3^k-x_3^*\right\|^2_{\mu_3 I_{n_3}-\beta A_3^TA_3}.\end{aligned} \tag{47}$$

This, together with the fact that $\alpha^*\ge 1/6$, completes the proof. □

Based on the above inequality, we are able to prove the following convergence result of the relaxed ADMM with three blocks. Since the proof is in line with that of Theorem 3, we omit it.

Theorem 6. *Let $\{v^k=(x_1^k,x_2^k,x_3^k,\lambda^k)\}$ be the sequence generated by the relaxed ADMM. Then $\{v^k\}$ converges to a KKT point of problem* (3) *under the conditions that $0<\beta\le\mu_3/\|A_3\|_2^2$ and A_1, A_2, and A_3 are of full column rank.*

5. Conclusion Remarks

In this paper, we take a step to investigate the ADMM for separable convex programming problems with three blocks. Based on the contractive analysis of the distance between the sequence and the solution set, we establish theoretical results to guarantee the global convergence of ADMM with three blocks under weaker conditions than those employed in [22]. By adopting the essential idea of [25], we also present a relaxed ADMM with an optimal step size to accelerate the ADMM and prove its convergence under mild assumptions.

Acknowledgment

The first author is supported by the Natural Science Foundation of Jiangsu Province and the National Natural Science Foundation of China under Project no. 71271112. The second author is supported by university natural science research fund of jiangsu province under grant no. 13KJD110002.

References

[1] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximations," *Computational Mathematics with Applications*, vol. 2, pp. 17–40, 1976.

[2] R. Glowinski and 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," vol. 9, no. R-2, pp. 41–76, 1975.

[3] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 3, pp. 293–318, 1992.

[4] D. Gabay, "Chapter ix applications of the method of multipliers to variational inequalities," *Studies in Mathematics and Its Applications*, vol. 15, pp. 299–331, 1983.

[5] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, NY, USA, 1984.

[6] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*, vol. 9, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 1989.

[7] B. He, L. Liao, D. Han, and H. Yang, "A new inexact alternating directions method for monotone variational inequalities," *Mathematical Programming*, vol. 92, no. 1, pp. 103–118, 2002.

[8] 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.

[9] C. Chen, B. He, and X. Yuan, "Matrix completion via an alternating direction method," *IMA Journal of Numerical Analysis*, vol. 32, no. 1, pp. 227–245, 2012.

[10] M. Fazel, T. K. Pong, D. F. Sun, and P. Tseng, "Hankel matrix rank minimization with applications to system identification and realization," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 946–977, 2012.

[11] 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.

[12] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for tvl1-l2 signal reconstruction from partial fourier data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 288–297, 2010.

[13] J. Yang and Y. Zhang, "Alternating direction method algorithms for l1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.

[14] 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.

[15] M. Tao and X. Yuan, "Recovering low-rank and sparse components of matrices from incomplete and noisy observations," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 57–81, 2011.

[16] J. Yang, D. Sun, and K. Toh, "A proximal point algorithm for logdeterminant optimization with group Lasso regularization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 857–293, 2013.

[17] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 763–770, 2010.

[18] K. Mohan, P. London, M. Fazel, D. Witten, and S. I. Lee, "Node-based learning of multiple gaussian graphical models," http://arxiv.org/abs/1303.5145.

[19] B. He, M. Tao, M. H. Xu, and X. M. Yuan, "Alternating directions based contraction method for generally separable linearly constrained convex programming problems," *Optimization*, vol. 62, no. 4, pp. 573–596, 2013.

[20] B. He, M. Tao, and X. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM Journal on Optimization*, vol. 22, no. 2, pp. 313–340, 2012.

[21] M. Hong and Z. Luo, "On the linear convergence of the alternating direction Method of multipliers," http://arxiv.org/abs/1208.3922.

[22] 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.

[23] R. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[24] R. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, no. 28, Princeton University Press, Princeton, NJ, USA, 1970.

[25] C. Ye and X.-M. Yuan, "A descent method for structured monotone variational inequalities," *Optimization Methods & Software*, vol. 22, no. 2, pp. 329–338, 2007.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 515902, 10 pages
http://dx.doi.org/10.1155/2013/515902

Research Article

The Strong Convergence of Prediction-Correction and Relaxed Hybrid Steepest-Descent Method for Variational Inequalities

Haiwen Xu[1,2]

[1] *School of Computer Science, Civil Aviation Flight University of China, Guanghan 618307, China*
[2] *School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*

Correspondence should be addressed to Haiwen Xu; xuhaiwen_dream@163.com

Received 22 June 2013; Accepted 19 August 2013

Academic Editor: Xu Minghua

We establish the strong convergence of prediction-correction and relaxed hybrid steepest-descent method (PRH method) for variational inequalities under some suitable conditions that simplify the proof. And it is to be noted that the proof is different from the previous results and also is not similar to the previous results. More importantly, we design a set of practical numerical experiments. The results demonstrate that the PRH method under some descent directions is more slightly efficient than that of the modified and relaxed hybrid steepest-descent method, and the PRH Method under some new conditions is more efficient than that under some old conditions.

1. Introduction

Let H be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, let K be a nonempty closed convex subset of H, and let $F : H \to H$ be an operator. Then the variational inequality problem VI(F, K) [1] is to find $x^* \in K$ such that

$$x^* \in K, \quad \langle x - x^*, F(x^*)\rangle \geq 0, \quad \forall x \in K. \tag{1}$$

The literature contains many methods for solving variational inequality problems; see [2–25] and references therein. According to the relationship between the variational inequality problems and a fixed point problem, we can obtain

$$\begin{aligned} &x^* \text{ is the solution of } \mathrm{VI}(F,K) \\ &\Longleftrightarrow x^* = P_K[x^* - \beta F(x^*)], \quad \beta > 0, \end{aligned} \tag{2}$$

where the projection operator P_K is the projection from H onto K, that is,

$$P_K(x) = \operatorname*{argmin}_{y\in K} \|x - y\|, \quad \forall x \in H. \tag{3}$$

In this paper, $F : H \to H$ is an operator with $F : \kappa$-Lipschtz and η-strongly monotone; that is, F satisfies the following conditions:

$$\begin{aligned} &\|F(x) - F(y)\| \leq \kappa \|x - y\|, \\ &\langle F(x) - F(y), x - y\rangle \geq \eta\|x - y\|^2, \quad \forall x, y \in K. \end{aligned} \tag{4}$$

If β is small enough, then P_K is a contraction. Naturally, the convergence of Picard iterates generated by the right-hand side of (2) is obtained by Banach's fixed point theorem. Such a method is called the projection method or more results about the projection method see [6, 8, 20] and so forth.

In fact, the projection P_K in the contraction methods may not be easy to compute, and a great effort is to compute the projection P_K in each iteration. Yamada and Deutsch have provided a hybrid steepest-descent method for solving the VI(F, K) [2, 3] in order to reduce the difficulty and complexity of computing the projection P_K. Subsequently, the convergence of hybrid steepest-descent methods was given out by Xu and Kim [4] and Zeng et al. [5]. Naturally, by analyzing several three-step iterative methods in each iteration by the fixed pointed equation, we can obtain the Noor iterations. Recently, Ding et al. [7] proposed a three-step relaxed hybrid steepest-descent method for variational

inequalities, and the simple proof of three-step relaxed hybrid steepest-descent methods under different conditions was introduced by Yao et al. [24]. The literature [14, 16] described a modified and relaxed hybrid steepest-descent (MRHSD) method and the different convergence of the MRHSD method under the different conditions. A set of practical numerical experiments in the literature [16] demonstrated that the MRHSD method has different efficiency under different conditions. Subsequently, the prediction-correction and relaxed hybrid steepest-descent method (PRH method) [15] makes more use of the history information and less decreases the loss of information than the methods [7, 14]. The PRH method introduced more descent directions than the MRHSD method [14, 16], and computing these descent directions only needs the history information.

In this paper, we will prove the strong convergence of PRH method under different and suitable restrictions imposed on parameters (Condition 12), which differs from that of [15]. Moreover, the proof of strong convergence is different from the previous proof in [15], which is not similar to that in [7] in Step 2. And more importantly, numerical experiments verify that the PRH method under Condition 12 is more efficient than that under Condition 10, and the PRH method under some descent directions is more slightly efficient than that of the MRHSD method [14, 16]. Furthermore, it is easy to obtain these descent directions.

The remainder of the paper is organized as follows. In Section 2, we review several lemmas and preliminaries. We prove the convergence theorem under Condition 12 in Section 3. In Section 4, we give out a series of numerical experiments, which demonstrated that the PRH method under Condition 12 is more efficient than under Condition 10. Section 5 concludes the paper.

2. Preliminaries

In order to proof the later convergence theorem, we introduce several lemmas and the main results in the following.

Lemma 1. *In a real Hilbert space H, there holds the inequality*

$$\|x+y\|^2 \le \|x\|^2 + 2\langle y, x+y\rangle, \quad \forall x, y \in H. \tag{5}$$

The lemma is a basic result of a Hilbert space with the inner product.

Lemma 2 (demiclosedness principle). *Assume that T is a nonexpansive self-mapping on a nonempty closed convex subset K of a Hilbert space H. If T has a fixed point, then $(I-T)$ is demiclosed. That is, whenever x_n is a sequence in K weakly converging to some $x \in K$ and the sequence $(I-T)x_n$ strongly converges to some $y \in H$, it follows that $(I-T)x = y$. Here I is the identity operator of H.*

The following lemma is an immediate result of a projection mapping onto a closed convex subset of a Hilbert space.

Lemma 3. *Let K be a nonempty closed convex subset of H. For all $x, y \in H$ and $z \in K$, then*

(1) $\langle P_K(x) - x, z - P_K(y)\rangle \ge 0$,

(2) $\|P_K(x) - P_K(y)\|^2 \le 1\|x-y\|^2 - \|P_K(x) - x + y - P_K(y)\|^2$.

Lemma 4 (see [13]). *Let $\{x_n\}$ and $\{y_n\}$ be bounded sequence in a Banach space X and let $\{\zeta_n\}$ be a sequence in $[0,1]$ with $0 < \liminf_{n\to\infty}\zeta_n \le \limsup_{n\to\infty}\zeta_n < 1$. Suppose $x_{n+1} = (1-\zeta_n)y_n + \zeta_n x_n$ for all integers $n \ge 0$ and $\limsup_{n\to\infty}(\|y_{n+1} - y_n\| - \|x_{n+1} - x_n\|) \le 0$. Then $\limsup_{n\to\infty}\|y_n - x_n\| = 0$.*

Lemma 5 ([5, 7]). *Let $\{s_n\}$ be a sequence of nonnegative real numbers satisfying the inequality*

$$s_{n+1} \le (1-\alpha_n)s_n + \alpha_n\tau_n + \gamma_n, \quad \forall n \ge 0, \tag{6}$$

where α_n, τ_n, and γ_n satisfy the following conditions:

(1) $\alpha_n \subset [0,1]$, $\sum_{n=0}^{\infty}\alpha_n = \infty$, *or* $\prod_{n=0}^{\infty}(1-\alpha_n) = 0$,

(2) $\lim_{n\to\infty}\sup\tau_n \le 0$,

(3) $\gamma_n \subset [0,\infty)$, $\sum_{n=0}^{\infty}\gamma_n < \infty$.

Then $\lim_{n\to\infty}s_n = 0$.

Since F is η-strongly monotone, $\mathrm{VI}(F,K)$ has a unique solution $x^* \in K$ [5]. Assume that $T : H \to H$ is a nonexpansive mapping with the fixed point set $\mathrm{Fix}(T) = K$. Obviously $\mathrm{Fix}(P_K) = K$.

For any given numbers $\lambda \in (0,1)$ and $\mu \in (0, 2\eta/\kappa^2)$, we define the mapping $T_\mu^\lambda : H \to H$ by

$$T_\mu^\lambda x : Tx - \lambda\mu F(Tx), \quad \forall x \in H. \tag{7}$$

Lemma 6 (see [5]). *If $0 < \mu < 2\eta/\kappa^2$ and $0 < \lambda < 1$, then T_μ^λ is a contraction. In fact,*

$$\left\|T_\mu^\lambda x - T_\mu^\lambda y\right\| \le (1-\lambda\delta)\|x-y\|, \tag{8}$$

where $\delta = 1 - \sqrt{1-\mu(2\eta-\mu\kappa^2)}$, for all $x, y \in H$.

Lemma 7 (see [7]). *Let $\{\alpha_n\}$ be a sequence of nonnegative numbers with $\limsup_{n\to\infty}\alpha_n < \infty$ and let $\{\beta_n\}$ be sequence of real numbers with $\limsup_{n\to\infty}\beta_n \le 0$. Then*

$$\limsup_{n\to\infty}\alpha_n\beta_n \le 0. \tag{9}$$

3. Convergence Theorem

Before analyzing the convergence theorem, we first review the PRH method and related results [15].

Algorithm 8 (see [15]). Take three fixed numbers $t, \rho, \gamma \in (0, 2\eta/\kappa^2)$, starting with arbitrarily chosen initial points $x_0 \in H$, compute the sequences $\{x_n\}, \{\overline{x}_n\}, \{\widetilde{x}_n\}, \{\widehat{x}_n\}$ such that;

Prediction

Step 1: $\overline{x}_n = \gamma_n x_n + (1-\gamma_n)[Tx_n - \lambda''_{n+1}\gamma F(Tx_n)]$,

Step 2: $\widetilde{x}_n = \beta_n x_n + (1-\beta_n)[T\overline{x}_n - \lambda'_{n+1}\rho F(T\overline{x}_n)]$,

Step 3: $\widehat{x}_n = \theta_n\overline{x}_n + (1-\theta_n)\widetilde{x}_n, 0 \le \theta_n \le 1$,

Correction

Step 4: $x_{n+1} = \alpha_n \overline{x}_n + (1-\alpha_n)[T\widehat{x}_n - \lambda_{n+1} tF(T\widehat{x}_n)]$,

where $T : H \to H$ is a nonexpansive mapping.

Let $\{\alpha_n\} \subset [0,1)$, $\{\beta_n\} \subset [0,1]$ and $\{\gamma_n\} \subset [0,1]$, $\{\lambda_n\}$, $\{\lambda'_n\}$, $\{\lambda''_n\} \subset (0,1)$ satisfy the following conditions.

Remark 9. In fact, the PRH method is the MRHSD method when $\theta_n \equiv 0$, for all n.

Condition 10. One has

$$\begin{aligned}
&(1)\quad \sum_1^\infty |\alpha_n - \alpha_{n-1}| < \infty, \qquad \sum_1^\infty |\beta_n - \beta_{n-1}| < \infty, \\
&\qquad \sum_1^\infty |\gamma_n - \gamma_{n-1}| < \infty, \\
&(2)\quad \lim_{n\to\infty} \alpha_n = 0, \qquad \lim_{n\to\infty} \beta_n = 1, \qquad \lim_{n\to\infty} \gamma_n = 1, \\
&(3)\quad \lim_{n\to\infty} \lambda_n = 0, \qquad \lim_{n\to\infty} \frac{\lambda_n}{\lambda_{n+1}} = 1, \qquad \sum_1^\infty \lambda_n = \infty, \\
&(4)\quad \lambda_n \geq \max\{\lambda'_n, \lambda''_n\}, \quad \forall n \geq 1.
\end{aligned} \tag{10}$$

Theorem 11 (see [15]). *In Condition 10, the sequence $\{x_n\}$ converges strongly to $x^* \in K$, and x^* is the unique solution of the $VI(F,K)$.*

We obtain the strong convergence theorem of PRH method for variational inequalities under different assumptions.

Condition 12. One has

$$\begin{aligned}
&(1)\quad 0 < \liminf_{n\to\infty} \alpha_n \leq \limsup_{n\to\infty} \alpha_n < 1, \\
&\qquad \lim_{n\to\infty} \beta_n = 1, \qquad \lim_{n\to\infty} \gamma_n = 1, \\
&(2)\quad \lim_{n\to\infty} \lambda_n = 0, \qquad \sum_1^\infty \lambda_n = \infty, \\
&(3)\quad \lambda_n \geq \max\{\lambda'_n, \lambda''_n\}, \quad \forall n \geq 1.
\end{aligned} \tag{11}$$

Theorem 13. *The sequence $\{x_n\}$ converges strongly to $x^* \in K$, and x^* is the unique solution of the $VI(F,K)$. Assume that $\{\alpha_n\}$, $\{\beta_n\}$ and $\{\gamma_n\}$, $\{\lambda_n\}$, $\{\lambda'_n\}$, $\{\lambda''_n\}$ satisfy Condition 12.*

Proof. We divide the proof into several steps.

Step 1. $\{x_n\}$, $\{\overline{x}_n\}$, $\{\widetilde{x}_n\}$, and $\{\widehat{x}_n\}$ are bounded. Since F is η-strongly monotone, VI(F,K) (1) has a unique solution $x^* \in K$, and $T_t^{\lambda_{n+1}} x^* = x^* - \lambda_{n+1} tF(x^*)$, $T_\rho^{\lambda'_{n+1}} x^* = x^* - \lambda_{n+1}\rho F(x^*)$, $T_\gamma^{\lambda''_{n+1}} x^* = x^* - \lambda_{n+1}\gamma F(x^*)$.

A series of computations yields

$$\begin{aligned}
\|x_{n+1} - x^*\| &= \left\|\alpha_n \overline{x}_n + (1-\alpha_n) T_t^{\lambda_{n+1}} \widehat{x} - x^*\right\| \\
&\leq \alpha_n \|\overline{x}_n - x^*\| + (1-\alpha_n) \left\|T_t^{\lambda_{n+1}} \widehat{x} - x^*\right\| \\
&\leq \alpha_n \|\overline{x}_n - x^*\| + (1-\alpha_n) \\
&\quad \times \left[\left\|T_t^{\lambda_{n+1}} \widehat{x} - T_t^{\lambda_{n+1}} x^*\right\| + \left\|T_t^{\lambda_{n+1}} x^* - x^*\right\|\right] \\
&\leq \alpha_n \|\overline{x}_n - x^*\| + (1-\alpha_n) \\
&\quad \times \left[(1-\lambda_{n+1}\tau)\|\widehat{x}_n - x^*\| + \lambda_{n+1} t \|F(x^*)\|\right],
\end{aligned} \tag{12}$$

where $\tau = 1 - \sqrt{1 - t(2\eta - t\kappa^2)} \in (0,1)$,

$$\begin{aligned}
\|\widetilde{x}_n - x^*\| &= \left\|\beta_n x_n + (1-\beta_n) T_\rho^{\lambda'_{n+1}} \overline{x}_n - x^*\right\| \\
&\leq \beta_n \|x_n - x^*\| + (1-\beta_n) \left\|T_\rho^{\lambda'_{n+1}} \overline{x}_n - x^*\right\| \\
&\leq \beta_n \|x_n - x^*\| + (1-\beta_n) \\
&\quad \times \left[\left\|T_\rho^{\lambda'_{n+1}} \overline{x}_n - T_\rho^{\lambda'_{n+1}} x^*\right\| + \left\|T_\rho^{\lambda'_{n+1}} x^* - x^*\right\|\right] \\
&\leq \beta_n \|x_n - x^*\| + (1-\beta_n) \\
&\quad \times \left[\left(1 - \lambda'_{n+1}\tau'\right)\|\overline{x}_n - x^*\| + \lambda'_{n+1}\rho \|F(x^*)\|\right] \\
&\leq \beta_n \|x_n - x^*\| + (1-\beta_n) \|\overline{x}_n - x^*\| \\
&\quad + (1-\beta_n) \lambda'_{n+1}\rho \|F(x^*)\|,
\end{aligned} \tag{13}$$

where $\tau' = 1 - \sqrt{1 - \rho(2\eta - t\kappa^2)} \in (0,1)$.

Moreover, we also obtain

$$\begin{aligned}
\|\widetilde{x}_n - x^*\| &= \left\|\gamma_n x_n + (1-\gamma_n) T_\gamma^{\lambda''_{n+1}} x_n - x^*\right\| \\
&\leq \gamma_n \|x_n - x^*\| + (1-\gamma_n) \left\|T_\gamma^{\lambda''_{n+1}} x_n - x^*\right\| \\
&\leq \gamma_n \|x_n - x^*\| + (1-\gamma_n) \\
&\quad \times \left[\left\|T_\gamma^{\lambda''_{n+1}} x_n - T_\gamma^{\lambda''_{n+1}} x^*\right\| + \left\|T_\gamma^{\lambda''_{n+1}} x^* - x^*\right\|\right] \\
&\leq \gamma_n \|x_n - x^*\| + (1-\gamma_n) \\
&\quad \times \left[\left(1 - \lambda''_{n+1}\tau''\right)\|x_n - x^*\| + \lambda''_{n+1}\gamma \|F(x^*)\|\right] \\
&\leq \|x_n - x^*\| + (1-\gamma_n) \lambda''_{n+1}\gamma \|F(x^*)\|,
\end{aligned} \tag{14}$$

where $\tau'' = 1 - \sqrt{1-\gamma(2\eta - t\kappa^2)} \in (0,1)$, subtituting; (14) into (13) and (14) into (12), we immediately obtain

$$\begin{aligned}\|\tilde{x}_n - x^*\| &= \beta_n \|x_n - x^*\| + (1-\beta_n) \\ &\quad \times \left[\left(1-\lambda'_{n+1}\tau'\right)\|\overline{x}_n - x^*\| + \lambda'_{n+1}\rho\|F(x^*)\|\right] \\ &\le \beta_n \|x_n - x^*\| + (1-\beta_n) \\ &\quad \times \left[\left(1-\lambda'_{n+1}\tau'\right)\|x_n - x^*\| \right. \\ &\quad \left. + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\| + \lambda'_{n+1}\rho\|F(x^*)\|\right] \\ &\le \|x_n - x^*\| + (1-\beta_n)\lambda_{n+1}(\gamma+\rho)\|F(x^*)\|.\end{aligned} \tag{15}$$

Furthermore,

$$\begin{aligned}\|\hat{x}_n - x^*\| &= \|\theta_n \overline{x}_n + (1-\theta_n)\tilde{x}_n - x^*\| \\ &\le \theta_n\|\overline{x}_n - x^*\| + (1-\theta_n)\|\tilde{x}_n - x^*\| \\ &\le \theta_n\left[\|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\|\right] \\ &\quad + (1-\theta_n)\left[\|x_n - x^*\| + (1-\beta_n)\lambda_{n+1}\right. \\ &\quad \left.\times(\gamma+\rho)\|F(x^*)\|\right] \\ &\le \|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\| \\ &\quad + (1-\beta_n)\lambda_{n+1}(\gamma+\rho)\|F(x^*)\|, \\ \|x_{n+1} - x^*\| &\le \alpha_n\|\overline{x}_n - x^*\| + (1-\alpha_n) \\ &\quad \times\left[(1-\lambda_{n+1}\tau)\|\hat{x}_n - x^*\| + \lambda_{n+1}t\|F(x^*)\|\right] \\ &\le \alpha_n\left[\|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\|\right] \\ &\quad + (1-\alpha_n)\left\{(1-\lambda_{n+1}\tau)\right. \\ &\quad \times\left[\|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\|\right. \\ &\quad \left.+ (1-\beta_n)\lambda_{n+1}(\gamma+\rho)\|F(x^*)\|\right] \\ &\quad \left.+ \lambda_{n+1}t\|F(x^*)\|\right\} \\ &\le \alpha_n\|x_n - x^*\| + \alpha_n(1-\gamma_n)\lambda_{n+1}\gamma\|F(x^*)\| + (1-\alpha_n) \\ &\quad \times\left[(1-\lambda_{n+1}\tau)\|x_n - x^*\| + \lambda_{n+1}(2\gamma+\rho+t)\|F(x^*)\|\right].\end{aligned} \tag{16}$$

It is easy to obtain the following by induction:

$$\|x_n - x^*\| \le M_0, \quad \forall n \ge 0, \tag{17}$$

where $M_0 = \max\{3\|x_0 - x^*\|, 3(\rho+\gamma+t)\|F(x^*)\|/\tau\}$,

$$\begin{aligned}\|\tilde{x}_n - x^*\| &\le \|x_n - x^*\| + (1-\beta_n)\lambda_{n+1}(\gamma+\rho)\|F(x^*)\| \\ &\le \left(1+\frac{\tau}{3}\right)M_0, \\ \|\overline{x}_n - x^*\| &\le \|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\| \\ &\le \left(1+\frac{\tau}{3}\right)M_0, \\ \|\hat{x}_n - x^*\| &\le \theta_n\|\overline{x}_n - x^*\| + (1-\theta_n)\|\tilde{x}_n - x^*\| \\ &\le 2\left(1+\frac{\tau}{3}\right)M_0.\end{aligned} \tag{18}$$

Hence

$$\begin{aligned}&\{Tx_n\}, \{T\overline{x}_n\}, \{T\tilde{x}_n\}, \{T\hat{x}_n\}, \\ &\{F(Tx_n)\}, \{F(T\overline{x}_n)\}, \{F(T\tilde{x}_n)\}, \{F(T\hat{x}_n)\}\end{aligned} \tag{19}$$

are also bounded.

Step 2. Consider $\|x_{n+1} - x_n\| \to 0$.

Indeed, by a series of computations, we have

$$\begin{aligned}&\|\overline{x}_n - \overline{x}_{n-1}\| \\ &\quad = \left\|\gamma_n x_n - \gamma_{n-1}x_{n-1} + (1-\gamma_n)T_\gamma^{\lambda''_{n+1}}x_n \right. \\ &\qquad \left. - (1-\gamma_{n-1})T_\gamma^{\lambda''_n}x_{n-1}\right\| \\ &\quad \le \|\gamma_n x_n - \gamma_{n-1}x_{n-1}\| \\ &\qquad + \left\|(1-\gamma_n)T_\gamma^{\lambda''_{n+1}}x_n - (1-\gamma_{n-1})T_\gamma^{\lambda''_n}x_{n-1}\right\| \\ &\quad \le \|x_n - x_{n-1}\| + \left|(1-\gamma_n)\lambda''_{n+1} - (1-\gamma_{n-1})\lambda''_n\right| \\ &\qquad \times\gamma\|F(Tx_{n-1})\| \\ &\qquad + |\gamma_n - \gamma_{n-1}|\left(\|x_{n-1}\| + \|Tx_{n-1}\|\right).\end{aligned} \tag{20}$$

According to (20) and the prediction step of Algorithm 8, we also obtain

$$\begin{aligned}&\|\tilde{x}_n - \tilde{x}_{n-1}\| \\ &\quad = \left\|\beta_n x_n - \beta_{n-1}x_{n-1} + (1-\beta_n)T_\rho^{\lambda'_{n+1}}\overline{x}_n \right. \\ &\qquad \left. - (1-\beta_{n-1})T_\rho^{\lambda'_n}\overline{x}_{n-1}\right\| \\ &\quad \le \|\beta_n x_n - \beta_{n-1}x_{n-1}\| \\ &\qquad + \left\|(1-\beta_n)T_\rho^{\lambda'_{n+1}}\overline{x}_n - (1-\beta_{n-1})T_\rho^{\lambda'_n}\overline{x}_{n-1}\right\|\end{aligned}$$

$$\begin{aligned}
&\leq \|x_n - x_{n-1}\| + \left|(1-\beta_n)\lambda'_{n+1} - (1-\beta_{n-1})\lambda'_n\right| \\
&\quad \times \rho \|F(T\overline{x}_{n-1})\| + (1-\beta_n)\left(1-\lambda'_{n+1}\tau'\right) \\
&\quad \times |\gamma_n - \gamma_{n-1}| (\|x_{n-1}\| + \|Tx_{n-1}\|) \\
&\quad + (1-\beta_n)\left(1-\lambda'_{n+1}\tau'\right)\left|(1-\gamma_n)\lambda''_{n+1} - \gamma_{n-1}\lambda''_n\right| \\
&\quad \times \gamma \|F(Tx_{n-1})\| + |\beta_n - \beta_{n-1}| \\
&\quad \times (\|x_{n-1}\| + \|T\overline{x}_{n-1}\| + \|T\overline{x}_{n-1}\|).
\end{aligned} \tag{21}$$

Also by the prediction step of Algorithm 8 and (20), (21), we have

$$\begin{aligned}
\|\widehat{x}_n - \widehat{x}_{n-1}\| &\leq \theta_n \|\overline{x}_n - \overline{x}_{n-1}\| + (1-\theta_n)\|\widetilde{x}_n - \widetilde{x}_{n-1}\| \\
&\leq \|x_n - x_{n-1}\| \\
&\quad + \left|(1-\gamma_n)\lambda''_{n+1} - (1-\gamma_{n-1})\lambda''_n\right| \gamma \|F(Tx_{n-1})\| \\
&\quad + |\gamma_n - \gamma_{n-1}| (\|x_{n-1}\| + \|Tx_{n-1}\|) \\
&\quad + \left|(1-\beta_n)\lambda'_{n+1} - (1-\beta_{n-1})\lambda'_n\right| \rho \|F(T\overline{x}_{n-1})\| \\
&\quad + (1-\beta_n)\left(1-\lambda'_{n+1}\tau'\right)|\gamma_n - \gamma_{n-1}| \\
&\quad \times (\|x_{n-1}\| + \|Tx_{n-1}\|) + (1-\beta_n)\left(1-\lambda'_{n+1}\tau'\right) \\
&\quad \times \left|(1-\gamma_n)\lambda''_{n+1} - \gamma_{n-1}\lambda''_n\right| \gamma \|F(Tx_{n-1})\| \\
&\quad + |\beta_n - \beta_{n-1}| (\|x_{n-1}\| + \|T\overline{x}_{n-1}\| + \|T\overline{x}_{n-1}\|).
\end{aligned} \tag{22}$$

Let

$$\widehat{y}_n = T_t^{\lambda_{n+1}}\widehat{x}_n = T\widehat{x}_n - \lambda_{n+1} t F(T\widehat{x}_n), \tag{23}$$

so we get

$$x_{n+1} = \alpha_n \overline{x}_n + (1-\alpha_n)\widehat{y}_n. \tag{24}$$

Furthermore,

$$\begin{aligned}
&\|\widehat{y}_n - \widehat{y}_{n-1}\| \\
&\quad = \|T\widehat{x}_n - T\widehat{x}_{n-1} + \lambda_n t F(T\widehat{x}_{n-1}) - \lambda_{n+1} t F(T\widehat{x}_n)\| \\
&\quad \leq \|T\widehat{x}_n - T\widehat{x}_{n-1}\| + \lambda_n t \|F(T\widehat{x}_{n-1})\| \\
&\qquad + \lambda_{n+1} t \|F(T\widehat{x}_n)\| \\
&\quad \leq \|\widehat{x}_n - \widehat{x}_{n-1}\| + \lambda_n t \|F(T\widehat{x}_{n-1})\| \\
&\qquad + \lambda_{n+1} t \|F(T\widehat{x}_n)\|.
\end{aligned} \tag{25}$$

Apply $\lim_{n\to\infty}\beta_n = 1, \lim_{n\to\infty}\lambda_n = 0$, and $\lim_{n\to\infty}\gamma_n = 1$ and (22), (25) to get

$$\begin{aligned}
&\|\widehat{y}_n - \widehat{y}_{n-1}\| - \|x_n - x_{n-1}\| \\
&\quad \leq \left|(1-\gamma_n)\lambda''_{n+1} - (1-\gamma_{n-1})\lambda''_n\right| \gamma \|F(Tx_{n-1})\| \\
&\qquad + |\gamma_n - \gamma_{n-1}| (\|x_{n-1}\| + \|Tx_{n-1}\|) \\
&\qquad + \left|(1-\beta_n)\lambda'_{n+1} - (1-\beta_{n-1})\lambda'_n\right| \rho \|F(T\overline{x}_{n-1})\| \\
&\qquad + (1-\beta_n)\left(1-\lambda'_{n+1}\tau'\right)|\gamma_n - \gamma_{n-1}| (\|x_{n-1}\| + \|Tx_{n-1}\|) \\
&\qquad + (1-\beta_n)\left(1-\lambda'_{n+1}\tau'\right) \\
&\qquad \times \left|(1-\gamma_n)\lambda''_{n+1} - \gamma_{n-1}\lambda''_n\right| \gamma \|F(T(x_{n-1}))\| \\
&\qquad + |\beta_n - \beta_{n-1}| (\|x_{n-1}\| + \|T\overline{x}_{n-1}\| + \|T\overline{x}_{n-1}\|) \\
&\qquad + \lambda_n t \|F(T\widehat{x}_{n-1})\| + \lambda_{n+1} t \|F(T\widehat{x}_n)\| \longrightarrow 0.
\end{aligned} \tag{26}$$

According to Lemma 4, we obtain

$$\lim_{n\to\infty} \|\widehat{y}_{n-1} - x_{n-1}\| = 0. \tag{27}$$

Furthermore, by $\lim_{n\to\infty}\gamma_n = 1$, we also get

$$\begin{aligned}
&\|\overline{x}_n - x_n\| \\
&\quad = \left\|-(1-\gamma_n)x_n + (1-\gamma_n)\left[Tx_n - \lambda'_{n+1}\gamma F(Tx_n)\right]\right\| \\
&\quad \leq (1-\gamma_n)\|x_n\| + (1-\gamma_n)\|Tx_n\| + \lambda'_{n+1}\gamma \|F(Tx_n)\| \longrightarrow 0.
\end{aligned} \tag{28}$$

By (27), (28) and the correction step of Algorithm 8, we immediately conclude that

$$\begin{aligned}
&\|x_n - x_{n-1}\| \\
&\quad = \|\alpha_{n-1}\overline{x}_{n-1} + (1-\alpha_{n-1})\widehat{y}_{n-1} - x_{n-1}\| \\
&\quad \leq \alpha_{n-1}\|\overline{x}_{n-1} - x_{n-1}\| + (1-\alpha_{n-1})\|\widehat{y}_{n-1} - x_{n-1}\| \longrightarrow 0,
\end{aligned} \tag{29}$$

so we get

$$\|x_{n+1} - x_n\| \longrightarrow 0. \tag{30}$$

Step 3. Consider $\|x_{n+1} - Tx_n\| \to 0$.

Indeed, by the prediction step of Algorithm 8, we have

$$\begin{aligned}
&\|\widetilde{x}_n - x_n\| \\
&\quad = \left\|-(1-\beta_n)x_n + (1-\beta_n)\left[T\overline{x}_n - \lambda'_{n+1}\rho F(T\overline{x}_n)\right]\right\| \\
&\quad \leq (1-\beta_n)\|x_n\| + (1-\beta_n)\left[\|T\overline{x}_n\| + \left\|\lambda'_{n+1}\rho F(T\overline{x}_n)\right\|\right].
\end{aligned} \tag{31}$$

According to the assumption $\lim_{n\to\infty}\beta_n = 1$ and $\lim_{n\to\infty}\lambda_n = 0$, then

$$\|\widetilde{x}_n - x_n\| \longrightarrow 0. \tag{32}$$

By (32), we immediately obtain

$$\|\widehat{x}_n - x_n\| \le \theta_n \|\overline{x}_n - x_n\| + (1-\theta_n)\|\widetilde{x}_n - x_n\| \longrightarrow 0. \quad (33)$$

By a series of computations, we can get

$$\begin{aligned}
&\|x_{n+1} - Tx_n\| \\
&\quad = \left\|\alpha_n(\overline{x}_n - Tx_n) + (1-\alpha_n)\left(T_t^{\lambda_{n+1}}\widehat{x} - Tx_n\right)\right\| \\
&\quad \le \alpha_n\|\overline{x}_n - Tx_n\| + (1-\alpha_n)\|T\widehat{x}_n - Tx_n\| \\
&\qquad + (1-\alpha_n)\lambda_{n+1}t\|F(T\widehat{x}_n)\| \\
&\quad \le \alpha_n\|\overline{x}_n - Tx_n\| + \|\widehat{x}_n - x_n\| + \lambda_{n+1}t\|F(T\widehat{x})\| \\
&\quad \le \alpha_n\|x_{n+1} - Tx_n\| + \alpha_n\|\overline{x}_n - x_{n+1}\| \\
&\qquad + \|\widehat{x}_n - x_n\| + \lambda_{n+1}t\|F(T\widehat{x})\|.
\end{aligned} \quad (34)$$

Hence, by (28), (33), and (34), we also obtain

$$\begin{aligned}
\|x_{n+1} - Tx_n\| &\le \frac{\alpha_n}{1-\alpha_n}\|\overline{x}_n - x_{n+1}\| \\
&\quad + \frac{\|\widehat{x}_n - x_n\|}{1-\alpha_n} + \frac{\lambda_{n+1}t\|F(T\widehat{x})\|}{1-\alpha_n} \longrightarrow 0.
\end{aligned} \quad (35)$$

Using Steps 2 and 3, it is easy to obtain the following corollary.

Corollary 14. Consider $\|x_n - Tx_n\| \to 0$.

Applying Steps 2 and 3, one gets

$$\|x_{n+1} - Tx_n\| \longrightarrow 0, \qquad \|x_{n+1} - x_n\| \longrightarrow 0, \quad (36)$$

so it is easy to see that

$$\|x_n - Tx_n\| \le \|x_{n+1} - Tx_n\| + \|x_{n+1} - x_n\| \longrightarrow 0. \quad (37)$$

Step 4. Consider $\lim_{n\to\infty}\sup\langle -F(x^*), T\widehat{x}_n - x^*\rangle \le 0$.

For some $\widehat{x} \in H$, here exits $\{Tx_{n_i}\} \to \widehat{x}$ weakly and such that

$$\begin{aligned}
&\lim_{n\to\infty}\sup\langle -F(x^*), Tx_n - x^*\rangle \\
&\quad = \lim_{n\to\infty}\sup\left\langle -F(x^*), Tx_{n_i} - x^*\right\rangle.
\end{aligned} \quad (38)$$

According to $\{Tx_{n_i}\} \to \widehat{x}$, we have

$$\widehat{x} \in \mathrm{Fix}(T) = K. \quad (39)$$

By x^* being the unique solution of VI(F, K), we can obtain

$$\begin{aligned}
&\lim_{n\to\infty}\sup\langle -F(x^*), Tx_n - x^*\rangle \\
&\quad = \lim_{n\to\infty}\sup\langle -F(x^*), \widehat{x} - x^*\rangle \\
&\quad \le 0.
\end{aligned} \quad (40)$$

Since $\|T\widehat{x}_n - Tx_n\| \le \|\widehat{x}_n - x_n\| \to 0$, we immediately conclude that

$$\begin{aligned}
&\lim_{n\to\infty}\sup\langle -F(x^*), T\widehat{x}_n - x^*\rangle \\
&\quad \le \lim_{n\to\infty}\sup\langle -F(x^*), T\widehat{x}_n - Tx_n\rangle \\
&\qquad + \lim_{n\to\infty}\sup\langle -F(x^*), Tx_n - x^*\rangle \\
&\quad \le \lim_{n\to\infty}\sup\langle -F(x^*), Tx_n - x^*\rangle \\
&\quad \le 0.
\end{aligned} \quad (41)$$

Step 5. By Step 1 and Lemma 1, we have

$$\begin{aligned}
&\|x_{n+1} - x^*\|^2 \\
&\quad = \left\|\alpha_n(\overline{x}_n - x^*) + (1-\alpha_n)\left(T_t^{\lambda_{n+1}}\widehat{x}_n - x^*\right)\right\|^2 \\
&\quad \le \|\alpha_n(\overline{x}_n - x^*)\|^2 + (1-\alpha_n) \\
&\qquad \times \left\|\left(T_t^{\lambda_{n+1}}\widehat{x}_n - T_t^{\lambda_{n+1}}x^* + T_t^{\lambda_{n+1}}x^* - x^*\right)\right\|^2 \\
&\quad \le \|\alpha_n(\overline{x}_n - x^*)\|^2 + (1-\alpha_n) \\
&\qquad \times \left[\left\|T_t^{\lambda_{n+1}}\widehat{x}_n - T_t^{\lambda_{n+1}}x^*\right\|^2\right. \\
&\qquad\quad \left. + 2\left\langle T_t^{\lambda_{n+1}}x^* - x^*, T_t^{\lambda_{n+1}}\widehat{x}_n - x^*\right\rangle\right] \\
&\quad \le \alpha_n\left[\|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\|\right]^2 \\
&\qquad + (1-\alpha_n)(1-\lambda_{n+1}\tau)^2 \\
&\qquad \times \left[\|x_n - x^*\| + (1-\gamma_n)\lambda''_{n+1}\gamma\|F(x^*)\|\right. \\
&\qquad\quad \left. + (1-\beta_n)\lambda_{n+1}(\gamma+\rho)\|F(x^*)\|\right]^2 \\
&\qquad + 2t\lambda_{n+1}\langle -F(x^*), T\widehat{x}_n - x^* - t\lambda_{n+1}F(T\widehat{x}_n)\rangle \\
&\quad \le \alpha_n\|x_n - x^*\|^2 + (1-\gamma_n)\lambda_{n+1}\gamma M \\
&\qquad + (1-\alpha_n)(1-\lambda_{n+1}\tau)^2\|x_n - x^*\|^2 \\
&\qquad + (1-\alpha_n)(1-\lambda_{n+1}\tau)^2(1-\beta_n)\lambda_{n+1}M \\
&\qquad + (1-\alpha_n)(1-\lambda_{n+1}\tau)^2(1-\gamma_n)\lambda_{n+1}\gamma M \\
&\qquad + 2t\lambda_{n+1}\langle -F(x^*), T\widehat{x}_n - x^* - t\lambda_{n+1}F(T\widehat{x}_n)\rangle \\
&\quad \le [1-(1-\alpha_n)\lambda_{n+1}\tau]\|x_n - x^*\|^2 \\
&\qquad + (1-\alpha_n)\lambda_{n+1}\tau w'_{n+1},
\end{aligned} \quad (42)$$

where

$$\begin{aligned} w'_{n+1} &= \frac{2t\left\langle -F\left(x^*\right), T\hat{x}_n - x^* - t\lambda_{n+1}F\left(T\hat{x}_n\right)\right\rangle}{\tau\left(1-\alpha_n\right)} \\ &\quad + \frac{\varphi_n}{\tau\left(1-\alpha_n\right)} + \frac{\xi_n}{\tau\left(1-\alpha_n\right)}, \\ \varphi_n &= \left(1-\gamma_n\right)\gamma M, \\ \xi_n &= \left(1-\alpha_n\right)\left(1-\lambda_{n+1}\tau\right)^2\left(1-\beta_n\right)M \\ &\quad + \left(1-\alpha_n\right)\left(1-\lambda_{n+1}\tau\right)^2\left(1-\gamma_n\right)\lambda_{n+1}\gamma M, \end{aligned} \tag{43}$$

and $M_0 \ll M < \infty$.

Denote

$$s'_{n+1} = \left\|x_{n+1} - x^*\right\|, \qquad u_n = \left(1-\alpha_n\right)\lambda_{n+1}\tau. \tag{44}$$

We can rewrite (42) as

$$s'_{n+1} \le \left(1-u_n\right)s'_n + u_n w'_n + 0. \tag{45}$$

In fact, u_n, w'_n satisfies Lemma 5; according to

$$\lim_{n\to\infty}\beta_n = 1, \qquad \lim_{n\to\infty}\gamma_n = 1, \qquad \lim_{n\to\infty}\lambda_n = 0, \tag{46}$$

we obtain

$$\begin{aligned} \frac{\varphi_n}{\tau\left(1-\alpha_n\right)} &\longrightarrow 0, \\ \frac{\xi_n}{\tau\left(1-\alpha_n\right)} &\longrightarrow 0. \end{aligned} \tag{47}$$

Moreover, by Step 4, we also obtain

$$\begin{aligned} &\lim_{n\to\infty}\frac{2t\left\langle -F\left(x^*\right), T\hat{x}_n - x^* - t\lambda_{n+1}F\left(T\hat{x}_n\right)\right\rangle}{\tau\left(1-\alpha_n\right)} \\ &\quad \le \frac{2t}{\tau}\lim_{n\to\infty}\sup\left\{\left\langle -F\left(x^*\right), T\hat{x}_n - x^*\right\rangle\right. \\ &\qquad\qquad \left. + \lambda_{n+1}\left\langle -F\left(x^*\right), -tF\left(T\hat{x}_n\right)\right\rangle\right\} \\ &\quad \le \frac{2t}{\tau}\lim_{n\to\infty}\sup\left\{\left\langle -F\left(x^*\right), T\hat{x}_n - x^*\right\rangle\right\} \\ &\qquad + \lim_{n\to\infty}\sup\left\{\lambda_{n+1}\left\langle -F\left(x^*\right), -tF\left(T\hat{x}_n\right)\right\rangle\right\} \\ &\quad \le 0 + 0 = 0. \end{aligned} \tag{48}$$

Furthermore, by (43), (47), and (48), it is easy to obtain

$$\lim_{n\to\infty}\sup w'_n \le 0. \tag{49}$$

Consequently apply Lemma 5 to obtain

$$\left\|x_n - x^*\right\| \longrightarrow 0. \tag{50}$$

□

4. Numerical Experiments

The problem considered in this section is

$$\min\left\{\frac{1}{2}\|X - C\|_F^2 \mid X \in K\right\}, \tag{51}$$

where $\|\cdot\|_F$ is the matrix Fröbenis norm; that is,

$$\|C\|_F = \left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left|C_{ij}\right|^2\right)^{1/2}. \tag{52}$$

Note that the matrix Fröbenis norm is induced by the inner product

$$\langle A, B\rangle = \operatorname{Trace}\left(A^T B\right). \tag{53}$$

The problems arise from finance and statistics, and we form the test problems similarly as in [9, 21].

Let $K = S_+^n \cap \mathfrak{B}$, where

$$\begin{aligned} S_+^n &= \left\{H \in \mathbb{R}^{n\times n} \mid H^T = H, H \succeq 0\right\}, \\ \mathfrak{B} &= \left\{H \in \mathbb{R}^{n\times n} \mid H^T = H, H_L \le H \le H_U\right\}. \end{aligned} \tag{54}$$

Let H_L, H_U be given $n \times n$ symmetric matrices, and C asymmetric which differs from previous approaches [9, 21], and it is to be noted that the extended contraction method (EC method) [9] has much difficulty in computing the examples when C is asymmetric, where $H_L \le H_U$ in element wise:

$$H_L \le H_U : \left(H_L\right)_{ij} \le \left(H_U\right)_{ij}, \quad \forall i, j \in 1, \ldots, n. \tag{55}$$

Then (51) is equivalent to the following variational inequality:

$$\left\langle X' - X, \nabla\left(\frac{1}{2}\|X - C\|^2\right)\right\rangle \ge 0, \quad \forall X' \in K. \tag{56}$$

So we get

$$\left\langle X' - X, X - C\right\rangle \ge 0, \quad \forall X' \in K. \tag{57}$$

According to Condition 10, we take the following parameter sequences, and let Condition 10 denote the parameter sequences:

$$\begin{aligned} \alpha_n &= \frac{1}{\ln n}, \\ \lambda_n = \lambda'_n = \lambda''_n &= \frac{1}{\ln(n+1)}, \\ \beta_n = \gamma_n &= 1 - \frac{1}{\ln n}, \\ \gamma = \rho = t &= c_0 > 0. \end{aligned} \tag{58}$$

According to Condition 12, we take the following parameter sequences, and let Condition 12 denote the parameter sequences:

$$\begin{aligned} \alpha_n &= 0.8 - \frac{1}{(10 * \ln n)}, \quad n = 2k, \\ \alpha_n &= 0.3 - \frac{1}{(10 * \ln n)}, \quad n = 2k - 1, \end{aligned}$$

Table 1: Numerical results for the PRH method and the EC method.

Asymmetric matrix	$c_0 = 0.1$, $\theta_n = 0.8$, tolerance $= 10^{-4}$						
	Condition 10		Condition 12		EC method		
n	It	cpu	It	cpu	It	cpu	tolerance
100	201	8.34	130	5.35	100	14.46	$8.289e+000$
200	333	75.44	208	47.14	100	94.30	$1.010e+002$
300	443	318.02	272	174.70	100	302.29	$4.899e+002$
400	543	789.16	330	446.00	100	686.83	$9.628e+002$
500	647	1747.70	388	972.18	100	1287.36	$1.756e+003$
1000	1082	19884.30	634	11502.13	100	9220.50	$9.826e+003$
2000	>2000	>150000	1052	128504.67	100	>74640.41	$>5.597e+003$

```
Matlab code:
C = zeros(n, n); HU = ones(n, n) * 0.1; HL = −HU;
for i = 1 : n
   for j = 1 : n
      t = mod(t * 42108 + 13846, 46273);
      C(i, j) = t * 2/46273 − 1;
   end;
end;
for i = 1 : n
   C(i, i) = abs(C(i, i)) * 2; HU(i, i) = 1; HL(i, i) = 1;
end;
```

Algorithm 1

Table 2: Numerical results for tolerance 10^{-4}.

Asymmetric matrix	$c_0 = 0.1$, $\theta_n = 0.8$			
	Condition 10		Condition 12	
n	It	cpu	It	cpu
100	204	8.78	130	5.45
200	330	76.08	208	47.72
300	445	323.20	272	175.89
400	548	867.56	330	450.59
500	663	1916.90	388	994.18

Table 3: Numerical results for tolerance 10^{-3}.

Asymmetric matrix	$c_0 = 0.1$, $\theta_n = 0.8$			
	Condition 10		Condition 12	
n	It	cpu	It	cpu
1000	193	3893.63	126	2280.74
2000	318	42981.02	200	28737.65

$$\lambda_n = \lambda_n' = \lambda_n'' = \frac{1}{\ln(n+1)},$$
$$\beta_n = 1 - \frac{1}{\ln n}, \quad n = 2k,$$
$$\beta_n = 1 - \frac{1}{\ln n}, \quad n = 2k-1,$$
$$\gamma_n = 1 - \frac{1}{\ln n}, \quad n = 2k,$$
$$\gamma_n = 1 - \frac{1}{\ln(2n)}, \quad n = 2k-1,$$
$$\gamma = \rho = t = c_0 > 0. \tag{59}$$

Obviously, we have much difficulty in computing the projection of $P_K[X]$, for all $x \in S^n$. In order to reduce the difficulty and complexity of computing the projection P_K, we define TX by

$$TX = H(G(X)), \tag{60}$$

where

$$G(X) = \min(H_U, \max(X, H_L)),$$
$$H(X) = P_{S_+^n}(X), \tag{61}$$

which can be computed without difficulty and the fixed point set of $\mathrm{Fix}(T) = K$. According to Theorems 11 and 13, the sequences generated by Algorithm 8 under Conditions 10 and 12 are convergent.

The computation begins with ones (n, n) in MATLAB and stops as soon as $\|x_{k+1} - x_k\| \le 10^{-3}$ or 10^{-2}. All codes were implemented in MATLAB 7.1 and ran at a Pentium R 1.70G processor, 2G Acer note computer.

We test the problems with $n = 100, 200, 300, 400, 500, 1000$, and 2000. The test results with the PRH method under

Table 4: Numerical results for tolerance 10^{-4}.

Asymmetric matrix	$\gamma = 0.1, \rho = 0.3, t = 0.1$									
	$\theta_n = 0$		$\theta_n = 0.2$		$\theta_n = 0.4$		$\theta_n = 0.6$		$\theta_n = 0.8$	
n	It	cpu	It	cpu	It	cpu	It	cpu	It	cpu
100	132	5.52	134	5.60	128	5.50	134	5.67	132	5.54
200	210	48.04	206	47.22	208	48.04	204	47.15	214	48.58
300	274	177.49	268	176.08	276	178.80	274	177.68	276	178.84
400	336	468.28	328	445.93	336	468.20	334	454.24	330	453.79
500	392	977.79	394	1012.57	378	948.44	386	953.91	390	971.10

```
Matlab code:
C = zeros(n, n); HU = ones(n, n) * 0.1; HL = −HU;
for i = 1 : n
   for j = 1 : n
      C = −1 + 2 * rand(n);
   end;
end;
for i = 1 : n
   C(i, i) = abs(C(i, i)) * 2; HU(i, i) = 1; HL(i, i) = 1;
end;
```

Algorithm 2

different conditions are reported in Tables 1, 2, 3, and 4. And the CPU time is in seconds. It is to be noted that the results of extended contraction method are only given out when the iteration step (It) is less than or equal to 100.

Test Examples 1. In this example we generate the data in a similar manner as in [9]. The entries of diagonal elements of C are randomly generated in the interval $(0, 2)$; the entries of off-diagonal elements of C are randomly generated in the interval $(-1, 1)$ (Algorithm 1):

$$\begin{gathered}(H_U)_{jj} = (H_L)_{jj} = 1,\\ (H_U)_{ij} = -(H_L)_{ij} = 0.1, \quad \forall i \neq j, i, j = 1, 2, \ldots, n.\end{gathered} \tag{62}$$

When $n \geq 1000$ and tolerance 10^{-4}, the computation time of the proposed method is too long, so the results of the PRH method give out approximate solution with $n \geq 1000$ and tolerance 10^{-3} in the following. And the extended contraction method (EC method) has much difficulty in computing the examples when C is asymmetric. Furthermore, by introducing auxiliary variable, the certain projection method or relaxed-PPA method [10] can be implemented by these tests.

Test Examples 2. We form the data of the second problems similarly as in the first test examples. The entries of diagonal elements of C are randomly generated in the interval $(0, 2)$; the entries of off-diagonal elements of C are generated from a uniform distribution in the same interval (Algorithm 2):

$$\begin{gathered}(H_U)_{jj} = (H_L)_{jj} = 1,\\ (H_U)_{ij} = -(H_L)_{ij} = 0.1, \quad \forall i \neq j, i, j = 1, 2, \ldots, n.\end{gathered} \tag{63}$$

From Tables 1 to 3, we found that the iteration numbers and CPU time of PRH under Condition 12 are more efficient than that under Condition 10. In Table 4 of our method, the tests' results give out that the PRH method under some descent directions is more slightly efficient than those of the MRHSD method [14, 16], and it is easy to obtain these descent directions. Furthermore, it is important to find γ, ρ, and t by Tables 2 and 4.

5. Conclusions

We have proved the strong convergence of PRH method under Condition 12, which differs from Condition 10. The result can be considered as an improvement and refinement of the previous results [14]. And more importantly, numerical experiments demonstrated that the PRH method under Condition 12 is more efficient than that under Condition 10, and the PRH method under some descent directions is more slightly efficient than that of the MRHSD method. How to select parameters of the PRH method for solving variational inequalities is worthy of further investigations in the future.

Acknowledgments

This research was supported by National Science and Technology Support Program (Grant no. 2011BAH24B06), Joint Fund of National Natural Science Foundation of China and Civil Aviation Administration of China (Grant no. U1233105), and Science Foundation of the Civil Aviation Flight University of China (Grant no. J2010-45).

References

[1] M. S. Gowda and Y. Song, "On semidefinite linear complementarity problems," *Mathematical Programming*, vol. 88, no. 3, pp. 575–587, 2000.

[2] I. Yamada, "The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings," in *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, D. Bumariu, Y. Censor, and S. Reich, Eds., vol. 8, pp. 473–504, North-Holland, Amsterdam, The Netherlands, 2001.

[3] F. Deutsch and I. Yamada, "Minimizing certain convex functions over the intersection of the fixed point sets of nonexpansive mappings," *Numerical Functional Analysis and Optimization*, vol. 19, no. 1-2, pp. 33–56, 1998.

[4] H. K. Xu and T. H. Kim, "Convergence of hybrid steepest-descent methods for variational inequalities," *Journal of Optimization Theory and Applications*, vol. 119, no. 1, pp. 185–201, 2003.

[5] L. C. Zeng, N. C. Wong, and J. C. Yao, "Convergence analysis of modified hybrid steepest-descent methods with variable parameters for variational inequalities," *Journal of Optimization Theory and Applications*, vol. 132, no. 1, pp. 51–69, 2007.

[6] L. C. Zeng, "On a general projection algorithm for variational inequalities," *Journal of Optimization Theory and Applications*, vol. 97, no. 1, pp. 229–235, 1998.

[7] X. P. Ding, Y. C. Lin, and J. C. Yao, "Three-step relaxed hybrid steepest-descent methods for variational inequalities," *Applied Mathematics and Mechanics*, vol. 28, no. 8, pp. 1029–1036, 2007.

[8] B. S. He, "A new method for a class of linear variational inequalities," *Mathematical Programming*, vol. 66, no. 2, pp. 137–144, 1994.

[9] B. S. He and M. H. Xu, "A general framework of contraction methods for monotone variational inequalities," *Pacific Journal of Optimization*, vol. 4, no. 2, pp. 195–212, 2008.

[10] B. S. He, "PPA-based contraction methods for general linearly constrained convex optimization," *Lectures of Contraction Methods for Convex Optimization and Monotone Variational Inequalities*, 06C, 2012, http://math.nju.edu.cn/~hebma/.

[11] N. J. Huang, X. X. Huang, and X. Q. Yang, "Connections among constrained continuous and combinatorial vector optimization," *Optimization*, vol. 60, no. 1-2, pp. 15–27, 2011.

[12] P. T. Harker and J. S. Pang, "A damped-Newton method for the linear complementarity problem," in *Computational Solution of Nonlinear Systems of Equations*, vol. 26, pp. 265–284, American Mathematical Society, Providence, RI, USA, 1990.

[13] T. Suzuki, "Strong convergence of Krasnoselskii and Mann's type sequences for one-parameter nonexpansive semigroups without Bochner integrals," *Journal of Mathematical Analysis and Applications*, vol. 305, no. 1, pp. 227–239, 2005.

[14] H. W. Xu, E. B. Song, H. P. Pan, H. Shao, and L. M. Sun, "The modified and relaxed hybrid steepestdescent methods for variational inequalities," in *Proceedings of the 1st International Conference on Modelling and Simulation*, vol. 2, pp. 169–174, World Academic Press, 2008.

[15] H. W. Xu, H. Shao, and Q. C. Zhang, "The Prediction-correction and relaxed hybrid steepest-descent method for variational inequalities," in *Proceedings of the International Symposium on Education and Computer Science*, vol. 1, pp. 252–256, IEEE Computer Society and Academy, 2009.

[16] H. W. Xu, "Efficient implementation of a modified and relaxed hybrid steepest-descent method for a type of variational inequality," *Journal of Inequalities and Applications*, vol. 2012, article 93, 2012.

[17] J. H. Hammond, *Solving asymmetric variational inequality problems and systems of equations with generalized nonlinear programming algorithms [Ph.D. dissertation]*, Department of Mathematics, MIT, Cambridge, Mass, USA, 1984.

[18] P. Tseng, "Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming," *Mathematical Programming*, vol. 48, no. 2, pp. 249–263, 1990.

[19] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[20] D. F. Sun, "A projection and contraction method for generalized nonlinear complementarity problems," *Mathematica Numerica Sinica*, vol. 16, no. 2, pp. 183–194, 1994.

[21] Y. Gao and D. F. Sun, "Calibrating least squares covariance matrix problems with equality and inequality constraints," Tech. Rep., Department of Mathematics, National University of Singapore, 2008.

[22] M. A. Noor, "Some recent advances in variational inequalities. I. Basic concepts," *New Zealand Journal of Mathematics*, vol. 26, no. 1, pp. 53–80, 1997.

[23] M. A. Noor, "New approximation schemes for general variational inequalities," *Journal of Mathematical Analysis and Applications*, vol. 251, no. 1, pp. 217–229, 2000.

[24] Y. Yao, M. A. Noor, R. Chen, and Y.-C. Liou, "Strong convergence of three-step relaxed hybrid steepest-descent methods for variational inequalities," *Applied Mathematics and Computation*, vol. 201, no. 1-2, pp. 175–183, 2008.

[25] "Advances in Equilibrium Modeling, Analysis, and Computation," in *Annals of Operations Research*, A. Nagurney, Ed., vol. 44, J. C. Baltzer AG Scientific Publishing, Basel, Switzerland, 1993.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 941861, 8 pages
http://dx.doi.org/10.1155/2013/941861

Research Article

A New Implementable Prediction-Correction Method for Monotone Variational Inequalities with Separable Structure

Feng Ma, Mingfang Ni, and Zhanke Yu

Institute 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; Accepted 27 August 2013

Academic Editor: Abdellah Bnouhachem

The monotone variational inequalities capture various concrete applications arising in many areas. In this paper, we develop a new prediction-correction method for monotone variational inequalities with separable structure. The new method can be easily implementable, and the main computational effort in each iteration of the method is to evaluate the proximal mappings of the involved operators. At each iteration, the algorithm also allows the involved subvariational inequalities to be solved in parallel. We establish the global convergence of the proposed method. Preliminary numerical results show that the new method can be competitive with Chen's proximal-based decomposition method in Chen and Teboulle (1994).

1. Introduction

The variational inequality (VI (Ω, F)) in the finite-dimensional space is to determine a vector $u \in \Omega$ such that

$$\left\langle u' - u, F(u) \right\rangle \geq 0, \quad \forall u' \in \Omega, \tag{1}$$

where $\Omega \in \mathfrak{R}^n$ is a nonempty closed convex subset and F is a continuous mapping from $\mathfrak{R}^n$ into itself. The VI (Ω, F) has found many efficient applications in a broad spectrum of areas such as traffic equilibrium [1] and network economic problems [2]. For solving (1), the proximal point algorithm (PPA), which was proposed by Martinet [3] and further studied by Rockafellar [4, 5], generates the new iterative point u^{k+1} via the following procedure:

$$\left\langle u' - u^{k+1}, F\left(u^{k+1}\right) + G\left(u^{k+1} - u^k\right) \right\rangle \geq 0, \quad \forall u' \in \Omega, \tag{2}$$

where $G \in \mathfrak{R}^{n\times n}$ is a positive definite matrix, playing the role of proximal regularization parameter. Note that the PPA has to solve systems of nonlinear equations in each iteration. In many cases, solving these equations is quite difficult. This difficulty has inspired the burst of approximate versions of the PPAs, in order to approximately solve (2) under certain "relative error." These new methods include well-known-extragradient type methods (EGM) as special cases. Assume that F is Lipschitz continuous; that is, there is $l \in (0, 1)$, such that

$$\beta \left\| F\left(u^k\right) - F\left(\tilde{u}^k\right) \right\| \leq l \left\| u^k - \tilde{u}^k \right\|. \tag{3}$$

Then at each iteration EGM takes the following general form:

$$\begin{gathered} \left\langle u' - \tilde{u}, \beta F\left(u^k\right) + \tilde{u} - u^k \right\rangle \geq 0, \quad \forall u' \in \Omega, \\ \left\langle u' - u^{k+1}, \beta F(\tilde{u}) + u^{k+1} - u^k \right\rangle \geq 0, \quad \forall u' \in \Omega. \end{gathered} \tag{4}$$

In this paper, we consider the following variational inequalities: find a vector $w \in \mathscr{D}$ such that

$$\left\langle w' - w, F(w) \right\rangle \geq 0, \quad \forall w' \in \mathscr{D}, \tag{5}$$

with

$$\begin{gathered} w := \begin{pmatrix} x \\ y \end{pmatrix}, \qquad F(w) := \begin{pmatrix} f(x) \\ g(y) \end{pmatrix}, \\ \mathscr{D} = \left\{ (x, y) \mid x \in \mathscr{X}, y \in \mathscr{Y}, Ax + By = b \right\}, \end{gathered} \tag{6}$$

where $\mathscr{D} \in \mathfrak{R}^{n+p}$ is a nonempty closed convex subset and $f : \mathscr{X} \to \mathfrak{R}^n$ and $g : \mathscr{Y} \to \mathfrak{R}^p$ are monotone operators. Problem (5) is referred to as a structured variational inequality (SVI) [6].

By attaching a Lagrange multiplier vector $\lambda \in \mathfrak{R}^m$ to the linear constraints $Ax + By = b$, the VI problem (5) is converted into the following form:

$$\left\langle \begin{matrix} x' - x, & f(x) - A^T\lambda \\ y' - y, & g(y) - B^T\lambda \\ \lambda' - \lambda, & Ax + By - b \end{matrix} \right\rangle \geq 0, \quad \forall u' \in \Omega, \tag{7}$$

where

$$\Omega = \mathcal{X} \times \mathcal{Y} \times \mathfrak{R}^m. \tag{8}$$

The compact form is

$$\left\langle u' - u, F(u) \right\rangle \geq 0, \quad \forall u' \in \Omega, \tag{9}$$

with

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

For the purpose of parallel computing, the proximal alternating directions method (PADM) generates $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \Omega$ as follows [7, 8]: first find an $\tilde{x}^k \in \mathcal{X}$ such that

$$\begin{aligned} &\left\langle x' - \tilde{x}^k, f\left(\tilde{x}^k\right) - A^T\left[\lambda^k - \beta\left(A\tilde{x}^k + y^k - b\right)\right] \right. \\ &\quad \left. + r\left(\tilde{x}^k - x^k\right)\right\rangle \geq 0, \quad \forall x \in \mathcal{X}. \end{aligned} \tag{11}$$

Then find an $\tilde{y}^k \in \mathcal{Y}$ such that

$$\begin{aligned} &\left\langle y' - \tilde{y}^k, g\left(\tilde{y}^k\right) - B^T\left[\lambda^k - \beta\left(A\tilde{x}^k + \tilde{y}^k - b\right)\right] \right. \\ &\quad \left. + s\left(\tilde{y}^k - y^k\right)\right\rangle \geq 0, \quad \forall y \in \mathcal{Y}. \end{aligned} \tag{12}$$

Finally, update $\tilde{\lambda}^k$ via

$$\tilde{\lambda}^k = \lambda^k - \beta\left(A\tilde{x}^k + B\tilde{y}^k - b\right). \tag{13}$$

Here $r \geq 0$ and $s \geq 0$ are given proximal parameters; $\beta \geq 0$ is a given penalty parameter for the linearly constraints. Note that when $r = s = 0$ in (11)-(12), the classical alternating directions method (ADM) is recovered. To make the PADM (11)–(13) more efficient and flexible, some strategies have been developed. For example, allow r, s, and β to vary from iteration to iteration according to certain strategies [8–10]; produce the new iterate based on the minor correction to the predictor. A simple and effective correction scheme is (see, e.g., [11, 12])

$$u^{k+1} = u^k - \alpha_k\left(u^k - \tilde{u}^k\right), \tag{14}$$

where $\alpha_k > 0$ is a chosen step size.

The PADM (11)–(13) is often easy to implement under the assumption that the decomposed subproblems have closed-form solutions or can be efficiently solved up to a high precision. However, in some cases, matrixes A and B are not identity matrices, and the two subproblems in PADM (11)-(12) are difficult to solve because the evaluation of $(A^TA + (1/\beta)f)^{-1}(Av)$ and $(B^TB + (1/\beta)g)^{-1}(Bv)$ could be costly. To overcome this difficulty, we propose a new implementable prediction-correction method for the SVI. At each iteration, we first decompose the problem to two small problems with respect to x and y, respectively. The two subproblems are all easy to solve under the assumption that the resolvent operators of f and g are easy to evaluate, where the resolvent operator of mapping T is defined as $(I + \lambda T)^{-1}(v)$. Then, we update the Lagrange multipliers and make a correction step to ensure the algorithm's convergence.

The SVI has been studied extensively both in the theoretical frameworks and applications. Recently, Han [13] proposed a hybrid entropic proximal decomposition method for the SVI. Han's method is based on logarithmic-quadratic functions and combined with self-adaptive strategy. He [14] presented a parallel splitting augmented Lagrangian method which can be extended to solve the system of equilibrium problems with three separable operators. Xu et al. [15] proposed two classes of correction methods for the SVI in which the mapping F does not have an explicit form. Besides, Xu and Wu [16] also studied a class of linearized proximal alternating direction methods and showed that the relaxation factor can have the same restriction region as for the general ADM. Yuan and Li [17] developed a logarithmic-quadratic-proximal- (LQP-) based decomposition method by applying the LQP terms to regularize the ADM subproblems; then Bnouhachem et al. [18] studied a new inexact LQP alternating direction method by solving a series of related systems of nonlinear equations.

The rest of this paper is organized as follows. In Section 2, we review some preliminaries which are useful for further analysis. In Section 3, we present the new implementable prediction-correction method for SVI, and the global convergence result is established. Numerical experiments and some conclusions are addressed in Sections 4 and 5, respectively.

2. Preliminaries

In this section, we make some standard assumptions and summarize some basic properties of VI which will be used in the subsequent discussions.

Assumption

(A1) $\mathcal{X}, \mathcal{Y}$ are simple closed convex sets.

A set which is said to be *simple* means that the projection onto the set is easy to compute, where the projection of a point v onto the closed convex set Ω, denoted by $P_\Omega(v)$, is defined as the nearest point $u \in \Omega$ to v; that is,

$$P_\Omega(v) = \arg\min\{\|u - v\| \mid u \in \Omega\}. \tag{15}$$

(A2) The mapping F is point-to-point, monotone, and continuous.

A mapping $F : \mathfrak{R}^n \to \mathfrak{R}^n$ is said to be *monotone* on Ω if

$$\langle u - v, F(u) - F(v)\rangle \geq 0, \quad \forall u, v \in \Omega. \tag{16}$$

(A3) The solution set of SVI (Ω, F), denoted by Ω^*, is nonempty.

Properties. Let G be a symmetric positive definite matrix; the G-norm of the vector u is denoted by $\|u\|_G := \sqrt{\langle u, Gu\rangle}$. In particular, when $G = I$, $\|u\| := \sqrt{\langle u, u\rangle}$ is the Euclidean norm of u. For a matrix A, $\|A\|$ denotes its norm $\|A\| := \max\{\|Ax\| : \|x\| \le 1\}$.

The following well-known properties of the projection operator will be used in the coming analysis.

Lemma 1. *Let $\Omega \in \mathfrak{R}^n$ be a nonempty closed convex set; let $P_\Omega(\cdot)$ be the projection operator onto Ω under the G-norm. Then*

$$\begin{gathered}\left\langle u' - P_\Omega\left(u'\right), G\left(u - P_\Omega\left(u'\right)\right)\right\rangle \le 0, \quad \forall u' \in \mathfrak{R}^n,\ \forall u \in \Omega,\\ \left\|P_\Omega(u) - P_\Omega\left(u'\right)\right\|_G \le \left\|u - u'\right\|_G, \quad \forall u,\ u' \in \mathfrak{R}^n,\\ \left\|u - P_\Omega\left(u'\right)\right\|_G^2 \le \left\|u - u'\right\|_G^2 - \left\|u' - P_\Omega\left(u'\right)\right\|_G^2,\\ \forall u' \in \mathfrak{R}^n,\ \forall u \in \Omega.\end{gathered} \tag{17}$$

For any arbitrary positive scalar β and $u \in \Omega$, let $e(u, \beta)$ denote the residual function associated with the mapping F; that is,

$$e(u, \beta) = u - P_\Omega\left[u - \beta F(u)\right]. \tag{18}$$

Lemma 2. *u^* is a solution of the SVI (Ω, F) if and only if $e(u^*, \beta) = 0$ for any given positive constant β (see [2, page 267]).*

Lemma 3. *Solving SVI (Ω, F) (7) is equivalent to find a zero point of the mapping*

$$\begin{aligned} e(u, \beta) &:= \begin{pmatrix} e_1(u, \beta) \\ e_2(u, \beta) \\ e_3(u, \beta) \end{pmatrix} \\ &= \begin{pmatrix} x - P_{\mathcal{X}}\left\{x - \beta\left[f(x) - A^T\lambda\right]\right\} \\ y - P_{\mathcal{Y}}\left\{y - \beta\left[g(y) - B^T\lambda\right]\right\} \\ \beta(Ax + By - b) \end{pmatrix}. \end{aligned} \tag{19}$$

3. The New Algorithm

In this section, we present a new prediction-correction method for SVI (Ω, F) and show its global convergence. But, at the beginning, to make the algorithm more succinct, we first define some matrices:

$$H = \begin{pmatrix} rI & 0 & 0 \\ 0 & sI & 0 \\ 0 & 0 & \frac{1}{\beta}I \end{pmatrix}, \qquad M = \begin{pmatrix} I & 0 & \frac{1}{r}A^T \\ 0 & I & \frac{1}{s}B^T \\ 0 & 0 & I \end{pmatrix}, \qquad Q = \begin{pmatrix} rI & 0 & A^T \\ 0 & sI & B^T \\ 0 & 0 & \frac{1}{\beta}I \end{pmatrix}. \tag{20}$$

Obviously, H is a symmetric positive definite matrix whenever $r > 0$, $s > 0$, and $\beta > 0$, and we also have $Q = HM$.

3.1. Description of the Algorithm

Algorithm 4. It is a prediction-correction-based algorithm for the SVI (Ω, F).

Phase 1 (initialization step). Given a small number $\epsilon > 0$, let $\gamma \in (0, 2)$; matrixes Q, M are defined in (20). Take $u^0 \in \mathfrak{R}^{n+p+m}$; set $k = 0$. Choose the parameters $r > 0$, $s > 0$, and $\beta > 0$ such that

$$r > 2\beta\left\|A^TA\right\|, \qquad s > 2\beta\left\|B^TB\right\|. \tag{21}$$

Phase 2 (prediction step). Generate the predictor $\tilde{x}^k$ via solving the following projection equation:

$$\tilde{x}^k = P_{\mathcal{X}}\left[x^k - \frac{1}{r}\left(f\left(\tilde{x}^k\right) - A^T\lambda^k\right)\right]. \tag{22}$$

Then find an $\tilde{y}^k \in \mathcal{Y}$ such that

$$\tilde{y}^k = P_{\mathcal{Y}}\left[y^k - \frac{1}{s}\left(g\left(\tilde{y}^k\right) - B^T\lambda^k\right)\right]. \tag{23}$$

Finally, update $\tilde{\lambda}^k$ via

$$\tilde{\lambda}^k = \lambda^k - \beta\left(A\tilde{x}^k + B\tilde{y}^k - b\right). \tag{24}$$

Phase 3 (correction step). Correct the predictor, and generate the new iterate u^{k+1} via

$$u^{k+1} = u^k - \alpha_k M\left(u^k - \tilde{u}^k\right), \tag{25}$$

where

$$\alpha_k = \gamma\alpha_k^*, \qquad \alpha_k^* = \frac{\left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle}{\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2}. \tag{26}$$

Phase 4 (convergence verification). If $\|u^k - u^{k+1}\| \le \epsilon$, stop; otherwise set $k := k + 1$; go to Phase 2.

Remark 5. Note that (22) does not involve $\tilde{y}^k$ and that (23) is independent on the $\tilde{x}^k$ generated by (22). Hence the two projections (22) and (23) are eligible for parallel computation.

Remark 6. It is easy to check that $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k)$ is a solution of SVI (Ω, F) if and only if $Ax^k = A\tilde{x}^k$, $By^k = B\tilde{y}^k$, and $\lambda^k = \tilde{\lambda}^k$. Thus, it is reasonable to take the magnitude of $\|u^k - u^{k+1}\| \le \epsilon$ as the stopping criterion.

Remark 7. The strategy of choosing the step size α_k in the correction step which coincides with the strategy in He's papers, see, for example, [19], will be explained in detail in the following section.

Remark 8. Our method and the methods proposed in [6, 15, 20] are all in the prediction-correction algorithmic framework, where at each iteration they make a prediction step to produce a predictor and a correction step to generate the new iterate via correcting this predictor.

3.2. Contractive Properties. Now, we start to prove some properties of the sequence $\{\tilde{u}^k\}$. The first lemma quantifies the discrepancy between the point $\tilde{u}^k$ and a solution point of SVI (Ω, F).

Lemma 9. *Let $\{\tilde{u}\}$ be generated by* (22)–(24)*, and let the matrix M be given in* (20)*. Then one has*

$$\left\langle u' - \tilde{u}^k, F(\tilde{u}) - Q\left(u^k - \tilde{u}^k\right)\right\rangle \ge 0, \quad \forall u' \in \Omega. \tag{27}$$

Proof. Note that $\tilde{u}^k$ generated by (22)–(24) are actually solutions of the following VIs:

$$\left\langle x' - \tilde{x}^k, f\left(\tilde{x}^k\right) - A^T\lambda^k - r\left(x^k - \tilde{x}^k\right)\right\rangle \ge 0, \quad \forall x \in \mathcal{X}, \tag{28}$$

$$\left\langle y' - \tilde{y}^k, g\left(\tilde{y}^k\right) - B^T\lambda^k - s\left(y^k - \tilde{y}^k\right)\right\rangle \ge 0, \quad \forall y \in \mathcal{Y}, \tag{29}$$

$$\left\langle \lambda' - \tilde{\lambda}^k, A\tilde{x}^k + B\tilde{y}^k - b - \frac{1}{\beta}\left(\lambda^k - \tilde{\lambda}^k\right)\right\rangle \ge 0, \quad \forall \lambda \in \mathfrak{R}^m. \tag{30}$$

Combining (28)–(30) together, we have

$$\left\langle \begin{array}{ll} x' - \tilde{x}^k, & f\left(\tilde{x}^k\right) - A^T\tilde{\lambda}^k - A^T\left(\lambda^k - \tilde{\lambda}^k\right) - r\left(x^k - \tilde{x}^k\right) \\ y' - \tilde{y}^k, & g\left(\tilde{y}^k\right) - B^T\tilde{\lambda}^k - B^T\left(\lambda^k - \tilde{\lambda}^k\right) - s\left(y^k - \tilde{y}^k\right) \\ \lambda' - \tilde{\lambda}, & A\tilde{x}^k + B\tilde{y}^k - b - \frac{1}{\beta}\left(\lambda^k - \tilde{\lambda}^k\right) \end{array}\right\rangle \ge 0, \quad \forall u' \in \Omega. \tag{31}$$

Using the notations of F (see (10)) and Q (see (20)), the earlier inequality can be rewritten into

$$\left\langle u' - \tilde{u}^k, F(\tilde{u}) - Q\left(u^k - \tilde{u}^k\right)\right\rangle \ge 0, \quad \forall u' \in \Omega. \tag{32}$$

The assertion (27) is thus proved. □

The following lemma plays a key role in proving the convergence of the algorithm.

Lemma 10. *Let matrixes Q, H be defined in* (20)*, if the parameters $r > 0, s > 0$, and $\beta > 0$ in* (22)–(24) *satisfy*

$$r > 2\beta\left\|A^TA\right\|, \qquad s > 2\beta\left\|B^TB\right\|. \tag{33}$$

Then for the matrix Q in (27)*, one has*

$$\langle u - \tilde{u}, Q(u - \tilde{u})\rangle \ge \left(1 - \frac{\mu}{2}\right)\|u - \tilde{u}\|_H^2 \qquad \forall u \ne \tilde{u} \in \mathfrak{R}^{n+p+m}, \tag{34}$$

with

$$\mu = \sqrt{\max\left\{\frac{2\beta\left\|A^TA\right\|}{r}, \frac{2\beta\left\|B^TB\right\|}{s}\right\}} \in (0, 1). \tag{35}$$

Proof. For any $u \ne \tilde{u}$, we have

$$\langle u - \tilde{u}, Q(u - \tilde{u})\rangle = \|u - \tilde{u}\|_H^2 + \left\langle \lambda - \tilde{\lambda}, A(x - \tilde{x})\right\rangle + \left\langle \lambda - \tilde{\lambda}, y - \tilde{y}\right\rangle. \tag{36}$$

According to the Cauchy-Schwarz inequality, we get

$$\begin{aligned} &\langle \lambda - \tilde{\lambda}, A(x - \tilde{x})\rangle + \langle \lambda - \tilde{\lambda}, y - \tilde{y}\rangle \\ &\quad = \frac{1}{2}\left(2\left\langle \lambda - \tilde{\lambda}, A(x - \tilde{x})\right\rangle + 2\left\langle \lambda - \tilde{\lambda}, B(y - \tilde{y})\right\rangle\right) \\ &\quad \ge -\frac{1}{2}\left\{\frac{2\beta}{\mu}\|A(x - \tilde{x})\|^2 + \frac{\mu}{2\beta}\left\|\lambda - \tilde{\lambda}\right\|^2\right\} \\ &\qquad - \frac{1}{2}\left\{\frac{2\beta}{\mu}\|B(y - \tilde{y})\|^2 + \frac{\mu}{2\beta}\left\|\lambda - \tilde{\lambda}\right\|^2\right\} \\ &\quad = -\frac{1}{2}\left\{\frac{2\beta}{\mu}\|A(x - \tilde{x})\|^2 + \frac{2\beta}{\mu}\right. \\ &\qquad \left. \times \|B(y - \tilde{y})\|^2 + \frac{\mu}{\beta}\left\|\lambda - \tilde{\lambda}\right\|^2\right\}. \end{aligned} \tag{37}$$

With the μ defined in (35), we have

$$\begin{aligned} \frac{2\beta}{\mu}\|A(x - \tilde{x})\|^2 &\le \mu r\|x - \tilde{x}\|^2, \\ \frac{2\beta}{\mu}\|B(y - \tilde{y})\|^2 &\le \mu s\|y - \tilde{y}\|^2. \end{aligned} \tag{38}$$

Substituting (38) into (37), combining (36), the assertion (34) is proved. □

Lemma 11. *Suppose that $u^* = (x^*, y^*, \lambda^*) \in \Omega$ is a solution point of* (9) *and the sequences $\{u^{k+1}\}$ are corrected by an undeterminate step size denoted by α instead of* (26)*; that is,*

$$u^{k+1} = u^k - \alpha M\left(u^k - \tilde{u}^k\right). \tag{39}$$

Then one has

$$\vartheta^k(\alpha) \geq \varphi^k(\alpha), \tag{40}$$

where

$$\vartheta^k(\alpha) = \left\|u^k - u^*\right\|_H^2 - \left\|u^{k+1} - u^*\right\|_H^2,$$
$$\varphi^k(\alpha) = 2\alpha\left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle - \alpha^2\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2. \tag{41}$$

Proof. One can see that

$$\begin{aligned}\vartheta^k(\alpha) &= \left\|u^k - u^*\right\|_H^2 - \left\|u^{k+1} - u^*\right\|_H^2 \\ &= \left\|u^k - u^*\right\|_H^2 - \left\|u^k - \alpha M\left(u^k - \tilde{u}^k\right) - u^*\right\|_H^2 \\ &= 2\alpha\left\langle u^k - u^*, HM\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\quad - \alpha^2\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2.\end{aligned} \tag{42}$$

On the other hand, since $Q = HM$, using the monotonicity of F and Lemma 9, we have

$$\begin{aligned}\left\langle u^k - u^*, HM\left(u^k - \tilde{u}^k\right)\right\rangle &= \left\langle u^k - u^*, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &= \left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\quad + \left\langle \tilde{u}^k - \tilde{u}^*, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\geq \left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\quad + \left\langle \tilde{u}^k - u^*, F\left(\tilde{u}^k\right)\right\rangle \\ &\geq \left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\quad + \left\langle \tilde{u}^k - u^*, F\left(u^*\right)\right\rangle \\ &\geq \left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle.\end{aligned} \tag{43}$$

Combining (42)-(43) together, we have

$$\begin{aligned}\vartheta^k(\alpha) &= 2\alpha\left\langle u^k - u^*, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\quad - \alpha^2\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2 \\ &\geq 2\alpha\left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\quad - \alpha^2\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2 \\ &= \varphi^k(\alpha).\end{aligned} \tag{44}$$

Thus, $\varphi^k(\alpha)$ is a lower bound of $\vartheta^k(\alpha)$ for any $\alpha > 0$. □

Remark 12. Note that $\varphi^k(\alpha)$ is a quadratic function of α and it reaches its maximum at

$$\alpha_k^* = \frac{\left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle}{\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2}. \tag{45}$$

Hence, it is reasonable to use the step size strategy (26). The parameter γ in (26) plays the role of a relaxation or scaling parameter. We can easily see that $\gamma \in (0, 2)$ can ensure convergence.

Now, we prove the Fejér monotonicity of the iterative sequence $\{u^k\}$ generated by the algorithm.

Theorem 13. *Suppose that $u^* = (x^*, y^*, \lambda^*) \in \Omega$ is a solution point of (9) and the sequences $\{u^k\}$ are generated by the algorithm. Then*

$$\begin{aligned}\left\|u^{k+1} - u^*\right\|_H^2 &\leq \left\|u^k - u^*\right\|_H^2 \\ &\quad - \frac{1}{2}r(2-r)\left(1 - \frac{\mu}{2}\right)\left\|u^k - \tilde{u}^k\right\|_H^2.\end{aligned} \tag{46}$$

Proof. According to Lemma 11,

$$\begin{aligned}\left\|u^{k+1} - u^*\right\|_H^2 &\leq \left\|u^k - u^*\right\|_H^2 - \varphi^k(\alpha_k) \\ &= \left\|u^k - u^*\right\|_H^2 - \left(2\alpha_k\left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle\right. \\ &\qquad \left. - \alpha_k^2\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2\right) \\ &= \left\|u^k - u^*\right\|_H^2 - \gamma(2-\gamma)\alpha_k^* \\ &\quad \times \left\langle u^k - \tilde{u}^k, Q\left(u^k - \tilde{u}^k\right)\right\rangle \\ &\leq \left\|u^k - u^*\right\|_H^2 - \gamma(2-\gamma)\alpha_k^*\left(1 - \frac{\mu}{2}\right)\|u - \tilde{u}\|_H^2.\end{aligned} \tag{47}$$

Moreover, it follows from (26) that the step size

$$\begin{aligned}\alpha_k^* &= \frac{\left\|u^k - \tilde{u}^k\right\|_{Q+Q^T}^2}{2\left\|M\left(u^k - \tilde{u}^k\right)\right\|_H^2} = \frac{\left\|u^k - \tilde{u}^k\right\|_{Q+Q^T}^2}{2\left\|u^k - \tilde{u}^k\right\|_{M^THM}^2} \\ &= \frac{1}{2}\left(\left(\left\|u^k - \tilde{u}^k\right\|_{M^THM}^2 + r\left\|x^k - \tilde{x}^k\right\|^2 + s\left\|y^k - \tilde{y}^k\right\|^2\right.\right. \\ &\qquad + \left(\frac{1}{\beta}\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 - \frac{1}{r}\left\|A^T\left(\lambda^k - \tilde{\lambda}^k\right)\right\|^2\right. \\ &\qquad \left.\left. - \frac{1}{s}\left\|B^T\left(\lambda^k - \tilde{\lambda}^k\right)\right\|^2\right)\right) \\ &\qquad \left.\times\left(\left\|u^k - \tilde{u}^k\right\|_{M^THM}^2\right)^{-1}\right) \\ &\geq \frac{1}{2}\left(\left(\left\|u^k - \tilde{u}^k\right\|_{M^THM}^2 + r\left\|x^k - \tilde{x}^k\right\|^2 + s\left\|y^k - \tilde{y}^k\right\|^2\right.\right. \\ &\qquad \left.+ \left(\frac{1}{\beta} - \frac{1}{r}\left\|A^TA\right\| - \frac{1}{s}\left\|B^TB\right\|\right)\left\|\lambda^k - \tilde{\lambda}^k\right\|^2\right) \\ &\qquad \left.\times\left(\left\|u^k - \tilde{u}^k\right\|_{M^THM}^2\right)^{-1}\right).\end{aligned} \tag{48}$$

Based on the conditions (33), we have

$$r\left\|x^k-\tilde{x}^k\right\|^2+s\left\|y^k-\tilde{y}^k\right\|^2 + \left(\frac{1}{\beta}-\frac{1}{r}\left\|A^TA\right\|-\frac{1}{s}\left\|B^TB\right\|\right)\left\|\lambda^k-\tilde{\lambda}^k\right\|^2\geq 0. \tag{49}$$

Hence,

$$\alpha_k^*\geq\frac{1}{2}\frac{\left\|u^k-\tilde{u}^k\right\|_{M^THM}^2}{\left\|u^k-\tilde{u}^k\right\|_{M^THM}^2}=\frac{1}{2}. \tag{50}$$

Substituting (50) into (47), we have

$$\left\|u^{k+1}-u^*\right\|_H^2\leq\left\|u^k-u^*\right\|_H^2 -\frac{1}{2}r(2-r)\left(1-\frac{\mu}{2}\right)\left\|u^k-\tilde{u}^k\right\|_H^2. \tag{51}$$

Thus, we obtain the assertion of this theorem. □

Based on the earlier results, we are now ready to prove the global convergence of the algorithm.

Theorem 14. *The sequence $\{u^k\}$ generated by the proposed algorithm converges to a solution of SVI (Ω, F).*

Proof. We prove the convergence of the proposed algorithm by following the standard analytic framework of contraction-type methods. It follows from (46) that $\{u^k\}$ is bounded, and we have that

$$\lim_{k\to\infty}\left\|u^k-\tilde{u}^k\right\|_H=0. \tag{52}$$

Consequently,

$$\lim_{k\to\infty}\left\|x^k-\tilde{x}^k\right\|=0,\qquad\lim_{k\to\infty}\left\|y^k-\tilde{y}^k\right\|=0, \lim_{k\to\infty}\left\|A\tilde{x}^k+B\tilde{y}^k-b\right\|=\lim_{k\to\infty}\left\|\frac{1}{\beta}\left(\lambda^k-\tilde{\lambda}^k\right)\right\|=0, \tag{53}$$

since (see (22) and (23))

$$\begin{aligned}\tilde{x}^k&=P_{\mathcal{X}}\left[\tilde{x}^k-\frac{1}{r}\left(f\left(\tilde{x}^k\right)-A^T\tilde{\lambda}^k\right)+\left(x^k-\tilde{x}^k\right)+\frac{1}{r}A^T\left(\lambda^k-\tilde{\lambda}^k\right)\right],\\ \tilde{y}^k&=P_{\mathcal{Y}}\left[\tilde{y}^k-\frac{1}{s}\left(g\left(\tilde{y}^k\right)-B^T\tilde{\lambda}^k\right)+\left(y^k-\tilde{y}^k\right)+\frac{1}{s}B^T\left(\lambda^k-\tilde{\lambda}^k\right)\right],\\ \tilde{\lambda}^k&=\lambda^k-\beta\left(A\tilde{x}^k+B\tilde{y}^k-b\right).\end{aligned} \tag{54}$$

It follows from (53) that

$$\begin{aligned}&\lim_{k\to\infty}\tilde{x}^k-P_{\mathcal{X}}\left[\tilde{x}^k-\frac{1}{r}\left(f\left(\tilde{x}^k\right)-A^T\tilde{\lambda}^k\right)\right]=0,\\ &\lim_{k\to\infty}\tilde{y}^k-P_{\mathcal{Y}}\left[\tilde{y}^k-\frac{1}{s}\left(g\left(\tilde{y}^k\right)-B^T\tilde{\lambda}^k\right)\right]=0,\\ &\lim_{k\to\infty}A\tilde{x}^k+B\tilde{y}^k-b=0.\end{aligned} \tag{55}$$

Because $\tilde{u}^k$ is also bounded, it has at least one cluster point. Let u^∞ be a cluster point of $\tilde{u}^k$, and let $\tilde{u}^{k_j}$ be the subsequence converging to u^∞. It follows from (55) that

$$\begin{aligned}&\lim_{j\to\infty}\tilde{x}^{k_j}-P_{\mathcal{X}}\left[\tilde{x}^{k_j}-\frac{1}{r}\left(f\left(\tilde{x}^{k_j}\right)-A^T\tilde{\lambda}^{k_j}\right)\right]=0,\\ &\lim_{j\to\infty}\tilde{y}^{k_j}-P_{\mathcal{Y}}\left[\tilde{y}^{k_j}-\frac{1}{s}\left(g\left(\tilde{y}^{k_j}\right)-B^T\tilde{\lambda}^{k_j}\right)\right]=0,\\ &\lim_{j\to\infty}A\tilde{x}^{k_j}+B\tilde{y}^{k_j}-b=0.\end{aligned} \tag{56}$$

Consequently,

$$\begin{aligned}&x^\infty-P_{\mathcal{X}}\left[x^\infty-\frac{1}{r}\left(f\left(x^\infty\right)-A^T\lambda^\infty\right)\right]=0,\\ &y^\infty-P_{\mathcal{Y}}\left[y^\infty-\frac{1}{s}\left(g\left(y^\infty\right)-B^T\lambda^\infty\right)\right]=0,\\ &Ax^\infty+By^\infty-b=0.\end{aligned} \tag{57}$$

Using the continuity of F and the projection operator $P_\Omega(\cdot)$, we have that u^∞ is a solution of SVI (Ω, F).

On the other hand, by taking limits over the subsequences in (52) and using $\lim_{j\to\infty}\tilde{u}^{k_j}=u^\infty$, we have that, for any $k>k_j$, it follows from (46) that

$$\left\|u^k-u^\infty\right\|_H\leq\left\|u^{k_j}-u^\infty\right\|_H. \tag{58}$$

Thus, the sequence $\{u^k\}$ converges to u^∞, which is a solution of SVI (Ω, F). □

4. Numerical Experiments

In this section, we present some numerical experiments results to show the effectiveness of the proposed algorithm. The codes are run on a notebook computer with Inter(R) Core(TM) 2 CPU 2.0 GHZ and RAM 2.00 GM under MATLAB Version 2009b.

We consider the following optimization problem:

$$\begin{aligned}\min_{x\in\mathfrak{R}^n,y\in\mathfrak{R}^p}\quad&\frac{1}{2}x^TPx+\frac{1}{2}y^TQy\\ \text{s.t.}\quad&Ax+By=b,\end{aligned} \tag{59}$$

where $P\in\mathfrak{R}^{n\times n}$, $Q\in\mathfrak{R}^{p\times p}$ are symmetric positive semidefinite matrixes, $A\in\mathfrak{R}^{m\times n}$, $B\in\mathfrak{R}^{m\times p}$, and $b\in\mathfrak{R}^m$.

Using the KKT condition, the problem (59) can be converted into the following variational inequality: find $w^*=(x^*,y^*,\lambda^*)\in\mathfrak{R}^{n+p+m}$ such that

$$\left\langle\begin{matrix}x'-x^*,\quad Px^*-A^T\lambda^*\\ y'-y^*,\quad Qy^*-B^T\lambda^*\\ \lambda'-\lambda^*,\quad Ax^*+By^*-b\end{matrix}\right\rangle\geq 0,\quad\forall w'\in\mathfrak{R}^{n+p+m}. \tag{60}$$

In this example, we randomly created the input data of the tested collection in the following manner.

Table 1: Numerical results for the example.

m	n	p	PDM			New algorithm		
			Iter.	CPU (s)	Error	Iter.	CPU (s)	Error
10	10	10	237	0.075	9.956×10^{-5}	237	0.075	9.895×10^{-5}
10	15	15	250	0.143	9.758×10^{-5}	250	0.086	9.815×10^{-5}
20	20	20	314	0.115	9.669×10^{-5}	314	0.112	9.728×10^{-5}
20	30	30	372	0.175	9.586×10^{-5}	372	0.178	9.597×10^{-5}
40	50	50	561	3.443	9.631×10^{-5}	561	1.340	9.625×10^{-5}
50	80	80	714	3.534	9.990×10^{-5}	715	1.963	9.892×10^{-5}
60	100	100	842	8.107	9.982×10^{-5}	842	7.274	9.996×10^{-5}
100	120	120	1065	9.773	9.926×10^{-5}	1065	11.786	9.938×10^{-5}
150	200	200	1661	24.451	9.942×10^{-5}	1661	21.366	9.947×10^{-5}
200	250	250	2055	38.037	9.907×10^{-5}	2055	35.020	9.911×10^{-5}
200	300	300	2445	66.520	9.964×10^{-5}	2445	61.673	9.970×10^{-5}

(i) P and Q were generated randomly with eigenvalues in $[5, 10]$ according to the following MATLAB scripts:

$$\begin{aligned}
&P = \text{rand}(n);\ [Q1, R1] = \text{qr}(P),\\
&S = 5 + 5 * \text{rand}(n, 1);\ P = Q1 * \text{diag}(S) * Q1',\\
&Q = \text{rand}(p);\ [Q2, R2] = \text{qr}(Q),\\
&S = 5 + 5 * \text{rand}(m, 1);\ P = Q2 * \text{diag}(S) * Q2'.
\end{aligned}$$

(ii) A and B were generated randomly with singular values in $[0, 3]$, and the maximum singular value is 3 according to the following MATLAB scripts:

$$\begin{aligned}
&A = \text{rand}(m, n);\ [U, S, V] = \text{svd}(A),\\
&S = S/S(1, 1) * 3;\ A = U * S * V',\\
&B = \text{rand}(m, p);\ [U, S, V] = \text{svd}(B),\\
&S = S/S(1, 1) * 3;\ B = U * S * V'.
\end{aligned}$$

(iii) b is generated randomly with $b = \text{rand}(m, 1) * 10$.

According to the data generation, we have $\|A^T A\| = 9$ and $\|B^T B\| = 9$.

To apply (22)–(25) to solve (59), instead of choosing the step length α_k judiciously as (24), we can simply choose $\alpha_k \equiv 1$ by takeing $r = 1/\alpha_k^*$ (since $\alpha_k^* > 1/2$ when $u \neq \tilde{u}$, we have $r \in (0, 2)$ which satisfies the requirement). Then, we obtain the following subproblems which are all easy enough to have closed-form solutions:

$$\begin{aligned}
\tilde{x}^k &= (rI + P)^{-1}\left(rx^k + A^T\lambda^k\right),\\
\tilde{y}^k &= (sI + Q)^{-1}\left(sy^k + B^T\lambda^k\right),\\
\tilde{\lambda}^k &= \lambda^k - \beta\left(A\tilde{x}^k + B\tilde{y}^k - b\right),\\
x^{k+1} &= \tilde{x}^k - \frac{1}{r}A^T\left(\lambda^k - \tilde{\lambda}^k\right),\\
y^{k+1} &= \tilde{y}^k - \frac{1}{s}B^T\left(\lambda^k - \tilde{\lambda}^k\right),\\
\lambda^{k+1} &= \tilde{\lambda}^k.
\end{aligned} \tag{61}$$

For comparison, we also solve it by the parallel decomposition method (denoted by PDM) that has been studied extensively in the literature (e.g., [21, 22]). For PDM, the restrictions on the proximal parameters are the same as our algorithm. By applying PDM to (59), we obtain the following subproblems which are also easy enough to have closed-form solutions:

$$\begin{aligned}
x^{k+1} &= (rI + P)^{-1}\left(rx^k - \beta A^T\left(Ax^k + By^k - b - \frac{1}{\beta}\lambda^k\right)\right),\\
y^{k+1} &= (sI + Q)^{-1}\left(sy^k - \beta B^T\left(Ax^k + By^k - b - \frac{1}{\beta}\lambda^k\right)\right),\\
\lambda^{k+1} &= \lambda^k - \beta\left(Ax^{k+1} + By^{k+1} - b\right).
\end{aligned} \tag{62}$$

We report the numerical experiments by building their performance profiles in terms of the number of iterations and the total of computational time. Here, we take $\beta = 3 + (n/10)$, $r = s = 20\beta$ for the two algorithms. We set the initial vector $(x^0, y^0, \lambda^0) = (0, 0, 0)$, and the stopping criterion is

$$\begin{aligned}
\text{Tol} = \max\Big\{&\left\|x^{k+1} - x^k\right\|_\infty,\\
&\left\|y^{k+1} - y^k\right\|_\infty, \left\|\lambda^{k+1} - \lambda^k\right\|_\infty\Big\} \le 10^{-4}.
\end{aligned} \tag{63}$$

The computational results are given in Table 1 for different choices of m, n, and p. We reported the number of iterations (Iter.) and the computing time in seconds (CPU(s)) when the mentioned stopping criterion is achieved.

The data in Table 1 indicates clearly that the proposed method is efficient compared with the classical PDM in [21, 22]. We can observe that the iteration numbers and the CPU time of the two algorithms are almost the same.

5. Conclusions

In this paper, we proposed a new implementable algorithm for solving the monotone variational inequalities with separable structure. At each iteration, the algorithm performs

two easily implementable projections parallelly to produce a predictor and then makes a simple correction to generate the new iterate. Under some mild conditions, we proved the global convergence of the new method. We also give some numerical experiments to show that the proposed method is applicable and valid.

References

[1] S. Dafermos, "Traffic equilibrium and variational inequalities," *Transportation Science*, vol. 14, no. 1, pp. 42–54, 1980.

[2] D. P. Bertsekas and E. M. Gafni, *Projection Methods for Variational Inequalities with Application to the Traffic Assignment Problem*, Springer, Berlin, Germany, 1982.

[3] B. Martinet, "Régularisation d'inéquations variationnelles par approximations successives," *Revue Française dInformatique et de Recherche Opérationelle*, vol. 4, pp. 154–158, 1970.

[4] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[5] R. T. Rockafellar, "Augmented lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 97–116, 1976.

[6] B. He, L. Liao, and M. Qian, "Alternating projection based prediction-correction methods for structured variational inequalities," *Journal of Computational Mathematics*, vol. 24, no. 6, pp. 693–710, 2006.

[7] P. Tseng, "Alternating projection-proximal methods for convex programming and variational inequalities," *SIAM Journal on Optimization*, vol. 7, no. 4, pp. 951–965, 1997.

[8] B. He, L. Liao, D. Han, and H. Yang, "A new inexact alternating directions method for monotone variational inequalities," *Mathematical Programming B*, vol. 92, no. 1, pp. 103–118, 2002.

[9] B. He, S. Wang, and H. Yang, "A modified variable-penalty alternating directions method for monotone variational inequalities," *Journal of Computational Mathematics*, vol. 21, pp. 495–504, 2003.

[10] B. He, H. Yang, and S. Wang, "Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities," *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 337–356, 2000.

[11] B. He, L. 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.

[12] B. He, Z. Yang, and X. Yuan, "An approximate proximal-extragradient type method for monotone variational inequalities," *Journal of Mathematical Analysis and Applications*, vol. 300, no. 2, pp. 362–374, 2004.

[13] D. Han, "A hybrid entropic proximal decomposition method with self-adaptive strategy for solving variational inequality problems," *Computers and Mathematics with Applications*, vol. 55, no. 1, pp. 101–115, 2008.

[14] B. He, "Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities," *Computational Optimization and Applications*, vol. 42, no. 2, pp. 195–212, 2009.

[15] M. Xu, J. Jiang, B. Li, and B. Xu, "An improved prediction-correction method for monotone variational inequalities with separable operators," *Computers and Mathematics with Applications*, vol. 59, no. 6, pp. 2074–2086, 2010.

[16] M. Xu and T. Wu, "A class of linearized proximal alternating direction methods," *Journal of Optimization Theory and Applications*, vol. 151, no. 2, pp. 321–337, 2011.

[17] 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.

[18] 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.

[19] B. He and X. Yuan, "Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective," *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 119–149, 2012.

[20] D. Han, "A generalized proximal-point-based prediction-correction method for variational inequality problems," *Journal of Computational and Applied Mathematics*, vol. 221, no. 1, pp. 183–193, 2008.

[21] B. He and X. Yuan, "The unified framework of some proximal-based decomposition methods for monotone variational inequalities with separable structure," *Pacific Journal of Optimization*, vol. 8, pp. 817–844, 2013.

[22] G. Chen and M. Teboulle, "A proximal-based decomposition method for convex minimization problems," *Mathematical Programming*, vol. 64, no. 1–3, pp. 81–101, 1994.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 680768, 8 pages
http://dx.doi.org/10.1155/2013/680768

Research Article

Convergence Analysis of Alternating Direction Method of Multipliers for a Class of Separable Convex Programming

Zehui Jia,[1] Ke Guo,[2] and Xingju Cai[1]

[1] *School of Mathematical Science and Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing, Jiangsu 210023, China*
[2] *College of Mathematics and Information, China West Normal University, Nanchong, Sichuan 637009, China*

Correspondence should be addressed to Xingju Cai; caixingju@njnu.edu.cn

Received 19 July 2013; Accepted 30 July 2013

Academic Editor: Xu Minghua

The purpose of this paper is extending the convergence analysis of Han and Yuan (2012) for alternating direction method of multipliers (ADMM) from the strongly convex to a more general case. Under the assumption that the individual functions are composites of strongly convex functions and linear functions, we prove that the classical ADMM for separable convex programming with two blocks can be extended to the case with more than three blocks. The problems, although still very special, arise naturally from some important applications, for example, route-based traffic assignment problems.

1. Introduction

In this paper, we consider the convex programming with separable functions:

$$\min\left\{\sum_{i=1}^{m} f_i(x_i) \mid \sum_{i=1}^{m} A_i x_i = b,\ x_i \in \mathcal{X}_i,\ i = 1,2,\ldots,m\right\}, \tag{1}$$

where $f_i : \mathcal{R}^{n_i} \to \mathcal{R} \cup \{+\infty\}$ $(i = 1,2,\ldots,m)$ are closed proper convex functions (not necessarily smooth); $A_i \in \mathcal{R}^{l\times n_i}$ $(i = 1,2,\ldots,m)$; $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$ $(i = 1,2,\ldots,m)$ are closed convex sets; $b \in \mathcal{R}^l$ and $\sum_{i=1}^{m} n_i = n$. Throughout the paper, we assume that the solution set of (1) is nonempty.

For the special case of (1) with $m = 2$,

$$\min\{f_1(x_1) + f_2(x_2) \mid A_1x_1 + A_2x_2 = b,\ x_i \in \mathcal{X}_i,\ i = 1,2\}, \tag{2}$$

the problem has been studied extensively. Among lots of numerical methods, one of the most popular methods is the alternating direction method of multipliers (ADMM) which was presented originally in [1, 2]. The iterative scheme of ADMM for (2) is as follows:

$$\begin{aligned}
x_1^{k+1} &= \arg\min\left\{f_1(x_1) - \left(\lambda^k\right)^T A_1 x_1 + \frac{\beta}{2}\left\|A_1x_1 + A_2x_2^k - b\right\|^2 \mid x_1 \in \mathcal{X}_i\right\};\\
x_2^{k+1} &= \arg\min\left\{f_2(x_2) - \left(\lambda^k\right)^T A_2 x_2 + \frac{\beta}{2}\left\|A_1x_1^{k+1} + A_2x_2 - b\right\|^2 \mid x_2 \in \mathcal{X}_2\right\};\\
\lambda^{k+1} &= \lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right),
\end{aligned} \tag{3}$$

where λ^k is Lagrange multiplier associated with the linear constraints and $\beta > 0$ is the penalty parameter. The convergence of ADMM for (2) was also established under the condition that the involved functions are convex and the constrained sets are convex too.

While there are diversified applications whose objective function is separable into $m \geq 3$ individual convex functions without coupled variables, such as traffic problems, the problem of recovering the low-rank, sparse components of matrices from incomplete and noisy observation in [3], the constrained total-variation image restoration and reconstruction problem in [4, 5], and the minimal surface PDE problem in [6], it is thus natural to extend ADMM from 2 blocks to m blocks, resulting in the iterative scheme:

$$\begin{aligned}
x_1^{k+1} &= \arg\min \Big\{ f_1(x_1) - \left(\lambda^k\right)^T A_1 x_1 \\
&\qquad + \frac{\beta}{2} \Big\| A_1 x_1 + A_2 x_2^k + \cdots \\
&\qquad\qquad + A_m x_m^k - b \Big\|^2 \mid x_1 \in \mathcal{X}_1 \Big\}; \\
x_2^{k+1} &= \arg\min \Big\{ f_2(x_2) - \left(\lambda^k\right)^T A_2 x_2 \\
&\qquad + \frac{\beta}{2} \Big\| A_1 x_1^{k+1} + A_2 x_2 + \cdots \\
&\qquad\qquad + A_m x_m^k - b \Big\|^2 \mid x_2 \in \mathcal{X}_2 \Big\}; \\
&\vdots \\
x_m^{k+1} &= \arg\min \Big\{ f_m(x_m) - \left(\lambda^k\right)^T A_m x_m \\
&\qquad + \frac{\beta}{2} \Big\| A_1 x_1^{k+1} + A_2 x_2^{k+1} \cdots \\
&\qquad\qquad + A_m x_m - b \Big\|^2 \mid x_m \in \mathcal{X}_m \Big\}; \\
\lambda^{k+1} &= \lambda^k - \beta \left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + \cdots + A_m x_m^{k+1} - b \right).
\end{aligned} \tag{4}$$

Unfortunately, the convergence of the natural extension is still open under convex assumption, and the recent convergence results [7] are under the assumption that all the functions involved in the objective functions are strongly convex. This lack of convergence has inspired some ADM-based methods, for example, prediction-correction type method [3, 8–11], that is, the iterate $x_1^{k+1}, x_2^{k+1}, \ldots, x_m^{k+1}$ is regarded as a prediction, and the next iterate is a correction for it. However, the numerical results show that the algorithm (4) always performs better than these variants. Recently, Han and Yuan [7] show that the global convergence of the extension of ADMM for $m \geq 3$ is valid if the involved functions are further assumed to be strongly convex. This result does not answer the open problem regarding the convergence of the extension of ADMM under the convex assumption, but it makes a key progress towards this objective.

In this paper, we consider the separable convex optimization problem (1) where each individual function f_i is the combination of a strongly convex function g_i and a linear transform B_i. That is, (1) takes the following form:

$$\min \left\{ \sum_{i=1}^m g_i(B_i x_i) \mid \sum_{i=1}^m A_i x_i = b,\ x_i \in \mathcal{X}_i,\ i = 1, 2, \ldots, m \right\}, \tag{5}$$

where $g_i : \mathcal{R}^{s_i} \to \mathcal{R} \cup \{+\infty\}$ $(i = 1, 2, \ldots, m)$ are closed proper strongly convex function with the modulus μ_i (not necessarily smooth); $A_i \in \mathcal{R}^{l \times n_i}$ $(i = 1, 2, \ldots, m)$; $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$ $(i = 1, 2, \ldots, m)$ are closed convex sets; $b \in \mathcal{R}^l$ and $\sum_{i=1}^m n_i = n$; $B_i \in \mathcal{R}^{s_i \times n_i}$ $(i = 1, 2, \ldots, m)$, where B_i may not have full column rank (if B_i has full column rank, the composite function is strongly convex and reduces to the case considered in [7]). Note that although (5) is very special, it arises frequently from many applications. One example is under the route-based traffic assignment problem [12], where g_i is the link traffic cost, B_i is the link-path incidence matrix, and x is the path follow vector.

In the following, we abuse a little the notation and still write g_i with f_i; that is, the problem under consideration is

$$\min \left\{ \sum_{i=1}^m f_i(B_i x_i) \mid \sum_{i=1}^m A_i x_i = b,\ x_i \in \mathcal{X}_i,\ i = 1, 2, \ldots, m \right\}, \tag{6}$$

where $f_i : \mathcal{R}^{s_i} \to \mathcal{R} \cup \{+\infty\}$ $(i = 1, 2, \ldots, m)$ are closed proper strongly convex function with the modulus μ_i (not necessarily smooth).

The rest of the paper is organized as follows. In the next section, we list some necessary preliminary results that will be used in the rest of the paper. We then describe the algorithm formally and analyze its global convergence under reasonable conditions in Section 3. We complete the paper with some conclusions in Section 4.

2. Preliminaries

In this section, we summarize some basic concepts and their properties that will be useful for further discussion.

Let $\|\cdot\|_p$ denote the standard definition of the l^p-norm, and particularly, let $\|\cdot\| = \|\cdot\|_2$ denote the Euclidean norm. For a symmetric and positive definite matrix G, we denote $\|\cdot\|_G$ the G-norm, that is, $\|x\|_G = \sqrt{x^T G x}$. If G is the product of a positive parameter β and the identity matrix I, that is, $G = \beta I$, we use the simpler notation: $\|\cdot\|_G = \|\cdot\|_\beta$.

Let $f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\}$. If the domain of f denoted by $\operatorname{dom} f = \{x \in \mathcal{R}^n \mid f(x) < +\infty\}$ is not empty, then f is said to be proper. If for any $x \in \mathcal{R}^n$ and $y \in \mathcal{R}^n$, we have

$$f(tx + (1-t)y) \leq t f(x) + (1-t) f(y), \quad \forall t \in [0, 1], \tag{7}$$

then f is said to be convex. Furthermore, f is said to be strongly convex with the modulus $\mu > 0$ if and only if

$$\begin{aligned}
f(tx + (1-t)y) &\leq t f(x) + (1-t) f(y) \\
&\quad - \frac{1}{2} \mu t (1-t) \|x - y\|^2, \quad \forall t \in [0, 1].
\end{aligned} \tag{8}$$

A set-valued operator T defined on $\mathscr{R}^n$ is said to be monotone if and only if

$$(u-\widetilde{u})^T(w-\widetilde{w})\geq 0,\quad \forall w\in Tu,\ \forall\widetilde{w}\in T\widetilde{u}, \tag{9}$$

and T is said to be strongly monotone with modulus $\mu>0$ if and only if

$$(u-\widetilde{u})^T(w-\widetilde{w})\geq \mu\|u-\widetilde{u}\|^2,\quad \forall w\in Tu,\ \forall\widetilde{w}\in T\widetilde{u}. \tag{10}$$

Let $\Gamma_0(\mathscr{R}^n)$ denote the set of closed proper convex functions from $\mathscr{R}^n$ to $\mathscr{R}\cup\{+\infty\}$. For any $f\in\Gamma_0(\mathscr{R}^n)$, the subdifferential of f which is the set-valued operator, defined by

$$\partial f: x\longmapsto\left\{\xi\in\mathscr{R}^n \mid (y-x)^T\xi+f(x)\leq f(y),\forall y\in \operatorname{dom} f\right\}, \tag{11}$$

is monotone. Moreover, if f is strongly convex function with the modulus μ, ∂f is strongly monotone with the modulus μ.

Let F be a mapping from a set $\Omega\subset\mathscr{R}^n\to\mathscr{R}^n$. Then F is said to be co-coercive on Ω with modulus $\gamma>0$, if

$$(u-v)^T(F(u)-F(v))\geq\gamma\|F(u)-F(v)\|^2,\quad \forall u,v\in\Omega. \tag{12}$$

Throughout the paper, we make the following assumptions.

Assumption 1. (i) $n_i\|B_ix_i\|\geq\|A_i\|\|x_i\|$, $\forall x_i\in\mathscr{R}^{n_i}$, $i\in\{1,2,\ldots,m\}$; (ii) the solution set of (1) is nonempty.

Remark 2. Assumption 1 is a little restrictive. However, some problems can satisfy it. A remarkable one is the following route-based traffic assignment problem.

Consider a transportation network $G(\mathscr{N},E)$, where $\mathscr{N}$ is the set of nodes. We denote the set of links by $\mathscr{A}$, and the number of the element of $\mathscr{A}$ by $N_{\mathscr{A}}$, respectively. Let RS denote the set of origin-destination (O-D) pairs. For an O-D pair rs $\in$ RS, let q^{rs} be its traffic demand; let P^{rs} be the set of routes connecting rs, and $p\in P^{\mathrm{rs}}$; $\mathscr{N}^{\mathrm{rs}}$ denotes the number of the routes connecting rs; let h_p^{rs} be the route flow on p. The feasible route flow vector $h=(p\in P^{\mathrm{rs}}\mid \mathrm{rs}\in\mathrm{RS})$ is thus given by

$$\begin{aligned}H&=\left\{h\mid\sum_{p\in P^{\mathrm{rs}}}h_p^{\mathrm{rs}}=q^{\mathrm{rs}},\ h_p^{\mathrm{rs}}\geq 0,\ \forall p\in P^{\mathrm{rs}},\ \mathrm{rs}\in\mathrm{RS}\right\}\\&=\left\{h\mid e^T\left(h_1^{\mathrm{rs}},h_2^{\mathrm{rs}},\ldots,h_{N^{\mathrm{rs}}}^{\mathrm{rs}}\right)=q^{\mathrm{rs}},\right.\\&\qquad\left.h_p^{\mathrm{rs}}\geq 0,\forall p\in P^{\mathrm{rs}},\ \mathrm{rs}\in\mathrm{RS}\right\}.\end{aligned} \tag{13}$$

Define E as the link-route incidence matrix such that

$$\delta_p^a=\begin{cases}1, & \text{if } p \text{ contains link } a\\ 0, & \text{otherwise.}\end{cases} \tag{14}$$

Then, link flow f_a can be written as

$$\begin{aligned}f_a&=\sum_{\mathrm{rs}\in\mathrm{RS}}\sum_{p\in P^{\mathrm{rs}}}\delta_p^a h_p^{\mathrm{rs}},\quad\forall a\in\mathscr{A},\\F&=EH=\{f\mid f=Eh,\ h\in H\}.\end{aligned} \tag{15}$$

By denoting the link cost function as $C_a(f)$ and for the additive case, the route cost function as $C_p(h)$, they can be related by

$$C_ph=\sum_{a\in\mathscr{A}}\delta_p^aC_a(f). \tag{16}$$

The user equilibrium traffic assignment problem can be formulated as a VI: find $f^*\in F$ such that

$$(f-f^*)^TC(f^*)\geq 0,\quad\forall f\in F, \tag{17}$$

or equivalently, find $h^*\in H$ such that

$$(h-h^*)^TE^TC(Eh^*)\geq 0,\quad\forall h\in H, \tag{18}$$

where $C=\{C_a\}$ is the vector of the link cost function.

In general, it is easy to show that e is a row of E and E is not a full column rank (if E is, then the above variational inequality is strongly monotone).

For simplicity, in the following, we only consider the case for $m=3$. Notice that for $m\geq 3$, it can be proved similarly following the processing of $m=3$.

3. The Method

In this section, we consider the following convex minimization problem with linear constraint, where the objective function is in the form of the sum of three individual functions without coupled variable:

$$\begin{aligned}&\min\quad f_1(B_1x_1)+f_2(B_2x_2)+f_3(B_3x_3)\\&\text{s.t.}\quad A_1x_1+A_2x_2+A_3x_3=b,\quad x_i\in\mathscr{X}_i,\ i=1,2,3,\end{aligned} \tag{19}$$

where $f_i:\mathscr{R}^{s_i}\to\mathscr{R}\cup\{+\infty\}$ $(i=1,2,3)$ are closed proper strongly convex function with the modulus μ_i (not necessarily smooth); $B_i\in\mathscr{R}^{s_i\times n_i}$ $(i=1,2,3)$, $A_i\in\mathscr{R}^{l\times n_i}$ $(i=1,2,3)$; $\mathscr{X}_i\subseteq\mathscr{R}^{n_i}$ $(i=1,2,3)$ are closed convex sets; $b\in\mathscr{R}^l$ and $\sum_{i=1}^3 n_i=n$.

The iterative scheme of ADMM for problem (19) is as follows:

$$\begin{aligned}
x_1^{k+1} &= \arg\min \Big\{ f_1(B_1x_1) - \left(\lambda^k\right)^T A_1x_1 \\
&\quad + \frac{\beta}{2}\Big\| A_1x_1 \\
&\quad + A_2x_2^k + A_3x_3^k - b\Big\|^2 \mid x_1 \in \mathcal{X}_1 \Big\}, \\
x_2^{k+1} &= \arg\min \Big\{ f_2(B_2x_2) - \left(\lambda^k\right)^T A_2x_2 \\
&\quad + \frac{\beta}{2}\Big\| A_1x_1^{k+1} \\
&\quad + A_2x_2 + A_3x_3^k - b\Big\|^2 \mid x_2 \in \mathcal{X}_2 \Big\}, \\
x_3^{k+1} &= \arg\min \Big\{ f_3(B_3x_3) - \left(\lambda^k\right)^T A_3x_3 \\
&\quad + \frac{\beta}{2}\Big\| A_1x_1^{k+1} \\
&\quad + A_2x_2^{k+1} + A_3x_3 - b\Big\|^2 \mid x_3 \in \mathcal{X}_3 \Big\}, \\
\lambda^{k+1} &= \lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^{k+1} - b\right),
\end{aligned} \tag{20}$$

where λ^k is the Lagrangian multiplier associated with the linear constraints and $\beta > 0$ is the penalty parameter.

4. Convergence

In this section, we prove the convergence of the extended ADMM for problem (19). As the assumptions aforementioned, by invoking the first-order necessary and sufficient condition for convex programming, we easily see that the problem (19) under the condition is characterized by the following variational inequality (VI): find $u^* \in \mathcal{U}$ and $\xi_i^* \in \partial f_i(B_ix_i^*)$ such that

$$\left(u - u^*\right)^T Q\left(u^*\right) \geq 0, \quad \forall u \in \mathcal{U}, \tag{21}$$

where

$$u := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix}, \qquad Q(u) := \begin{pmatrix} B_1^T\xi_1 - A_1^T\lambda \\ B_2^T\xi_2 - A_2^T\lambda \\ B_3^T\xi_3 - A_3^T\lambda \\ A_1x_1 + A_2x_2 + A_3x_3 - b \end{pmatrix},$$
$$\mathcal{U} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{R}^l. \tag{22}$$

We denote the VI (21)-(22) by MVI$(\mathcal{U}, Q)$.

Similarly, in [7], we propose an easily implementable stopping criterion for executing (20):

$$\max\left\{ \max_{1\leq i\leq 3}\left\| A_ix_i^k - A_ix_i^{k+1}\right\|, \left\| \sum_{i=1}^3 A_ix_i^k - b\right\| \right\} \leq \epsilon, \tag{23}$$

and its rationale can be seen in the following lemma.

Lemma 3 (see [7]). *If $\sum_{i=1}^3 A_ix_i^k - b = 0$ and $A_ix_i^k = A_ix_i^{k+1}$ $(i = 1, 2, 3)$, then $(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}, \lambda^{k+1})$ is a solution of MVI$(\mathcal{U}, Q)$.*

Lemma 3 implies that the iterate $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ is a solution of MVI$(\mathcal{U}, Q)$ when the inequality (23) holds with $\epsilon = 0$. Some techniques of establishing the error bounds in [13] can help us analyze how precisely the iterate satisfies the optimality conditions when the proposed stopping criterion is satisfied with a tolerance $\epsilon > 0$.

Lemma 4. *Let $(x_1^*, x_2^*, x_3^*, \lambda^*)$ be the solution of the problem (19), and let λ^* be a corresponding Lagrange multiplier associated with the linear constraint. Then, the sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ generated by (20) satisfies*

$$\begin{aligned}
&\left(\lambda^k - \lambda^*\right)^T \left(\sum_{i=1}^3 A_ix_i^{k+1} - b\right) \\
&\quad \geq \sum_{i=1}^3 \left(x_i^{k+1} - x_i^*\right)^T \left(B_i^T\xi_i^{k+1} - B_i^T\xi_i^*\right) \\
&\qquad + \beta\left\| \sum_{i=1}^3 A_ix_i^{k+1} - b\right\|^2 + \beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T \\
&\qquad \times \left[A_2x_2^k - A_2x_2^{k+1} + \left(A_3x_3^k - A_3x_3^{k+1}\right)\right] \\
&\qquad + \beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T \left(A_3x_3^k - A_3x_3^{k+1}\right).
\end{aligned} \tag{24}$$

Proof. By invoking the first-order optimality condition for the x_i^{k+1}-related subproblem in (20), for any $x_i \in \mathcal{X}_i$, $i = 1, 2, 3$, we get

$$\begin{aligned}
&\left(x_1 - x_1^{k+1}\right)^T \Big\{B_1^T\xi_1^{k+1} \\
&\qquad - A_1^T\left[\lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^k + A_3x_3^k - b\right)\right]\Big\} \\
&\quad \geq 0, \\
&\left(x_2 - x_2^{k+1}\right)^T \Big\{B_2^T\xi_2^{k+1} - A_2^T \\
&\qquad \times \left[\lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} + A_3x_3^k - b\right)\right]\Big\} \\
&\quad \geq 0,
\end{aligned}$$

$$\left(x_3 - x_3^{k+1}\right)^T \left\{B_3^T \xi_3^{k+1} - A_3^T \times \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b\right)\right]\right\} \geq 0. \tag{25}$$

Setting $x_i = x_i^*$ $(i = 1, 2, 3)$ in (25), we have

$$\begin{aligned}
&\left(x_1^* - x_1^{k+1}\right)^T \left\{B_1^T \xi_1^{k+1} - A_1^T \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^k - b\right)\right]\right\} \geq 0,\\
&\left(x_2^* - x_2^{k+1}\right)^T \left\{B_2^T \xi_2^{k+1} - A_2^T \times \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^k - b\right)\right]\right\} \geq 0,\\
&\left(x_3^* - x_3^{k+1}\right)^T \left\{B_3^T \xi_3^{k+1} - A_3^T \times \left[\lambda^k - \beta\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b\right)\right]\right\} \geq 0.
\end{aligned} \tag{26}$$

On the other hand, setting $(x_1, x_2, x_3) = (x_1^{k+1}, x_2^{k+1}, x_3^{k+1})$ in (21), it follows that

$$\begin{pmatrix} x_1^{k+1} - x_1^* \\ x_2^{k+1} - x_2^* \\ x_3^{k+1} - x_3^* \end{pmatrix}^T \begin{pmatrix} B_1^T \xi_1^* - A_1^T \lambda^* \\ B_2^T \xi_2^* - A_2^T \lambda^* \\ B_3^T \xi_3^* - A_3^T \lambda^* \end{pmatrix} \geq 0. \tag{27}$$

Adding (26) and (27), we obtain

$$\begin{aligned}
&\left(x_1^{k+1} - x_1^*\right)^T \left\{\left(B_1^T \xi_1^* - B_1^T \xi_1^{k+1}\right) - A_1^T\left(\lambda^* - \lambda^k\right) - \beta A_1^T\left(A_1 x_1^{k+1} + A_2 x_2^k + A_3 x_3^k - b\right)\right\} \geq 0,\\
&\left(x_2^{k+1} - x_2^*\right)^T \left\{\left(B_2^T \xi_2^* - B_2^T \xi_2^{k+1}\right) - A_2^T\left(\lambda^* - \lambda^k\right) - \beta A_2^T\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^k - b\right)\right\} \geq 0,\\
&\left(x_3^{k+1} - x_3^*\right)^T \left\{\left(B_3^T \xi_3^* - B_3^T \xi_3^{k+1}\right) - A_3^T\left(\lambda^* - \lambda^k\right) - \beta A_3^T\left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 x_3^{k+1} - b\right)\right\} \geq 0.
\end{aligned} \tag{28}$$

With the rearrangement of the above inequalities, we derive that

$$\begin{aligned}
&\left(x_1^{k+1} - x_1^*\right)^T A_1^T \left(\lambda^k - \lambda^*\right)\\
&\quad \geq \left(x_1^{k+1} - x_1^*\right)^T \left(B_1^T \xi_1^{k+1} - B_1^T \xi_1^*\right)\\
&\qquad + \beta\left(A_1 x_1^{k+1} - A_1 x_1^*\right)^T \left(\sum_{i=1}^3 A_i x_i^{k+1} - b\right)\\
&\qquad + \beta\left(A_1 x_1^{k+1} - A_1 x_1^*\right)^T \times \left[\left(A_2 x_2^k - A_2 x_2^{k+1}\right) + \left(A_3 x_3^k - A_3 x_3^{k+1}\right)\right],\\
&\left(x_2^{k+1} - x_2^*\right)^T A_2^T \left(\lambda^k - \lambda^*\right)\\
&\quad \geq \left(x_2^{k+1} - x_2^*\right)^T \left(B_2^T \xi_2^{k+1} - B_2^T \xi_2^*\right)\\
&\qquad + \beta\left(A_2 x_2^{k+1} - A_2 x_2^*\right)^T \left(\sum_{i=1}^3 A_i x_i^{k+1} - b\right)\\
&\qquad + \beta\left(A_2 x_2^{k+1} - A_2 x_2^*\right)^T \left(A_3 x_3^k - A_3 x_3^{k+1}\right),\\
&\left(x_3^{k+1} - x_3^*\right)^T A_3^T \left(\lambda^k - \lambda^*\right)\\
&\quad \geq \left(x_3^{k+1} - x_3^*\right)^T \left(B_3^T \xi_3^{k+1} - B_3^T \xi_3^*\right)\\
&\qquad + \beta\left(A_3 x_3^{k+1} - A_3 x_3^*\right)^T \left(\sum_{i=1}^3 A_i x_i^{k+1} - b\right).
\end{aligned} \tag{29}$$

Adding the above inequalities (29), we have

$$\begin{aligned}
&\left(\lambda^k - \lambda^*\right)^T \left(\sum_{i=1}^3 A_i x_i^{k+1} - b\right)\\
&\quad \geq \sum_{i=1}^3 \left(x_i^{k+1} - x_i^*\right)^T \left(B_i^T \xi_i^{k+1} - B_i^T \xi_i^*\right) + \beta\left\|\sum_{i=1}^3 A_i x_i^{k+1} - b\right\|^2\\
&\qquad + \beta\left(A_1 x_1^{k+1} - A_1 x_1^*\right)^T \left[\left(A_2 x_2^k - A_2 x_2^{k+1}\right) + \left(A_3 x_3^k - A_3 x_3^{k+1}\right)\right]\\
&\qquad + \beta\left(A_2 x_2^{k+1} - A_2 x_2^*\right)^T \left(A_3 x_3^k - A_3 x_3^{k+1}\right).
\end{aligned} \tag{30}$$

The proof is complete. □

Hereafter, we define a matrix which will make the notation of proof more succinct. More specifically, let

$$M = \begin{pmatrix} 2\beta A_1^T A_1 & 0 & 0 & 0 \\ 0 & 2\beta A_2^T A_2 & 0 & \\ 0 & 0 & 2\beta A_3^T A_3 & 0 \\ 0 & 0 & 0 & \frac{1}{\beta} I \end{pmatrix}. \tag{31}$$

Obviously, M is a positive semidefinite matrix, only for analysis convenience; we denote

$$\|u\|_M^2 = 2\beta\left(\|A_1x_1\|^2 + \|A_2x_2\|^2 + \|A_3x_3\|^2\right) + \|\lambda\|_{1/\beta}^2. \tag{32}$$

Lemma 5. *Let $u^* = (x_1^*, x_2^*, x_3^*, \lambda^*)$ be a solution of $MVI(\mathcal{U}, Q)$, and let the sequence $\{(x_1^k, x_2^k, x_3^k, \lambda^k)\}$ be generated by* (20). *Then, one has*

$$\begin{aligned}\left\|u^{k+1} - u^*\right\|_M^2 \le{}& \left\|u^k - u^*\right\|_M^2 \\ &+ \sum_{i=1}^{3} 3\beta\left\|A_ix_i^{k+1} - A_ix_i^*\right\|^2 \\ &- 2\sum_{i=1}^{3}\mu_i\left\|B_ix_i^{k+1} - B_ix_i^*\right\|^2.\end{aligned} \tag{33}$$

Proof. From (20) and Lemma 4, we have

$$\begin{aligned}\left\|\lambda^{k+1} - \lambda^*\right\|_{1/\beta}^2 &= \left\|\lambda^k - \lambda^* - \beta\left(\sum_{i=1}^{3}A_ix_i^{k+1} - b\right)\right\|_{1/\beta}^2 \\ &= \left\|\lambda^k - \lambda^*\right\|_{1/\beta}^2 \\ &\quad - 2\left(\lambda^k - \lambda^*\right)^T\left(\sum_{i=1}^{3}A_ix_i^{k+1} - b\right) \\ &\quad + \beta\left\|\sum_{i=1}^{3}A_ix_i^{k+1} - b\right\|^2 \\ &\le \left\|\lambda^k - \lambda^*\right\|_{1/\beta}^2 \\ &\quad - 2\sum_{i=1}^{3}\left(x_i^{k+1} - x_i^*\right)^T\left(B_i^T\xi_i^{k+1} - B_i^T\xi_i^*\right) \\ &\quad - \beta\left\|\sum_{i=1}^{3}A_ix_i^{k+1} - b\right\|^2 \\ &\quad - 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T \\ &\quad \times\left(\sum_{i=2}^{3}\left(A_ix_i^k - A_ix_i^{k+1}\right)\right) \\ &\quad - 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^k - A_3x_3^{k+1}\right).\end{aligned} \tag{34}$$

Since

$$\begin{aligned}&\left\|\sum_{i=1}^{3}\left(A_ix_i^{k+1} - A_ix_i^*\right)\right\|^2 \\ &\quad = \sum_{i=1}^{3}\left\|A_i\left(x_i^{k+1} - x_i^*\right)\right\|^2 \\ &\qquad + \sum_{i\ne j}\left(A_i\left(x_i^{k+1} - x_i^*\right)\right)^TA_j\left(x_j^{k+1} - x_j^*\right),\end{aligned} \tag{35}$$

and $A_1x_1^* + A_2x_2^* + A_3x_3^* = b$, we can get

$$\begin{aligned}&-\beta\left\|\sum_{i=1}^{3}A_ix_i^{k+1} - b\right\|^2 \\ &\quad = -\beta\sum_{i=1}^{3}\left\|A_i\left(x_i^{k+1} - x_i^*\right)\right\|^2 \\ &\qquad - \beta\sum_{i\ne j}\left(A_i\left(x_i^{k+1} - x_i^*\right)\right)^TA_j\left(x_j^{k+1} - x_j^*\right).\end{aligned} \tag{36}$$

Using Cauchy-Schwarz inequality, we have

$$\begin{aligned}&-2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(\sum_{i=2}^{3}\left(A_ix_i^k - A_ix_i^{k+1}\right)\right) \\ &\quad - 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^k - A_3x_3^{k+1}\right) \\ &= -2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_2x_2^k - A_2x_2^*\right) \\ &\quad + 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_2x_2^{k+1} - A_2x_2^*\right) \\ &\quad - 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_3x_3^k - A_3x_3^*\right) \\ &\quad + 2\beta\left(A_1x_1^{k+1} - A_1x_1^*\right)^T\left(A_3x_3^{k+1} - A_3x_3^*\right) \\ &\quad - 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^k - A_3x_3^*\right) \\ &\quad + 2\beta\left(A_2x_2^{k+1} - A_2x_2^*\right)^T\left(A_3x_3^{k+1} - A_3x_3^*\right) \\ &\le 2\beta\left\|A_1x_1^{k+1} - A_1x_1^*\right\|^2 + \beta\left\|A_2x_2^{k+1} - A_2x_2^*\right\|^2 \\ &\quad + \beta\left\|A_2\left(x_2^k - x_2^*\right)\right\|^2 + 2\beta\left\|A_3\left(x_3^k - x_3^*\right)\right\|^2 \\ &\quad + \beta\sum_{i\ne j}\left(A_i\left(x_i^{k+1} - x_i^*\right)\right)^TA_j\left(x_j^{k+1} - x_j^*\right).\end{aligned} \tag{37}$$

Substituting (36) and (37) into (34), we get

$$\begin{aligned}\left\|\lambda^{k+1}-\lambda^*\right\|_{1/\beta}^2 \le \left\|\lambda^k-\lambda^*\right\|_{1/\beta}^2 &+ 2\beta\sum_{i=1}^{3}\left\|A_i\left(x_i^k-x_i^*\right)\right\|^2 \\ &+ \beta\sum_{i=1}^{3}\left\|A_i\left(x_i^{k+1}-x_i^*\right)\right\|^2 \\ &- 2\sum_{i=1}^{3}\left(x_i^{k+1}-x_i^*\right)^T\left(B_i^T\xi_i^{k+1}-B_i^T\xi_i^*\right).\end{aligned} \tag{38}$$

Since f_i is strongly convex, from the strong monotonicity of the subdifferential mapping ∂f_i (with the modulus μ_i), then we have

$$\begin{aligned}&\left(x_i^{k+1}-x_i^*\right)^T\left(B_i^T\xi_i^{k+1}-B_i^T\xi_i^*\right) \\ &\quad=\left(B_ix_i^{k+1}-B_ix_i^*\right)^T\left(\xi_i^{k+1}-\xi_i^*\right)\ge\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2,\end{aligned} \tag{39}$$

where $\xi_i^*\in\partial f_i(B_ix_i^*)$, $\xi_i^{k+1}\in\partial f_i(B_ix_i^{k+1})$, for any $i\in\{1,2,3\}$.

By using the notion of $\|u^{k+1}-u^*\|_M^2$, from (38) we have

$$\begin{aligned}&\left\|u^{k+1}-u^*\right\|_M^2 \\ &\quad=\left\|\lambda^{k+1}-\lambda^*\right\|_{1/\beta}^2 \\ &\qquad+2\beta\sum_{i=1}^{3}\left\|A_i(x_i^{k+1}-x_i^*)\right\|^2 \\ &\quad\le\left\|\lambda^k-\lambda^*\right\|_{1/\beta}^2+2\beta\sum_{i=1}^{3}\left\|A_i(x_i^k-x_i^*)\right\|^2 \\ &\qquad+3\beta\sum_{i=1}^{3}\left\|A_i(x_i^{k+1}-x_i^*)\right\|^2-2\sum_{i=1}^{3}\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2 \\ &\quad\le\left\|u^k-u^*\right\|_M^2+\sum_{i=1}^{3}3\beta\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 \\ &\qquad-2\sum_{i=1}^{3}\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2.\end{aligned} \tag{40}$$

The proof is complete. □

Theorem 6. *Under Assumption 1, for any*

$$0<\beta<\min_{1\le i\le 3}\left\{\frac{2\mu_i}{3n_i^2}\right\}, \tag{41}$$

the sequence $\{(x_1^k,x_2^k,x_3^k,\lambda^k)\}$ *generated by* (20) *converges to a solution of* $MVI(\mathcal{U},Q)$.

Proof. From Lemma 5, we have

$$\begin{aligned}\left\|u^{k+1}-u^*\right\|_M^2 \le \left\|u^k-u^*\right\|_M^2 &+ \sum_{i=1}^{3}3\beta\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 \\ &- 2\sum_{i=1}^{3}\mu_i\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2,\end{aligned} \tag{42}$$

where

$$0<\beta<\min_{1\le i\le 3}\left\{\frac{2\mu_i}{3n_i^2}\right\}. \tag{43}$$

From Assumption 1, it follows that

$$\begin{aligned}\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2 &\le \left\|A_i\right\|^2\left\|x_i^{k+1}-x_i^*\right\|^2 \\ &\le n_i^2\left\|B_ix_i^{k+1}-B_ix_i^*\right\|^2,\quad i=1,2,3.\end{aligned} \tag{44}$$

Consequently,

$$\begin{aligned}\left\|u^{k+1}-u^*\right\|_M^2 &\le \left\|u^k-u^*\right\|_M^2 \\ &\quad+\sum_{i=1}^{3}\left(3\beta-\frac{2\mu_i}{n_i^2}\right)\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2.\end{aligned} \tag{45}$$

From (45), we have

$$\left\|u^{k+1}-u^*\right\|_M^2\le\left\|u^k-u^*\right\|_M^2\le\cdots\le\left\|u^0-u^*\right\|_M^2<+\infty, \tag{46}$$

which means that the generated sequence $\{u^k\}$ is bounded.

Furthermore, it follows that

$$\begin{aligned}&\sum_{k=0}^{+\infty}\left\{\sum_{i=1}^{3}\left(2\frac{\mu_i}{n_i^2}-3\beta\right)\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2\right\} \\ &\quad\le\sum_{k=0}^{+\infty}\left\{\left\|u^k-u^*\right\|_M-\left\|u^{k+1}-u^*\right\|_M\right\}<+\infty,\end{aligned} \tag{47}$$

which means that

$$\lim_{k\to+\infty}\sum_{i=1}^{3}\left\|A_ix_i^{k+1}-A_ix_i^*\right\|^2=0. \tag{48}$$

Therefore, we have

$$\lim_{k\to+\infty}\left\|\sum_{i=1}^{3}A_ix_i^{k+1}-b\right\|^2=0. \tag{49}$$

Since $\|A_i\|$ is nonzero and bounded, from (48) we have

$$\lim_{k\to+\infty}\left\|x_i^{k+1}-x_i^*\right\|=0,\quad\forall i=1,2,3. \tag{50}$$

Since $\{u^k\}$ is bounded, $\{\lambda^k\}$ has at least one cluster point, say $\overline{\lambda}$. Let $\{\lambda^{k_j}\}$ be the corresponding subsequence that converges

to $\overline{\lambda}$. Taking a limit along this subsequence in (25) and (49), we obtain $\xi_i^* \in \partial f_i(B_i x_i^*)$,

$$\begin{gathered}(x_i - x_i^*)^T \left(B_i^T \xi_i^* - A_i^T \overline{\lambda}\right) \geq 0, \quad \forall x_i \in \mathcal{X}_i,\ i = 1, 2, 3, \\ \sum_{i=1}^{3} A_i x_i^* - b = 0,\end{gathered} \tag{51}$$

which follows that $\overline{\lambda}$ is an optimal Lagrange multiplier. Since λ^* is arbitrary, we can set $\lambda^* = \overline{\lambda}$ in (46) and conclude that the whole generated sequence converges to a solution of MVI$(\mathcal{U}, Q)$. □

5. Conclusions

In this paper, we extend the convergence analysis of the ADMM for the separable convex optimization problem with strongly convex functions to the case in which the individual functions are composites of strongly convex functions with a linear transform. Under further assumptions, we established the global convergence of the algorithm.

It should be admitted that although some problems arising from applications such as traffic assignment fall into our analysis, the problems considered here are too special. Thus, it is far away to solve the open problem of convergence of the ADMM with more than three blocks.

Acknowledgments

Xingju Cai was supported by the National Natural Science Foundation of China (NSFC) Grants nos. 11071122 and 11171159 and by the Doctoral Fund of Ministry of Education of China no. 20103207110002.

References

[1] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers and Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[2] D. Gabay, "Applications of the method of multipliers to variational inequalities," in *Augmented Lagrangian Methods: Applications to Numerical Solution of Boundary-Value Problems*, M. Fortin and R. Glowinski, Eds., pp. 299–331, North-Holland Publisher, Amsterdam, The Netherland, 1983.

[3] M. Tao and X. Yuan, "Recovering low-rank and sparse components of matrices from incomplete and noisy observations," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 57–81, 2011.

[4] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2710–2736, 2010.

[5] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[6] 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.

[7] D. Han and X. Yuan, "A note on the alternating direction method of multipliers," *Journal of Optimization Theory and Applications*, vol. 155, pp. 227–238, 2012.

[8] D. R. Han, X. M. Yuan, W. X. Zhang, and X. J. Cai, "An ADM-based splitting method for separable convex programming," *Computational Optimization and Applications*, vol. 54, pp. 343–369, 2013.

[9] B. S. He, M. Tao, and X. M. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM Journal on Optimization*, vol. 22, pp. 313–340, 2012.

[10] B. S. He, M. Tao, M. H. Xu, and X. M. Yuan, "Alternating directions based contraction method for generally separable linearly constrained convex programming problems," *Optimization*, vol. 62, pp. 573–596, 2013.

[11] B. S. He, M. Tao, and X. M. Yuan, "A splitting method for separable convex programming," *IMA Journal of Numerical Analysis*. In press.

[12] D. Han and H. K. Lo, "Solving non-additive traffic assignment problems: a descent method for co-coercive variational inequalities," *European Journal of Operational Research*, vol. 159, no. 3, pp. 529–544, 2012.

[13] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementary Problems. Volume I and II*, Springer, New York, NY, USA, 2003.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 492305, 13 pages
http://dx.doi.org/10.1155/2013/492305

Research Article

A General Self-Adaptive Relaxed-PPA Method for Convex Programming with Linear Constraints

Xiaoling Fu

Institute of Systems Engineering, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Xiaoling Fu; fuxlnju@hotmail.com

Received 27 June 2013; Accepted 21 July 2013

Academic Editor: Abdellah Bnouhachem

We present an efficient method for solving linearly constrained convex programming. Our algorithmic framework employs an implementable proximal step by a slight relaxation to the subproblem of proximal point algorithm (PPA). In particular, the stepsize choice condition of our algorithm is weaker than some elegant PPA-type methods. This condition is flexible and effective. Self-adaptive strategies are proposed to improve the convergence in practice. We theoretically show under mild conditions that our method converges in a global sense. Finally, we discuss applications and perform numerical experiments which confirm the efficiency of the proposed method. Comparisons of our method with some state-of-the-art algorithms are also provided.

1. Introduction

In this paper, we consider the following generic convex programming:

$$\min\{\theta(x) \mid Ax = b \,(\text{or } Ax \geq b),\ x \in \mathcal{X}\}, \tag{1}$$

where $\theta(x) : \Re^n \to \Re$ is a convex (not necessary smooth) function, $A \in \Re^{m\times n}$, $b \in \Re^m$, and $\mathcal{X} \subset \Re^n$ is a closed convex set. Problem (1) generalizes a wide range of problems that frequently arise in signal and image processing and reconstruction, mechanics, statistics, operations research, and other fields, for example, basis pursuit [1–4], nearest correlation matrix [5–7], matrix completion problem [2, 3, 8–10], and so forth. Before we begin, some assumptions should be presented for problem (1).

Assumption 1. The solution set of (1) is denoted by $\mathcal{X}^*$, and it is assumed to be nonempty.

Assumption 2. The objective function is simple. This means that, for a given constant ϱ, the following proximal problem admits a closed-form solution or can be solved efficiently with high precision:

$$\min_{x\in\mathcal{X}} \theta(x) + \frac{\varrho}{2}\|x - a\|^2, \tag{2}$$

where a is any given vector. At first sight, this assumption seems to be quite restrictive, but this is indeed for many problems in practice. For example, nuclear norm function in matrix completion problem, l_1-norm function in basis pursuit problem, and so forth.

Many fundamental methods have been developed over the past decades to solve problem (1). Proximal point algorithm (PPA) is one of the leading approaches for solving convex optimization problems. It is earlier used for regularized linear equations and has been applied to convex optimization by Martinet [11]. There are some significant theoretical achievements [12–19] in the field of PPA for convex optimization and monotone variational inequalities (VIs). Nowadays, it is still the object of intensive investigation [20] and leads to a variety of primal and dual methods. Common to PPA and its variants is the difficulty of their subproblems; this restricts the practical interest. Augmented Lagrangian method (ALM) [21] is a powerful method for linearly constrained problems. It can be regarded as a variant of PPA applied to the dual problem of (1). However, with the additional regularized term $\|Ax - b\|^2$, its subproblems require an inverse operator of the form $(I+(1/s)A^TA)^{-1}$ which is hard to implement in some cases. Particularly, A^TA is general or large scale, so the computation of inverse operator may fail. Hence,

ALM is not sufficiently competitive when the objective function $\theta(x)$ is "simple." Extragradient method (EGM) [22] is a practical method for (1) which employs the information of current iteration. In fact, EGM is an explicit type method and requires two calls to the gradient per iteration; therefore, it might suffer slow convergence. Recently, He and Yuan [23] proposed a contraction method based on PPA (PPA-CM) to solve (1), which is elegant and simple. Inspired by PPA-CM, a Lagrangian PPA-based (LPPA) contraction method is presented in [24] which employs an asymmetrical proximal term [25]. These two PPA-based methods have nice convergence properties that are similar in many ways to PPA, but they both require a quite restrictive condition for convergence and are sensitive to the initial choice of stepsizes.

In this paper, we focus on development of PPA-type method for solving (1). Based on LPPA, we propose a general self-adaptive relaxed-PPA method (SRPPA) which is simple yet efficient. Our algorithm capitalizes certain features of PPA, hence, we term it relaxed-PPA. The proposed algorithm has several nice fronts. First, our method is a PPA-type method with asymmetrical linear term, which is clearly a different nature to classical PPA. It relaxes the jointly structure of subproblem to a tractable one. Second, we provide two simple search directions for new iterate. Third, the stepsize choice is flexible, and the condition for convergence guarantee is weaker than both PPA-CM and LPPA. Finally, simple adaptive strategies are employed to choose stepsize, and this appealing aspect is significantly important in practice. We also demonstrate that our method is relevant for various applications whose practical success is made possible by our efficient algorithm.

This paper is organized as follows. In Section 2, we provide some notations and preliminaries which are useful for subsequent analysis. In Section 3, we review some related works. The general relaxed-PPA and its variant are formally presented in Section 4. Self-adaptive strategies to choose stepsize are also described. Next, in Section 5, the convergence analysis is provided. In Section 6, we present some concrete applications of (1) and elaborate on the implementation of our method; preliminary numerical results are also reported to verify the efficiency of our proposed method. Finally, in Section 7, we conclude the paper with a discussion about the future research directions.

2. Preliminaries

In this section, we first establish some important notations used throughout this paper. Then, we describe the variational inequality formulation of (1) which is convenient for the convergence analysis.

$\partial\theta(x)$ denotes the subdifferential set of the convex function $\theta(x)$:

$$\partial\theta(x) := \left\{d \in \Re^n \mid \theta(y) - \theta(x) \geq d^T(y-x), \forall y \in \Re^n\right\}, \tag{3}$$

and $d \in \partial\theta(x)$ is called a subgradient of $\theta(x)$, see [26]. Let $f(x) \in \partial\theta(x)$ and $f(y) \in \partial\theta(y)$, by the convexity of the function θ, we have

$$(x-y)^T(f(x) - f(y)) \geq 0, \quad \forall x, y \in \Re^n, \tag{4}$$

which indicates that the mapping f is monotone.

Now, we show that (1) can be characterized by a variational inequality; see, for example, [27]. By attaching a Lagrange multiplier vector $\lambda \in \Lambda$ to the linear constraint $Ax = b$ (or $Ax \geq b$), the Lagrangian function of (1) is

$$L(x,\lambda) = \theta(x) - \lambda^T(Ax - b); \tag{5}$$

here,

$$\Lambda = \begin{cases} \Re^m, & \text{for the equality constraints } Ax = b, \\ \Re^m_+, & \text{for the inequality constraints } Ax \geq b, \end{cases} \tag{6}$$

and $L(x,\lambda)$ is defined on $\mathcal{X} \times \Lambda$. Then, by the optimality condition, we can easily see that (1) amounts to finding a pair of (x^*, λ^*) which satisfies

$$\begin{gathered} x^* \in \mathcal{X}, \quad (x - x^*)^T\left\{f(x^*) - A^T\lambda^*\right\} \geq 0, \quad \forall x \in \mathcal{X}, \\ \lambda^* \in \Lambda, \quad (\lambda - \lambda^*)^T(Ax^* - b) \geq 0, \quad \forall \lambda \in \Lambda, \end{gathered} \tag{7}$$

where $f(x^*) \in \partial\theta(x^*)$. Denoting

$$u = \begin{pmatrix} x \\ \lambda \end{pmatrix}, \quad F(u) = \begin{pmatrix} f(x) - A^T\lambda \\ Ax - b \end{pmatrix}, \qquad \Omega = \mathcal{X} \times \Lambda, \tag{8}$$

the system (7) can be characterized by the following variational inequality denoted by VI(Ω, F):

$$u^* \in \Omega, \quad (u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega. \tag{9}$$

Recalling the monotonicity of f, it is easy to get that VI(Ω, F) (9) is monotone. Since the solution set of (1) is assumed to be nonempty, the solution set of VI(Ω, F), denoted by Ω^*, is also nonempty. Our analysis will be built upon this equivalent VI formulation.

3. The Existing Related Methods

There are basically two lines of research for VI(Ω, F) (9), either deal with it by implicit methods that are in general computationally intractable or concentrate on relaxing it with explicit methods. In this section, we first briefly review the well-known classical PPA and EGM. And then, PPA-CM [23] and LPPA [24] are discussed, which will provide motivation for our general self-adaptive relaxed-PPA.

3.1. Classical PPA for the Equivalent Variational Inequality. PPA and its variants are implicit methods which have fast asymptotical convergence rate and produce highly accurate solutions. At each iteration, the subproblem of classical PPA consists of a regularized term, which can be expressed as

follows: given any iterate $u^k = (x^k, \lambda^k)$, find $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$\tilde{u}^k \in \Omega, \quad \left(u - \tilde{u}^k\right)^T \left\{F\left(\tilde{u}^k\right) + r\left(\tilde{u}^k - u^k\right)\right\} \geq 0, \quad \forall u \in \Omega. \tag{10}$$

Then, the update step is taken as follows:

$$u^{k+1} = u^k - \gamma\left(u^k - \tilde{u}^k\right), \quad \gamma \in (0, 2). \tag{11}$$

PPA has a nice convergence property

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left\|u^k - u^*\right\|^2 - \gamma(2-\gamma)\left\|u^k - \tilde{u}^k\right\|^2. \tag{12}$$

Although classical PPA is conceptually appealing, subproblem (10) presents certain computational challenges. More specifically, primal variable x and dual variable λ are tied together, and their subproblems are treated as a joint problem. In most cases, this joint subproblem may be as difficult as the original problem (9). As a result, PPA is "conceptual" rather than implementable. It lacks capability in exploiting potential decomposable/specific structure of subproblem. Variants of classical PPA have been explored in the literature, in order to solve the proximal subproblem (10), inexactly, see, for example, [14, 15, 17, 19]. Unfortunately, inexact variants take expensive computation for obtaining approximative solutions.

3.2. The Methods Based on the Simplest Relaxation. To overcome the drawbacks of the classical PPA, it is instinctive to relax subproblem (10) to a solvable one. The most straightforward and simplest relaxation is to replace $F(\tilde{u}^k)$ with $F(u^k)$ in the proximal subproblem (10), which amounts to the following subproblem:

$$\tilde{u}^k \in \Omega, \quad \left(u - \tilde{u}^k\right)^T \left\{F\left(u^k\right) + r\left(\tilde{u}^k - u^k\right)\right\} \geq 0, \quad \forall u \in \Omega. \tag{13}$$

The solution of the relaxed problem (13) is given by $\tilde{u}^k = P_\Omega[u^k - (1/r)F(u^k)]$. It is clear that methods with such relaxation are explicit type methods. However, $\tilde{u}^k$ cannot be accepted directly as the new iterate. Using the terminology "*predictor-corrector*," such point can be viewed as a *predictor*. Here, we list two simple methods which employ *predictor* $\tilde{u}^k$ to obtain *corrector* as the new iterate.

(i) The extragradient method (EGM) updates the new iterate (corrector) by

$$u^{k+1} = P_\Omega\left[u^k - \frac{1}{r}F\left(\tilde{u}^k\right)\right]. \tag{14}$$

(ii) The projection and contraction methods (PCM) [28–30] perform update as follows:

$$u^{k+1} = u^k - \gamma\alpha_k^* d^k \quad \text{or} \quad u^{k+1} = P_\Omega\left[u^k - \frac{\gamma\alpha_k^*}{r}F\left(\tilde{u}^k\right)\right], \tag{15}$$

where

$$d^k = \left(u^k - \tilde{u}^k\right) - \frac{1}{r}\left(F\left(u^k\right) - F\left(\tilde{u}^k\right)\right),$$
$$\alpha_k^* = \frac{\left(u^k - \tilde{u}^k\right)^T d^k}{\left\|d^k\right\|^2}. \tag{16}$$

The sequence $\{u^k\}$ generated by the above mentioned EGM or PCM satisfies

$$\left\|u^{k+1} - u^*\right\|^2 \leq \left\|u^k - u^*\right\|^2 - c\left\|u^k - \tilde{u}^k\right\|^2, \quad c > 0, \tag{17}$$

which is similar to PPA. Both EGM and PCM use the simplest relaxation to obtain $\tilde{u}^k$ in kth iteration, hence are computationally practical. These methods have appealing practical aspects; however, such simplest relaxation does not exploit the inner structure of VI(Ω, F) (9). This observation prompts the need for dedicated relaxations.

3.3. PPA-Type Contraction Method. The algorithms that are closely related to ours are PPA-CM [23] and LPPA [24]. The PPA-CM obtains the predictor $\tilde{u}^k$ by solving the following subproblem: find $(\tilde{x}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$\tilde{u}^k \in \Omega, \quad \left(u - \tilde{u}^k\right)^T \left\{F\left(\tilde{u}^k\right) + S\left(\tilde{u}^k - u^k\right)\right\} \geq 0, \quad \forall u \in \Omega, \tag{18}$$

where

$$S = \begin{pmatrix} rI_n & -A^T \\ -A & sI_m \end{pmatrix}. \tag{19}$$

And perform the update

$$u^{k+1} = u^k - \gamma\left(u^k - \tilde{u}^k\right), \quad \gamma \in (0, 2). \tag{20}$$

The framework of LPPA is as follows:

$$\tilde{u}^k \in \Omega, \quad \left(u - \tilde{u}^k\right)^T \left\{F\left(\tilde{u}^k\right) + M\left(\tilde{u}^k - u^k\right)\right\} \geq 0, \quad \forall u \in \Omega, \tag{21}$$

where

$$M = \begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix}. \tag{22}$$

And the new iterate is defined by

$$u^{k+1} = u^k - \gamma\alpha M\left(u^k - \tilde{u}^k\right), \quad \gamma \in (0, 2). \tag{23}$$

Both procedures are simple and can solve subproblem efficiently; but their nice convergence results require a quite restrictive condition, that is; $rs > \|A^T A\|$ in PPA-CM and $rs > (1/2)\|A^T A\|$ in LPPA, respectively. The stepsizes r, s are directly determined by such condition; hence, it is important to estimate $\|A^T A\|$. Overestimation may lead to poor stepsizes and slow convergence, while underestimation may result in divergence. In addition, they are both quite sensitive to the choice of r, s. To overcome those drawbacks, we propose a general self-adaptive relaxed-PPA, and as mentioned earlier, it can provide improved guarantee for convergence and has potential progress in the choice of stepsize. Furthermore, self-adaptive strategies are designed to improve performance.

Step 1. Initialization. Let $\gamma \in (0,2)$ and pick $(x^0, \lambda^0) \in \Re^n \times \Lambda$, set $k = 0$.
Step 2. Predictor. Let

$$\tilde{x}^k = \operatorname{Argmin}\left\{\theta(x) + \frac{r}{2}\left\|x - \left[x^k + \frac{1}{r}A^T\lambda^k\right]\right\|^2 \mid x \in \mathcal{X}\right\}, \tag{$*$}$$

and

$$\tilde{\lambda}^k = P_\Lambda\left[\lambda^k - \frac{1}{s}(A\tilde{x}^k - b)\right]. \tag{$**$}$$

Step 3. If the stepsizes r and s are chosen satisfying

$$\varphi(u^k, \tilde{u}^k) \geq \frac{1}{4}\left\|d(u^k, \tilde{u}^k)\right\|_G^2, \tag{$\#$}$$

then go to Step 4. Otherwise, increase r, s and go back to Step 2.
Step 4. Corrector and updating

$$u^{k+1} = u^k - \alpha_k G^{-1} M(u^k - \tilde{u}^k). \tag{\$}$$

Step 5. Adjustment

$$(r,s) = \begin{cases} \left(\frac{r}{2}, \frac{s}{2}\right), & \text{if } \varphi(u^k, \tilde{u}^k) \geq \kappa * \left\|d(u^k, \tilde{u}^k)\right\|_G^2; \\ (r,s), & \text{otherwise.} \end{cases} \quad \text{Here } \kappa > 4.$$

Set $k := k + 1$.

Algorithm 1: General primal-dual relaxed-PPA method.

Step 1. Initialization. Let $\gamma \in (0,2)$ and pick $(x^0, \lambda^0) \in \Re^n \times \Lambda$, set $k = 0$.
Step 2. Predictor. Let

$$\tilde{\lambda}^k = P_\Lambda\left[\lambda^k - \frac{1}{s}(Ax^k - b)\right], \tag{$\dagger$}$$

and

$$\tilde{x}^k = \operatorname{Argmin}\left\{\theta(x) + \frac{r}{2}\left\|x - \left[x^k + \frac{1}{r}A^T\tilde{\lambda}^k\right]\right\|^2 \mid x \in \mathcal{X}\right\}. \tag{$\dagger\dagger$}$$

Step 3. If the stepsize r and s are chosen satisfying

$$\varphi\left(u^k, \tilde{u}^k\right) \geq \frac{1}{4}\left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2, \tag{$\ddagger$}$$

then go to Step 4. Otherwise, increase r, s and go back to Step 2.
Step 4. Corrector and updating

$$u^{k+1} = u^k - \alpha_k G^{-1} M(u^k - \tilde{u}^k). \tag{$\wedge$}$$

Step 5. Adjustment

$$(r,s) = \begin{cases} \left(\frac{r}{2}, \frac{s}{2}\right), & \text{if } \varphi\left(u^k, \tilde{u}^k\right) \geq \kappa * \left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2; \\ (r,s), & \text{otherwise.} \end{cases} \quad \text{Here } \kappa > 4.$$

Set $k := k + 1$.

Algorithm 2: General dual-primal relaxed-PPA method.

4. The General Self-Adaptive Relaxed PPA-Method

In this section, we weave together the ideas of the previous section to present general self-adaptive relaxed-PPA method (SRPPA) which is mostly inspired by LPPA [24]. At first sight, the predictor applied in SRPPA is much the same as LPPA, but the stepsize choice condition for convergence is quite different; moreover, we prove that it is weaker than LPPA. Self-adaptive strategies are elaborately designed to ensure the robustness of our algorithm. Two simple yet efficient constructions of new iterate are also presented which will provide some inspirations for designing various search directions.

4.1. General Relaxed-PPA Method. The general primal-dual relaxed-PPA method with implementable structure for (1) is summarized in Algorithm 1. Note that the order of x and λ can be changed to obtain a variant, which is summarized in Algorithm 2. Our relaxed-PPA is intended to blend the implementable properties of EGM (or PCM) with the fast convergence performance of PPA. Now, it is helpful to introduce additional notations that will be used in the rest of this paper. Let G be a positive symmetry definite matrix (we will specify it later),

$$M = \begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix}, \qquad H = \begin{pmatrix} rI_n & 0 \\ 0 & sI_m \end{pmatrix}, \tag{24}$$

$$d\left(u^k, \tilde{u}^k\right) = G^{-1}M\left(u^k - \tilde{u}^k\right), \tag{25}$$

$$\varphi\left(u^k, \tilde{u}^k\right) = \left(u^k - \tilde{u}^k\right)^T M\left(u^k - \tilde{u}^k\right). \tag{26}$$

The relaxed-PPA described here involves two steps. First, we solve the relaxed subproblem $(*)$, $(**)$ to obtain predictor, which is nice and efficient for the nature of the problem under study. Note that the x-predictor step $(*)$ involves minimizing θ plus a convex quadratic function, and under Assumption 2, it can be efficiently solved or it admits a closed form solution. And then, λ-predictor step $(**)$ is just a projection onto Λ which is tractable and computationally efficient. It is clear that the prediction step employs a Gauss-Seidel manner to update information efficiently. The correction step ($) only involves matrix-vector multiplication which is very simple and straightforward.

Remark 3. We first make some insight into the correction step in Algorithm 1. The obtained $\tilde{u}^k$ plays no direct role as the new iterate. Instead, we need some kind of "corrector" defined in ($). Although matrix G in ($) is just a required positive symmetry definite, our goal here is to fully integrate the information of u^k and $\tilde{u}^k$ to construct effective, simple search direction $G^{-1}M(u^k - \tilde{u}^k)$ for the corrector. Based on this consideration, we elaborately provide two simple choices of G.

Case 1. It is natural to set $G = H$ to induce a simple update form. Then, it is easy to get that

$$\begin{pmatrix} x^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ \lambda^k \end{pmatrix} - \alpha_k \begin{pmatrix} x^k - \tilde{x}^k + \frac{1}{r} A^T \left(\lambda^k - \tilde{\lambda}^k\right) \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix}. \quad (27)$$

Case 2. Let $G = MH^{-1}M^T$. This case is a little less intuitive, but it can lead to a simple update form as well as Case 1. The underlying derivation is a little more complicate. Applying G in ($), we get

$$\begin{aligned} u^{k+1} &= u^k - \alpha_k M^{-T} H M^{-1} M \left(u^k - \tilde{u}^k\right) \\ &= u^k - \alpha_k M^{-T} H \left(u^k - \tilde{u}^k\right). \end{aligned} \quad (28)$$

Recalling M is a lower triangular matrix, by the fact that its inverse is also a lower triangular, we have

$$G^{-1}M = M^{-T}H = \begin{pmatrix} I_n & 0 \\ -\frac{A}{s} & I_m \end{pmatrix}. \quad (29)$$

Plugging the previous relationship to (31), we have

$$\begin{pmatrix} x^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ \lambda^k \end{pmatrix} - \alpha_k \begin{pmatrix} x^k - \tilde{x}^k \\ \lambda^k - \tilde{\lambda}^k - \frac{1}{s} A \left(x^k - \tilde{x}^k\right) \end{pmatrix}. \quad (30)$$

In fact, this is a scheme of Gaussian back substitution.

Both cases only involve one matrix-vector multiplication which makes the update form simple. And the computational cost is usually inexpensive.

Remark 4. We now study the subproblem described in $(*)$, $(**)$, and the stepsize choice condition (#). For easy analysis, we characterize $(*)$, $(**)$ as the following VI formulation.

Find $(\tilde{x}^k, \tilde{\lambda}^k) \in \Omega$ such that

$$\begin{aligned} &\begin{pmatrix} x - \tilde{x}^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} f\left(\tilde{x}^k\right) - A^T \tilde{\lambda}^k \\ A\tilde{x}^k - b \end{pmatrix} + \begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix} \begin{pmatrix} \tilde{x}^k - x^k \\ \tilde{\lambda}^k - \lambda^k \end{pmatrix} \right\} \\ &\quad \geq 0, \quad \forall \begin{pmatrix} x \\ \lambda \end{pmatrix} \in \Omega, \end{aligned} \quad (31)$$

and its compact form

$$\tilde{u}^k \in \Omega, \quad \left(u - \tilde{u}^k\right)^T \left\{F\left(\tilde{u}^k\right) + M\left(\tilde{u}^k - u^k\right)\right\} \geq 0, \quad \forall u \in \Omega. \quad (32)$$

We observe that subproblem (32) is similar to (10) in PPA, except for the construction of asymmetry matrix M. As mentioned before, (32) is the same as the prediction subproblem in [24]. Even though they are closely related, the stepsize choice here is quite different. We provide more specific and weaker condition for stepsize r, s. It is clear that condition (#) does not need prior knowledge of matrix A. Furthermore, it only involves matrix-vector multiplication, and so, it is easy to verify, and it is amenable to large-scale A. If r, s fail to meet this convergence condition (#), one should appropriately increase r, s. In the following subsection, we will elaborate on the self-adaptive strategies to increase the stepsizes. At this point, condition (#) may be seen somewhat unmotivated. Some insight into this will be provided later, as we proceed with the convergence analysis. The convergence condition in [24] has a quite different feature: r, s satisfy

$$rs \geq \frac{1}{2} \left\| A^T A \right\|. \quad (33)$$

It is stronger than condition (#). The following lemma is devoted to the proof of this result.

Lemma 5. *Let $\{u^k\}$ be the sequence generated by Algorithm 1, H, $d(u^k, \tilde{u}^k)$, and $\varphi(u^k, \tilde{u}^k)$ defined in* (24), (25), *and* (26), *respectively. Suppose that condition* (33) *is satisfied. Then, condition (#) holds immediately.*

Proof. Note that

$$\begin{aligned} \varphi\left(u^k, \tilde{u}^k\right) &= r\left\|x^k - \tilde{x}^k\right\|^2 + s\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 \\ &\quad + \left(\lambda^k - \tilde{\lambda}^k\right)^T A \left(x^k - \tilde{x}^k\right). \end{aligned} \quad (34)$$

Recall that matrix G described in Algorithm 1 can be designed in two different cases.

Case 1. If $G = H$, we immediately have

$$\begin{aligned} \left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2 &= \left\|u^k - \tilde{u}^k\right\|_H^2 + \frac{1}{r}\left\|A^T(\lambda^k - \tilde{\lambda}^k)\right\|^2 \\ &\quad + 2\left(\lambda^k - \tilde{\lambda}^k\right)^T A \left(x^k - \tilde{x}^k\right). \end{aligned} \quad (35)$$

According to Cauchy-Schwarz inequality, we get

$$\begin{aligned}
&\varphi\left(u^k,\tilde{u}^k\right)-\frac{1}{4}\left\|d\left(u^k,\tilde{u}^k\right)\right\|_G^2\\
&\quad=\frac{3}{4}\left\|u^k-\tilde{u}^k\right\|_H^2-\frac{1}{4r}\left\|A^T\left(\lambda^k-\tilde{\lambda}^k\right)\right\|^2\\
&\qquad+\frac{1}{2}\left(\lambda^k-\tilde{\lambda}^k\right)^TA\left(x^k-\tilde{x}^k\right)\\
&\quad\geq\frac{3}{4}\left\|u^k-\tilde{u}^k\right\|_H^2-\frac{1}{4r}\left\|A^T(\lambda^k-\tilde{\lambda}^k)\right\|^2\\
&\qquad-\frac{1}{4}\left(2r\left\|x^k-\tilde{x}^k\right\|^2+\frac{1}{2r}\left\|A^T\left(\lambda^k-\tilde{\lambda}^k\right)\right\|^2\right)\\
&\quad=\frac{r}{4}\left\|x^k-\tilde{x}^k\right\|^2+\frac{3s}{4}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2\\
&\qquad-\frac{3}{8r}\left\|A^T(\lambda^k-\tilde{\lambda}^k)\right\|^2\geq\frac{r}{4}\left\|x^k-\tilde{x}^k\right\|^2.
\end{aligned}\tag{36}$$

The last inequality follows directly from condition (33).

Case 2. If $G=MH^{-1}M^T$, by the definition of $d(u^k,\tilde{u}^k)$, we obtain

$$\begin{aligned}
\left\|d\left(u^k,\tilde{u}^k\right)\right\|_G^2&=\left(u^k-\tilde{u}^k\right)^TM^TG^{-1}M\left(u^k-\tilde{u}^k\right)\\
&=\left\|u^k-\tilde{u}^k\right\|_H^2.
\end{aligned}\tag{37}$$

Then, we get

$$\begin{aligned}
&\varphi\left(u^k,\tilde{u}^k\right)-\frac{1}{4}\left\|d\left(u^k,\tilde{u}^k\right)\right\|_G^2\\
&\quad=\frac{3}{4}\left\|u^k-\tilde{u}^k\right\|_H^2+\left(\lambda^k-\tilde{\lambda}^k\right)^TA\left(x^k-\tilde{x}^k\right)\\
&\quad\geq\frac{3}{4}\left\|u^k-\tilde{u}^k\right\|_H^2-\frac{1}{2}\frac{4r}{3}\left\|x^k-\tilde{x}^k\right\|^2\\
&\qquad-\frac{1}{2}\frac{3}{4r}\left\|A^T\left(\lambda^k-\tilde{\lambda}^k\right)\right\|^2\\
&\quad=\frac{r}{12}\left\|x^k-\tilde{x}^k\right\|^2+\frac{3s}{4}\left\|\lambda^k-\tilde{\lambda}^k\right\|^2-\frac{3}{8r}\left\|A^T(\lambda^k-\tilde{\lambda}^k)\right\|^2\\
&\quad\geq\frac{r}{12}\left\|x^k-\tilde{x}^k\right\|^2.
\end{aligned}\tag{38}$$

The first inequality follows from the Cauchy-Schwarz inequality, and the last one follows directly from condition (33).

Note that, in both cases, we have that condition (#) holds if $rs\geq(1/2)\|A^TA\|$. □

Condition (33) is not only stronger than Condition (#), but it also requires that matrix M is positive semidefinite, while condition (#) does not. Furthermore, condition (33) may require the explicit expression of A or knowledge of $\|A^TA\|$. Despite these drawbacks, condition (33) is appealing to the problems in which $\|A^TA\|$ is known beforehand or easy to compute/obtain. For instance, A is small scale, an identity matrix or a projection operator, and so forth. It is clear that both condition (#) and (33) are more flexible than the one in PPA-CM [23]. The most aggressive condition (#) may lead to further improvement in stepsize choice. Moreover, it is worthwhile to notice that condition (#) is elegantly designed and provides $\varphi(u^k,\tilde{u}^k)$ with favourable property. In fact, for general matrix G, condition (#) also can guarantee convergence.

Remark 6. The update stepsize α_k plays an important role here. In fact, it can be regarded as an optimal stepsize which will be illustrated in the following section.

Remark 7. We should restrict the adjustment in Step 5 of Algorithm 1 within a limited number to avoid divergence.

In Algorithm 1, we carry out the x-predictor before performing λ-predictor. The roles of x and λ are symmetric; hence, sweeping the order will not break the Gauss-Seidel structure. We switch x and λ and obtain a variant of relaxed-PPA with the order of the x-predictor step and λ-predictor step reversed. This variant is illustrated in Algorithm 2. However, there is no a priori information to know which algorithm is superior. Here, we let

$$M=\begin{pmatrix}rI_n&0\\-A&sI_m\end{pmatrix}.\tag{39}$$

4.2. Adaptive Enhancements. Both PPA-CM and LPPA employ fixed stepsizes r,s. Experiments reveal that they will suffer slow convergence when stepsizes r,s are chosen inappropriately. A natural question is, how to choose the proper initial stepsizes r,s. Here, we propose self-adaptive strategies with the goal of improving the convergence in practice, as well as making performance less dependent on the initial choice of stepsizes. Our strategies dynamically incorporate the information of the current iteration to perform more informative choice of stepsizes for the next iteration [31]. When doing so, the algorithm will be adaptive and free from the initial choice. Denote

$$\begin{pmatrix}d_x^k\\d_\lambda^k\end{pmatrix}=\begin{pmatrix}\left(x^k-\tilde{x}^k\right)+\frac{1}{r}A^T\left(\lambda^k-\tilde{\lambda}^k\right)\\\lambda^k-\tilde{\lambda}^k\end{pmatrix},\tag{40}$$

and then, (31) can be rewritten as

$$\begin{pmatrix}x-\tilde{x}^k\\\lambda-\tilde{\lambda}^k\end{pmatrix}^T\left\{\begin{pmatrix}f\left(\tilde{x}^k\right)-A^T\tilde{\lambda}^k\\A\tilde{x}^k-b\end{pmatrix}-\begin{pmatrix}rI_n&0\\0&sI_m\end{pmatrix}\begin{pmatrix}d_x^k\\d_\lambda^k\end{pmatrix}\right\}\geq0,\quad\forall\begin{pmatrix}x\\\lambda\end{pmatrix}\in\Omega.\tag{41}$$

Under H-norm, the quantity d_x^k can be viewed as a residual for the dual feasibility condition, and d_λ^k can be viewed as a primal residual. These two residuals converge to zero as relaxed-PPA proceeds. Note that

$$\begin{aligned}
\left\|d_x^k\right\|_r^2&=r\left\|x^k-\tilde{x}^k\right\|^2+2\left(x^k-\tilde{x}^k\right)^TA^T\left(\lambda^k-\tilde{\lambda}^k\right)\\
&\quad+\frac{1}{r}\left\|A^T\left(\lambda^k-\tilde{\lambda}^k\right)\right\|^2,\\
\left\|d_\lambda^k\right\|_s^2&=s\left\|\lambda^k-\tilde{\lambda}^k\right\|^2.
\end{aligned}\tag{42}$$

If $\left\|d_x^k\right\|_r^2 \geq \tau_1 \left\|d_\lambda^k\right\|_s^2$
$r := r; \quad s := s * 2;$
else if $\tau_2 \left\|d_x^k\right\|_r^2 \leq \left\|d_\lambda^k\right\|_s^2$
$r := r * 2; \quad s := s;$
else
$r := r * 1.5; s := s * 1.5.$

Algorithm 3: Adaptation-I.

if $r\left\|x^k - \tilde{x}^k\right\|^2 \geq \tau_1 \left(s\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 + \frac{1}{r}\left\|A^T(\lambda^k - \tilde{\lambda}^k)\right\|^2\right)$
$r := \frac{r}{2}; s = \frac{\mu\left\|AA^T\right\|}{r};$
else if $\tau_2 r\left\|x^k - \tilde{x}^k\right\|^2 \leq \left(s\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 + \frac{1}{r}\left\|A^T(\lambda^k - \tilde{\lambda}^k)\right\|^2\right)$
$s := \frac{s}{2}; r = \frac{\mu\left\|AA^T\right\|}{s};$
else
$r := r; \quad s := s.$

Algorithm 4: Adaptation-II.

And this implies that small values of s tend to reduce the primal residual but have potential to enlarge violations of dual feasibility and tend to produce larger dual residual. This observation motivates us to balance primal and dual residuals. When condition (#) fails, we increase stepsizes r, s properly according to the adaptation shown in Algorithm 3.

Here, $\tau_1 > 1$, $\tau_2 > 1$. This adaptation strategy increases s when the dual residual $\|d_x^k\|_r^2$ appears large compared to the primal residual $\|d_\lambda^k\|_s^2$ and increases r when the dual residual $\|d_x^k\|_r^2$ seems too small relative to the primal residual $\|d_\lambda^k\|_s^2$.

As mentioned, condition (33) is stronger than condition (#). If one chooses condition (33), our RPPA also converges. It must have predetermined stepsizes satisfying $rs = \mu\|A^TA\|$ (here, $\mu \geq 0.5$). However, there is no priority knowledge of the choice of individual r or s. Here, we can also adjust r, s automatically when choosing condition (33). Intuitively, we consider expansion of the entire residual under H-norm:

$$\begin{aligned}\left\|d_x^k\right\|_r^2 + \left\|d_\lambda^k\right\|_s^2 &= r\left\|x^k - \tilde{x}^k\right\|^2 + s\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 \\ &\quad + \frac{1}{r}\left\|A^T\left(\lambda^k - \tilde{\lambda}^k\right)\right\|^2 + 2\left(x^k - \tilde{x}^k\right)^T A^T\left(\lambda^k - \tilde{\lambda}^k\right);\end{aligned} \tag{43}$$

there are three terms involving r or s, and we intend to balance these terms. For fixed μ, take $s = (\mu/r)\|AA^T\|$; then

$$s\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 = \frac{\mu}{r}\left\|AA^T\right\|\left\|\lambda^k - \tilde{\lambda}^k\right\|^2. \tag{44}$$

Applying (44) into (43), clearly we have

$$\begin{aligned}\left\|d_x^k\right\|_r^2 + \left\|d_\lambda^k\right\|_s^2 &= r\left\|x^k - \tilde{x}^k\right\|^2 \\ &\quad + \frac{1}{r}\left(\mu\left\|AA^T\right\|\left\|\lambda^k - \tilde{\lambda}^k\right\|^2 + \left\|A^T\left(\lambda^k - \tilde{\lambda}^k\right)\right\|^2\right) \\ &\quad + 2\left(x^k - \tilde{x}^k\right)^T A^T\left(\lambda^k - \tilde{\lambda}^k\right).\end{aligned} \tag{45}$$

Now, we consider adjusting stepsize to balance $r\|x^k - \tilde{x}^k\|^2$ and $(1/r)(\mu\|AA^T\|\|\lambda^k - \tilde{\lambda}^k\|^2 + \|A^T(\lambda^k - \tilde{\lambda}^k)\|^2)$ and obtain another adaptation strategy (see Algorithm 4).

It is worth noting that too many adjustments of stepsizes by Algorithm 4 might cause the algorithm to diverge in practice, and we therefore restrict these adaptations within a limited number of iterations. If one chooses Algorithm 4, there is no need to carry out Step 5 in Algorithm 1 (or Algorithm 2). These techniques embedded into relaxed-PPA automatically choose a "better" stepsize for the next iteration.

5. Convergence Analysis

In this section, we analyze convergence of our primal-dual relaxed-PPA. The convergence analysis of dual-primal scheme can follow a similar procedure.

Let $u^* = (x^*, \lambda^*)$ be any solution point, setting $u = u^*$ in (32) yields

$$\left(\tilde{u}^k - u^*\right)^T M\left(u^k - \tilde{u}^k\right) \geq \left(\tilde{u}^k - u^*\right)^T F\left(\tilde{u}^k\right). \tag{46}$$

Since $\tilde{u}^k \in \Omega$, we have $(\tilde{u}^k - u^*)^T F(u^*) \geq 0$. Consequently, by using the monotonicity of F, the right hand side of (46) is nonnegative, and thus

$$\left(\tilde{u}^k - u^*\right)^T M\left(u^k - \tilde{u}^k\right) \geq 0. \tag{47}$$

Now, we are writing the update as

$$u(\alpha) = u^k - \alpha d\left(u^k, \tilde{u}^k\right) \tag{48}$$

and

$$\begin{aligned} \vartheta(\alpha) &:= \left\|u^k - u^*\right\|_G^2 - \left\|u(\alpha) - u^*\right\|_G^2, \\ q(\alpha) &= 2\alpha\left(u^k - \tilde{u}^k\right)^T G d\left(u^k, \tilde{u}^k\right) - \alpha^2\left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2. \end{aligned} \tag{49}$$

Here, $\vartheta(\alpha)$ can be viewed as a progress function. The following lemma shows that $q(\alpha)$ is a lower bound of $\vartheta(\alpha)$.

Lemma 8. *Let $\vartheta(\alpha)$ and $q(\alpha)$ be defined in* (49)*; then one has*

$$\vartheta(\alpha) \geq q(\alpha). \tag{50}$$

Proof. Let u^* be any solution, from the definition of $u(\alpha)$, we have

$$\begin{aligned} \vartheta(\alpha) &= \left\|u^k - u^*\right\|_G^2 - \left\|\left(u^k - u^*\right) - \alpha d\left(u^k, \tilde{u}^k\right)\right\|_G^2 \\ &= 2\alpha\left(u^k - u^*\right)^T G d\left(u^k, \tilde{u}^k\right) - \alpha^2\left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2. \end{aligned} \tag{51}$$

Applying (47) to the first term of (51) gives

$$\begin{aligned} \left(u^k - u^*\right)^T G d\left(u^k, \tilde{u}^k\right) &= \left(u^k - \tilde{u}^k\right)^T G d\left(u^k, \tilde{u}^k\right) \\ &\quad + \left(\tilde{u}^k - u^*\right)^T G d\left(u^k, \tilde{u}^k\right) \\ &\geq \left(u^k - \tilde{u}^k\right)^T G d\left(u^k, \tilde{u}^k\right). \end{aligned} \tag{52}$$

Substituting (52) into (51), we immediately obtain the assertion. □

We note that $q(\alpha)$ is a quadratic function of α and it is natural to maximize $q(\alpha)$ to obtain an "optimal" α:

$$\alpha_k^* = \frac{\left(u^k - \tilde{u}^k\right)^T G d\left(u^k, \tilde{u}^k\right)}{\left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2}. \tag{53}$$

We now show that the "optimal" α_k^* is bounded above from zero in the following Lemma.

Lemma 9. *Let sequence $\{u^k\}$ be produced by Algorithm 1, α_k^* defined in* (53)*; then, one has*

$$\alpha_k^* \geq \frac{1}{4} > 0. \tag{54}$$

Proof. Using the definition of α_k^* in (53), we have, for all k,

$$\alpha_k^* = \frac{\varphi\left(u^k, \tilde{u}^k\right)}{\left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2} \geq \frac{1}{4}. \tag{55}$$

The inequality follows from condition (#). □

Setting $\alpha = \alpha_k = \alpha_k^* \gamma$ in (50) yields

$$\begin{aligned} \left\|u^{k+1} - u^*\right\|_G^2 &\leq \left\|u^k - u^*\right\|_G^2 \\ &\quad - \gamma(2-\gamma)\alpha_k^*\left(u^k - \tilde{u}^k\right)^T G d\left(u^k, \tilde{u}^k\right). \end{aligned} \tag{56}$$

Combining Lemmas 8 and 9, we immediately obtain the following convergence theorem.

Theorem 10. *Let sequence $\{u^k\}$ be produced by Algorithm 1; then one gets*

$$\left\|u^{k+1} - u^*\right\|_G^2 \leq \left\|u^k - u^*\right\|_G^2 - \frac{\gamma(2-\gamma)}{16}\left\|d\left(u^k, \tilde{u}^k\right)\right\|_G^2. \tag{57}$$

Theorem 11. *Let sequence $\{u^k\}$ be generated by Algorithm 1. Then, $\{u^k\}$ converges to some u^∞ which is a solution of $VI(\Omega, F)$* (9).

Proof. First, for each $u \in \Omega$, we have

$$\left(u - \tilde{u}^k\right)^T F\left(\tilde{u}^k\right) \geq \left(u - \tilde{u}^k\right)^T M\left(u^k - \tilde{u}^k\right). \tag{58}$$

It follows from (57) that $\{u^k\}$ is a bounded sequence and

$$\lim_{k\to\infty}\left\|u^k - \tilde{u}^k\right\|_G = 0. \tag{59}$$

Consequently, $\{\tilde{u}^k\}$ is also bounded. Since $\lim_{k\to\infty}\|u^k - \tilde{u}^k\|_G = 0$, it follows from (58) that

$$\lim_{k\to\infty}\left(u - \tilde{u}^k\right)^T F\left(\tilde{u}^k\right) \geq 0, \quad \forall u \in \Omega. \tag{60}$$

Because $\{\tilde{u}^k\}$ is bounded, it has at least a cluster point. Let u^∞ be a cluster point of $\{\tilde{u}^k\}$ and let the subsequence $\{\tilde{u}^{k_j}\}$ converge to u^∞. It follows that

$$\lim_{j\to\infty}\left(u - \tilde{u}^{k_j}\right)^T F\left(\tilde{u}^{k_j}\right) \geq 0, \quad \forall u \in \Omega, \tag{61}$$

and consequently,

$$\left(u - u^\infty\right)^T F\left(u^\infty\right) \geq 0. \quad \forall u \in \Omega. \tag{62}$$

This means that u^∞ is a solution of $\mathrm{VI}(\Omega, F)$. Note that the inequality (57) is true for all solution points of $\mathrm{VI}(\Omega, F)$, and hence, we have

$$\left\|u^{k+1} - u^\infty\right\|_G^2 \leq \left\|u^k - u^\infty\right\|_G^2, \quad \forall k \geq 0. \tag{63}$$

Since $\tilde{u}^{k_j} \to u^\infty$ $(j \to \infty)$ and $u^k - \tilde{u}^k \to 0$ $(k \to \infty)$, for any given $\varepsilon > 0$, there exists an integer $l > 0$ such that

$$\left\|\tilde{u}^{k_l} - u^\infty\right\|_G < \frac{\varepsilon}{2}, \qquad \left\|u^{k_l} - \tilde{u}^{k_l}\right\|_G < \frac{\varepsilon}{2}. \tag{64}$$

Therefore, for any $k \geq k_l$, it follows from (63) and (64) that

$$\begin{aligned} \left\|u^k - u^\infty\right\|_G &\leq \left\|u^{k_l} - u^\infty\right\|_G \\ &\leq \left\|u^{k_l} - \tilde{u}^{k_l}\right\|_G + \left\|\tilde{u}^{k_l} - u^\infty\right\|_G \leq \varepsilon. \end{aligned} \tag{65}$$

This implies that the sequence $\{u^k\}$ converges to u^∞, which is a solution of $\mathrm{VI}(\Omega, F)$ (9). □

Table 1: Comparison of behaviours of four SRPPAs.

A	SPDRPPAG1-I		SDPRPPAG1-I		SPDRPPAG2-I		SDPRPPAG2-I	
n	It.	CPU	It.	CPU	It.	CPU	It.	CPU
500	360	0.13	241	0.13	513	0.19	875	0.26
1500	379	1.32	259	0.89	387	1.72	594	2.04
2500	501	4.39	343	3.01	429	4.93	648	5.72
3500	399	6.65	325	5.57	663	14.44	1089	18.02
4500	389	12.13	457	14.47	493	20.61	991	31.26
5500	515	24.21	462	21.97	497	30.88	813	38.43
6500	449	25.01	423	23.41	456	33.42	932	51.62
7500	520	45.72	473	41.79	724	63.06	886	77.91

6. Applications and Preliminary Numerical Experiments

The general self-adaptive relaxed-PPA (SRPPA) offers a flexible framework for solving many interesting problems. We illustrate our algorithm with different applications: basis pursuit problem, nearest correlation matrix problem. In this section, we describe the results of experiments whose goal is to demonstrate the efficiency of general relaxed-PPA (RPPA) and its self-adaptive version. To that end, we compare RPPA with certain state-of-the-art algorithms on different problems. Our experiments focus on efficiency and speed of convergence and evaluate the methods in terms of their number of iterations and computational times.

All the codes were written by Matlab R2009b version, and all the numerical experiments were performed on a Lenovo desktop computer with Intel (R) Core (TM) i5 CPU with 3.2 GHz and 3.5 GB RAM.

6.1. Basis Pursuit Problem. Basis pursuit (BP) finds signal representations in overcomplete dictionaries by equality-constrained l_1 minimization problem. Formally, one solves the problem

$$\min\left\{\|x\|_1 Ax = b,\ x \in \Re^n\right\}. \tag{66}$$

And here, $\|\cdot\|_1$ denotes the l_1 norm defined as $\|x\|_1 := \sum_{i=1}^{n} |x_i|$. BP is a fundamental problem in image processing and modern statistical signal processing, particularly the theory of compressed sensing; see, for example, [1–4] for intensive study. We now discuss our approach to BP problem of over-complete representations. Our experiments in this subsection use synthetic data which were mainly designed to illustrate the nice performance of our RPPA. The synthetic problem that we test here is similar to the one employed in [32]. We generate the data as follows: matrix A is a random $m \times n$ matrix, with Gaussian i.i.d. entries of zero mean and variance 1, with $m = n/2$. $x_{\text{original}} \in R^n$ is the original sparse signal, constructed with $m/5$ nonzero values, randomly from standard normal distribution. We use x_{original} to generate the measurements as $b = Ax_{\text{original}}$. It is desirable to use test problems that have a precisely known solution. In fact, when x_{original} is very sparse, it is the solution to the minimization problem (66). Hence, in our synthetic problem, x_{original} is exactly the solution.

In our first experiment, we compared general RPPA using two different G's mentioned in Section 4.1. For BP problem, we use condition (#) and Algorithm 3. Since A constructed here is a general random matrix, and when A is large scale, $\|A^T A\|$ might be obtained costly. A simple stopping criterion

$$\text{err}. = \left\|x^k - x_{\text{original}}\right\| \le \text{Tol} \tag{67}$$

was used in this experiment, and the stopping tolerance Tol was set to 10^{-10}. In all the tests, initial stepsizes were set as $s = 10$, $r = 1$, the primal variable x^0 was initialized as zeros$(n, 1)$, and the dual λ^0 was ones$(m, 1)$ in Matlab. Table 1 summarizes the performance of general SRPPA. Here, SPDRPPAGi-I(SPDRPPAGi-II) denotes self-adaptive primal-dual RPPA with Algorithm 3 (Algorithm 4), $G = H$, if $i = 1$, and $G = MH^{-1}M^T$, if $i = 2$. DPRPPA (DPRPPA) denotes dual-primal RPPA version.

Basically, SRPPAs converge very quickly and achieved tight error 10^{-10} in a few hundred iterations. For this experiment, one can see that SDPRPPAG1-I is fastest in all cases. Both SDPRPPAG1-I and SPDRPPAG1-I are Gaussian type methods, with $G = H$, and they exhibit very similar performance. SDPRPPAG2-I and SPDRPPAG2-I with $G = MH^{-1}M^T$ are Gaussian back substitute form methods and perform a little slower than Gaussian type methods. We also plot a figure to graphically illustrate the performance of four SRPPAs. Figure 1 shows the results from the test with $n = 1000$ and $n = 6000$, depicting error versus CPU time. Quality-wise, SPDRPPAG1-I was on par with SDPRPPAG1-I.

In the second experiment, we compare the performance of SPDRPPAG1-I with TFOCS (source code can be found at http://cvxr.com/tfocs/) [32], ADMM (source code can be found at http://www.stanford.edu/~boyd/papers/admm/) [33], and PPA-CM. To make the comparison independent of the stopping criterion for each algorithm, we first run TFOCS to get its solution x_{TFOCS} and set a benchmark error

$$\text{benchmark err}. = \left\|x_{\text{TFOCS}} - x_{\text{original}}\right\|_2 \tag{68}$$

and then run other algorithms until they obtain smaller errors than this benchmark. TFOCS was stopped upon

$$\frac{\left\|x^{k+1} - x^k\right\|}{\max\left\{1, \left\|x^{k+1}\right\|\right\}} \le \text{Tol}. \tag{69}$$

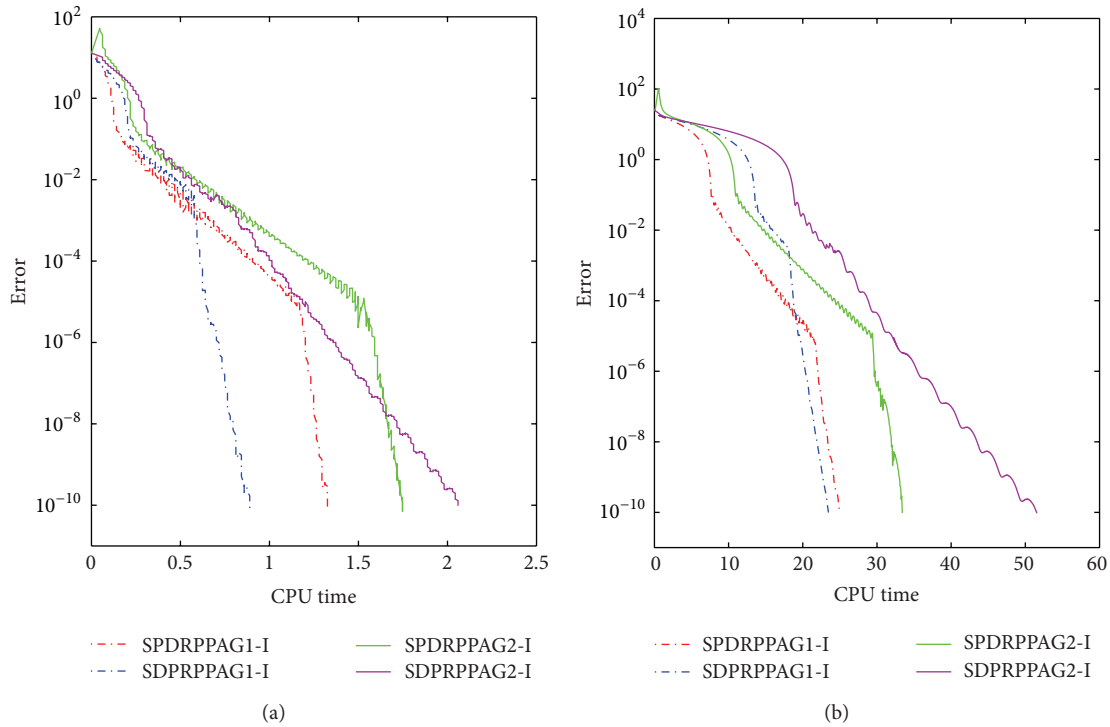

Figure 1: Comparing SRPPAs applied to BP problem with $n = 1000$ (a) and $n = 6000$ (b). The horizontal axis gives the CPU time; the vertical axis gives the error between the solution and the original.

Table 2: Performance of different iterative methods.

A	TFOCS		ADMM		PPA-CM		SPDRPPAG1-I	
n	It./CPU	Err.	It./CPU	Err.	It./CPU	Err.	It./CPU	Err.
512	1681/4.16	$7.5e-10$	492/0.67	$5.0e-10$	1027/1.41	$6.2e-10$	391/0.10	$4.0e-10$
1024	406/1.68	$3.7e-10$	274/0.53	$3.0e-10$	848/0.76	$2.3e-10$	214/0.28	$3.0e-10$
2048	507/4.01	$4.4e-9$	607/3.93	$2.9e-9$	804/3.65	$4.0e-9$	239/1.49	$3.9e-9$
4096	933/19.50	$2.6e-9$	1070/27.47	$2.1e-9$	845/15.09	$2.3e-9$	422/9.84	$2.3e-9$
4200	461/9.91	$2.1e-9$	916/25.48	$1.8e-9$	868/16.27	$2.1e-9$	391/9.44	$1.5e-9$
4300	451/9.74	$2.2e-9$	464/20.37	$1.6e-9$	884/16.91	$2.1e-9$	429/10.54	$2.1e-9$
4600	505/12.37	$1.36e-9$	2155/56.45	$1.2e-9$	863/18.95	$1.2e-9$	425/11.89	$1.1e-9$
4700	801/20.15	$1.3e-11$	1517/45.37	$8.2e-12$	1102/34.27	$1.2e-11$	425/14.20	$1.1e-11$
5500	407/13.64	$2.4e-6$	—/—	—	546/20.13	$1.9e-6$	308/12.40	$2.0e-6$
6500	1257/56.11	$1.5e-5$	—/—	—	508/26.48	$1.4e-5$	308/17.22	$1.3e-5$
7500	801/81.42	$1.1e-12$	—/—	—	1313/813.52	$1.1e-12$	522/69.52	$1.1e-12$
8500	842/107.18	$2.3e-7$	—/—	—	724/100.36	$2.2e-7$	542/83.07	$2.2e-7$

Since we found that $\mathrm{Tol} = 10^{-12}$ is small enough to guarantee very high accuracy, we set $\mathrm{Tol} = 10^{-12}$ in TFOCS. The parameters of TFOCS and ADMM were taken with their defaults. To guarantee the convergence, fixed stepsizes r, s were set to $s = 100$, $r = 1.01 * \|AA^T\|/s$ for PPA-CM. In SPDRPPAG1-I, we also choose the same convergence condition (#) and initial step size $s = 10$, $r = 1$ as the previous experiment. We varied the size of A from $n = 512$ ($m = n/2$) to $n = 8500$. The results of this experiment are summarized in Table 2. There, we report the run time in seconds, the number of iterations, and the error of the recovery solution. In Table 2, "—" means "out of memory."

We observe from Table 2 that four algorithms reach high accuracy around 10^{-9}. SPDRPPAG1-I is about two times faster than the first-order method implemented in the TFOCS package, and moreover, it usually outperforms TFOCS in terms of iterations. For medium size problems, SPDRPPAG1-I is clearly faster than ADMM. Even for small size problems, SPDRPPAG1-I shows its superior performance. The main reason lies in that ADMM computed $(I + AA^T)^{-1}$ to solve its subproblem exactly which would take expensive computational cost. Not surprisingly, the general SPDRPPAG1-I performs better than the primary PPA-CM. Here, the total iterations of SPDRPPAG1-I are less than 50%

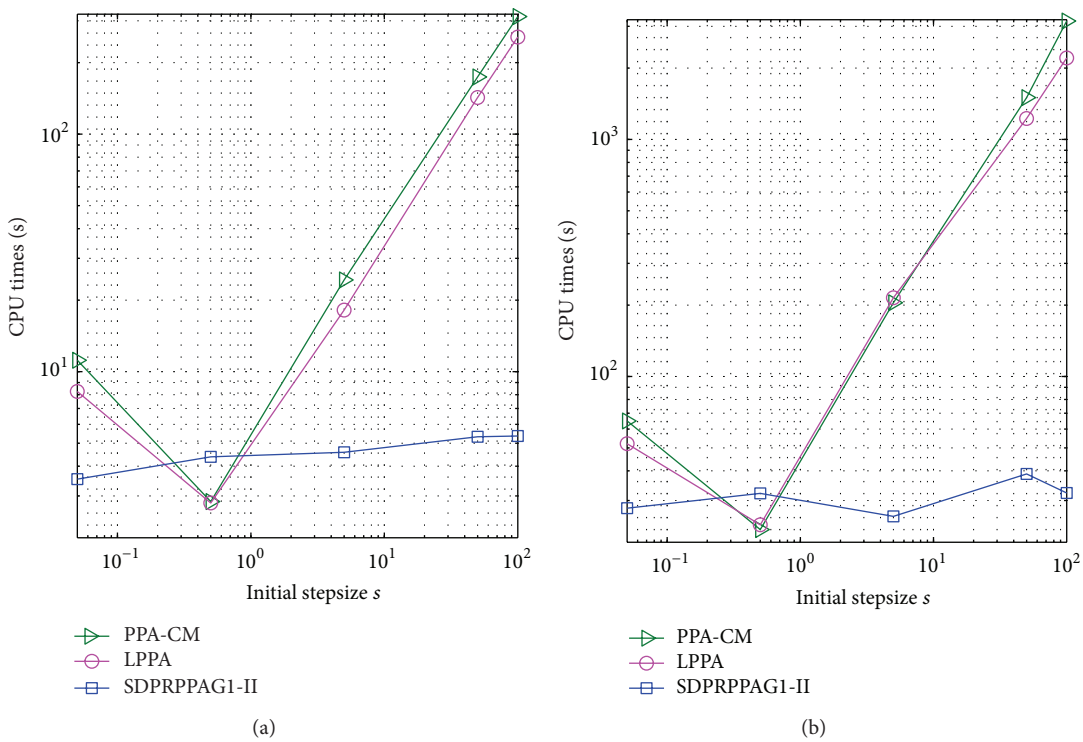

Figure 2: CPU times as a function of the initial stepsize s for PPA-CM, LPPA, and SDPRPPAG1-II. The plot on the left is for $n = 500$, while the plot on the right is for $n = 1000$.

of PPA-CM. As we have mentioned, "optimal" update stepsize α_k and more flexible condition for convergence may provide SPDRPPAG1-I improved performance. SPDRPPAG1-I is faster than PPA-CM in terms of CPU times. However, the superiority of CPU time is not so significant as iteration number. For the cases $n = 4300$, it is just about 62% of PPA-CM. This is not particularly surprising; compared to PPA-CM, SPDRPPAG1-I has to take extra computation for convergence condition and "optimal" α_k in each iteration.

6.2. Nearest Correlation Matrix Problem. The nearest correlation matrix problem is solving the problem

$$\min\left\{\frac{1}{2}\|X - C\|_F^2 \mid \operatorname{diag}(X) = e,\ X \in S_+^n\right\}, \qquad (70)$$

where $e \in \Re^n$ is the vector whose entries are all 1s, S_+^n denotes the cone of positive definite symmetric matrices, diag(X) is the vector of diagonal elements of X, and $\|\cdot\|_F$ denotes the matrix Fröbenius norm $\|X\|_F = \operatorname{trace}(X^T X)^{1/2}$.

Here, we apply PPA-CM, LPPA, and SDPRPPA1-II for solving (70). The standard Matlab Mex interface mexsvd is used to conduct the eigenvalue decomposition. We constructed test data sets and stopping criterion like those of [24]. As mentioned in the prequel, we expect our SRPPA to produce robust performance. To assess the effectiveness of the adaptive strategies proposed in Section 6, we now move on to the description of experiments that focus on the consequences of the initial stepsizes. For investigating, we used dimensions $n \in \{500, 1000\}$ and varied s from 0.05 to 100, and initial points were set to 0 in all cases. Note that $A = I$; we fixed $r = 1.01/s$ for PPA-CM, $r = 0.65/s$ for LPPA and chose $r = 0.65/s$ as initial start for SDPRPPAG1-II. Since the experiments with other values of n give qualitatively similar results, we therefore do not plot those results to avoid clutter in the figures. The respective numerical results are plotted in Figure 2.

It is clear that, for PPA-CM and LPPA, the convergence performance was a result of the stepsize selection. They are both fairly sensitive to initial stepsize s (or r). The results confirm that, with inappropriate stepsizes, both PPA-CM and LPPA become significantly slow. SDPRPPAG1-II yields significantly robust performance with adaptive strategy. And it is independent of the initial stepsizes and illustrates its superior performance. Furthermore, SDPRPPAG1-II yields competitive results even when PPA-CM and LPPA chose the "good" initial stepsize. This underlies the importance of adaptive strategy in producing good performance. Of course, care should be taken. For instance, the cost of computing optimal stepsize α_k here is negligible, compared to the computation of SVD; when they are more costly, general LPPA will be expected to perform slower than PPA-CM.

7. Conclusions

In this paper, we proposed an efficient general self-adaptive relaxed-PPA method for linearly constrained convex programming and provided theoretical convergence analysis for

this method. The stepsizes choice condition is flexible and simple. Self-adaptive strategies are provided to make our method more efficient and robust. Experiments of the method in comparison to other state-of-art methods are provided to confirm the efficiency of the proposed method. Our numerical results suggest that SRPPA is effective and simple to implement. There are a few directions for further research, but we list here only two. The first is the question of whether we may modify the algorithm to work with more general constrained convex problems. Second, we aim to provide various relaxations of the subproblem for the practical purpose.

Acknowledgments

The author is grateful to Caihua Chen, Wenxing Zhang, and Xiangfeng Wang for interesting discussions on nearest correlation matrix problem. Xiaoling Fu was supported by the NSFC Grant 70901018.

References

[1] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[3] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[5] R. Borsdorf and N. J. Higham, "A preconditioned Newton algorithm for the nearest correlation matrix," *IMA Journal of Numerical Analysis*, vol. 30, no. 1, pp. 94–107, 2010.

[6] S. Boyd and L. Xiao, "Least-squares covariance matrix adjustment," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 2, pp. 532–546, 2005.

[7] N. J. Higham, "Computing the nearest correlation matrix—a problem from finance," *IMA Journal of Numerical Analysis*, vol. 22, no. 3, pp. 329–343, 2002.

[8] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[9] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[10] E. J. Candès and T. Tao, "The power of convex relaxation: near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[11] B. Martinet, "Régularisation d'inéquations variationnelles par approximations successives," *Revue Franqaise dInformatique et de Recherche Operationelle*, vol. 4, pp. 154–158, 1970.

[12] A. Bnouhachem and M. A. Noor, "Inexact proximal point method for general variational inequalities," *Journal of Mathematical Analysis and Applications*, vol. 324, no. 2, pp. 1195–1212, 2006.

[13] A. Bnouhachem and M. A. Noor, "An interior proximal point algorithm for nonlinear complementarity problems," *Nonlinear Analysis: Hybrid Systems*, vol. 4, no. 3, pp. 371–380, 2010.

[14] J. V. Burke and M. Qian, "A variable metric proximal point algorithm for monotone operators," *SIAM Journal on Control and Optimization*, vol. 37, no. 2, pp. 353–375, 1999.

[15] J. Eckstein, "Approximate iterations in Bregman-function-based proximal algorithms," *Mathematical Programming*, vol. 83, no. 1, pp. 113–123, 1998.

[16] J. Eckstein and P. J. S. Silva, "Proximal methods for nonlinear programming: double regularization and inexact subproblems," *Computational Optimization and Applications*, vol. 46, no. 2, pp. 279–304, 2010.

[17] O. Güler, "On the convergence of the proximal point algorithm for convex minimization," *SIAM Journal on Control and Optimization*, vol. 29, no. 2, pp. 403–419, 1991.

[18] M. R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, pp. 303–320, 1969.

[19] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[20] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 1–108, 2013.

[21] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 97–116, 1976.

[22] G. M. Korpelevič, "An extragradient method for finding saddle points and for other problems," *Èkonomika i Matematicheskie Metody*, vol. 12, no. 4, pp. 747–756, 1976.

[23] B. S. He and X. M. Yuan, "A contraction method with implementable proximal regularization for linearly constrained convex programming," *Optimization Online*, pp. 1–14, 2010.

[24] Y. F. You, X. L. Fu, and B. S. He, "Lagrangian PPA-based contractionmethods for linearly constrained convex optimization," *Pacific Journal of Optimization*. In press.

[25] B. S. He, X. L. Fu, and Z. K. Jiang, "Proximal-point algorithm using a linear proximal term," *Journal of Optimization Theory and Applications*, vol. 141, no. 2, pp. 299–319, 2009.

[26] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, no. 28, Princeton University Press, Princeton, NJ, USA, 1970.

[27] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, NY, USA, 2003.

[28] B. He, "A class of projection and contraction methods for monotone variational inequalities," *Applied Mathematics and Optimization*, vol. 35, no. 1, pp. 69–76, 1997.

[29] B. S. He and L. Z. Liao, "Improvements of some projection methods for monotone nonlinear variational inequalities," *Journal of Optimization Theory and Applications*, vol. 112, no. 1, pp. 111–128, 2002.

[30] B. He, X. Yuan, and J. J. Z. Zhang, "Comparison of two kinds of prediction-correction methods for monotone variational inequalities," *Computational Optimization and Applications*, vol. 27, no. 3, pp. 247–267, 2004.

[31] X. Fu and B. He, "Self-adaptive projection-based prediction-correction method for constrained variational inequalities," *Frontiers of Mathematics in China*, vol. 5, no. 1, pp. 3–21, 2010.

[32] S. R. Becker, E. J. Candès, and M. C. Grant, "Templates for convex cone problems with applications to sparse signal

recovery," *Mathematical Programming Computation*, vol. 3, no. 3, pp. 165–218, 2011.

[33] 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.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 938724, 9 pages
http://dx.doi.org/10.1155/2013/938724

Research Article

The Study of Fixed Point Theory for Various Multivalued Non-Self-Maps

Wei-Shih Du,[1] Erdal Karapınar,[2] and Naseer Shahzad[3]

[1] *Department of Mathematics, National Kaohsiung Normal University, Kaohsiung 824, Taiwan*
[2] *Department of Mathematics, Atilim University, İncek, 06836 Ankara, Turkey*
[3] *Department of Mathematics, King Abdulaziz University, P.O. Box 80203, Jeddah 21859, Saudi Arabia*

Correspondence should be addressed to Erdal Karapınar; erdalkarapinar@yahoo.com

Received 24 April 2013; Accepted 19 July 2013

Academic Editor: Abdellah Bnouhachem

The basic motivation of this paper is to extend, generalize, and improve several fundamental results on the existence (and uniqueness) of coincidence points and fixed points for well-known maps in the literature such as Kannan type, Chatterjea type, Mizoguchi-Takahashi type, Berinde-Berinde type, Du type, and other types from the class of self-maps to the class of non-self-maps in the framework of the metric fixed point theory. We establish some fixed/coincidence point theorems for multivalued non-self-maps in the context of complete metric spaces.

1. Introduction

During the last few decades, the celebrated Banach contraction principle, also known as the Banach fixed point theorem [1], has become one of the core topics of applied mathematical analysis. As a consequence, a number of generalizations, extensions, and improvement of the praiseworthy Banach contraction principle in various direction have been explored and reported by various authors; see, for example, [2–30] and the references therein. In parallel with the Banach contraction principle, Kannan [5] and Chatterjea [6] created, respectively, different type, fixed point theorems as follows.

Theorem 1 (Kannan). *Let (X,d) be a complete metric space, $T : X \to X$ is a single-valued map, and $\gamma \in [0,1/2)$. Assume that*

$$d(Tx,Ty) \le \gamma [d(x,Tx) + d(y,Ty)] \quad \forall x, y \in X. \tag{1}$$

Then T has a unique fixed point in X.

Theorem 2 (Chatterjea). *Let (X,d) be a complete metric space, $T : X \to X$ is a single-valued map, and $\gamma \in [0,1/2)$. Assume that*

$$d(Tx,Ty) \le \gamma [d(x,Ty) + d(y,Tx)] \quad \forall x, y \in X. \tag{2}$$

Then T has a unique fixed point in X.

The characterization of the renowned Banach fixed point theorem in the setting of multivalued maps is one of the most outstanding ideas of research in fixed point theory. The remarkable examples in this trend were given by Nadler [2], Mizoguchi and Takahashi [3], and M. Berinde and V. Berinde [4]. On the other hand, investigation of the existence of a fixed point of non-self-maps under certain condition is an interesting research subject of metric fixed point theory, see, for example, [19–27], and references therein.

The following attractive result was reported by M. Berinde and V. Berinde [4] in 2007.

Theorem 3 (M. Berinde and V. Berinde). *Let (X,d) be a complete metric space, $T : X \to \mathscr{CB}(X)$ a multivalued map, $\varphi : [0,\infty) \to [0,1)$ an $\mathscr{MT}$-function (i.e., $\limsup_{s\to t^+}\varphi(s) < 1$ for all $t \in [0,\infty)$), and $L \ge 0$. Assume that*

$$\mathscr{H}(Tx,Ty) \le \varphi(d(x,y))\, d(x,y) + L d(y,Tx) \quad \forall x, y \in X. \tag{3}$$

Then T has a fixed point in X.

If we take $L = 0$ in Theorem 3, then we conclude the remarkable result of Mizoguchi and Takahashi [3] which is a partial answer of problem 9 in [8].

Theorem 4 (Mizoguchi and Takahashi). *Let (X,d) be a complete metric space, $T : X \to \mathcal{CB}(X)$ a multivalued map, and $\varphi : [0,\infty) \to [0,1)$ an $\mathcal{MT}$-function. Assume that*

$$\mathcal{H}(Tx,Ty) \le \varphi(d(x,y))\, d(x,y) \quad \forall x,y \in X. \tag{4}$$

Then T has a fixed point in X.

Recently, Du [12] established the following theorem which is an extension of Theorem 3 and hence Theorem 4.

Theorem 5 (Du). *Let (X,d) be a complete metric space, $T : X \to \mathcal{CB}(X)$ a multivalued map, $\varphi : [0,\infty) \to [0,1)$ an $\mathcal{MT}$-function and $h : X \to [0,\infty)$ a function. Assume that*

$$\mathcal{H}(Tx,Ty) \le \varphi(d(x,y))\, d(x,y) + h(y)\, d(y,Tx) \quad \forall x,y \in X. \tag{5}$$

Then T has a fixed point in X.

The basic objective of this paper is to investigate the existence of coincidence and fixed points of multivalued non-self-maps under the certain conditions in the setting of metric spaces. The presented results generalize, improve, and extend several crucial and notable results that examine the existence of the coincidence/fixed point of well-known maps such as Kannan type, Chatterjea type, Mizoguchi-Takahashi type, Berinde-Berinde type, Du type, and other types in the context of complete metric spaces.

2. Preliminaries

Let (X,d) be a metric space. For each $x \in X$ and $A \subseteq X$, let $d(x,A) = \inf_{y\in A} d(x,y)$. Denote by $\mathcal{N}(X)$ the class of all nonempty subsets of X and $\mathcal{CB}(X)$ the family of all nonempty closed and bounded subsets of X. A function $\mathcal{H} : \mathcal{CB}(X) \times \mathcal{CB}(X) \to [0,\infty)$ defined by

$$\mathcal{H}(A,B) = \max\left\{\sup_{x\in B} d(x,A), \sup_{x\in A} d(x,B)\right\} \tag{6}$$

is said to be the Hausdorff metric on $\mathcal{CB}(X)$ induced by the metric d on X. It is also known that $(\mathcal{CB}(X),\mathcal{H})$ is a complete metric space whenever (X,d) is a complete metric space.

Let K be a nonempty subset of X, $g : K \to X$ a single-valued map, and $T : K \to \mathcal{N}(X)$ a multivalued map. A point x in X is a *coincidence point* of g and T if $gx \in Tx$. If $g = id$ is the identity map, then $x = gx \in Tx$ and call x a *fixed point* of T. The set of fixed points of T and the set of coincidence point of g and T are denoted by $\mathcal{F}_K(T)$ and $\mathcal{COP}_K(g,T)$, respectively. In particular, if $K \equiv X$, we use $\mathcal{F}(T)$ and $\mathcal{COP}(g,T)$ instead of $\mathcal{F}_K(T)$ and $\mathcal{COP}_K(g,T)$, respectively. Throughout this paper, we denote by $\mathbb{N}$, and $\mathbb{R}$, the set of positive integers and real numbers, respectively.

Let f be a real-valued function defined on $\mathbb{R}$. For $c \in \mathbb{R}$, we recall that

$$\limsup_{x\to c^+} f(x) = \inf_{\varepsilon>0} \sup_{c<x<c+\varepsilon} f(x). \tag{7}$$

Definition 6 (see [9–18]). A function $\varphi : [0,\infty) \to [0,1)$ is said to be an *$\mathcal{MT}$-function* (or *$\mathcal{R}$-function*) if $\limsup_{s\to t^+}\varphi(s) < 1$ for all $t \in [0,\infty)$.

It is evident that if $\varphi : [0,\infty) \to [0,1)$ is a nondecreasing function or a nonincreasing function, then φ is an $\mathcal{MT}$-function. So the set of $\mathcal{MT}$-functions is a rich class. An example which is not an $\mathcal{MT}$-function is given as follows. Let $\varphi : [0,\infty) \to [0,1)$ be defined by

$$\varphi(t) := \begin{cases} \dfrac{\sin t}{t}, & \text{if } t \in \left(0, \dfrac{\pi}{2}\right] \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

We note that φ is not an $\mathcal{MT}$-function, since $\limsup_{s\to 0^+}\varphi(s) = 1$.

In what follows that, we recall some characterizations of $\mathcal{MT}$ -functions proved first by Du [12].

Theorem 7 (see [12]). *Let $\varphi : [0,\infty) \to [0,1)$ be a function. Then the following statements are equivalent.*

(a) *φ is an $\mathcal{MT}$-function.*

(b) *For each $t \in [0,\infty)$, there exist $r_t^{(1)} \in [0,1)$ and $\varepsilon_t^{(1)} > 0$ such that $\varphi(s) \le r_t^{(1)}$ for all $s \in (t, t+\varepsilon_t^{(1)})$.*

(c) *For each $t \in [0,\infty)$, there exist $r_t^{(2)} \in [0,1)$ and $\varepsilon_t^{(2)} > 0$ such that $\varphi(s) \le r_t^{(2)}$ for all $s \in [t, t+\varepsilon_t^{(2)}]$.*

(d) *For each $t \in [0,\infty)$, there exist $r_t^{(3)} \in [0,1)$ and $\varepsilon_t^{(3)} > 0$ such that $\varphi(s) \le r_t^{(3)}$ for all $s \in (t, t+\varepsilon_t^{(3)}]$.*

(e) *For each $t \in [0,\infty)$, there exist $r_t^{(4)} \in [0,1)$ and $\varepsilon_t^{(4)} > 0$ such that $\varphi(s) \le r_t^{(4)}$ for all $s \in [t, t+\varepsilon_t^{(4)})$.*

(f) *For any nonincreasing sequence $\{x_n\}_{n\in\mathbb{N}}$ in $[0,\infty)$, one has $0 \le \sup_{n\in\mathbb{N}}\varphi(x_n) < 1$.*

(g) *φ is a function of contractive factor; that is, for any strictly decreasing sequence $\{x_n\}_{n\in\mathbb{N}}$ in $[0,\infty)$, we have $0 \le \sup_{n\in\mathbb{N}}\varphi(x_n) < 1$.*

3. Existence Theorems of Coincidence Points and Fixed Points for Multivalued Non-Self-Maps of Kannan Type and Chatterjea Type

In this section, we prove the existence of coincidence points and fixed points of multivalued non-self-maps of Kannan type and Chatterjea type. For this purpose, we first established a new intersection theorem of $\mathcal{COP}_K(g,T)$ and $\mathcal{F}_K(T)$ for multivalued non-self-maps in complete metric spaces.

Theorem 8. *Let (X,d) be a complete metric space, K a nonempty closed subset of X, $T : K \to \mathcal{CB}(X)$ a multivalued map and $g : K \to X$ a continuous self-map. Suppose that*

(D1) *$Tx \cap K \ne \emptyset$ for all $x \in K$,*

(D2) *$Tx \cap K$ is g-invariant (i.e., $g(Tx \cap K) \subseteq Tx \cap K$) for each $x \in K$,*

(D3) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}&\mathscr{H}(Tx, Ty\cap K)\\&\quad\le \gamma\left[d(x, Tx\cap K) + d(y, Tx\cap K) + d(y, Ty\cap K)\right]\\&\qquad + h(y)\, d(gy, Tx\cap K)\\&\qquad\qquad \forall x, y\in K.\end{aligned} \tag{9}$$

Then $\mathscr{COP}_K(g,T)\cap\mathscr{F}_K(T)\neq\emptyset$.

Proof. Since K a nonempty closed subset of X and X is complete, (K,d) is also a complete metric space. Let $x\in K$. Put $k=\gamma/(1-\gamma)$ and $\lambda=(1+k)/2$. So $0\le k<\lambda<1$. Let $y\in Tx\cap K$ be arbitrary. Then $d(y,Tx\cap K))=0$. By (D2), we have $d(gy,Tx\cap K)=0$. Hence (9) implies

$$\begin{aligned}&\mathscr{H}(Tx, Ty\cap K)\\&\quad\le \gamma\left[d(x, Tx\cap K) + \mathscr{H}(Tx, Ty\cap K)\right]\\&\qquad\qquad \forall y\in Tx\cap K.\end{aligned} \tag{10}$$

Inequality (10) shows that

$$\begin{aligned}d(y, Ty\cap K) &\le \mathscr{H}(Tx, Ty\cap K)\\&\le kd(x, Tx\cap K) < \lambda d(x,y)\\&\qquad \forall y\in Tx\cap K.\end{aligned} \tag{11}$$

Let $x\in K$ be given. Take $x_1=x$. By (D1), $Tx_1\cap K\neq\emptyset$. Choose $x_2\in Tx_1\cap K$. If $x_2=x_1$, then $x_1\in\mathscr{F}_K(T)$ and hence $gx_1\in Tx_1$ from (D2). Hence $x_1\in\mathscr{COP}_K(g,T)\cap\mathscr{F}_K(T)$ and the proof is finished. Otherwise, if $x_2\neq x_1$, then $d(x_1,x_2)>0$. By (11), we have

$$d(x_2, Tx_2\cap K) < \lambda d(x_1,x_2), \tag{12}$$

which implies that there exists $x_3\in Tx_2\cap K$ such that

$$d(x_2,x_3) < \lambda d(x_1,x_2). \tag{13}$$

Next, by (11) again, there exists $x_4\in Tx_3\cap K$ such that

$$d(x_3,x_4) < \lambda d(x_2,x_3). \tag{14}$$

By induction, we can obtain a sequence $\{x_n\}$ in K satisfying

$$x_{n+1}\in Tx_n\cap K, \tag{15}$$

$$d(x_{n+1},x_{n+2}) < \lambda d(x_n,x_{n+1}). \tag{16}$$

By (16), we have

$$\begin{aligned}d(x_{n+1},x_{n+2}) &< \lambda d(x_n,x_{n+1})\\&< \lambda^2 d(x_{n-1},x_n)\\&< \cdots\\&< \lambda^n d(x_1,x_2), \quad \text{for } n\in\mathbb{N}.\end{aligned} \tag{17}$$

Let $\rho_n=(\lambda^{n-1}/(1-\lambda))d(x_1,x_2)$, $n\in\mathbb{N}$. For $m,n\in\mathbb{N}$ with $m>n$, we have

$$d(x_n,x_m)\le\sum_{j=n}^{m-1} d(x_j,x_{j+1}) < \rho_n. \tag{18}$$

Since $0<\lambda<1$, $\lim_{n\to\infty}\rho_n=0$ and hence $\lim_{n\to\infty}\sup\{d(x_n,x_m) : m>n\}=0$. This proves that $\{x_n\}$ is a Cauchy sequence in K. By the completeness of K, there exists $v\in K$ such that $x_n\to v$ as $n\to\infty$. By (15) and (D2), we have

$$gx_{n+1}\in Tx_n\cap K \quad \text{for each } n\in\mathbb{N}. \tag{19}$$

Since g is continuous and $\lim_{n\to\infty}x_n=v$, we have

$$\lim_{n\to\infty} gx_n = gv. \tag{20}$$

Since the function $x\mapsto d(x,Tv)$ is continuous, by (9), (15), (19), and (20), we get

$$\begin{aligned}&d(v, Tv\cap K)\\&\quad=\lim_{n\to\infty} d(x_{n+1}, Tv\cap K)\\&\quad\le\lim_{n\to\infty}\mathscr{H}(Tx_n, Tv\cap K)\\&\quad\le\lim_{n\to\infty}\{\gamma\left[d(x_n, Tx_n\cap K) + d(v, Tx_n\cap K)\right.\\&\qquad\qquad \left.+d(v, Tv\cap K)\right]\\&\qquad\qquad +h(v)\, d(gv, Tx_n\cap K)\}\\&\quad\le\lim_{n\to\infty}\{\gamma\left[d(x_n, x_{n+1})\right.\\&\qquad\qquad \left.+d(v, x_{n+1}) + d(v, Tv\cap K)\right]\\&\qquad\qquad +h(v)\, d(gv, gx_{n+1})\}\\&\quad=\gamma d(v, Tv\cap K),\end{aligned} \tag{21}$$

which implies $d(v,Tv\cap K)=0$. By the closedness of Tv, we have $v\in Tv\cap K$. From (D2), $gv\in Tv\cap K\subseteq Tv$. Hence we verify $v\in\mathscr{COP}_K(g,T)\cap\mathscr{F}_K(T)$. The proof is complete. □

Theorem 9. *In Theorem 8, if condition (D3) is replaced with one of the following conditions:*

(K1) *there exist a function* $h : K\to[0,\infty)$ *and* $\gamma\in[0,1/2)$ *such that*

$$\begin{aligned}&\mathscr{H}(Tx, Ty\cap K)\\&\quad\le\gamma\left[d(x, Tx\cap K)\right.\\&\qquad\qquad \left.+d(y, Tx\cap K) + d(y, Ty\cap K)\right]\\&\qquad + h(y)\, d(gy, Tx) \quad \forall x, y\in K,\end{aligned} \tag{22}$$

(K2) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx\cap K) + d(y, Tx) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx\cap K) \quad \forall x, y \in K,\end{aligned} \tag{23}$$

(K3) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx\cap K) + d(y, Tx) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y \in K,\end{aligned} \tag{24}$$

(K4) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx) + d(y, Tx\cap K) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx\cap K) \quad \forall x, y \in K,\end{aligned} \tag{25}$$

(K5) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx) + d(y, Tx\cap K) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y \in K,\end{aligned} \tag{26}$$

(K6) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx) + d(y, Tx) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx\cap K) \quad \forall x, y \in K,\end{aligned} \tag{27}$$

(K7) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx) + d(y, Tx) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y \in K.\end{aligned} \tag{28}$$

Then $\mathcal{COP}_K(g,T) \cap \mathcal{F}_K(T) \neq \emptyset$.

Proof. It is obvious that any of these conditions (K1)–(K7) implies condition (D3) as in Theorem 8. So the desired conclusion follows from Theorem 8 immediately. □

The following fixed point theorem for multivalued non-self-maps of generalized Kannan type can be established immediately from Theorem 9 for $g \equiv id$ (the identity mapping).

Theorem 10. *Let* (X,d) *be a complete metric space,* K *a nonempty closed subset of* X*, and* $T : K \to \mathcal{CB}(X)$ *a multivalued map. Suppose that* $Tx \cap K \neq \emptyset$ *for all* $x \in K$ *and one of the following conditions holds:*

(P1) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx\cap K) + d(y, Tx\cap K) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y \in K,\end{aligned} \tag{29}$$

(P2) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx\cap K) + d(y, Tx\cap K) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(y, Tx) \quad \forall x, y \in K,\end{aligned} \tag{30}$$

(P3) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx\cap K) + d(y, Tx) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y \in K,\end{aligned} \tag{31}$$

(P4) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx\cap K) + d(y, Tx) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(y, Tx) \quad \forall x, y \in K,\end{aligned} \tag{32}$$

(P5) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx) + d(y, Tx\cap K) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y \in K,\end{aligned} \tag{33}$$

(P6) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathcal{H}(Tx, Ty\cap K) &\le \gamma\left[d(x, Tx) + d(y, Tx\cap K) + d(y, Ty\cap K)\right] \\ &\quad + h(y)\, d(y, Tx) \quad \forall x, y \in K,\end{aligned} \tag{34}$$

(P7) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \gamma\left[d(x,Tx)+d(y,Tx)+d(y,Ty\cap K)\right] \\ &\quad +h(y)\,d(y,Tx\cap K) \quad \forall x,y\in K,\end{aligned} \tag{35}$$

(P8) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \gamma\left[d(x,Tx)+d(y,Tx)+d(y,Ty\cap K)\right] \\ &\quad +h(y)\,d(y,Tx) \quad \forall x,y\in K.\end{aligned} \tag{36}$$

Then $\mathscr{F}_K(T)\neq\emptyset$.

As a consequence of Theorem 10, we obtain the following generalized Kannan type fixed point theorems for multivalued maps.

Corollary 11. *Let* (X,d) *be a complete metric space,* K *a nonempty closed subset of* X*, and* $T : K \to \mathscr{CB}(X)$ *a multivalued map. Suppose that* $Tx\cap K\neq\emptyset$ *for all* $x\in K$ *and there exists* $\gamma\in[0,1/2)$ *such that*

$$\mathscr{H}(Tx,Ty\cap K)\le\gamma\left[d(x,Tx\cap K)+d(y,Ty\cap K)\right] \quad \forall x,y\in K. \tag{37}$$

Then $\mathscr{F}_K(T)\neq\emptyset$.

Remark 12. (a) If $K = X$ in Corollary 11, then we can obtain a multivalued version of Kannan's fixed point theorem [5].

(b) Theorems 8–10 and Corollary 11 all extend and generalize Kannan's fixed point theorem.

Theorem 13. *Let* (X,d) *be a complete metric space,* K *a nonempty closed subset of* X*,* $T : K \to \mathscr{CB}(X)$ *a multivalued map, and* $g : K \to X$ *a continuous self-map. Suppose that conditions (D1) and (D2) as in Theorem 8 hold. If there exist* $h : K \to [0,\infty)$ *and* $\gamma\in[0,1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx,Ty\cap K) &\le \alpha\left[d(x,Ty\cap K)+d(y,Tx\cap K)\right] \\ &\quad +h(y)\,d(gy,Tx\cap K) \quad \forall x,y\in K.\end{aligned} \tag{38}$$

Then $\mathscr{COP}_K(g,T)\cap\mathscr{F}_K(T)\neq\emptyset$.

Proof. Let $x\in K$. Since $\alpha\in[0,1/2)$, by the denseness of $\mathbb{R}$, we can find $\beta>0$ such that $\alpha<\beta<1/2$. Let $y\in Tx\cap K$ be arbitrary. Then $d(y,Tx\cap K)=0$. By (D2), we have $d(gy,Tx\cap K)=0$. Hence (38) has been reduced to

$$\begin{aligned}d(y,Ty\cap K) &\le \mathscr{H}(Tx,Ty\cap K) \\ &\le \alpha d(x,Ty\cap K) \\ &< \beta d(x,Ty\cap K) \quad \forall y\in Tx\cap K.\end{aligned} \tag{39}$$

Let $x\in K$ be given. Take $x_1 = x$. By (D1), $Tx_1\cap K\neq\emptyset$. Choose $x_2\in Tx_1\cap K$. If $x_2 = x_1$, then $x_1\in\mathscr{F}_K(T)$ and hence $gx_1\in Tx_1$ from (D2). Hence $x_1\in\mathscr{COP}_K(g,T)\cap\mathscr{F}_K(T)$ and the proof is finished. Otherwise, if $x_2\neq x_1$, then $d(x_1,x_2)>0$. By (39), we have

$$d(x_2,Tx_2\cap K)<\beta d(x_1,Tx_2\cap K), \tag{40}$$

which implies that there exists $x_3\in Tx_2\cap K$ such that

$$\begin{aligned}d(x_2,x_3) &< \beta d(x_1,Tx_2\cap K) \\ &\le \beta d(x_1,x_3) \\ &\le \beta\left[d(x_1,x_2)+d(x_2,x_3)\right].\end{aligned} \tag{41}$$

Let $\gamma=\beta/(1-\beta)$. Then $\gamma\in(0,1)$ and the last inequality implies

$$d(x_2,x_3)<\gamma d(x_1,x_2). \tag{42}$$

Continuing in this way, we can construct inductively a sequence $\{x_n\}_{n\in\mathbb{N}}$ in K satisfying

$$\begin{gathered}x_{n+1}\in Tx_n\cap K, \\ d(x_{n+1},x_{n+2})<\gamma d(x_n,x_{n+1})\end{gathered} \tag{43}$$

for each $n\in\mathbb{N}$. Using a similar argument as in the proof of Theorem 8, we have the following:

(i) $x_{n+1}\in Tx_n\cap K$;

(ii) $\{x_n\}$ is a Cauchy sequence in K;

(iii) there exists $v\in K$ such that $x_n\to v$ as $n\to\infty$;

(iv) $gx_{n+1}\in Tx_n\cap K$ for each $n\in\mathbb{N}$;

(v) $\lim_{n\to\infty}gx_n = gv$.

By (38), we get

$$\begin{aligned}d(v,Tv\cap K) &= \lim_{n\to\infty} d(x_{n+1},Tv\cap K) \\ &\le \lim_{n\to\infty}\mathscr{H}(Tx_n,Tv\cap K) \\ &\le \lim_{n\to\infty}\{\alpha\left[d(x_n,Tv\cap K)+d(v,Tx_n\cap K)\right] \\ &\qquad +h(v)\,d(gv,Tx_n\cap K)\} \\ &\le \lim_{n\to\infty}\{\alpha\left[d(x_n,Tv\cap K)+d(v,x_{n+1})\right] \\ &\qquad +h(v)\,d(gv,gx_{n+1})\} \\ &= \alpha d(v,Tv\cap K),\end{aligned} \tag{44}$$

which implies $d(v,Tv\cap K)=0$. By the closedness of Tv, we have $v\in Tv\cap K$. By (D2), $gv\in Tv\cap K\subseteq Tv$ and hence $v\in\mathscr{COP}_K(g,T)\cap\mathscr{F}_K(T)$. The proof is complete. □

Theorem 14. *In Theorem 13, if inequality* (38) *is replaced with one of the following inequalities:*

(C1)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y\in K,\end{aligned} \tag{45}$$

(C2)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{46}$$

(C3)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y\in K,\end{aligned} \tag{47}$$

(C4)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx)\right] \\ &\quad + h(y)\, d(gy, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{48}$$

(C5)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y\in K,\end{aligned} \tag{49}$$

(C6)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx)\right] \\ &\quad + h(y)\, d(gy, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{50}$$

(C7)

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx)\right] \\ &\quad + h(y)\, d(gy, Tx) \quad \forall x, y\in K,\end{aligned} \tag{51}$$

then $\mathscr{COP}_K(g, T)\cap \mathscr{F}_K(T)\neq\emptyset$.

Applying Theorem 14, we can prove the following fixed point theorems for multivalued maps of generalized Chatterjea type.

Theorem 15. *Let* (X, d) *a complete metric space,* K *a nonempty closed subset of* X*, and* $T : K \to \mathscr{CB}(X)$ *a multivalued map. Suppose that* $Tx\cap K\neq\emptyset$ *for all* $x\in K$ *and one of the following conditions holds:*

(Q1) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{52}$$

(Q2) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(y, Tx) \quad \forall x, y\in K,\end{aligned} \tag{53}$$

(Q3) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{54}$$

(Q4) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx\cap K)\right] \\ &\quad + h(y)\, d(y, Tx) \quad \forall x, y\in K,\end{aligned} \tag{55}$$

(Q5) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{56}$$

(Q6) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty\cap K) + d(y, Tx)\right] \\ &\quad + h(y)\, d(y, Tx) \quad \forall x, y\in K,\end{aligned} \tag{57}$$

(Q7) *there exist a function* $h : K \to [0, \infty)$ *and* $\gamma\in[0, 1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty\cap K) &\le \alpha\left[d(x, Ty) + d(y, Tx)\right] \\ &\quad + h(y)\, d(y, Tx\cap K) \quad \forall x, y\in K,\end{aligned} \tag{58}$$

(Q8) *there exist a function* $h : K \to [0,\infty)$ *and* $\gamma \in [0,1/2)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty \cap K) &\le \alpha[d(x,Ty) + d(y,Tx)] \\ &\quad + h(y)\, d(y,Tx) \quad \forall x, y \in K.\end{aligned} \tag{59}$$

Then $\mathscr{F}_K(T) \neq \emptyset$.

The following result is a generalized Chatterjea's type fixed point theorem for multivalued maps in complete metric spaces.

Corollary 16. *Let* (X,d) *be a complete metric space,* K *a nonempty closed subset of* X*, and* $T : K \to \mathscr{BC}(X)$ *a multivalued map. Suppose that* $Tx \cap K \neq \emptyset$ *for all* $x \in K$ *and there exists* $\gamma \in [0,1/2)$ *such that*

$$\mathscr{H}(Tx, Ty \cap K) \le \alpha[d(x,Ty) + d(y,Tx)] \quad \forall x, y \in X. \tag{60}$$

Then $\mathscr{F}_K(T) \neq \emptyset$.

Remark 17. (a) If $K = X$ in Corollary 16, then we can obtain a multivalued version of Chatterjea's fixed point theorem [6].

(b) Theorems 13–15 and Corollary 16 all improve and generalize Chatterjea's fixed point theorem.

4. New Coincidence and Fixed Point Results for Various Multivalued Non-Self-Maps: Mizoguchi-Takahashi Type, Berinde-Berinde Type, and Du Type

In this section, we prove some coincidence and fixed point theorems for multivalued non-self-maps of Mizoguchi-Takahashi type, Berinde-Berinde type, and Du type.

Recall first the following auxiliary result.

Lemma 18 (see [9, Lemma 2.1]). *Let* $\varphi : [0,\infty) \to [0,1)$ *be an* $\mathscr{MT}$*-function. Suppose that* $\kappa : [0,\infty) \to [0,1)$ *is defined by* $\kappa(t) = (1+\varphi(t))/2$*. Then,* κ *is also an* $\mathscr{MT}$*-function.*

Theorem 19. *Let* (X,d) *be a complete metric space,* K *a nonempty closed subset of* X*,* $T : K \to \mathscr{BC}(X)$ *a multivalued map, and* $g : K \to X$ *be a continuous self-map. Suppose that conditions (D1) and (D2) as in Theorem 8 hold. If there exist an* $\mathscr{MT}$*-function* $\varphi : [0,\infty) \to [0,1)$ *and a function* $h : K \to [0,\infty)$ *such that*

$$\begin{aligned}\mathscr{H}(Tx, Ty \cap K) &\le \varphi(d(x,y))\, d(x,y) \\ &\quad + h(y)\, d(gy, Tx \cap K) \quad \forall x, y \in K,\end{aligned} \tag{61}$$

then $\mathscr{COP}_K(g,T) \cap \mathscr{F}_K(T) \neq \emptyset$.

Proof. Since K is a nonempty closed subset of X and X is complete, (K,d) is also a complete metric space. Note first that for each $x \in K$, by (D2), we have $d(gy, Tx \cap K) = 0$ for all $y \in Tx \cap K$. So, for each $x \in K$, by (61), we obtain

$$d(y, Ty \cap K) \le \varphi(d(x,y))\, d(x,y) \quad \forall y \in Tx \cap K. \tag{62}$$

Define $\kappa : [0,\infty) \to [0,1)$ by $\kappa(t) = (1+\varphi(t))/2$. Then, by Lemma 18, κ is also an $\mathscr{MT}$-function. Let $x \in K$ be given. Take $x_1 = x$. Since $Tx_1 \cap K \neq \emptyset$ from (D1), we can choose $x_2 \in Tx_1 \cap K$. If $x_2 = x_1$, then $x_1 \in \mathscr{F}_K(T)$ and hence $gx_1 \in Tx_1$ from (D2). Thus, $x_1 \in \mathscr{COP}_K(g,T) \cap \mathscr{F}_K(T)$ and hence we achieved the result. Now, suppose that $x_2 \neq x_1$; that is, $d(x_1,x_2) > 0$. By (62), we have

$$\begin{aligned}d(x_2, Tx_2 \cap K) &\le \varphi(d(x_1,x_2))\, d(x_1,x_2) \\ &< \kappa(d(x_1,x_2))\, d(x_1,x_2),\end{aligned} \tag{63}$$

which implies that there exists $x_3 \in Tx_2 \cap K$ such that

$$d(x_2,x_3) < \kappa(d(x_1,x_2))\, d(x_1,x_2). \tag{64}$$

Next, by (62) again, there exists $x_4 \in Tx_3 \cap K$ such that

$$d(x_3,x_4) < \kappa(d(x_2,x_3))\, d(x_2,x_3). \tag{65}$$

Iteratively, we can obtain a sequences $\{x_n\}$ in K satisfying

$$x_{n+1} \in Tx_n \cap K, \tag{66}$$

$$d(x_{n+1},x_{n+2}) < \kappa(d(x_n,x_{n+1}))\, d(x_n,x_{n+1}) \tag{67}$$

for each $n \in \mathbb{N}$. Since $\kappa(t) < 1$ for all $t \in [0,\infty)$, by (ii), we know that $\{d(x_n,x_{n+1})\}$ is strictly decreasing in $[0,\infty)$. Since κ is an $\mathscr{MT}$-function, by (g) of Theorem 7, we obtain

$$0 < d(x_1,x_2) \le \sup_{n\in\mathbb{N}} \kappa(d(x_n,x_{n+1})) < 1. \tag{68}$$

Let $\gamma := \sup_{n\in\mathbb{N}} \kappa(d(x_n,x_{n+1}))$. So $\gamma \in (0,1)$. By (67), we have

$$\begin{aligned}d(x_{n+1},x_{n+2}) &< \kappa(d(x_n,x_{n+1}))\, d(x_n,x_{n+1}) \\ &\le \gamma d(x_n,x_{n+1}) \\ &< \gamma^2 d(x_{n-1},x_n) \\ &< \cdots \\ &< \gamma^n d(x_1,x_2), \quad \text{for } n \in \mathbb{N}.\end{aligned} \tag{69}$$

Let $\alpha_n = (\gamma^{n-1}/(1-\gamma))d(x_1,x_2)$, $n \in \mathbb{N}$. For $m, n \in \mathbb{N}$ with $m > n$, we have

$$d(x_n,x_m) \le \sum_{j=n}^{m-1} d(x_j,x_{j+1}) < \alpha_n. \tag{70}$$

Since $0 < \gamma < 1$, $\lim_{n\to\infty}\alpha_n = 0$ and hence $\lim_{n\to\infty}\sup\{d(x_n,x_m) : m > n\} = 0$. This proves that $\{x_n\}$ is a Cauchy sequence in K. By the completeness of K, there exists $v \in K$ such that $x_n \to v$ as $n \to \infty$. Thanks to (66) and (D2), we have

$$gx_{n+1} \in Tx_n \cap K \text{ for each } n \in \mathbb{N}. \tag{71}$$

Since g is continuous and $\lim_{n\to\infty} x_n = v$, we have

$$\lim_{n\to\infty} gx_n = gv. \tag{72}$$

Since the function $x \mapsto d(x, Tv)$ is continuous, by (61), (66), and (72), we get

$$\begin{aligned} & d(v, Tv \cap K) \\ & \quad = \lim_{n\to\infty} d(x_{n+1}, Tv \cap K) \\ & \quad \le \lim_{n\to\infty} \mathscr{H}(Tx_n, Tv \cap K) \\ & \quad \le \lim_{n\to\infty} \{\varphi(d(x_n, v))\, d(x_n, v) + h(v)\, d(gv, gx_{n+1})\} = 0, \end{aligned} \tag{73}$$

which implies $d(v, Tv \cap K) = 0$. By the closedness of Tv, we have $v \in Tv \cap K$. By (D2), $gv \in Tv \cap K \subseteq Tv$ and hence $v \in \mathscr{COP}_K(g, T) \cap \mathscr{F}_K(T)$. The proof is complete. □

Theorem 20. *In Theorem 19, if inequality* (61) *is replaced with the following inequality:*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + h(y)\, d(gy, Tx) \\ & \qquad \forall x, y \in K. \end{aligned} \tag{74}$$

Then $\mathscr{COP}_K(g, T) \cap \mathscr{F}_K(T) \neq \emptyset$.

Corollary 21. *Let* (X, d) *be a complete convex metric space,* K *a nonempty closed subset of* X, $T : K \to \mathscr{BC}(X)$ *a multivalued map, and* $g : K \to X$ *a continuous self-map. Suppose that*

(i) $Tx \cap K \neq \emptyset$ *for all* $x \in K$,

(ii) $Tx \cap K$ *is* g*-invariant (i.e.,* $g(Tx \cap K) \subseteq Tx \cap K$*) for each* $x \in K$,

(iii) *there exist an* $\mathscr{MT}$*-function* $\varphi : [0, \infty) \to [0, 1)$ *and* $L \ge 0$ *such that*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + L d(gy, Tx \cap K) \quad \forall x, y \in K. \end{aligned} \tag{75}$$

Then $\mathscr{COP}_K(g, T) \cap \mathscr{F}_K(T) \neq \emptyset$.

Corollary 22. *Let* (X, d) *be a complete convex metric space,* K *a nonempty closed subset of* X, $T : K \to \mathscr{BC}(X)$ *a multivalued map, and* $g : K \to X$ *a continuous self-map. Suppose that*

(i) $Tx \cap K \neq \emptyset$ *for all* $x \in K$,

(ii) $Tx \cap K$ *is* g*-invariant (i.e.* $g(Tx \cap K) \subseteq Tx \cap K$*) for each* $x \in K$,

(ii) *there exist an* $\mathscr{MT}$*-function* $\varphi : [0, \infty) \to [0, 1)$ *and* $L \ge 0$ *such that*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + L d(gy, Tx) \quad \forall x, y \in K. \end{aligned} \tag{76}$$

Then $\mathscr{COP}_K(g, T) \cap \mathscr{F}_K(T) \neq \emptyset$.

As a direct consequence of Theorems 19 and 20, we obtain the following fixed point result for multivalued non-self-maps of Du type in complete metric spaces.

Theorem 23. *Let* (X, d) *be a complete convex metric space,* K *a nonempty closed subset of* X, *and* $T : K \to \mathscr{BC}(X)$ *a multivalued map. Suppose that* $Tx \cap K \neq \emptyset$ *for all* $x \in K$, *and one of the following conditions holds:*

(W1) *there exist an* $\mathscr{MT}$*-function* $\varphi : [0, \infty) \to [0, 1)$ *and a function* $h : K \to [0, \infty)$ *such that*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + h(y)\, d(y, Tx \cap K) \quad \forall x, y \in K, \end{aligned} \tag{77}$$

(W2) *there exist an* $\mathscr{MT}$*-function* $\varphi : [0, \infty) \to [0, 1)$ *and a function* $h : K \to [0, \infty)$ *such that*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + h(y)\, d(y, Tx) \quad \forall x, y \in K. \end{aligned} \tag{78}$$

Then $\mathscr{F}_K(T) \neq \emptyset$.

Proof. Let $g = id$ be the identity map. It is easy to verify that all the conditions of Theorem 19 (or Theorem 20) are satisfied. Hence the conclusion follows from Theorem 19 (or Theorem 20). □

The following fixed point theorems for multivalued non-self-maps of generalized Berinde-Berinde type and generalized Mizoguchi-Takahashi type are established immediately from Theorem 23.

Corollary 24. *Let* (X, d) *be a complete convex metric space,* K *a nonempty closed subset of* X, *and* $T : K \to \mathscr{BC}(X)$ *a multivalued map. Suppose that*

(i) $Tx \cap K \neq \emptyset$ *for all* $x \in K$,

(ii) *there exist an* $\mathscr{MT}$*-function* $\varphi : [0, \infty) \to [0, 1)$ *and* $L \ge 0$ *such that*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + L d(y, Tx \cap K) \quad \forall x, y \in K. \end{aligned} \tag{79}$$

Then $\mathscr{F}_K(T) \neq \emptyset$.

Corollary 25. *Let* (X, d) *be a complete convex metric space,* K *a nonempty closed subset of* X, *and* $T : K \to \mathscr{BC}(X)$ *a multivalued map. Suppose that*

(i) $Tx \cap K \neq \emptyset$ *for all* $x \in K$,

(ii) *there exist an* $\mathscr{MT}$*-function* $\varphi : [0, \infty) \to [0, 1)$ *and* $L \ge 0$ *such that*

$$\begin{aligned} & \mathscr{H}(Tx, Ty \cap K) \\ & \quad \le \varphi(d(x, y))\, d(x, y) + L d(y, Tx) \quad \forall x, y \in K. \end{aligned} \tag{80}$$

Then $\mathscr{F}_K(T) \neq \emptyset$.

Corollary 26. *Let (X, d) be a complete convex metric space, K a nonempty closed subset of X, and $T : K \to \mathscr{BC}(X)$ a multivalued map. Suppose that*

(i) *$Tx \cap K \neq \emptyset$ for all $x \in K$,*

(ii) *there exists an $\mathscr{MT}$-function $\varphi : [0, \infty) \to [0, 1)$ such that*

$$\begin{aligned} &\mathscr{H}(Tx, Ty \cap K) \\ &\quad \leq \varphi(d(x, y))\, d(x, y) \quad \forall x, y \in K. \end{aligned} \tag{81}$$

Then $\mathscr{F}_K(T) \neq \emptyset$.

Remark 27. (a) If $K = X$ in Theorem 23, then we can obtain Du's fixed point theorem [12, Theorem 2.6].

(b) Theorems 19, 20 and 23, and Corollaries 21–26 all generalize and improve Du's fixed point theorem, Berinde-Berinde's fixed point theorem, Mizoguchi-Takahashi's fixed point theorem, Nadler's fixed point theorem, and Banach's contraction principle.

Acknowledgment

The first author was supported partially by grant no. NSC 101-2115-M-017-001 of the National Science Council of the Republic of China.

References

[1] S. Banach, "Sur les operations dans les ensembles abstraits et leur application aux equations integerales," *Fundamenta Mathematicae*, vol. 3, pp. 133–181, 1922.

[2] S. B. Nadler, Jr., "Multi-valued contraction mappings," *Pacific Journal of Mathematics*, vol. 30, pp. 475–488, 1969.

[3] N. Mizoguchi and W. Takahashi, "Fixed point theorems for multivalued mappings on complete metric spaces," *Journal of Mathematical Analysis and Applications*, vol. 141, no. 1, pp. 177–188, 1989.

[4] M. Berinde and V. Berinde, "On a general class of multi-valued weakly Picard mappings," *Journal of Mathematical Analysis and Applications*, vol. 326, no. 2, pp. 772–782, 2007.

[5] R. Kannan, "Some results on fixed points. II," *The American Mathematical Monthly*, vol. 76, pp. 405–408, 1969.

[6] S. K. Chatterjea, "Fixed-point theorems," *Comptes Rendus de l'Académie Bulgare des Sciences*, vol. 25, pp. 727–730, 1972.

[7] N. Shioji, T. Suzuki, and W. Takahashi, "Contractive mappings, Kannan mappings and metric completeness," *Proceedings of the American Mathematical Society*, vol. 126, no. 10, pp. 3117–3124, 1998.

[8] S. Reich, "Some problems and results in fixed point theory," in *Topological Methods in Nonlinear Functional Analysis (Toronto, Ont., 1982)*, vol. 21 of *Contemp. Math.*, pp. 179–187, American Mathematical Society, Providence, RI, USA, 1983.

[9] W.-S. Du, "Some new results and generalizations in metric fixed point theory," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 73, no. 5, pp. 1439–1446, 2010.

[10] Z. He, W.-S. Du, and I.-J. Lin, "The existence of fixed points for new nonlinear multivalued maps and their applications," *Fixed Point Theory and Applications*, vol. 2011, article 84, 13 pages, 2011.

[11] W.-S. Du, "On generalized weakly directional contractions and approximate fixed point property with applications," *Fixed Point Theory and Applications*, vol. 2012, article 6, 22 pages, 2012.

[12] W.-S. Du, "On coincidence point and fixed point theorems for nonlinear multivalued maps," *Topology and Its Applications*, vol. 159, no. 1, pp. 49–56, 2012.

[13] W.-S. Du, "On approximate coincidence point properties and their applications to fixed point theory," *Journal of Applied Mathematics*, vol. 2012, Article ID 302830, 17 pages, 2012.

[14] W.-S. Du, Z. He, and Y.-L. Chen, "New existence theorems for approximate coincidence point property and approximate fixed point property with applications to metric fixed point theory," *Journal of Nonlinear and Convex Analysis*, vol. 13, no. 3, pp. 459–474, 2012.

[15] W. S. Du, "New cone fixed point theorems for nonlinear multivalued maps with their applications," *Applied Mathematics Letters*, vol. 24, no. 2, pp. 172–178, 2011.

[16] W.-S. Du and S.-X. Zheng, "Nonlinear conditions for coincidence point and fixed point theorems," *Taiwanese Journal of Mathematics*, vol. 16, no. 3, pp. 857–868, 2012.

[17] W.-S. Du and H. Lakzian, "Nonlinear conditions for the existence of best proximity points," *Journal of Inequalities and Applications*, vol. 2012, article 206, 7 pages, 2012.

[18] W.-S. Du, "New existence results and generalizations for coincidence points and fixed points without global completeness," *Abstract and Applied Analysis*, vol. 2013, Article ID 214230, 12 pages, 2013.

[19] F. E. Browder, "Nonexpansive nonlinear operators in a Banach space," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 54, pp. 1041–1044, 1965.

[20] F. E. Browder and W. V. Petryshyn, "Construction of fixed points of nonlinear mappings in Hilbert space," *Journal of Mathematical Analysis and Applications*, vol. 20, pp. 197–228, 1967.

[21] W. A. Kirk, "Remarks on pseudo-contractive mappings," *Proceedings of the American Mathematical Society*, vol. 25, pp. 820–823, 1970.

[22] W. A. Kirk, "Fixed point theorems for nonlinear nonexpansive and generalized contraction mappings," *Pacific Journal of Mathematics*, vol. 38, pp. 89–94, 1971.

[23] N. A. Assad and W. A. Kirk, "Fixed point theorems for set-valued mappings of contractive type," *Pacific Journal of Mathematics*, vol. 43, pp. 553–562, 1972.

[24] S. Reich, "Fixed points of condensing functions," *Journal of Mathematical Analysis and Applications*, vol. 41, pp. 460–467, 1973.

[25] N. A. Assad, "On some nonself nonlinear contractions," *Mathematica Japonica*, vol. 33, no. 1, pp. 17–26, 1988.

[26] N. A. Assad, "On some nonself mappings in Banach spaces," *Mathematica Japonica*, vol. 33, no. 4, pp. 501–515, 1988.

[27] N. A. Assad, "A fixed point theorem for some non-self-mappings," *Tamkang Journal of Mathematics*, vol. 21, no. 4, pp. 387–393, 1990.

[28] W. Sintunavarat and P. Kumam, "Weak condition for generalized multi-valued (f, α, β)-weak contraction mappings," *Applied Mathematics Letters*, vol. 24, no. 4, pp. 460–465, 2011.

[29] W. Sintunavarat and P. Kumam, "Common fixed point theorem for hybrid generalized multi-valued contraction mappings," *Applied Mathematics Letters*, vol. 25, no. 1, pp. 52–57, 2012.

[30] W. Sintunavarat and P. Kumam, "Common fixed point theorem for cyclic generalized multi-valued contraction mappings," *Applied Mathematics Letters*, vol. 25, no. 11, pp. 1849–1855, 2012.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 218964, 7 pages
http://dx.doi.org/10.1155/2013/218964

Research Article

Algorithmic Approach to the Split Problems

Ming Ma

Tianjin and Education Ministry, Key Laboratory of Advanced Composite Materials, Tianjin 300387, China

Correspondence should be addressed to Ming Ma; maming@tjpu.edu.cn

Received 6 June 2013; Accepted 26 June 2013

Academic Editor: Abdellah Bnouhachem

This paper deals with design algorithms for the split variational inequality and equilibrium problems. Strong convergence theorems are demonstrated.

1. Introduction

Let $\mathbb{H}$ be a real Hilbert space. Let $\mathbb{C}$ and $\mathbb{Q}$ be two nonempty closed convex subsets of $\mathbb{H}$. Consider the following problem.

Problem 1. Find a point $u^{\S} \in \mathbb{C}$ such that

$$\Psi\left(u^{\S}\right) \in \mathbb{Q}. \tag{1}$$

This problem is called split feasibility problem when Ψ is a bounded linear operator. In this case, Problem 1 can be applied to many practical problems such as signal processing and image reconstruction. Specifically, we can find the prototype of Problem 1 in intensity-modulated radiation therapy; see, for example, [1–3]. Based on this relation, many mathematicians were devoted to study the split feasibility problem and develop its iterative algorithms. Related works can be found in [4–8] and the references therein.

Let $\mathbb{A}, \Psi : \mathbb{C} \to \mathbb{H}$ be two mappings. Consider the variational inequality of finding $u^{\dagger} \in \mathbb{C}, \Psi(u^{\dagger}) \in \mathbb{C}$ such that

$$\left\langle \mathbb{A}u^{\dagger}, \Psi(u) - \Psi\left(u^{\dagger}\right)\right\rangle \geq 0, \tag{2}$$

for all $\Psi(u) \in \mathbb{C}$. We use $\mathrm{VI}(\mathbb{A}, \Psi)$ to denote the set of solutions of (2). Variational inequality problems have important applications in many fields such as elasticity, optimization, economics, transportation, and structural analysis, and various numerical methods have been studied by many researchers; see, for instance, [9–17].

Let $\varrho : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ be an equilibrium bifunction; that is, $\varrho(u, u) = 0$ for each $u \in \mathbb{C}$. Consider the equilibrium problem which is to find $u^* \in \mathbb{C}$ such that

$$\varrho\left(u^*, v\right) \geq 0, \quad \forall v \in \mathbb{C}. \tag{3}$$

Denote the set of solutions of (3) by $\mathrm{EP}(\varrho, \mathbb{C})$. The equilibrium problems include fixed point problems, optimization problems, and variational inequality problems as special cases. Some algorithms have been proposed to solve the equilibrium problems; see, for example, [18–22]. Thus it is an interesting topic associated with algorithmic approach to the variational inequality and equilibrium problems. In this paper, our main purpose is to study the following split problem involved in the variational inequality and equilibrium problems. Find a point $x^{\natural}$ such that

$$\begin{gathered} x^{\natural} \in \mathrm{VI}(\mathbb{A}, \Psi), \\ \Psi\left(x^{\natural}\right) \in \mathrm{EP}(\varrho, \mathbb{C}). \end{gathered} \tag{4}$$

We are devoted to study (4) with operator Ψ being a nonlinear mapping. For this purpose, we develop an iterative algorithm for solving the split problem (4). We can compute $x^{\natural}$ iteratively by using our algorithm. Convergence analysis is given under some mild assumptions.

2. Basic Concepts

Let $\mathbb{C}$ be a nonempty closed convex subset of a real Hilbert space $\mathbb{H}$. An operator $\mathbb{B} : \mathbb{C} \to \mathbb{H}$ is said to be

(i) monotone $\twoheadrightarrow \langle u - v, \mathbb{B}u - \mathbb{B}v\rangle \geq 0$ for all $u, v \in \mathbb{C}$;

(ii) strongly monotone $\twoheadrightarrow$ $\langle u-v, \mathbb{B}u-\mathbb{B}v\rangle \geq \zeta\|u-v\|^2$ for some constant $\zeta > 0$ and for all $u, v \in \mathbb{C}$;

(iii) inverse-strongly monotone $\twoheadrightarrow$ $\langle u-v, \mathbb{B}u-\mathbb{B}v\rangle \geq \varsigma\|\mathbb{B}u-\mathbb{B}v\|^2$ for some $\varsigma > 0$ and for all $u, v \in \mathbb{C}$; in this case, $\mathbb{B}$ is called ς-inverse strongly monotone;

(iv) ς-inverse strongly θ-monotone $\twoheadrightarrow$ $\langle \theta(u)-\theta(v), \mathbb{B}u-\mathbb{B}v\rangle \geq \varsigma\|\mathbb{B}u-\mathbb{B}v\|^2$ for all $u, v \in \mathbb{C}$ and for some $\varsigma > 0$, where $\theta : \mathbb{C} \to \mathbb{C}$ is a mapping.

A mapping $\vartheta : \mathbb{C} \to \mathbb{H}$ is said to be

(i) nonexpansive $\twoheadrightarrow$ $\|\vartheta u-\vartheta v\| \leq \|u-v\|$ for all $u, v \in \mathbb{C}$;

(ii) firmly nonexpansive $\twoheadrightarrow$ $\|\vartheta u-\vartheta v\|^2 \leq \langle u-v, \vartheta u-\vartheta v\rangle$ for all $u, v \in \mathbb{C}$;

(iii) L-Lipschitz continuous $\twoheadrightarrow$ $\|\vartheta u-\vartheta v\| \leq L\|u-v\|$ for some constant $L > 0$ and for all $u, v \in \mathbb{C}$. In such a case, ϑ is said to be L-Lipschitz continuous.

In the sequel, we use $\mathrm{Fix}(\vartheta)$ to denote the set of fixed points of ϑ.

Let $\mathbb{A} : \mathbb{H} \to 2^{\mathbb{H}}$ be a multivalued mapping. The effective domain of $\mathbb{A}$ is denoted by $\mathrm{dom}(\mathbb{A})$. $\mathbb{A}$ is said to be

(i) monotone $\twoheadrightarrow$ $\langle x-y, u-v\rangle \geq 0$ for all $x, y \in \mathrm{dom}(\mathbb{A})$, $u \in \mathbb{A}x$, and $v \in \mathbb{A}y$;

(ii) maximal monotone $\twoheadrightarrow$ $\mathbb{A}$ is monotone and its graph is not strictly contained in the graph of any other monotone operator on $\mathbb{H}$.

A function $f : \mathbb{H} \to \mathbb{R}$ is said to be convex if for any $u, v \in \mathbb{H}$ and for any $\tau \in [0,1]$, $f(\tau u+(1-\tau)v) \leq \tau f(u)+(1-\tau)f(v)$.

Let $\mathrm{proj}_{\mathbb{C}} : \mathbb{C} \to \mathbb{H}$ be the metric projection from $\mathbb{H}$ onto $\mathbb{C}$. It is known that $\mathrm{proj}_{\mathbb{C}}$ satisfies the following inequality:

$$\langle x-\mathrm{proj}_{\mathbb{C}}x, y-\mathrm{proj}_{\mathbb{C}}x\rangle \leq 0. \tag{5}$$

for all $x \in \mathbb{H}$ and $y \in \mathbb{C}$. From this characteristic inequality, we can deduce that $\mathrm{proj}_{\mathbb{C}}$ is firmly nonexpansive.

3. Useful Lemmas

In this section, we present several lemmas which will be used in the next section.

Lemma 2 (see [19]). *Let $\mathbb{C}$ be a nonempty closed convex subset of a real Hilbert space $\mathbb{H}$. Let $\varrho : \mathbb{C}\times\mathbb{C} \to \mathbb{R}$ be a bifunction. Assume that ϱ satisfies the following conditions:*

($\mathfrak{F}1$) *$\varrho(u,u) = 0$ for all $u \in \mathbb{C}$;*

($\mathfrak{F}2$) *ϱ is monotone, that is, $\varrho(u,v)+\varrho(v,u) \leq 0$ for all $u, v \in \mathbb{C}$;*

($\mathfrak{F}3$) *for each $u, v, w \in \mathbb{C}$, $\lim_{t\downarrow 0}\varrho(tw+(1-t)u, v) \leq \varrho(u,v)$;*

($\mathfrak{F}4$) *for each $u \in \mathbb{C}$, $v \mapsto \varrho(u,v)$ is convex and lower semicontinuous.*

Let $\varpi > 0$ and $u \in \mathbb{C}$. Then there exists $w \in \mathbb{C}$ such that

$$\varrho(w,v)+\frac{1}{\varpi}\langle v-w, w-u\rangle \geq 0, \quad \forall v \in \mathbb{C}. \tag{6}$$

Set $F_{\varpi}(u) = \{w \in \mathbb{C} : \varrho(w,v)+(1/\varpi)\langle v-w, w-u\rangle \geq 0$ for all $v \in \mathbb{C}\}$. Then one have the following:

(i) *F_{ϖ} is single valued and F_{ϖ} is firmly nonexpansive,*

(ii) *$\mathrm{EP}(\varrho,\mathbb{C})$ is closed and convex and $\mathrm{EP}(\varrho,\mathbb{C}) = \mathrm{Fix}(F_{\varpi})$.*

Lemma 3 (see [23]). *Let $\mathbb{C}$ be a nonempty closed convex subset of a real Hilbert space $\mathbb{H}$. For $x \in \mathbb{H}$, let the mapping F_{ϖ} be the same as in Lemma 2. Then for $\mu, \nu > 0$ and $x \in \mathbb{H}$, one has*

$$\left\|F_{\mu}(x)-F_{\nu}(x)\right\|^2 \leq \frac{\mu-\nu}{\mu}\left\langle F_{\mu}(x)-F_{\nu}(x), F_{\mu}(x)-x\right\rangle. \tag{7}$$

Lemma 4 (see [24]). *Let $\{u_n\}$ and $\{v_n\}$ be two bounded sequences in a Banach space $\mathbb{E}$, and let $\{\kappa_n\}$ be a sequence in $[0,1]$ satisfying $0 < \liminf_{n\to\infty}\kappa_n \leq \limsup_{n\to\infty}\kappa_n < 1$. Suppose $u_{n+1} = (1-\kappa_n)v_n+\kappa_n u_n$ for all $n \geq 0$ and $\limsup_{n\to\infty}(\|v_{n+1}-v_n\|-\|u_{n+1}-u_n\|) \leq 0$. Then, $\lim_{n\to\infty}\|u_n-v_n\| = 0$.*

Lemma 5 (see [25]). *Let $\mathbb{C}$ be a nonempty closed convex subset of a real Hilbert space $\mathbb{H}$. Let $\mathbb{S} : \mathbb{C} \to \mathbb{C}$ be a nonexpansive mapping with $\mathrm{Fix}(\mathbb{S}) \neq \emptyset$. Then $\mathbb{S}$ is demiclosed on $\mathbb{C}$.*

Lemma 6 (see [26]). *Let $\{a_n\} \subset [0,\infty)$ be a sequence. Assume that $a_{n+1} \leq (1-\gamma_n)a_n+\delta_n\gamma_n$, where $\{\gamma_n\}$ is a sequence in $(0,1)$, and $\{\delta_n\}$ is a sequence satisfying $\sum_{n=1}^{\infty}\gamma_n = \infty$ and $\limsup_{n\to\infty}\delta_n \leq 0$ (or $\sum_{n=1}^{\infty}|\delta_n\gamma_n| < \infty$). Then $\lim_{n\to\infty}a_n = 0$.*

4. Main Results

In this section, we firstly present our problem and algorithm constructed. Consequently, we give the convergence analysis of the presented algorithm.

Problem 7. Let $\mathbb{C}$ be a nonempty closed convex subset of a real Hilbert space $\mathbb{H}$. Assume that

(1) $\Psi : \mathbb{C} \to \mathbb{C}$ is a weakly continuous and ζ-strongly monotone mapping such that $R(\Psi) = \mathbb{C}$;

(2) $\mathbb{A} : \mathbb{C} \to \mathbb{H}$ is an ς-inverse strongly Ψ-monotone mapping;

(3) $\varrho : \mathbb{C}\times\mathbb{C} \to \mathbb{R}$ is a bifunction satisfying conditions ($\mathfrak{F}1$)–($\mathfrak{F}4$) in Lemma 2.

Our objective is to

$$\text{find } x^{\natural} \in \mathrm{VI}(\mathbb{A},\Psi) \text{ such that } \Psi\left(x^{\natural}\right) \in \mathrm{EP}(\varrho,\mathbb{C}). \tag{8}$$

We use Υ to denote the set of solutions of (8). In the following, we assume that Υ is nonempty. For solving Problem 7, we introduce the following algorithm.

Algorithm 8.
Step 0 (initialization). Let

$$u_0 \in \mathbb{C}. \tag{9}$$

Step 1. For given $\{u_n\}$, let the sequence $\{v_n\}$ be generated iteratively by

$$v_n = \mathrm{proj}_{\mathbb{C}}\left(\Psi\left(u_n\right) - \mu_n \mathbb{A} u_n\right), \quad n \geq 0, \tag{10}$$

where $\mathrm{proj}_{\mathbb{C}}$ is the metric projection and $\{\mu_n\}$ is a real number sequence.

Step 2. For given $\{v_n\}$, find $\{z_n\}$ such that

$$\varrho\left(z_n, y\right) + \frac{1}{\varpi_n}\left\langle y - z_n, z_n - \left(1 - \alpha_n\right) v_n\right\rangle \geq 0, \quad \forall y \in \mathbb{C}, \tag{11}$$

where $\{\varpi_n\} \subset (0, \infty)$ and $\{\alpha_n\} \subset [0, 1]$ are two real number sequences.

Step 3. For the previous sequences $\{u_n\}$ and $\{z_n\}$, let the $(n+1)$th sequence $\{u_{n+1}\}$ be generated by

$$\Psi\left(u_{n+1}\right) = \kappa_n \Psi\left(u_n\right) + \left(1 - \kappa_n\right) z_n, \quad n \geq 0, \tag{12}$$

where $\{\kappa_n\} \subset [0, 1]$ is a real number sequence.

Theorem 9. *Assume that the following conditions are satisfied:*

(C1) $\lim_{n\to\infty}\alpha_n = 0$ *and* $\sum_n \alpha_n = \infty$;

(C2) $0 < \liminf_{n\to\infty}\kappa_n \leq \limsup_{n\to\infty}\kappa_n < 1$;

(C3) $\varpi_n \in (\eta_1, \eta_2) \subset (0, \infty)$, $\mu_n \in (\xi_1, \xi_2) \subset (0, 2\varsigma)$, *and* $\zeta \in (0, 2\varsigma)$;

(C4) $\lim_{n\to\infty}(\mu_{n+1} - \mu_n) = 0$ *and* $\lim_{n\to\infty}(\varpi_{n+1} - \varpi_n) = 0$.

Then the sequence $\{u_n\}$ *generated by Algorithm 8 converges strongly to* $x^* \in \Upsilon$.

Proof. Let $\check{x} \in \Upsilon$. Hence $\check{x} \in \mathrm{VI}(\mathbb{A}, \Psi)$ and $\Psi(\check{x}) \in \mathrm{EP}(\varrho, \mathbb{C})$, noting that $\check{x} \in \mathrm{VI}(\mathbb{A}, \Psi)$ implies $\Psi(\check{x}) = \mathrm{proj}_{\mathbb{C}}(\Psi(\check{x}) - \nu\mathbb{A}\check{x})$ for all $\nu > 0$. Hence $\Psi(\check{x}) = \mathrm{proj}_{\mathbb{C}}(\Psi(\check{x}) - \mu_n\mathbb{A}\check{x})$ for all $n \geq 0$. Thus, from (10), we have

$$\begin{aligned}
&\left\|v_n - \Psi(\check{x})\right\|^2 \\
&\quad = \left\|\mathrm{proj}_{\mathbb{C}}\left(\Psi\left(u_n\right) - \mu_n \mathbb{A} u_n\right) - \mathrm{proj}_{\mathbb{C}}\left(\Psi(\check{x}) - \mu_n \mathbb{A}\check{x}\right)\right\|^2 \\
&\quad \leq \left\|\left(\Psi\left(u_n\right) - \mu_n \mathbb{A} u_n\right) - \left(\Psi(\check{x}) - \mu_n \mathbb{A}\check{x}\right)\right\|^2 \\
&\quad = \left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\|^2 - 2\mu_n\left\langle \mathbb{A} u_n - \mathbb{A}\check{x}, \Psi\left(u_n\right) - \Psi(\check{x})\right\rangle \\
&\qquad + \mu_n^2\left\|\mathbb{A} u_n - \mathbb{A}\check{x}\right\|^2 \\
&\quad \leq \left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\|^2 \\
&\qquad -2\mu_n\varsigma\left\|\mathbb{A} u_n - \mathbb{A}\check{x}\right\|^2 + \mu_n^2\left\|\mathbb{A} u_n - \mathbb{A}\check{x}\right\|^2 \\
&\quad \leq \left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\|^2 + \mu_n\left(\mu_n - 2\varsigma\right)\left\|\mathbb{A} u_n - \mathbb{A}\check{x}\right\|^2.
\end{aligned} \tag{13}$$

Condition (C3) and (13) imply that

$$\left\|v_n - \Psi(\check{x})\right\| \leq \left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\|. \tag{14}$$

From Lemma 2 and (11), we get $z_n = \digamma_{\varpi_n}(1 - \alpha_n)v_n$ for all $n \geq 0$. Since $\Psi(\check{x}) \in \mathrm{EP}(\varrho, \mathbb{C})$, from Lemma 2 we deduce that $\Psi(\check{x}) = \digamma_{\varpi_n}\Psi(\check{x})$ for all $n \geq 0$. So,

$$\begin{aligned}
&\left\|z_n - \Psi(\check{x})\right\| \\
&\quad = \left\|\digamma_{\varpi_n}\left(1 - \alpha_n\right) v_n - \digamma_{\varpi_n}\Psi(\check{x})\right\| \\
&\quad \leq \left\|\left(1 - \alpha_n\right) v_n - \Psi(\check{x})\right\| \\
&\quad \leq \left(1 - \alpha_n\right)\left\|v_n - \Psi(\check{x})\right\| + \alpha_n\left\|\Psi(\check{x})\right\| \\
&\quad \text{by (14)} \\
&\quad \leq \left(1 - \alpha_n\right)\left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\| + \alpha_n\left\|\Psi(\check{x})\right\|.
\end{aligned} \tag{15}$$

It follows that

$$\begin{aligned}
&\left\|\Psi\left(u_{n+1}\right) - \Psi(\check{x})\right\| \\
&\quad \leq \kappa_n\left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\| + \left(1 - \kappa_n\right)\left\|z_n - \Psi(\check{x})\right\| \\
&\quad \leq \kappa_n\left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\| \\
&\qquad + \left(1 - \kappa_n\right)\left(1 - \alpha_n\right)\left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\| \\
&\qquad + \left(1 - \kappa_n\right)\alpha_n\left\|\Psi(\check{x})\right\| \\
&\quad = \left[1 - \left(1 - \kappa_n\right)\alpha_n\right]\left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\| \\
&\qquad + \left(1 - \kappa_n\right)\alpha_n\left\|\Psi(\check{x})\right\|.
\end{aligned} \tag{16}$$

By induction

$$\left\|\Psi\left(u_n\right) - \Psi(\check{x})\right\| \leq \max\left\{\left\|\Psi\left(u_0\right) - \Psi(\check{x})\right\|, \left\|\Psi(\check{x})\right\|\right\}. \tag{17}$$

Hence, $\{\Psi(u_n)\}$ is bounded. Since Ψ is ζ-strongly monotone, we can get $\zeta\|u_n - \check{x}\| \leq \|\Psi(u_n) - \Psi(\check{x})\|$. So, $\|u_n - \check{x}\| \leq (1/\zeta)\|\Psi(u_n) - \Psi(\check{x})\| \leq (1/\zeta)\max\{\|\Psi(u_0) - \Psi(\check{x})\|, \|\Psi(\check{x})\|\}$. This implies that $\{u_n\}$ is bounded. Next, we show $\|u_{n+1} - u_n\| \to 0$. From $z_n = \digamma_{\varpi_n}(1 - \alpha_n)v_n$, we have

$$\begin{aligned}
&\left\|z_{n+1} - z_n\right\| \\
&\quad = \left\|\digamma_{\varpi_{n+1}}\left(1 - \alpha_{n+1}\right) v_{n+1} - \digamma_{\varpi_n}\left(1 - \alpha_n\right) v_n\right\| \\
&\quad \leq \left\|\digamma_{\varpi_{n+1}}\left(1 - \alpha_{n+1}\right) v_{n+1} - \digamma_{\varpi_{n+1}}\left(1 - \alpha_n\right) v_n\right\| \\
&\qquad + \left\|\digamma_{\varpi_{n+1}}\left(1 - \alpha_n\right) v_n - \digamma_{\varpi_n}\left(1 - \alpha_n\right) v_n\right\| \\
&\quad \leq \left\|\left(1 - \alpha_{n+1}\right) v_{n+1} - \left(1 - \alpha_n\right) v_n\right\| \\
&\qquad + \left\|\digamma_{\varpi_{n+1}}\left(1 - \alpha_n\right) v_n - \digamma_{\varpi_n}\left(1 - \alpha_n\right) v_n\right\|.
\end{aligned} \tag{18}$$

Using Lemma 3, we obtain

$$\begin{aligned}
&\left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - F_{\varpi_n}(1-\alpha_n)v_n\right\|^2 \\
&\quad\le \frac{\varpi_{n+1}-\varpi_n}{\varpi_{n+1}} \\
&\qquad\times \left\langle F_{\varpi_{n+1}}(1-\alpha_n)v_n \right. \\
&\qquad\quad - F_{\varpi_n}(1-\alpha_n)v_n, F_{\varpi_{n+1}}(1-\alpha_n)v_n \\
&\qquad\quad \left. -(1-\alpha_n)v_n \right\rangle \\
&\quad\le \frac{|\varpi_{n+1}-\varpi_n|}{\varpi_{n+1}}\left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - F_{\varpi_n}(1-\alpha_n)v_n\right\| \\
&\qquad\times \left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - (1-\alpha_n)v_n\right\|.
\end{aligned} \tag{19}$$

Then

$$\begin{aligned}
&\left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - F_{\varpi_n}(1-\alpha_n)v_n\right\| \\
&\quad\le \frac{|\varpi_{n+1}-\varpi_n|}{\varpi_{n+1}}\left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - (1-\alpha_n)v_n\right\|.
\end{aligned} \tag{20}$$

By condition (C3), we have $\varpi_n > \eta_1 > 0$. So,

$$\begin{aligned}
&\|z_{n+1}-z_n\| \\
&\quad\le \|(1-\alpha_{n+1})v_{n+1} - (1-\alpha_n)v_n\| \\
&\qquad + \frac{|\varpi_{n+1}-\varpi_n|}{\varpi_{n+1}} \\
&\qquad\times \left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - (1-\alpha_n)v_n\right\| \\
&\quad\le (1-\alpha_{n+1})\|v_{n+1}-v_n\| \\
&\qquad + |\alpha_{n+1}-\alpha_n|\|v_n\| + \frac{|\varpi_{n+1}-\varpi_n|}{\eta_1} \\
&\qquad\times \left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - (1-\alpha_n)v_n\right\|.
\end{aligned} \tag{21}$$

From (10), we have

$$\begin{aligned}
&\|v_{n+1}-v_n\| \\
&\quad= \|\mathrm{proj}_{\mathbb{C}}(\Psi(u_{n+1}) - \mu_{n+1}\mathbb{A}u_{n+1}) \\
&\qquad -\mathrm{proj}_{\mathbb{C}}(\Psi(u_n) - \mu_n\mathbb{A}u_n)\| \\
&\quad\le \|\Psi(u_{n+1}) - \mu_{n+1}\mathbb{A}u_{n+1} \\
&\qquad - (\Psi(u_n) - \mu_{n+1}\mathbb{A}u_n)\| + |\mu_{n+1}-\mu_n|\|\mathbb{A}(u_n)\| \\
&\quad\le \|\Psi(u_{n+1}) - \Psi(u_n)\| + |\mu_{n+1}-\mu_n|\|\mathbb{A}(u_n)\|.
\end{aligned} \tag{22}$$

Therefore,

$$\begin{aligned}
\|z_{n+1}-z_n\| &\le (1-\alpha_{n+1})\|\Psi(u_{n+1}) - \Psi(u_n)\| \\
&\quad + |\alpha_{n+1}-\alpha_n|\|v_n\| + |\mu_{n+1}-\mu_n|\|\mathbb{A}(u_n)\| \\
&\quad + \frac{1}{\eta_1}|\varpi_{n+1}-\varpi_n|\left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - (1-\alpha_n)v_n\right\|.
\end{aligned} \tag{23}$$

It follows that

$$\begin{aligned}
&\|z_{n+1}-z_n\| - \|\Psi(u_{n+1}) - \Psi(u_n)\| \\
&\quad\le |\alpha_{n+1}-\alpha_n|\|v_n\| + |\mu_{n+1}-\mu_n|\|\mathbb{A}(u_n)\| \\
&\qquad + \frac{1}{\eta_1}|\varpi_{n+1}-\varpi_n|\left\|F_{\varpi_{n+1}}(1-\alpha_n)v_n - (1-\alpha_n)v_n\right\|.
\end{aligned} \tag{24}$$

Since $\lim_{n\to\infty}\alpha_n = 0$, $\lim_{n\to\infty}(\mu_{n+1}-\mu_n) = 0$, and $\lim_{n\to\infty}(\varpi_{n+1}-\varpi_n) = 0$ and the sequences $\{\Psi(u_n)\}$, $\{z_n\}$, $\{v_n\}$, and $\{\mathbb{A}u_n\}$ are bounded, we deduce that

$$\limsup_{n\to\infty}\left(\|z_{n+1}-z_n\| - \|\Psi(u_{n+1}) - \Psi(u_n)\|\right) \le 0. \tag{25}$$

Applying Lemma 4, we obtain

$$\lim_{n\to\infty}\|z_n - \Psi(u_n)\| = 0. \tag{26}$$

Thus,

$$\lim_{n\to\infty}\|\Psi(u_{n+1}) - \Psi(u_n)\| = \lim_{n\to\infty}(1-\kappa_n)\|z_n - \Psi(u_n)\| = 0. \tag{27}$$

This together with the ζ-strong monotonicity of Ψ implies that

$$\lim_{n\to\infty}\|u_{n+1}-u_n\| = 0. \tag{28}$$

From (13) and (16), we derive

$$\begin{aligned}
&\|\Psi(u_{n+1}) - \Psi(\check{x})\|^2 \\
&\quad\le \kappa_n\|\Psi(u_n) - \Psi(\check{x})\|^2 + (1-\kappa_n)\|z_n - \Psi(\check{x})\|^2 \\
&\quad\le \kappa_n\|\Psi(u_n) - \Psi(\check{x})\|^2 + (1-\kappa_n) \\
&\qquad\times \left[(1-\alpha_n)\|v_n - \Psi(\check{x})\|^2 + \alpha_n\|\Psi(\check{x})\|^2\right]
\end{aligned}$$

$$\begin{aligned}
&\le (1-\kappa_n) \\
&\quad \times \left[(1-\alpha_n)\left\|(\Psi(u_n)-\mu_n\mathbb{A}u_n)-(\Psi(\check{x})-\mu_n\mathbb{A}\check{x})\right\|^2\right. \\
&\qquad \left.+\alpha_n\|\Psi(\check{x})\|^2\right]+\kappa_n\|\Psi(u_n)-\Psi(\check{x})\|^2 \\
&\le \kappa_n\|\Psi(u_n)-\Psi(\check{x})\|^2+(1-\kappa_n)(1-\alpha_n) \\
&\quad \times\left(\|\Psi(u_n)-\Psi(\check{x})\|^2+\mu_n(\mu_n-2\varsigma)\|\mathbb{A}u_n-\mathbb{A}\check{x}\|^2\right) \\
&\quad +\alpha_n\|\Psi(\check{x})\|^2 \\
&\le \|\Psi(u_n)-\Psi(\check{x})\|^2 \\
&\quad +(1-\kappa_n)(1-\alpha_n)\mu_n(\mu_n-2\varsigma)\|\mathbb{A}u_n-\mathbb{A}\check{x}\|^2 \\
&\quad +\alpha_n\|\Psi(\check{x})\|^2.
\end{aligned} \tag{29}$$

Hence,

$$\begin{aligned}
&(1-\kappa_n)(1-\alpha_n)\mu_n(2\varsigma-\mu_n)\|\mathbb{A}u_n-\mathbb{A}\check{x}\|^2 \\
&\quad \le \|\Psi(u_n)-\Psi(\check{x})\|^2-\|\Psi(u_{n+1})-\Psi(\check{x})\|^2 \\
&\qquad +\alpha_n\|\Psi(\check{x})\|^2 \\
&\quad \le \left(\|\Psi(u_n)-\Psi(\check{x})\|+\|\Psi(u_{n+1})-\Psi(\check{x})\|\right) \\
&\qquad \times\|\Psi(u_{n+1})-\Psi(u_n)\|+\alpha_n\|\Psi(\check{x})\|^2.
\end{aligned} \tag{30}$$

Since $\alpha_n \to 0$, $\|\Psi(u_{n+1})-\Psi(u_n)\| \to 0$, and $\liminf_{n\to\infty}(1-\kappa_n)(1-\alpha_n)\mu_n(2\varsigma-\mu_n)>0$, we obtain

$$\lim_{n\to\infty}\|\mathbb{A}u_n-\mathbb{A}\check{x}\|=0. \tag{31}$$

Set $y_n=\Psi(u_n)-\mu_n\mathbb{A}u_n-(\Psi(\check{x})-\mu_n\mathbb{A}\check{x})$ for all n. By using the firm nonexpansivity of projection, we get

$$\begin{aligned}
&\|v_n-\Psi(\check{x})\|^2 \\
&\quad =\|\mathrm{proj}_{\mathbb{C}}(\Psi(u_n)-\mu_n\mathbb{A}u_n) \\
&\qquad -\mathrm{proj}_{\mathbb{C}}(\Psi(\check{x})-\mu_n\mathbb{A}\check{x})\|^2 \\
&\quad \le \langle y_n, v_n-\Psi(\check{x})\rangle \\
&\quad =\frac{1}{2}\left\{\|y_n\|^2+\|v_n-\Psi(\check{x})\|^2-\|y_n-v_n+\Psi(\check{x})\|^2\right\} \\
&\quad \le \frac{1}{2}\left\{\|\Psi(u_n)-\Psi(\check{x})\|^2+\|v_n-\Psi(\check{x})\|^2\right. \\
&\qquad \left.-\|\Psi(u_n)-v_n-\mu_n(\mathbb{A}u_n-\mathbb{A}\check{x})\|^2\right\} \\
&\quad =\frac{1}{2}\left\{\|\Psi(u_n)-\Psi(\check{x})\|^2+\|v_n-\Psi(\check{x})\|^2\right. \\
&\qquad -\|\Psi(u_n)-v_n\|^2-\mu_n^2\|\mathbb{A}u_n-\mathbb{A}\check{x}\| \\
&\qquad \left.+2\mu_n\langle\Psi(u_n)-v_n,\mathbb{A}u_n-\mathbb{A}\check{x}\rangle\right\}.
\end{aligned} \tag{32}$$

It follows that

$$\begin{aligned}
\|v_n-\Psi(\check{x})\|^2 &\le \|\Psi(u_n)-\Psi(\check{x})\|^2-\|\Psi(u_n)-v_n\|^2 \\
&\quad +2\mu_n\|\Psi(u_n)-v_n\|\|\mathbb{A}u_n-\mathbb{A}\check{x}\|.
\end{aligned} \tag{33}$$

From (29) and (32), we have

$$\begin{aligned}
&\|\Psi(u_{n+1})-\Psi(\check{x})\|^2 \\
&\quad \le \kappa_n\|\Psi(u_n)-\Psi(\check{x})\|^2+(1-\kappa_n) \\
&\qquad \times\left[(1-\alpha_n)\|v_n-\Psi(\check{x})\|^2+\alpha_n\|\Psi(\check{x})\|^2\right] \\
&\quad \le \kappa_n\|\Psi(u_n)-\Psi(\check{x})\|^2+(1-\alpha_n)(1-\kappa_n) \\
&\qquad \times\|\Psi(u_n)-\Psi(\check{x})\|^2-(1-\kappa_n)\|\Psi(u_n)-v_n\|^2 \\
&\qquad +(1-\kappa_n)\alpha_n\|\Psi(\check{x})\|^2+2\mu_n(1-\kappa_n) \\
&\qquad \times\|\Psi(u_n)-v_n\|\|\mathbb{A}u_n-\mathbb{A}\check{x}\| \\
&\quad \le \|\Psi(u_n)-\Psi(\check{x})\|^2-(1-\kappa_n)\|\Psi(u_n)-v_n\|^2 \\
&\qquad +2\mu_n\|\Psi(u_n)-v_n\|\|\mathbb{A}u_n-\mathbb{A}\check{x}\|+\alpha_n\|\Psi(\check{x})\|^2.
\end{aligned} \tag{34}$$

Then, we obtain

$$\begin{aligned}
&(1-\kappa_n)\|\Psi(u_n)-v_n\|^2 \\
&\quad \le \left(\|\Psi(u_n)-\Psi(\check{x})\|+\|\Psi(u_{n+1})-\Psi(\check{x})\|\right) \\
&\qquad \times\|\Psi(u_{n+1})-\Psi(x_n)\| \\
&\qquad +2\mu_n\|\Psi(u_n)-v_n\|\|\mathbb{A}u_n-\mathbb{A}\check{x}\|+\alpha_n\|\Psi(\check{x})\|^2.
\end{aligned} \tag{35}$$

Since $\lim_{n\to\infty}\alpha_n=0$, $\lim_{n\to\infty}\|\Psi(u_{n+1})-\Psi(u_n)\|=0$, and $\lim_{n\to\infty}\|\mathbb{A}u_n-\mathbb{A}\check{x}\|=0$, we deduce that

$$\lim_{n\to\infty}\|\Psi(u_n)-v_n\|=0. \tag{36}$$

Next, we prove $\limsup_{n\to\infty}\langle\Psi(x^*),v_n-\Psi(x^*)\rangle\ge 0$, where x^* satisfies (GVI): $\langle\Psi(x^*),\Psi(x)-\Psi(x^*)\rangle\ge 0$, for all $x\in\Upsilon$ (note that Ψ is ζ-strongly monotone; we can easily deduce that the solution of (GVI) is unique). We take a subsequence $\{v_{n_i}\}$ of $\{v_n\}$ such that

$$\begin{aligned}
&\limsup_{n\to\infty}\langle\Psi(x^*),v_n-\Psi(x^*)\rangle \\
&\quad =\lim_{i\to\infty}\left\langle\Psi(x^*),v_{n_i}-\Psi(x^*)\right\rangle \\
&\quad =\lim_{i\to\infty}\left\langle\Psi(x^*),\Psi(u_{n_i})-\Psi(x^*)\right\rangle.
\end{aligned} \tag{37}$$

By the boundedness of $\{u_{n_i}\}$, we can choose a subsequence $\{u_{n_{i_j}}\}$ of $\{u_{n_i}\}$ such that $u_{n_{i_j}}\to z$ weakly. For the convenience, we may assume that $u_{n_i}\rightharpoonup z$. This implies that $\Psi(u_{n_i})\rightharpoonup\Psi(z)$ due to the weak continuity of Ψ. Now, we show $z\in\Upsilon$. We firstly show $\Psi(z)\in\mathrm{EP}(\varrho,\mathbb{C})$.

Note that $\varpi_n \in (\eta_1, \eta_2)$. Then we choose a subsequence $\{\varpi_{n_i}\}$ of $\{\varpi_n\}$ such that $\lim_{i\to\infty}\varpi_{n_i} = \varpi \in (\eta_1, \eta_2)$. From (26) and (36), we deduce that $\|z_n - v_n\| = \|F_{\varpi_n}(1-\alpha_n)v_n - v_n\| \to 0$. Thus, $\|z_{n_i} - v_{n_i}\| = \|F_{\varpi_{n_i}}(1-\alpha_{n_i})v_{n_i} - v_{n_i}\| \to 0$. From Lemma 2, we know that F_ϖ is nonexpansive. By demiclosed principle (Lemma 5), we get immediately that $\Psi(z) \in \mathrm{Fix}(F_\varpi) = \mathrm{EP}(\varrho, \mathbb{C})$.

Next we prove $z \in \mathrm{VI}(\mathbb{A}, \Psi)$. Set

$$Rv = \begin{cases} \mathbb{A}v + N_C(v), & v \in C, \\ \emptyset, & v \notin C. \end{cases} \tag{38}$$

By [27], we know that R is maximal Ψ-monotone. Let $(v, w) \in G(R)$. Since $w - \mathbb{A}v \in N_C(v)$ and $u_n \in C$, we have $\langle \Psi(v) - \Psi(u_n), w - \mathbb{A}v\rangle \geq 0$. Noting that $v_n = \mathrm{proj}_{\mathbb{C}}(\Psi(u_n) - \mu_n \mathbb{A}u_n)$, we get

$$\langle \Psi(v) - v_n, v_n - (\Psi(u_n) - \mu_n \mathbb{A}u_n)\rangle \geq 0. \tag{39}$$

It follows that

$$\left\langle \Psi(v) - v_n, \frac{v_n - \Psi(u_n)}{\mu_n} + \mathbb{A}u_n \right\rangle \geq 0. \tag{40}$$

Then,

$$\begin{aligned}
&\left\langle \Psi(v) - \Psi(u_{n_i}), w\right\rangle \\
&\quad \geq \left\langle \Psi(v) - \Psi(u_{n_i}), \mathbb{A}v\right\rangle \\
&\quad \geq \left\langle \Psi(v) - \Psi(u_{n_i}), \mathbb{A}v\right\rangle \\
&\qquad - \left\langle \Psi(v) - v_{n_i}, \frac{v_{n_i} - \Psi(u_{n_i})}{\mu_{n_i}}\right\rangle \\
&\qquad - \left\langle \Psi(v) - v_{n_i}, \mathbb{A}u_{n_i}\right\rangle \\
&\quad = \left\langle \Psi(v) - \Psi(u_{n_i}), \mathbb{A}v - \mathbb{A}u_{n_i}\right\rangle \\
&\qquad + \left\langle \Psi(v) - \Psi(u_{n_i}), \mathbb{A}u_{n_i}\right\rangle \\
&\qquad - \left\langle \Psi(v) - v_{n_i}, \frac{v_{n_i} - \Psi(u_{n_i})}{\mu_{n_i}}\right\rangle \\
&\qquad - \left\langle \Psi(v) - v_{n_i}, \mathbb{A}u_{n_i}\right\rangle \\
&\quad \geq - \left\langle \Psi(v) - v_{n_i}, \frac{v_{n_i} - \Psi(u_{n_i})}{\mu_{n_i}}\right\rangle \\
&\qquad - \left\langle \Psi(u_{n_i}) - v_{n_i}, \mathbb{A}u_{n_i}\right\rangle.
\end{aligned} \tag{41}$$

Since $\|\Psi(u_{n_i}) - v_{n_i}\| \to 0$ and $\Psi(u_{n_i}) \rightharpoonup \Psi(z)$, we deduce that $\langle \Psi(v) - \Psi(z), w\rangle \geq 0$ by taking $i \to \infty$ in (41). Thus, $z \in R^{-1}0$ by the maximal Ψ-monotonicity of R. Hence, $z \in \mathrm{VI}(\mathbb{A}, \Psi)$. Therefore, $z \in \Upsilon$. From (37), we obtain

$$\begin{aligned}
&\limsup_{n\to\infty} \langle \Psi(x^*), v_n - \Psi(x^*)\rangle \\
&\quad = \lim_{i\to\infty} \left\langle \Psi(x^*), \Psi(u_{n_i}) - \Psi(x^*)\right\rangle \\
&\quad = \langle \Psi(x^*), \Psi(z) - \Psi(x^*)\rangle \geq 0.
\end{aligned} \tag{42}$$

From (12), we have

$$\begin{aligned}
&\|\Psi(u_{n+1}) - \Psi(x^*)\|^2 \\
&\quad \leq \kappa_n \|\Psi(u_n) - \Psi(x^*)\|^2 \\
&\qquad + (1-\kappa_n)\|(1-\alpha_n)v_n - \Psi(x^*)\|^2 \\
&\quad \leq \kappa_n \|\Psi(u_n) - \Psi(x^*)\|^2 + (1-\kappa_n) \\
&\qquad \times \Big[(1-\alpha_n)\|v_n - \Psi(x^*)\|^2 \\
&\qquad\quad - 2\alpha_n(1-\alpha_n)\langle \Psi(x^*), v_n - \Psi(x^*)\rangle \\
&\qquad\quad + \alpha_n^2 \|\Psi(x^*)\|^2\Big] \\
&\quad \leq \kappa_n \|\Psi(u_n) - \Psi(x^*)\|^2 + (1-\kappa_n) \\
&\qquad \times \Big[(1-\alpha_n)\|\Psi(u_n) - \Psi(x^*)\|^2 - 2\alpha_n(1-\alpha_n) \\
&\qquad\quad \times \langle \Psi(x^*), v_n - \Psi(x^*)\rangle + \alpha_n^2\|\Psi(x^*)\|^2\Big] \\
&\quad = [1 - (1-\kappa_n)\alpha_n]\|\Psi(u_n) - \Psi(x^*)\|^2 + (1-\kappa_n)\alpha_n \\
&\qquad \times \{2(1-\alpha_n)\langle -\Psi(x^*), v_n - \Psi(x^*)\rangle \\
&\qquad\quad + \alpha_n\|\Psi(x^*)\|^2\}.
\end{aligned} \tag{43}$$

Using Lemma 6, we conclude that $\Psi(u_n) \to \Psi(x^*)$, and hence $u_n \to x^*$. This completes the proof. □

References

[1] Y. Censor and T. Elfving, "A multiprojection algorithm using Bregman projections in a product space," *Numerical Algorithms*, vol. 8, no. 2, pp. 221–239, 1994.

[2] C. Byrne, "A unified treatment of some iterative algorithms in signal processing and image reconstruction," *Inverse Problems*, vol. 20, no. 1, pp. 103–120, 2004.

[3] Y. Censor, T. Bortfeld, B. Martin, and A. Trofimov, "A unified approach for inversion problems in intensity modulated radiation therapy," *Physics in Medicine and Biology*, vol. 51, pp. 2353–2365, 2006.

[4] Q. Yang, "The relaxed CQ algorithm solving the split feasibility problem," *Inverse Problems*, vol. 20, no. 4, pp. 1261–1266, 2004.

[5] B. Qu and N. Xiu, "A note on the *CQ* algorithm for the split feasibility problem," *Inverse Problems*, vol. 21, no. 5, pp. 1655–1665, 2005.

[6] H.-K. Xu, "Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces," *Inverse Problems*, vol. 26, no. 10, Article ID 105018, 17 pages, 2010.

[7] Y. Yao, Y.-C. Liou, and N. Shahzad, "A strongly convergent method for the split feasibility problem," *Abstract and Applied Analysis*, vol. 2012, Article ID 125046, 15 pages, 2012.

[8] L.-C. Ceng, Q. H. Ansari, and J.-C. Yao, "Relaxed extragradient methods for finding minimum-norm solutions of the split feasibility problem," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 75, no. 4, pp. 2116–2125, 2012.

[9] G. Stampacchia, "Formes bilinéaires coercitives sur les ensembles convexes," vol. 258, pp. 4413–4416, 1964.

[10] G. M. Korpelevič, "An extragradient method for finding saddle points and for other problems," *Èkonomika I Matematicheskie Metody*, vol. 12, pp. 747–756, 1976.

[11] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, NY, USA, 1984.

[12] A. N. Iusem, "An iterative algorithm for the variational inequality problem," *Computational and Applied Mathematics*, vol. 13, pp. 103–114, 1994.

[13] M. A. Noor, "Some developments in general variational inequalities," *Applied Mathematics and Computation*, vol. 152, no. 1, pp. 199–277, 2004.

[14] A. Bnouhachem, M. Li, M. Khalfaoui, and Z. Sheng, "A modified inexact implicit method for mixed variational inequalities," *Journal of Computational and Applied Mathematics*, vol. 234, no. 12, pp. 3356–3365, 2010.

[15] A. Bnouhachem, M. A. Noor, and Z. Hao, "Some new extragradient iterative methods for variational inequalities," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 70, no. 3, pp. 1321–1329, 2009.

[16] A. Bnouhachem and M. A. Noor, "Inexact proximal point method for general variational inequalities," *Journal of Mathematical Analysis and Applications*, vol. 324, no. 2, pp. 1195–1212, 2006.

[17] A. Bnouhachem, M. A. Noor, and Th. M. Rassias, "Three-steps iterative algorithms for mixed variational inequalities," *Applied Mathematics and Computation*, vol. 183, no. 1, pp. 436–446, 2006.

[18] E. Blum and W. Oettli, "From optimization and variational inequalities to equilibrium problems," *The Mathematics Student*, vol. 63, pp. 123–145, 1994.

[19] P. L. Combettes and S. A. Hirstoaga, "Equilibrium programming in Hilbert spaces," *Journal of Nonlinear and Convex Analysis*, vol. 6, no. 1, pp. 117–136, 2005.

[20] M. A. Noor and W. Oettli, "On general nonlinear complementarity problems and quasi-equilibria," *Le Matematiche*, vol. 49, no. 2, pp. 313–331, 1994.

[21] Y. Yao, M. A. Noor, and Y.-C. Liou, "On iterative methods for equilibrium problems," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 70, no. 1, pp. 497–509, 2009.

[22] M. A. Noor, "Fundamentals of equilibrium problems," *Mathematical Inequalities & Applications*, vol. 9, no. 3, pp. 529–566, 2006.

[23] S. Takahashi and W. Takahashi, "Strong convergence theorem for a generalized equilibrium problem and a nonexpansive mapping in a Hilbert space," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 69, no. 3, pp. 1025–1033, 2008.

[24] T. Suzuki, "Strong convergence theorems for infinite families of nonexpansive mappings in general Banach spaces," *Fixed Point Theory and Applications*, vol. 2005, no. 1, pp. 103–123, 2005.

[25] K. Goebel and W. A. Kirk, *Topics in Metric Fixed Point Theory*, vol. 28 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, UK, 1990.

[26] H.-K. Xu, "Iterative algorithms for nonlinear operators," *Journal of the London Mathematical Society*, vol. 66, no. 1, pp. 240–256, 2002.

[27] L. J. Zhang, J. M. Chen, and Z. B. Hou, "Viscosity approximation methods for nonexpansive mappings and generalized variational inequalities," *Acta Mathematica Sinica*, vol. 53, no. 4, pp. 691–698, 2010.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 634371, 12 pages
http://dx.doi.org/10.1155/2013/634371

Research Article

Multidimensional Fixed-Point Theorems in Partially Ordered Complete Partial Metric Spaces under (ψ, φ)-Contractivity Conditions

A. Roldán, J. Martínez-Moreno, C. Roldán, and E. Karapınar

University of Jaén, Campus Las Lagunillas s/n, 23071 Jaén, Spain

Correspondence should be addressed to A. Roldán; afroldan@ujaen.es

Received 17 March 2013; Accepted 17 June 2013

Academic Editor: Abdelouahed Hamdi

We study the existence and uniqueness of coincidence point for nonlinear mappings of any number of arguments under a weak (ψ, φ)-contractivity condition in partial metric spaces. The results we obtain generalize, extend, and unify several classical and very recent related results in the literature in metric spaces (see Aydi et al. (2011), Berinde and Borcut (2011), Gnana Bhaskar and Lakshmikantham (2006), Berzig and Samet (2012), Borcut and Berinde (2012), Choudhury et al. (2011), Karapınar and Luong (2012), Lakshmikantham and Ćirić (2009), Luong and Thuan (2011), and Roldán et al. (2012)) and in partial metric spaces (see Shatanawi et al. (2012)).

1. Introduction

The notion of *coupled fixed point* was introduced by Guo and Lakshmikantham [1] in 1987. In a recent paper, Gnana Bhaskar and Lakshmikantham [2] introduced the concept *mixed monotone property* for contractive operators of the form $F : X \times X \to X$, where X is a partially ordered metric space, and then established some coupled fixed-point theorems. After that, many results appeared on coupled fixed-point theory in different contexts (see, e.g., [3–6]). Later, Berinde and Borcut [7] introduced the concept of *tripled fixed point* and proved tripled fixed-point theorems using mixed monotone mappings (see also [8–10]).

Very recently, Roldán et al. [11] proposed the notion of *coincidence point* between mappings in any number of variables and showed some existence and uniqueness theorems that extended the mentioned previous results for this kind of nonlinear mappings, not necessarily permuted or ordered, in the framework of partially ordered complete metric spaces, using a weaker contraction condition, that also generalized other works by Berzig and Samet [12], Karapınar and Berinde [13].

Partial metric spaces were firstly introduced by Matthews in [14] as an attempt to generalize the metric spaces by establishing the condition that the distance between a point to itself (which is not necessarily zero) is less or equal than the distance between that point and another point of the space. In the mentioned papers, Matthews studied topological properties of partial metric spaces and stated a modified version of a Banach contraction mapping principle on this kind of spaces. After Matthews' pioneering work, the theory of partial metric spaces and particularly the field of fixed-point theorems have expansively been developed due to the increasing interest in this area and motivated by its possible applications (see [15, 16] and references therein).

In this paper, our main aim is to study a weaker contractivity condition for nonlinear mappings of any number of arguments. This condition can be particularized in a variety of forms that let us extend the previously mentioned results and other recent ones in this field (see [2, 5, 7, 9, 11, 12, 16–20]). We also notice that our results cannot be obtained by the very recent paper of Haghi et al. [21] (for more details see Remark 26).

2. Preliminaries

Preliminaries and notation about coincidence points can also be found in [11]. Let n be a positive integer. Henceforth, X will denote a nonempty set, and X^n will denote the product space

$X^n = X \times X \times \overset{n}{\cdots} \times X$. Throughout this paper, m and k will denote nonnegative integers and $i, j, s \in \{1, 2, \ldots, n\}$. Unless otherwise stated, "for all m" will mean "for all $m \geq 0$", and "for all i" will mean "for all $i \in \{1, 2, \ldots, n\}$". Let $\mathbb{R}_0^+ = [0, \infty[$.

A *metric on* X is a mapping $d : X \times X \to \mathbb{R}$ satisfying, for all $x, y, z \in X$:

(i) $d(x, y) = 0$ if, and only if, $x = y$;

(ii) $d(x, y) \leq d(z, x) + d(z, y)$.

From these properties, we can easily deduce that $d(x, y) \geq 0$ and $d(y, x) = d(x, y)$ for all $x, y \in X$. The last requirement is called the *triangle inequality*. If d is a metric on X, we say that (X, d) is a *metric space* (for short, an MS).

Definition 1 (see [22]). A triple $(X, d, \leq)$ is called *a partially ordered metric space* if (X, d) is a MS and $\leq$ is a partial order on X.

Definition 2 (see [2]). An ordered MS $(X, d, \leq)$ is said to have the *sequential g-monotone property* if it verifies

(i) if $\{x_m\}$ is a nondecreasing sequence and $\{x_m\} \xrightarrow{d} x$, then $gx_m \leq gx$ for all m;

(ii) if $\{y_m\}$ is a nonincreasing sequence and $\{y_m\} \xrightarrow{d} y$, then $gy_m \geq gy$ for all m.

If g is the identity mapping, then X is said to have the *sequential monotone property*.

Henceforth, fix a partition $\{A, B\}$ of two non-empty subsets of $\Lambda_n = \{1, 2, \ldots, n\}$; that is, $A \cup B = \Lambda_n$ and $A \cap B = \emptyset$. We will denote

$$\begin{aligned} \Omega_{A,B} &= \{\sigma : \Lambda_n \to \Lambda_n : \sigma(A) \subseteq A \text{ and } \sigma(B) \subseteq B\}, \\ \Omega'_{A,B} &= \{\sigma : \Lambda_n \to \Lambda_n : \sigma(A) \subseteq B \text{ and } \sigma(B) \subseteq A\}. \end{aligned} \tag{1}$$

If $(X, \leq)$ is a partially ordered space, $x, y \in X$, and $i \in \Lambda_n$, we will use the following notation:

$$x \leq_i y \Longleftrightarrow \begin{cases} x \leq y, & \text{if } i \in A, \\ x \geq y, & \text{if } i \in B. \end{cases} \tag{2}$$

Let $F : X^n \to X$ and $g : X \to X$ be two mappings.

Definition 3 (see [11]). One says that F and g are commuting if $gF(x_1, \ldots, x_n) = F(gx_1, \ldots, gx_n)$ for all $x_1, \ldots, x_n \in X$.

Definition 4 (see [11]). Let $(X, \leq)$ be a partially ordered space. One says that F has the *mixed g-monotone property* (with respect to $\{A, B\}$) if F is g-monotone nondecreasing in arguments of A and g-monotone nonincreasing in arguments of B; that is, for all $x_1, x_2, \ldots, x_n, y, z \in X$ and all i,

$$\begin{aligned} & gy \leq gz \\ & \Longrightarrow F(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \\ & \quad \leq_i F(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n). \end{aligned} \tag{3}$$

Henceforth, let $\sigma_1, \sigma_2, \ldots, \sigma_n, \tau : \Lambda_n \to \Lambda_n$ be $n + 1$ mappings from Λ_n into itself, and let Φ be the $(n + 1)$-tuple $(\sigma_1, \sigma_2, \ldots, \sigma_n, \tau)$.

Definition 5 (see [11]). A point $(x_1, x_2, \ldots, x_n) \in X^n$ is called a *Φ-coincidence point of the mappings Fand g* if

$$F\left(x_{\sigma_i(1)}, x_{\sigma_i(2)}, \ldots, x_{\sigma_i(n)}\right) = gx_{\tau(i)} \quad \forall i. \tag{4}$$

If g is the identity mapping on X, then $(x_1, x_2, \ldots, x_n) \in X^n$ is called a *Φ-fixed point of the mapping F*.

Remark 6. If F and g are commuting and $(x_1, x_2, \ldots, x_n) \in X^n$ is a Φ-coincidence point of F and g, then $(gx_1, gx_2, \ldots, gx_n)$ also is a Φ-coincidence point of F and g.

Definition 7 (see [14]). A *partial metric on* X is a mapping $p : X \times X \to \mathbb{R}_0^+$ verifying, for all $x, y, z \in X$:

$(P1)$ $p(x, x) \leq p(x, y)$;

$(P2)$ $p(x, x) = p(x, y) = p(y, y) \Rightarrow x = y$;

$(P3)$ $p(x, y) = p(y, x)$;

$(P4)$ $p(x, z) + p(y, y) \leq p(x, y) + p(y, z)$.

In this case, (X, p) is a *partial metric space* (for short, a PMS).

Example 8 (see, e.g., [14]). Let $X = \mathbb{R}_0^+$, and define p on X by $p(x, y) = \max\{x, y\}$ for all $x, y \in X$. Then, (X, p) is a partial metric space.

Example 9 (see [14]). Let $X = \{[a, b] : a, b \in \mathbb{R}, a \leq b\}$, and define $p([a, b], [c, d]) = \max\{b, d\} - \min\{a, c\}$. Then, (X, p) is a partial metric space.

Example 10 (see [14]). Let $X = [0, 1] \cup [2, 3]$, and define $p : X \times X \to \mathbb{R}_0^+$ by

$$p(x, y) = \begin{cases} \max\{x, y\}, & \text{if } \{x, y\} \cap [2, 3] \neq \emptyset, \\ |x - y|, & \text{if } \{x, y\} \subset [0, 1]. \end{cases} \tag{5}$$

Then, (X, p) is a partial metric space.

Example 11 (see, e.g., [23, 24]). Let (X, d) and (X, p) be a metric space and a partial metric space, respectively. Functions $\rho_i : X \times X \to \mathbb{R}_0^+$ $(i \in \{1, 2, 3\})$ given by

$$\begin{aligned} \rho_1(x, y) &= d(x, y) + p(x, y), \\ \rho_2(x, y) &= d(x, y) + \max\{u(x), u(y)\}, \\ \rho_3(x, y) &= d(x, y) + a, \end{aligned} \tag{6}$$

define partial metrics on X, where $u : X \to \mathbb{R}_0^+$ is an arbitrary function and $a \geq 0$.

Obviously, if (X, d) is a MS and we define $p = d$, then (X, p) is a PMS. Indeed, a partial metric p on X verifies

(i) $p(x, y) = 0 \Rightarrow x = y$;

(ii) $p(x, y) = p(y, x)$;

(iii) $p(x, z) \leq p(x, y) + p(y, z)$,

but the condition $p(x,x)=0$ does not necessarily hold. For a partial metric p on X, the mappings $d_p, d_m : X\times X \to \mathbb{R}_0^+$ given by

$$\begin{aligned} d_p(x,y) &= 2p(x,y)-p(x,x)-p(y,y), \\ d_m(x,y) &= \max\{p(x,y)-p(x,x), p(y,y)-p(y,y)\} \\ &= p(x,y)-\min\{p(x,x), p(y,y)\}, \end{aligned} \tag{7}$$

for all $x,y\in X$, are (usual) metrics on X. On a PMS, the concepts of convergence, Cauchy sequences, completeness, and continuity are defined as follows.

Definition 12 (see [14, 25, 26]). Let $\{x_m\}$ be a sequence on a PMS (X,p).

(i) $\{x_m\}$ *p-converges* to $x\in X$ (and one will write $\{x_m\}\xrightarrow{p} x$) if $p(x,x)=\lim_{m\to\infty}p(x,x_m)$.

(ii) $\{x_m\}$ is called *p-Cauchy* if $\lim_{m,m'\to\infty}p(x_m,x_{m'})$ exists (and it is finite).

(iii) (X,p) is said to be *p-complete* if every p-Cauchy sequence $\{x_m\}$ in X p-converges to a point $x\in X$ such that $p(x,x)=\lim_{m,m'\to\infty}p(x_m,x_{m'})$.

(iv) A mapping $f:X\to X$ is said to be *p-continuous at* $x_0\in X$ if, for every $\varepsilon>0$, there exists $\delta>0$ such that $f(B_p(x_0,\delta))\subseteq B_p(f(x_0),\varepsilon)$.

We have used the previous notation because we need to distinguish between p-convergence and d_p-convergence on X and usual convergence for real sequences.

Lemma 13 (see [14, 25, 26]). *Let $\{x_m\}$ be a sequence on a PMS (X,p).*

(1) *$\{x_m\}$ is p-Cauchy if, and only if, it is d_p-Cauchy.*

(2) *$\{x_m\}\xrightarrow{d_p} x$ if, and only if, $\{x_m\}\xrightarrow{p} x$ and $p(x,x)=\lim_{m,m'\to\infty}p(x_m,x_{m'})$; that is,*

$$\begin{aligned} &\{d_p(x_m,x)\}\longrightarrow 0 \Longleftrightarrow p(x,x) \\ &\quad = \lim_{m\to\infty}p(x,x_m)=\lim_{m,m'\to\infty}p(x_m,x_{m'}). \end{aligned} \tag{8}$$

(3) *(X,p) is complete if, and only if, the MS (X,d_p) is complete.*

(4) *If $\{x_m\}\xrightarrow{p} x$ and $p(x,x)=0$, then $\lim_{m\to\infty}p(x_m,y)=p(x,y)$ for all $y\in X$.*

3. Auxiliary Results

We will use the following results about real sequences in the proof of our main theorems.

Lemma 14. *Let $\{a_m^1\}_{m\in\mathbb{N}},\dots,\{a_m^n\}_{m\in\mathbb{N}}$ be n real lower bounded sequences such that $\{\max(a_m^1,\dots,a_m^n)\}_{m\in\mathbb{N}}\to\delta$. Then, there exists $i_0\in\{1,2,\dots,n\}$ and a subsequence $\{a_{m(k)}^{i_0}\}_{k\in\mathbb{N}}$ such that $\{a_{m(k)}^{i_0}\}_{k\in\mathbb{N}}\to\delta$.*

Proof. Let $b_m=\max(a_m^1,a_m^2,\dots,a_m^n)$ for all m. As $\{b_m\}$ is convergent, it is bounded. As $a_m^i\le b_m$ for all m and i, then every $\{a_m^i\}$ is bounded. As $\{a_m^1\}_{m\in\mathbb{N}}$ is a real bounded sequence, it has a convergent subsequence $\{a_{\sigma_1(m)}^1\}_{m\in\mathbb{N}}\to a_1$. Consider the subsequences $\{a_{\sigma_1(m)}^2\}_{m\in\mathbb{N}},\{a_{\sigma_1(m)}^3\}_{m\in\mathbb{N}},\dots,\{a_{\sigma_1(m)}^n\}_{m\in\mathbb{N}}$; that are $n-1$ real bounded sequences and the sequence $\{b_{\sigma_1(m)}\}_{m\in\mathbb{N}}$ that also converges to δ. As $\{a_{\sigma_1(m)}^2\}_{m\in\mathbb{N}}$ is a real bounded sequence, it has a convergent subsequence $\{a_{\sigma_2\sigma_1(m)}^2\}_{m\in\mathbb{N}}\to a_2$. Then, the sequences $\{a_{\sigma_2\sigma_1(m)}^3\}_{m\in\mathbb{N}}$, $\{a_{\sigma_2\sigma_1(m)}^4\}_{m\in\mathbb{N}},\dots,\{a_{\sigma_2\sigma_1(m)}^n\}_{m\in\mathbb{N}}$ also are $n-2$ real bounded sequences, $\{a_{\sigma_2\sigma_1(m)}^1\}_{m\in\mathbb{N}}\to a_1$, and $\{b_{\sigma_2\sigma_1(m)}\}_{m\in\mathbb{N}}\to\delta$. Repeating this process n times, we can find n subsequences $\{a_{\sigma(m)}^1\}_{m\in\mathbb{N}},\{a_{\sigma(m)}^2\}_{m\in\mathbb{N}},\dots,\{a_{\sigma(m)}^n\}_{m\in\mathbb{N}}$ (where $\sigma=\sigma_n\cdots\sigma_1$) such that $\{a_{\sigma(m)}^i\}_{m\in\mathbb{N}}\to a_i$ for all i. And $\{b_{\sigma(m)}\}_{m\in\mathbb{N}}\to\delta$. But

$$\begin{aligned} \{b_{\sigma(m)}\}_{m\in\mathbb{N}} &= \left\{\max\left(a_{\sigma(m)}^n,\dots,a_{\sigma(m)}^n\right)\right\}_{m\in\mathbb{N}} \\ &\longrightarrow \max(a_1,\dots,a_n), \end{aligned} \tag{9}$$

so $\delta=\max(a_1,\dots,a_n)$, and there exists $i_0\in\{1,2,\dots,n\}$ such that $a_{i_0}=\delta$. Therefore, there exists $i_0\in\{1,2,\dots,n\}$ and a subsequence $\{a_{\sigma(m)}^{i_0}\}_{m\in\mathbb{N}}$ such that $\{a_{\sigma(m)}^{i_0}\}_{m\in\mathbb{N}}\to a_{i_0}=\delta$. □

Lemma 15. *Let $\{a_m\}_{m\in\mathbb{N}}$ be a sequence of nonnegative real numbers which has not any subsequence converging to zero. Then, for all $\varepsilon>0$, there exist $\delta\in]0,\varepsilon[$ and $m_0\in\mathbb{N}$ such that $a_m\ge\delta$ for all $m\ge m_0$.*

Proof. Suppose that the conclusion is not true. Then, there exists $\varepsilon_0>0$ such that, for all $\delta\in]0,\varepsilon_0[$, there exists $m_0\in\mathbb{N}$ verifying $a_{m_0}<\delta$. Let $k_0\in\mathbb{N}$ be such that $1/k_0<\varepsilon_0$. For all $k\in\mathbb{N}$, take $\delta_k=1/(k+k_0)\in]0,\varepsilon_0[$. Then, there exists $m(k)\in\mathbb{N}$ verifying $0\le a_{m(k)}<\delta_k=1/(k+k_0)$. Taking limit when $k\to\infty$, we deduce that $\lim_{k\to\infty}a_{m(k)}=0$. Then, $\{a_m\}$ has a subsequence converging to zero (maybe, reordering $\{a_{m(k)}\}$), but this is a contradiction. □

Lemma 16. *If $\{x_m\}_{m\in\mathbb{N}}$ is a sequence in a MS (X,d) that is not Cauchy, then there exist $\varepsilon_0>0$ and two subsequences $\{x_{m(k)}\}_{k\in\mathbb{N}}$ and $\{x_{n(k)}\}_{k\in\mathbb{N}}$ such that, for all $k\in\mathbb{N}$,*

$$\begin{gathered} k<m(k)<n(k)<m(k+1), \\ d(x_{m(k)},x_{n(k)})\ge\varepsilon_0, \qquad d(x_{m(k)},x_{n(k)-1})<\varepsilon_0. \end{gathered} \tag{10}$$

Proof. We know that

$$\begin{aligned} &\{x_m\} \text{ is Cauchy} \\ &\Longleftrightarrow [\forall\varepsilon>0,\exists n_0\in\mathbb{N}:(m,n\ge n_0\Longrightarrow d(x_m,x_n)<\varepsilon)]. \end{aligned} \tag{11}$$

If this condition is not true, then

$$\exists\varepsilon_0>0:(\forall n_0\in\mathbb{N},\exists m,n\ge n_0 \text{ such that } d(x_m,x_n)\ge\varepsilon_0). \tag{12}$$

Let $n_0 = 2$. Then, there exists $m_1, n_1 \in \mathbb{N}$ such that $m_1, n_1 \geq n_0$ and $d(x_{m_1}, x_{n_1}) \geq \varepsilon_0$. Let $m(1) = \min(m_1, n_1) \geq n_0 = 2 > 1$, and consider the numbers

$$d\left(x_{m(1)}, x_{m(1)+1}\right), \quad d\left(x_{m(1)}, x_{m(1)+2}\right), \ldots, d\left(x_{m(1)}, x_{\max(m_1,n_1)}\right). \tag{13}$$

Since $d(x_{m(1)}, x_{\max(m_1,n_1)}) = d(x_{m_1}, x_{n_1}) \geq \varepsilon_0$, between the previous numbers there exists a first nonnegative integer $n(1) \in \{m(1)+1, m(1)+2, \ldots, \max(m_1, n_1)\}$ such that $d(x_{m(1)}, x_{n(1)}) \geq \varepsilon_0$ but $d(x_{m(1)}, x_j) < \varepsilon_0$ for all $j \in \{m(1), m(1)+1, \ldots, n(1)-1\}$. In particular, $d(x_{m(1)}, x_{n(1)-1}) < \varepsilon_0$.

Now, let $n_0 = n(1)+1$. Then, there exists $m_2, n_2 \in \mathbb{N}$ such that $m_2, n_2 \geq n(1)+1$ and $d(x_{m_2}, x_{n_2}) \geq \varepsilon_0$. Let $m(2) = \min(m_2, n_2) \geq n_0 = n(1)+1 > n(1)$, and consider the numbers

$$d\left(x_{m(2)}, x_{m(2)+1}\right), \quad d\left(x_{m(2)}, x_{m(2)+2}\right), \ldots, d\left(x_{m(2)}, x_{\max(m_2,n_2)}\right). \tag{14}$$

Since $d(x_{m(2)}, x_{\max(m_2,n_2)}) = d(x_{m_2}, x_{n_2}) \geq \varepsilon_0$, between the previous numbers there exists a first nonnegative integer $n(2) \in \{m(2)+1, m(2)+2, \ldots, \max(m_2, n_2)\}$ such that $d(x_{m(2)}, x_{n(2)}) \geq \varepsilon_0$ but $d(x_{m(2)}, x_j) < \varepsilon_0$ for all $j \in \{m(2), m(2)+1, \ldots, n(2)-1\}$. In particular, $d(x_{m(2)}, x_{n(2)-1}) < \varepsilon_0$.

Repeating this process, we can find two subsequences $\{x_{m(k)}\}$ and $\{x_{n(k)}\}$ such that, for all $k \in \mathbb{N}$:

$$k < m(k) < n(k) < m(k+1), \quad d\left(x_{m(k)}, x_{n(k)}\right) \geq \varepsilon_0, \qquad d\left(x_{m(k)}, x_{n(k)-1}\right) < \varepsilon_0. \tag{15}$$

□

Definition 17. Let Ψ be the family of all continuous, nondecreasing mappings $\psi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ such that $\psi(t) = 0$ if, and only if, $t = 0$.

These mappings are known as *altering distance functions* (see [27]). Note that every selected $\psi \in \Psi$ commutes with max; that is, $\psi(\max(s_1, s_2, \ldots, s_N)) = \max(\psi(s_1), \psi(s_2), \ldots, \psi(s_N))$ for all $s_1, s_2, \ldots, s_N \in [0, \infty)$.

Lemma 18. *If $\psi \in \Psi$ and $\lim_{m\to\infty}\psi(a_m) = 0$, then $\lim_{m\to\infty} a_m = 0$.*

Proof. As there exists $\psi(a_m)$, then $a_m \in \operatorname{dom}\psi = [0, \infty[$. If the conclusion is not true, there exists $\varepsilon_0 > 0$ such that, for all $n_0 \in \mathbb{N}$, there exists $n \geq n_0$ verifying $a_n \geq \varepsilon_0$. This means that $\{a_m\}$ has a subsequence $\{a_{m(k)}\}_k$ such that $a_{m(k)} \geq \varepsilon_0$. As ψ is nondecreasing, $\psi(\varepsilon_0) \leq \psi(a_{m(k)})$ for all $k \in \mathbb{N}$. Therefore, $\{\psi(a_m)\}_m$ has a subsequence $\{\psi(a_{m(k)})\}_k$ lower bounded by $\psi(\varepsilon_0) > 0$, but this is impossible since $\lim_{m\to\infty}\psi(a_m) = 0$. □

With regards to coincidence points, it is possible to consider the following simplification. If τ is a permutation of Λ_n, and we reorder (4), then we deduce that every coincidence point may be seen as a coincidence point associated to the identity mapping on Λ_n (see, for instance, [28]).

Lemma 19. *Let τ be a permutation of Λ_n, and let $\Phi = (\sigma_1, \sigma_2, \ldots, \sigma_n, \tau)$ and $\Phi' = (\sigma_{\tau^{-1}(1)}, \sigma_{\tau^{-1}(2)}, \ldots, \sigma_{\tau^{-1}(n)}, I_{\Lambda_n})$. Then, a point $(x_1, x_2, \ldots, x_n) \in X^n$ is a Φ-coincidence point of the mappings F and g if, and only if, $(x_1, x_2, \ldots, x_n)$ is a Φ'-coincidence point of the mappings F and g.*

Therefore, in the sequel, without loss of generality, we will only consider Υ-coincidence points where $\Upsilon = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, that is, that verify $F(x_{\sigma_i(1)}, x_{\sigma_i(2)}, \ldots, x_{\sigma_i(n)}) = gx_i$ for all i. We also show some preliminary results on PMS.

Lemma 20. *Let $\{x_m\}$ be a sequence on a PMS (X, p), and let $x \in X$.*

(1) *If $\{x_m\} \xrightarrow{p} x$ and $p(x, x) = 0$, then $\{x_m\} \xrightarrow{d_p} x$, $\{d_p(x_m, y)\} \to d_p(x, y)$ and $\{p(x_m, y)\} \to p(x, y)$ for all $y \in X$.*

(2) *If $\{x_m\} \xrightarrow{d_p} x$ and $\{p(x_m, x_m)\} \to 0$, then $p(x, x) = 0$.*

Proof. (1) Since $0 \leq p(x_m, x_m) \leq p(x, x_m)$ and $\lim_{m\to\infty} p(x, x_m) = p(x, x) = 0$, then $\lim_{m\to\infty} p(x_m, x_m) = 0$. Therefore, $\lim_{m\to\infty} d_p(x, x_m) = \lim_{m\to\infty}(2p(x, x_m) - p(x, x) - p(x_m, x_m)) = 0$, so $\{x_m\} \xrightarrow{d_p} x$. Since d_p is continuous, then $\{d_p(x_m, y)\} \to d_p(x, y)$ for all $y \in X$, and item 4 of Lemma 13 implies that $\{p(x_m, y)\} \to p(x, y)$.

(2) Item 2 of Lemma 13 shows that $p(x, x) = \lim_{m,m'\to\infty} p(x_m, x_{m'}) = \lim_{m\to\infty} p(x_m, x_m) = 0$. □

Remark 21. Although the limit in a MS is unique, the p-limit in a PMS is not necessarily unique. For instance, let (X, p) as in Example 10. Then, (X, p) is a complete PMS (see [14]). Consider $x_m = 2.5 - 1/(2m)$ for all $m \in \mathbb{N}$. Then, $\{x_m\} \xrightarrow{d_p} 2.5$ but $\{x_m\} \xrightarrow{p} x_0$ whenever $x_0 \in [2.5, 3]$.

Definition 22. Let $N \in \mathbb{N}$, let (X, p) be a PMS, let $G : X^N \to X$ be a mapping, and let $Y_0 = (x_1, x_2, \ldots, x_N) \in X^N$. We will say that G is *α_p-continuous* at Y_0 if, for all sequences $\{x_m^1\}, \{x_m^2\}, \ldots, \{x_m^N\}$ on X such that $\{x_m^i\} \xrightarrow{p} x_i$ for all $i \in \{1, 2, \ldots, N\}$, $p(x_i, x_i) = 0$ for all $i \in \{1, 2, \ldots, N\}$ and $\{p(G(x_m^1, x_m^2, \ldots, x_m^N), G(x_m^1, x_m^2, \ldots, x_m^N))\} \to 0$, we have that $\{G(x_m^1, x_m^2, \ldots, x_m^N)\} \xrightarrow{p} G(Y_0)$ and $p(G(Y_0), G(Y_0)) = 0$. One will say that G is *α_p-continuous* if it is continuous at every point $Y_0 \in X^N$.

Lemma 23. *If (X, p) is a PMS, and $G : X^N \to X$ is d_p-continuous at $Y_0 \in X^N$, then G is α_p-continuous at Y_0.*

Proof. Let $\{x_m^1\}, \{x_m^2\}, \ldots, \{x_m^N\}$ sequences on X such that $\{x_m^i\} \xrightarrow{p} x_i$ for all $i \in \{1, 2, \ldots, N\}$, $p(x_i, x_i) = 0$ for all $i \in \{1, 2, \ldots, N\}$, and $\{p(G(x_m^1, x_m^2, \ldots, x_m^N), G(x_m^1, x_m^2, \ldots, x_m^N))\} \to 0$. Item 1

of Lemma 20 implies that $\{x_m^i\} \xrightarrow{d_p} x_i$ for all $i \in \{1, 2, \ldots, N\}$. Since G is d_p-continuous at $Y_0 = (x_1, x_2, \ldots, x_N)$, then $\{G(x_m^1, x_m^2, \ldots, x_m^N)\} \xrightarrow{d_p} G(x_1, x_2, \ldots, x_N)$. Item 2 of Lemma 13 assures us that $\{G(x_m^1, \ldots, x_m^N)\} \xrightarrow{p} G(x_1, \ldots, x_N)$ and

$$\begin{aligned} &p\left(G\left(x_1, x_2, \ldots, x_N\right), G\left(x_1, x_2, \ldots, x_N\right)\right) \\ &\quad = \lim_{m,m' \to \infty} p\left(G\left(x_m^1, x_m^2, \ldots, x_m^N\right), G\left(x_{m'}^1, x_{m'}^2, \ldots, x_{m'}^N\right)\right) \\ &\quad = \lim_{m \to \infty} p\left(G\left(x_m^1, x_m^2, \ldots, x_m^N\right), G\left(x_m^1, x_m^2, \ldots, x_m^N\right)\right) = 0. \end{aligned} \tag{16}$$

Then, G is α_p-continuous at Y_0. □

4. Main Results

In the following result, we show sufficient conditions to ensure the existence of Υ-coincidence points, where $\Upsilon = (\sigma_1, \sigma_2, \ldots, \sigma_n)$.

Theorem 24. *Let (X, p) be a complete PMS, and let $\leq$ a partial order on X. Let $\Upsilon = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be an n-tuple of mappings from $\{1, 2, \ldots, n\}$ into itself verifying $\sigma_i \in \Omega_{A,B}$ if $i \in A$ and $\sigma_i \in \Omega'_{A,B}$ if $i \in B$. Let $F : X^n \to X$ and $g : X \to X$ be two mappings such that F has the mixed g-monotone property on X, $F(X^n) \subseteq g(X)$ and g is α_p-continuous and commuting with F. Assume that there exist $\psi, \varphi \in \Psi$ such that*

$$\begin{aligned} &\psi\left(p\left(F\left(x_1, x_2, \ldots, x_n\right), F\left(y_1, y_2, \ldots, y_n\right)\right)\right) \\ &\quad \leq \psi\left(\max_{1 \leq i \leq n} p\left(gx_i, gy_i\right)\right) - \varphi\left(\max_{1 \leq i \leq n} p\left(gx_i, gy_i\right)\right), \end{aligned} \tag{17}$$

for which $gx_i \leq_i gy_i$ for all i. Suppose either F is α_p-continuous or $(X, d_p, \leq)$ has the sequential g-monotone property. If there exist $x_0^1, x_0^2, \ldots, x_0^n \in X$ verifying $gx_0^i \leq_i F(x_0^{\sigma_i(1)}, x_0^{\sigma_i(2)}, \ldots, x_0^{\sigma_i(n)})$ for all i, then F and g have, at least, one Υ-coincidence point.

Proof. The proof is divided into seven steps. The first two steps are the same as in the proof of Theorem 9 in [11], since the contractivity condition does not play any role in these parts of the proof.

Step 1. There exist n sequences $\{x_m^1\}_{m \geq 0}, \{x_m^2\}_{m \geq 0}, \ldots, \{x_m^n\}_{m \geq 0}$ such that $gx_{m+1}^i = F(x_m^{\sigma_i(1)}, x_m^{\sigma_i(2)}, \ldots, x_m^{\sigma_i(n)})$ for all m and all i.

Step 2. $gx_m^i \leq_i gx_{m+1}^i$ for all m and all i.

Step 3. We claim that $\{p(gx_m^i, gx_{m+1}^i)\}_{m \geq 0} \to 0$ for all i (i.e., $\{\max_{1 \leq j \leq n} p(gx_m^j, gx_{m+1}^j)\}_{m \geq 0} \to 0$).

Indeed, define $\delta_m = \max_{1 \leq j \leq n} p(gx_m^j, gx_{m+1}^j)$ for all m. As $gx_m^i \leq_i gx_{m+1}^i$ for all m and all i, then condition (17) implies that, for all $m \geq 1$ and all i:

$$\begin{aligned} &\psi\left(p\left(gx_m^i, gx_{m+1}^i\right)\right) \\ &\quad = \psi\left(p\left(F\left(x_{m-1}^{\sigma_i(1)}, x_{m-1}^{\sigma_i(2)}, \ldots, x_{m-1}^{\sigma_i(n)}\right), F\left(x_m^{\sigma_i(1)}, x_m^{\sigma_i(2)}, \ldots, x_m^{\sigma_i(n)}\right)\right)\right) \\ &\quad \leq \psi\left(\max_{1 \leq j \leq n} p\left(gx_{m-1}^{\sigma_i(j)}, gx_m^{\sigma_i(j)}\right)\right) - \varphi\left(\max_{1 \leq j \leq n} p\left(gx_{m-1}^{\sigma_i(j)}, gx_m^{\sigma_i(j)}\right)\right) \\ &\quad \leq \psi\left(\max_{1 \leq j \leq n} p\left(gx_{m-1}^j, gx_m^j\right)\right) = \psi\left(\delta_{m-1}\right). \end{aligned} \tag{18}$$

Therefore, for all $m \geq 1$, $\psi(\delta_m) = \psi(\max_{1 \leq i \leq n} p(gx_m^i, gx_{m+1}^i)) = \max_{1 \leq i \leq n} \psi(p(gx_m^i, gx_{m+1}^i)) \leq \psi(\delta_{m-1})$. This means that the sequence $\{\psi(\delta_m)\}_{m \geq 1}$ is nonincreasing and lower bounded. Hence, it is convergent; that is, there exists $\Delta \geq 0$ such that $\{\psi(\delta_m)\}_{m \geq 1} \to \Delta$. We are going to show that $\Delta = 0$. Since

$$\begin{aligned} &\left\{\max_{1 \leq i \leq n} \psi\left(p\left(gx_m^i, gx_{m+1}^i\right)\right)\right\}_m \\ &\quad = \left\{\psi\left(\max_{1 \leq i \leq n} p\left(gx_m^i, gx_{m+1}^i\right)\right)\right\}_m \\ &\quad = \left\{\psi\left(\delta_m\right)\right\}_m \longrightarrow \Delta, \end{aligned} \tag{19}$$

Lemma 14 assures that there exist $i_0 \in \{1, 2, \ldots, n\}$ and a subsequence $\{\psi(p(gx_{m(k)}^{i_0}, gx_{m(k)+1}^{i_0}))\}_k$ such that $\{\psi(p(gx_{m(k)}^{i_0}, gx_{m(k)+1}^{i_0}))\}_k \to \Delta$. Repeating (18), for all $k \geq 1$,

$$\begin{aligned} &\psi\left(p\left(gx_{m(k)}^{i_0}, gx_{m(k)+1}^{i_0}\right)\right) \\ &\quad \leq \psi\left(\max_{1 \leq j \leq n} p\left(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}\right)\right) \\ &\qquad - \varphi\left(\max_{1 \leq j \leq n} p\left(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}\right)\right). \end{aligned} \tag{20}$$

Consider the sequence

$$\left\{\max_{1 \leq j \leq n} p(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)})\right\}_{k \geq 1}. \tag{21}$$

Suppose that this sequence has no subsequence converging to zero. Using $\varepsilon = 1$, Lemma 15 assures us that there exists $\delta' \in]0, 1[$ and $k_0 \in \mathbb{N}$ such that $\max_{1 \leq j \leq n} p(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}) \geq \delta'$ for all $k \geq k_0$. It follows that

$$-\varphi\left(\max_{1 \leq j \leq n} p\left(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}\right)\right) \leq -\varphi\left(\delta'\right) \quad \forall k \geq k_0. \tag{22}$$

Then, (20) says to us

$$\begin{aligned}
&\psi\left(p\left(gx_{m(k)}^{i_0}, gx_{m(k)+1}^{i_0}\right)\right)\\
&\quad\le \psi\left(\max_{1\le j\le n} p\left(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}\right)\right)\\
&\qquad - \varphi\left(\max_{1\le j\le n} p\left(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}\right)\right)\\
&\quad\le \psi\left(\max_{1\le j\le n} p\left(gx_{m(k)-1}^{\sigma_{i_0}(j)}, gx_{m(k)}^{\sigma_{i_0}(j)}\right)\right) - \varphi\left(\delta'\right)\\
&\quad\le \psi\left(\delta_{m(k)-1}\right) - \varphi\left(\delta'\right).
\end{aligned} \tag{23}$$

Taking limit in k, we deduce that $\Delta \le \Delta - \varphi(\delta') < \Delta$, which is impossible. Therefore, the sequence in (21) must have a subsequence $\{\max_{1\le j\le n} p(gx_{m'(k)-1}^{\sigma_{i_0}(j)}, gx_{m'(k)}^{\sigma_{i_0}(j)})\}_{k\ge1}$ converging to zero. Since ψ and φ are continuous, taking limit when $k \to \infty$ in (20) using this subsequence, we deduce that $0 \le \Delta \le \psi(0) - \varphi(0) = 0$, so $\Delta = 0$. Then, we have just proved that $\Delta = 0$. Therefore, $\{\psi(\delta_m)\}_{m\ge1} \to \Delta = 0$, and Lemma 18 assures that $\{\delta_m\}_{m\ge1} \to 0$, which means that $\{p(gx_m^j, gx_{m+1}^j)\} \to 0$ for all j since $0 \le p(gx_m^j, gx_{m+1}^j) \le \delta_m$ for all m and all j.

Step 4. $\{p(gx_m^i, gx_m^i)\}_{m\ge0} \to 0$ for all i (i.e., $\{\max_{1\le j\le n} p(gx_m^j, gx_m^j)\}_{m\ge0} \to 0$). It is the same proof of Step 3.

Since $d_p(gx_m^i, gx_{m+1}^i) = 2p(gx_m^i, gx_{m+1}^i) - p(gx_m^i, gx_m^i) - p(gx_{m+1}^i, gx_{m+1}^i)$ for all m and i, joining Steps 3 and 4, it follows that

$$\left\{d_p\left(gx_m^i, gx_{m+1}^i\right)\right\} \longrightarrow 0 \quad \forall i. \tag{24}$$

Step 5. Every sequence $\{gx_m^i\}_{m\ge0}$ is d_p-Cauchy. We reason by contradiction. Suppose that $\{gx_m^{i_1}\}_{m\ge0}, \ldots, \{gx_m^{i_s}\}_{m\ge0}$ are not d_p-Cauchy ($s \ge 1$) and $\{gx_m^{i_{s+1}}\}_{m\ge0}, \ldots, \{gx_m^{i_n}\}_{m\ge0}$ are d_p-Cauchy, being $\{i_1, \ldots, i_n\} = \{1, \ldots, n\}$. By Lemma 16, for all $r \in \{1, 2, \ldots, s\}$, there exists $\varepsilon_r > 0$ and subsequences $\{gx_{m_r(k)}^{i_r}\}_{k\in\mathbb{N}}$ and $\{gx_{n_r(k)}^{i_r}\}_{k\in\mathbb{N}}$ such that

$$\begin{gathered}
k < m_r(k) < n_r(k),\\
d_p\left(gx_{m_r(k)}^{i_r}, gx_{n_r(k)}^{i_r}\right) \ge \varepsilon_r,\\
d_p\left(gx_{m_r(k)}^{i_r}, gx_{n_r(k)-1}^{i_r}\right) < \varepsilon_r, \quad \forall k \in \mathbb{N}.
\end{gathered} \tag{25}$$

Now, let $\varepsilon_0 = \max(\varepsilon_1, \ldots, \varepsilon_s) > 0$ and $\varepsilon_0' = \min(\varepsilon_1, \ldots, \varepsilon_s) > 0$. Since $\{gx_m^{i_{s+1}}\}_{m\ge0}, \ldots, \{gx_m^{i_n}\}_{m\ge0}$ are d_p-Cauchy, for all $j \in \{i_{s+1}, \ldots, i_n\}$, there exists $n_1^j \in \mathbb{N}$ such that if $m, m' \ge n_1^j$, then $d_p(gx_m^j, gx_{m'}^j) < \varepsilon_0'/8$. Since $\{p(gx_m^j, gx_m^j)\} \to 0$ by Step 4, there exists $n_2^j \in \mathbb{N}$ such that if $m \ge n_2^j$, then $p(gx_m^j, gx_m^j) < \varepsilon_0'/8$. Define $n_0 = \max_{j\in\{i_{s+1},\ldots,i_n\}}(n_1^j, n_2^j)$. If $m, m' \ge n_0$, then

$$\begin{aligned}
0 &\le p\left(gx_m^j, gx_{m'}^j\right)\\
&= \frac{d_p\left(gx_m^j, gx_{m'}^j\right) + p\left(gx_m^j, gx_m^j\right) + p\left(gx_{m'}^j, gx_{m'}^j\right)}{2}\\
&< \frac{\varepsilon_0'/8 + \varepsilon_0'/8 + \varepsilon_0'/8}{2} = \frac{3\varepsilon_0'}{16} < \frac{\varepsilon_0'}{4}.
\end{aligned} \tag{26}$$

Therefore, we have proved that there exists $n_0 \in \mathbb{N}$ such that if $m, m' \ge n_0$, then

$$d_p\left(gx_m^j, gx_{m'}^j\right) < \frac{\varepsilon_0'}{4}, \qquad p\left(gx_m^j, gx_{m'}^j\right) < \frac{\varepsilon_0'}{4}, \qquad \forall j \in \{i_{s+1}, \ldots, i_n\}. \tag{27}$$

Next, let $q \in \{1, 2, \ldots, s\}$ such that $\varepsilon_q = \varepsilon_0 = \max(\varepsilon_1, \ldots, \varepsilon_s)$. Let $k_1 \in \mathbb{N}$ such that $n_0 < m_q'(k_1)$, and define $m(1) = m_q(k_1)$. Consider the numbers $m(1) + 1, m(1) + 2, \ldots, n_q(k_1)$ until finding the first positive integer $n(1) > m(1)$ verifying

$$\max_{1\le r\le s} d_p\left(gx_{m(1)}^{i_r}, gx_{n(1)}^{i_r}\right) \ge \varepsilon_0, \quad d_p\left(gx_{m(1)}^{i_j}, gx_{n(1)-1}^{i_j}\right) < \varepsilon_0, \qquad \forall j \in \{1, 2, \ldots, s\}. \tag{28}$$

Now let $k_2 \in \mathbb{N}$ such that $n(1) < m_q(k_2)$, and define $m(2) = m_q(k_2)$. Consider the numbers $m(2) + 1, m(2) + 2, \ldots, n_q(k_2)$ until finding the first positive integer $n(2) > m(2)$ verifying

$$\max_{1\le r\le s} d_p\left(gx_{m(2)}^{i_r}, gx_{n(2)}^{i_r}\right) \ge \varepsilon_0, \quad d_p\left(gx_{m(2)}^{i_j}, gx_{n(2)-1}^{i_j}\right) < \varepsilon_0, \qquad \forall j \in \{1, 2, \ldots, s\}. \tag{29}$$

Repeating this process, we can find sequences such that, for all $k \ge 1$,

$$\begin{gathered}
n_0 < m(k) < n(k) < m(k+1),\\
\max_{1\le r\le s} d_p\left(gx_{m(k)}^{i_r}, gx_{n(k)}^{i_r}\right) \ge \varepsilon_0,\\
d_p\left(gx_{m(k)}^{i_j}, gx_{n(k)-1}^{i_j}\right) < \varepsilon_0, \quad \forall j \in \{1, 2, \ldots, s\}.
\end{gathered} \tag{30}$$

Note that by (27), $d_p(gx_{m(k)}^{i_r}, gx_{n(k)}^{i_r}), d_p(gx_{m(k)}^{i_r}, gx_{n(k)-1}^{i_r}) < \varepsilon_0'/4 < \varepsilon_0/2$ for all $r \in \{s+1, s+2, \ldots, n\}$, so

$$\begin{gathered}
\max_{1\le j\le n} d_p\left(gx_{m(k)}^j, gx_{n(k)}^j\right) = \max_{1\le r\le s} d_p\left(gx_{m(k)}^{i_r}, gx_{n(k)}^{i_r}\right) \ge \varepsilon_0,\\
d_p\left(gx_{m(k)}^i, gx_{n(k)-1}^i\right) < \varepsilon_0,
\end{gathered} \tag{31}$$

for all $i \in \{1, 2, \ldots, n\}$ and all $k \geq 1$. Furthermore, for all j,

$$\begin{aligned} &2p\left(gx^j_{m(k)-1}, gx^j_{n(k)-1}\right) - p\left(gx^j_{m(k)-1}, gx^j_{m(k)-1}\right) \\ &\quad - p\left(gx^j_{n(k)-1}, gx^j_{n(k)-1}\right) \\ &\qquad = d_p\left(gx^j_{m(k)-1}, gx^j_{n(k)-1}\right) \\ &\qquad \leq d_p\left(gx^j_{m(k)-1}, gx^j_{m(k)}\right) + d_p\left(gx^j_{m(k)}, gx^j_{n(k)-1}\right) \\ &\qquad \leq d_p\left(gx^j_{m(k)-1}, gx^j_{m(k)}\right) + \varepsilon_0. \end{aligned} \tag{32}$$

Therefore, for all j and all k,

$$\begin{aligned} &p\left(gx^j_{m(k)-1}, gx^j_{n(k)-1}\right) \\ &\quad \leq \left(\varepsilon_0 + d_p\left(gx^j_{m(k)-1}, gx^j_{m(k)}\right) + p\left(gx^j_{m(k)-1}, gx^j_{m(k)-1}\right)\right. \\ &\qquad \left. + p\left(gx^j_{n(k)-1}, gx^j_{n(k)-1}\right)\right) \times 2^{-1}. \end{aligned} \tag{33}$$

Next, for all k, let $i(k) \in \{1, 2, \ldots, s\}$ be an index such that

$$\begin{aligned} d_p\left(gx^{i(k)}_{m(k)}, gx^{i(k)}_{n(k)}\right) &= \max_{1\leq r\leq s} d_p\left(gx^{i_r}_{m(k)}, gx^{i_r}_{n(k)}\right) \\ &= \max_{1\leq j\leq n} d_p\left(gx^j_{m(k)}, gx^j_{n(k)}\right) \geq \varepsilon_0. \end{aligned} \tag{34}$$

Then, for all k,

$$\begin{aligned} &p\left(gx^{i(k)}_{m(k)}, gx^{i(k)}_{n(k)}\right) \\ &\quad = \left(d_p\left(gx^{i(k)}_{m(k)}, gx^{i(k)}_{n(k)}\right) + p\left(gx^{i(k)}_{m(k)}, gx^{i(k)}_{m(k)}\right)\right. \\ &\qquad \left. + p\left(gx^{i(k)}_{n(k)}, gx^{i(k)}_{n(k)}\right)\right) \times 2^{-1} \geq \frac{\varepsilon_0}{2}. \end{aligned} \tag{35}$$

Applying the contractivity condition (17), it follows, for all k,

$$\begin{aligned} 0 &< \psi\left(\frac{\varepsilon_0}{2}\right) \\ &\leq \psi\left(p\left(gx^{i(k)}_{m(k)}, gx^{i(k)}_{n(k)}\right)\right) \\ &\leq \psi\left(\max_{1\leq j\leq n} p\left(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}\right)\right) \\ &\quad - \varphi\left(\max_{1\leq j\leq n} p\left(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}\right)\right). \end{aligned} \tag{36}$$

Consider the sequence:

$$\left\{\max_{1\leq j\leq n} p\left(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}\right)\right\}_{k\geq 1}. \tag{37}$$

If this sequence has a subsequence that converges to zero, then we can take limit when $k \to \infty$ in (36) using this subsequence, so that we would have $0 < \psi(\varepsilon_0/2) \leq \psi(0) - \varphi(0) = 0$, which is impossible since $\varepsilon_0 > 0$. Therefore, the sequence (37) has no subsequence converging to zero. In this case, taking $\varepsilon_0 > 0$ in Lemma 15, there exist $\delta \in]0, \varepsilon_0[$ and $k_0 \in \mathbb{N}$ such that $\max_{1\leq j\leq n} p(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}) \geq \delta$, for all $k \geq k_0$. It follows that, for all $k \geq k_0$, $-\varphi(\max_{1\leq j\leq n} p(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1})) \leq -\varphi(\delta)$. Thus, by (36),

$$\begin{aligned} 0 &< \psi\left(\frac{\varepsilon_0}{2}\right) \\ &\leq \psi\left(\max_{1\leq j\leq n} p\left(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}\right)\right) \\ &\quad - \varphi\left(\max_{1\leq j\leq n} p\left(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}\right)\right) \\ &\leq \psi\left(\max_{1\leq j\leq n} p\left(gx^{\sigma_{i(k)}(j)}_{m(k)-1}, gx^{\sigma_{i(k)}(j)}_{n(k)-1}\right)\right) - \varphi(\delta) \\ &\leq \psi\left(\max_{1\leq j\leq n} p\left(gx^{j}_{m(k)-1}, gx^{j}_{n(k)-1}\right)\right) - \varphi(\delta). \end{aligned} \tag{38}$$

Fix any $\gamma > 0$ and we are going to prove that $\psi(\varepsilon_0/2) + \varphi(\delta) \leq \psi(\varepsilon_0/2 + \gamma)$. Indeed, by Step 3 and (24), since

$$\begin{gathered} \left\{\max_{1\leq i\leq n} p\left(gx^i_{m(k)-1}, gx^i_{m(k)-1}\right)\right\}, \\ \left\{\max_{1\leq i\leq n} p\left(gx^i_{n(k)-1}, gx^i_{n(k)-1}\right)\right\}, \\ \left\{\max_{1\leq i\leq n} d_p\left(gx^i_{m(k)-1}, gx^i_{m(k)}\right)\right\} \end{gathered} \tag{39}$$

are sequences converging to zero, we can find $m_1 \in \mathbb{N}$ such that if $m(k) \geq m_1$, then

$$\begin{gathered} \max_{1\leq i\leq n} p\left(gx^i_{m(k)-1}, gx^i_{m(k)-1}\right) \leq \frac{\gamma}{2}, \\ \max_{1\leq i\leq n} p\left(gx^i_{n(k)-1}, gx^i_{n(k)-1}\right) \leq \frac{\gamma}{2}, \\ \max_{1\leq i\leq n} d_p\left(gx^i_{m(k)-1}, gx^i_{m(k)}\right) \leq \frac{\gamma}{2}. \end{gathered} \tag{40}$$

Therefore, (33) implies that, for all j and for all k such that $m(k) > m_1$,

$$\begin{aligned} &p\left(gx^j_{m(k)-1}, gx^j_{n(k)-1}\right) \\ &\quad \leq \left(\varepsilon_0 + d_p\left(gx^j_{m(k)-1}, gx^j_{m(k)}\right) + p\left(gx^j_{m(k)-1}, gx^j_{m(k)-1}\right)\right. \\ &\qquad \left. + p\left(gx^j_{n(k)-1}, gx^j_{n(k)-1}\right)\right) \times 2^{-1} \\ &\quad \leq \frac{\varepsilon_0 + \gamma/2 + \gamma/2 + \gamma/2}{2} = \frac{\varepsilon_0}{2} + \frac{3\gamma}{4} < \frac{\varepsilon_0}{2} + \gamma. \end{aligned} \tag{41}$$

Then, (38) guarantees that $0 < \psi(\varepsilon_0/2) \leq \psi(\max_{1\leq j\leq n} p(gx^j_{m(k)-1}, gx^j_{n(k)-1})) - \varphi(\delta) \leq \psi(\varepsilon_0/2 + \gamma) - \varphi(\delta)$. This means that $\psi(\varepsilon_0/2) + \varphi(\delta) \leq \psi(\varepsilon_0/2 + \gamma)$ for all $\gamma > 0$. If we take $\gamma = 1/m > 0$ (where $m \in \mathbb{N}$), we deduce that $\psi(\varepsilon_0/2) + \varphi(\delta) \leq \psi(\varepsilon_0/2 + 1/m)$ for all $m \in \mathbb{N}$. Since ψ is continuous, we have that $\psi(\varepsilon_0/2) + \varphi(\delta) \leq \psi(\varepsilon_0/2)$, which is

impossible since $\varphi(\delta) > 0$. This contradiction finally proves that every sequence $\{gx_m^i\}_{m\geq 0}$ is d_p-Cauchy.

Since X is p-complete, then X is d_p-complete (item 3 of Lemma 13). Then, there exist $x_1, x_2, \ldots, x_n \in X$ such that $\{gx_m^i\} \xrightarrow{d_p} x_i$ for all i. Furthermore, $p(x_i, x_i) = \lim_{m,m'\to\infty} p(gx_m^i, gx_{m'}^i) = \lim_{m\to\infty} p(gx_m^i, gx_m^i) = 0$ for all i. Since g is α_p-continuous, then $\{ggx_m^i\} \xrightarrow{p} gx_i$ and $p(gx_i, gx_i) = 0$ for all i. Item 1 of Lemma 20 shows that $\{ggx_m^i\} \xrightarrow{d_p} gx_i$ for all i. Therefore, for all i, $\lim_{m\to\infty} p(ggx_{m+1}^i, ggx_{m+1}^i) = \lim_{m,m'\to\infty} p(ggx_{m+1}^i, ggx_{m'+1}^i) = p(gx_i, gx_i) = 0$. Moreover, for all m and all i, $ggx_{m+1}^i = gF(x_m^{\sigma_i(1)}, x_m^{\sigma_i(2)}, \ldots, x_m^{\sigma_n(n)}) = F(gx_m^{\sigma_i(1)}, gx_m^{\sigma_i(2)}, \ldots, gx_m^{\sigma_i(n)})$.

Step 6. Suppose that F is α_p-continuous. In this case, we know that $\{gx_m^i\} \xrightarrow{p} x_i$ and $p(x_i, x_i) = 0$ for all i and

$$\begin{aligned}&\left\{p\left(F\left(gx_m^{\sigma_i(1)}, gx_m^{\sigma_i(2)}, \ldots, gx_m^{\sigma_i(n)}\right),\right.\right.\\&\qquad\left.\left.F\left(gx_m^{\sigma_i(1)}, gx_m^{\sigma_i(2)}, \ldots, gx_m^{\sigma_i(n)}\right)\right)\right\}\\&\quad=\left\{p\left(ggx_{m+1}^i, ggx_{m+1}^i\right)\right\} \longrightarrow 0,\end{aligned} \tag{42}$$

which implies that $\{F(gx_m^{\sigma_i(1)}, gx_m^{\sigma_i(2)}, \ldots, gx_m^{\sigma_i(n)})\} \xrightarrow{p} F(x_{\sigma_i(1)}, x_{\sigma_i(2)}, \ldots, x_{\sigma_i(n)})$ and $p(F(x_{\sigma_i(1)}, \ldots, x_{\sigma_i(n)}), F(x_{\sigma_i(1)}, \ldots, x_{\sigma_i(n)})) = 0$ for all i. Item 1 of Lemma 20 assures us that, for all i,

$$\begin{aligned}\left\{ggx_{m+1}^i\right\} &= \left\{F\left(gx_m^{\sigma_i(1)}, gx_m^{\sigma_i(2)}, \ldots, gx_m^{\sigma_i(n)}\right)\right\}\\&\xrightarrow{d_p} F\left(x_{\sigma_i(1)}, x_{\sigma_i(2)}, \ldots, x_{\sigma_i(n)}\right).\end{aligned} \tag{43}$$

Since the limit in a MS is unique, we deduce that $F(x_{\sigma_i(1)}, x_{\sigma_i(2)}, \ldots, x_{\sigma_i(n)}) = gx_i$ for all i, so $(x_1, x_2, \ldots, x_n)$ is a Υ-coincidence point of F and g.

Step 7. Suppose that $(X, d_p, \leq)$ has the sequential g-monotone property. In this case, by Step 2, we know that $gx_m^i \leq_i gx_{m+1}^i$ for all m and all i. This means that the sequence $\{gx_m^i\}_{m\geq 0}$ is monotone. As $\{gx_m^i\} \xrightarrow{d_p} x_i$, we deduce that $ggx_m^i \leq_i gx_i$ for all m and all i. This condition implies that, for all m and all j,

$$\text{either} \quad \left[ggx_m^{\sigma_j(i)} \leq_i gx_{\sigma_j(i)}\ \forall i\right] \quad \text{or} \quad \left[gx_{\sigma_j(i)} \leq_i ggx_m^{\sigma_j(i)} \forall i\right] \tag{44}$$

(the first case occurs when $j \in A$ and the second one when $j \in B$). Then, by (17), for all j,

$$\begin{aligned}&\psi\left(p\left(ggx_{m+1}^j, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)\right)\\&\quad=\psi\left(p\left(F\left(gx_m^{\sigma_j(1)}, gx_m^{\sigma_j(2)}, \ldots, gx_m^{\sigma_j(n)}\right),\right.\right.\\&\qquad\left.\left.F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)\right)\\&\quad\leq\psi\left(\max_{1\leq i\leq n} p\left(ggx_m^{\sigma_j(i)}, gx_{\sigma_j(i)}\right)\right)\\&\qquad-\varphi\left(\max_{1\leq i\leq n} p\left(ggx_m^{\sigma_j(i)}, gx_{\sigma_j(i)}\right)\right)\\&\quad\leq\psi\left(\max_{1\leq i\leq n} p\left(ggx_m^i, gx_i\right)\right).\end{aligned} \tag{45}$$

Since $\{ggx_m^i\} \xrightarrow{d_p} gx_i$ for all i, then

$$\lim_{m\to\infty} p\left(ggx_m^i, gx_i\right) = p\left(gx_i, gx_i\right) = 0 \quad \forall i. \tag{46}$$

Therefore, $\lim_{m\to\infty}(\max_{1\leq i\leq n} p(ggx_m^i, gx_i)) = 0$. Taking limit when $m \to \infty$ in (45), we deduce that $\lim_{m\to\infty}\psi(p(ggx_{m+1}^j, F(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}))) = 0$ for all j. As $\psi \in \Psi$, Lemma 18 guarantees that

$$\lim_{m\to\infty} p\left(ggx_{m+1}^j, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right) = 0 \quad \forall j. \tag{47}$$

Finally, for all j,

$$\begin{aligned}&d_p\left(gx_j, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)\\&\quad=2p\left(gx_j, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)-p\left(gx_j, gx_j\right)\\&\qquad-p\left(F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right),\right.\\&\qquad\qquad\left.F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)\\&\quad\leq 2p\left(gx_j, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)\\&\quad\leq 2\left[p\left(ggx_m^i, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)p\left(gx_j, ggx_m^i\right)\right.\\&\qquad\left.+p\left(ggx_m^i, F\left(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)}\right)\right)\right].\end{aligned} \tag{48}$$

Using (46) and (47), we conclude that $d_p(gx_j, F(x_{\sigma_j(1)}, x_{\sigma_j(2)}, \ldots, x_{\sigma_j(n)})) = 0$ for all j. □

Remark 25. In the previous theorem, if the image Im d of the metric d is not the whole set $[0, \infty[$, then ψ and φ can only be defined on Im d, and we can consider a wider range of mappings since it is only necessary to impose that they are continuous and nondecreasing on Im d.

Remark 26. We notice also that our paper cannot be deduced from the recent interesting paper of Haghi et al. [21] on partial metric space. In fact, we use a partial order $\leq$. Then, we only suppose (17) for which $gx_i \leq_i gy_i$ for all i (not necessarily on points which are not comparable). Further, we use a self-map $g : X \to X$ which implies that

$$P(A,B) = \max_{1\leq i\leq n} p(ga_i, gb_i), \quad A = (a_1, a_2, \ldots, a_n), \quad B = (b_1, b_2, \ldots, b_n) \in X^n \tag{49}$$

is not necessarily a partial metric on X^n. For instance, let $X = \mathbb{R}_0^+ = [0,\infty)$ provided with its usual partial order and the partial metric $p(x,y) = \max(x,y)$. Consider

$$gx = \begin{cases} 0, & \text{if } 0 \leq x \leq 1, \\ x-1, & \text{if } x > 1. \end{cases} \tag{50}$$

Then, g is continuous, but

$$\begin{aligned} &P((0,0,\ldots,0),(0,0,\ldots,0)) \\ &\quad = P((1,1,\ldots,1),(0,0,\ldots,0)) \\ &\quad = P((1,1,\ldots,1),(1,1,\ldots,1)) = 0, \end{aligned} \tag{51}$$

but $(0,0,\ldots,0) \neq (1,1,\ldots,1)$. Then, P does not verify the axiom $p(x,x) = p(x,y) = p(y,y) \Rightarrow x = y$. Therefore, we cannot apply Theorem 2.4 on Haghi et al. [21].

As a result, we cannot use Theorem 2.7 in [21] since T has an influence in $-\varphi(\max\{p(x,y), p(y,Ty)\})$, and our mapping F has not a role in the left side of (17).

5. Consequences

Remark 27. Theorem 9 in [11] is an easy consequence of Theorem 24 if we take $p = d$, $\psi(t) = t$, and $\varphi(t) = (1-k)t$ for all $t \in \mathbb{R}_0^+$.

In the next result, let Γ_0 be the family of all nondecreasing on each argument, continuous mappings $\phi : [0,\infty[^n \to \mathbb{R}_0^+$ verifying $\phi(x_1, x_2, \ldots, x_n) = 0$ if, and only if, $x_1 = x_2 = \cdots = x_n = 0$. Examples of such mappings are the following, where $k > 0$, $\alpha_i > 0$ and $n_i \in \mathbb{N}$ for all i.

(i) $\phi(x_1, \ldots, x_n) = k\max_{1\leq i\leq n} x_i$.

(ii) $\phi(x_1, \ldots, x_n) = \sum_{i=1}^n \alpha_i x_i^{n_i}$.

(iii) $\phi(x_1, \ldots, x_n) = \sqrt[n]{\alpha_1 x_1^2 + \cdots + \alpha_n x_n^2}$.

Lemma 28. *Let $\phi \in \Gamma_0$, and define $\varphi : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ as $\varphi(t) = \min(\phi(te_1), \phi(te_2), \ldots, \phi(te_n))$ for all $t \geq 0$, where $\{e_1, e_2, \ldots, e_n\}$ is the usual basis of $\mathbb{R}^n$. Then, $\varphi \in \Psi$ and $\varphi(\max_{1\leq i\leq n} x_i) \leq \phi(x_1, x_2, \ldots, x_n)$ for all $x_1, x_2, \ldots, x_n \in \mathbb{R}_0^+$.*

Proof. First part is clear. If $x_{i_0} = \max_{1\leq i\leq n} x_i$, then $\phi(x_{i_0} e_{i_0}) = \phi(0,0,\ldots,0,x_{i_0},0,\ldots,0) \leq \phi(x_1, x_2, \ldots, x_{i_0-1}, x_{i_0}, x_{i_0+1}, \ldots, x_n) = \phi(x_1, x_2, \ldots, x_n)$. Therefore, $\varphi(\max_{1\leq i\leq n} x_i) = \varphi(x_{i_0}) \leq \phi(x_{i_0} e_{i_0}) \leq \phi(x_1, x_2, \ldots, x_n)$. □

Corollary 29. *Thesis of Theorem 24 also holds if one replaces the contractivity condition* (17) *by any of the following list (for which $gx_i \leq_i gy_i$ for all i).*

(A) *This condition can be found in [11] and [12], there exist $\psi \in \Psi$ and $\phi \in \Gamma_0$ such that*

$$\begin{aligned} &\psi(p(F(x_1, x_2, \ldots, x_n), F(y_1, y_2, \ldots, y_n))) \\ &\quad \leq \psi\left(\max_{1\leq i\leq n} p(gx_i, gy_i)\right) \\ &\qquad - \phi(p(gx_1, gy_1), \ldots, p(gx_n, gy_n)). \end{aligned} \tag{52}$$

(B) *In [17], there exist $\psi, \varphi \in \Psi$ and $\beta_1, \beta_2, \ldots, \beta_n \in [0,1]$ such that $\beta_1 + \beta_2 + \cdots + \beta_n \leq 1$ and*

$$\begin{aligned} &\psi(p(F(x_1, x_2, \ldots, x_n), p(y_1, y_2, \ldots, y_n))) \\ &\quad \leq \psi\left(\sum_{i=1}^n \beta_i p(gx_i, gy_i)\right) - \varphi\left(\max_{1\leq i\leq n} p(gx_i, gy_i)\right). \end{aligned} \tag{53}$$

(C) *There exist $\psi, \varphi \in \Psi$ and $\alpha_1, \alpha_2, \ldots, \alpha_n > 0$ such that*

$$\begin{aligned} &\psi(p(F(x_1, x_2, \ldots, x_n), F(y_1, y_2, \ldots, y_n))) \\ &\quad \leq \psi\left(\max_{1\leq i\leq n} p(gx_i, gy_i)\right) - \varphi\left(\sum_{i=1}^n \alpha_i p(gx_i, gy_i)\right). \end{aligned} \tag{54}$$

(D) *In [2, 7, 9], there exist $\psi \in \Psi$, $\alpha_1, \alpha_2, \ldots, \alpha_n > 0$, and $\beta_1, \beta_2, \ldots, \beta_n \geq 0$ such that $\beta_1 + \beta_2 + \cdots + \beta_n \leq 1$ and*

$$\begin{aligned} &\psi(p(F(x_1, x_2, \ldots, x_n), F(y_1, y_2, \ldots, y_n))) \\ &\quad \leq \psi\left(\sum_{i=1}^n \beta_i p(gx_i, gy_i)\right) - \sum_{i=1}^n \alpha_i p(gx_i, gy_i). \end{aligned} \tag{55}$$

(E) *In [5], there exist $\psi \in \Psi$ and $\alpha_1, \alpha_2, \ldots, \alpha_n > 0$ such that*

$$\begin{aligned} &\psi(p(F(x_1, x_2, \ldots, x_n), F(y_1, y_2, \ldots, y_n))) \\ &\quad \leq \psi\left(\frac{1}{n}\sum_{i=1}^n p(gx_i, gy_i)\right) - \sum_{i=1}^n \alpha_i p(gx_i, gy_i). \end{aligned} \tag{56}$$

(F) *In [19, 20], there exist $\psi, \varphi \in \Psi$ such that ψ is subadditive ($\psi(s+t) \leq \psi(s) + \psi(t)$ for all $t, s \in [0,\infty)$) and*

$$\begin{aligned} &\psi(p(F(x_1, x_2, \ldots, x_n), F(y_1, y_2, \ldots, y_n))) \\ &\quad \leq \frac{1}{n}\psi\left(\sum_{i=1}^n p(gx_i, gy_i)\right) - \varphi\left(\max_{1\leq i\leq n} p(gx_i, gy_i)\right). \end{aligned} \tag{57}$$

Of course, it is also interesting to particularize all the previous items to the following cases: $\psi(t) = \lambda t$ (where $\lambda > 0$), $\varphi(t) = \mu t$ (where $\mu > 0$), or $gx = x$ for all $x \in X$.

Proof. (A) By Lemma 28, there exists $\varphi \in \Psi$ such that $-\phi(x_1, x_2, \ldots, x_n) \leq -\varphi(\max_{1\leq i\leq n} x_i)$ for all $x_1, x_2, \ldots, x_n \in [0,\infty[$, so (52) implies (17). (B) It is obvious that $\sum_{i=1}^n \beta_i p(gx_i, gy_i) \leq (\sum_{i=1}^n \beta_i)\max_{1\leq j\leq n} p(gx_j, gy_j) \leq \max_{1\leq j\leq n} p(gx_j, gy_j)$, so (53) implies (17). (C) We only

take $\phi(x_1, x_2, \ldots, x_n) = \varphi(\sum_{i=1}^{n} \alpha_i x_i)$ in item (A). (D) It is a mixture of (B) and (C). (E) It is a particular case of (D) where $\beta_i = 1/n$ for all i. (F) If ψ is subadditive, then $(1/n)\psi(t) \leq \psi(t/n)$ for all $t \geq 0$, so we may choose $\beta_i = 1/n$ for all i in (B). □

6. Uniqueness of Υ-Coincidence Points

Consider on the product space X^n the following partial order: for $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in X^n$,

$$(x_1, x_2, \ldots, x_n) \leq (y_1, y_2, \ldots, y_n) \Longleftrightarrow x_i \leq_i y_i, \quad \forall i. \quad (58)$$

We say that $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ are *comparable* if $(x_1, x_2, \ldots, x_n) \leq (y_1, y_2, \ldots, y_n)$ or $(x_1, x_2, \ldots, x_n) \geq (y_1, y_2, \ldots, y_n)$.

Theorem 30. *Under the hypothesis of Theorem 24, assume that for all Υ-coincidence points $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n) \in X^n$ of F and g there exists $(u_1, u_2, \ldots, u_n) \in X^n$ such that $(gu_1, gu_2, \ldots, gu_n)$ is comparable, at the same time, to $(gx_1, gx_2, \ldots, gx_n)$ and to $(gy_1, gy_2, \ldots, gy_n)$.*

Then, F and g have a unique Υ-coincidence point $(z_1, z_2, \ldots, z_n) \in X^n$ such that $gz_i = z_i$ for all i.

Proof. From Theorem 24, the set of Υ-coincidence points of F and g is nonempty. The proof is divided into two steps.

Step 1. We claim that if $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in X^n$ are two Υ-coincidence points of F and g, then

$$gx_i = gy_i \quad \forall i. \quad (59)$$

Let $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in X^n$ be two Υ-coincidence points of F and g, and let $(u_1, u_2, \ldots, u_n) \in X^n$ be a point such that $(gu_1, gu_2, \ldots, gu_n)$ is comparable, at the same time, to $(gx_1, gx_2, \ldots, gx_n)$ and to $(gy_1, gy_2, \ldots, gy_n)$. Using $(u_1, u_2, \ldots, u_n)$, define the following sequences. Let $u_0^i = u_i$ for all i. Reasoning as in Theorem 24, we can determine sequences $\{u_m^1\}_{m\geq 0}, \{u_m^2\}_{m\geq 0}, \ldots, \{u_m^n\}_{m\geq 0}$ such that $gu_{m+1}^i = F(u_m^{\sigma_i(1)}, u_m^{\sigma_i(2)}, \ldots, u_m^{\sigma_i(n)})$ for all m and all i. We are going to prove that $gx_i = \lim_{m\to 0}^{d_p} gu_m^i = gy_i$ for all i, so (59) will be true.

Firstly, we reason with $(gu_1, gu_2, \ldots, gu_n)$ and $(gx_1, gx_2, \ldots, gx_n)$, and the same argument holds for $(gu_1, gu_2, \ldots, gu_n)$ and $(gy_1, gy_2, \ldots, gy_n)$. As $(gu_1, gu_2, \ldots, gu_n)$ and $(gx_1, gx_2, \ldots, gx_n)$ are comparable, we can suppose that $(gu_1, gu_2, \ldots, gu_n) \leq (gx_1, gx_2, \ldots, gx_n)$ (the other case is similar); that is, $gu_0^i = gu_i \leq_i gx_i$ for all i. Using that F has the mixed g-monotone property and reasoning as in Theorem 24, it is possible to prove that $gu_m^i \leq_i gx_i$ for all $m \geq 1$ and all i. This condition implies that, for all j and all $m \geq 1$

$$\text{either} \quad \left[gu_m^{\sigma_j(i)} \leq_i gx_{\sigma_j(i)} \forall i\right] \quad \text{or} \quad \left[gx_{\sigma_j(i)} \leq_i gu_m^{\sigma_j(i)} \forall i\right]. \quad (60)$$

Define $\beta_m = \max_{1\leq i\leq n} p(gu_m^i, gx_i)$ for all m. Reasoning as in Theorem 24, it is not difficult to prove that $\{\beta_m\}_{m\geq 1} \to 0$ which means that $\lim_{m\to\infty}\beta_m = \lim_{m\to\infty}(\max_{1\leq i\leq n} p(gu_m^i, gx_i)) = 0$. As $0 \leq p(gu_m^i, gx_i) \leq \beta_m$ for all m and all i, we deduce that $\{p(gu_m^i, gx_i)\}_{m\geq 1} \to 0 = p(gx_i, gx_i)$ for all i; that is, $\{gu_m^i\} \xrightarrow{p} gx_i$ for all i. Item 1 of Lemma 20 shows that

$$\left\{gu_m^i\right\} \xrightarrow{d_p} gx_i \quad \forall i. \quad (61)$$

If we had supposed that $(gx_1, gx_2, \ldots, gx_n) \leq (gu_1, gu_2, \ldots, gu_n)$, we would have obtained the same property (61). And as $(gu_1, gu_2, \ldots, gu_n)$ also is comparable to $(gy_1, gy_2, \ldots, gy_n)$, we can reason in the same way to prove that $\{gu_m^i\} \xrightarrow{d_p} gy_i$ for all i. Since the limit in a MS is unique, $gx_i = gy_i$ for all i.

Let $(x_1, x_2, \ldots, x_n) \in X^n$ be a Υ-coincidence point of F and g, and define $z_i = gx_i$ for all i. As $(z_1, z_2, \ldots, z_n) = (gx_1, gx_2, \ldots, gx_n)$, Remark 6 assures us that $(z_1, z_2, \ldots, z_n)$ also is a Υ-coincidence point of F and g.

Step 2. We claim that $(z_1, z_2, \ldots, z_n)$ is the unique Υ-coincidence point of F and g such that $gz_i = z_i$ for all i. It is similar to Step 2 in Theorem 11 in [11]. □

It is natural to say that g is *injective* on the set of all Υ-coincidence points of F and g when $gx_i = gy_i$ for all i implies $x_i = y_i$ for all i when $(x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \in X^n$ are two Υ-coincidence points of F and g. For example, this is true is g is injective on X.

Corollary 31. *In addition to the hypotheses of Theorem 30, suppose that g is injective on the set of all Υ-coincidence points of F and g. Then, F and g have a unique Υ-coincidence point.*

Proof. If $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ are two Υ-coincidence points of F and g, we have proved in (59) that $gx_i = gy_i$ for all i. As g is injective on these points, then, $x_i = y_i$ for all i. □

Corollary 32. *In addition to the hypotheses of Theorem 30, suppose that $(z_{\sigma_i(1)}, z_{\sigma_i(2)}, \ldots, z_{\sigma_i(n)})$ is comparable to $(z_{\sigma_j(1)}, z_{\sigma_j(2)}, \ldots, z_{\sigma_j(n)})$ for all i, j. Then, $z_1 = z_2 = \cdots = z_n$.*

In particular, there exists a unique $z \in X$ such that $F(z, z, \ldots, z) = z$, which verifies $gz = z$.

Proof. Let $M = \max_{1\leq i,j\leq n} p(z_i, z_j)$, let $j_0, s_0 \in \{1, 2, \ldots, n\}$ such that $p(z_{j_0}, z_{s_0}) = M$, and let

$$\Lambda = \max_{1\leq i\leq n} p\left(z_{\sigma_{j_0}(i)}, z_{\sigma_{s_0}(i)}\right) \leq M. \quad (62)$$

Fix $j, s \in \{1, 2, \ldots, n\}$. As $(z_{\sigma_j(1)}, z_{\sigma_j(2)}, \ldots, z_{\sigma_j(n)})$ is comparable to $(z_{\sigma_s(1)}, z_{\sigma_s(2)}, \ldots, z_{\sigma_s(n)})$, then either $z_{\sigma_j(i)} \leq_i z_{\sigma_s(i)}$ for all i or $z_{\sigma_s(i)} \leq_i z_{\sigma_j(i)}$ for all i. Since $gz_i = z_i$ for all i, we know that either $gz_{\sigma_j(i)} \leq_i gz_{\sigma_s(i)}$ for all i or $gz_{\sigma_s(i)} \leq_i gz_{\sigma_j(i)}$ for all i. In any case, applying (17),

$$\begin{aligned}
\psi(M) &= \psi\left(p\left(z_{j_0}, z_{s_0}\right)\right) \\
&= \psi\left(p\left(gz_{j_0}, gz_{s_0}\right)\right) \\
&= \psi\left(p\left(F\left(z_{\sigma_{j_0}(1)}, z_{\sigma_{j_0}(2)}, \ldots, z_{\sigma_{j_0}(n)}\right),\right.\right. \\
&\qquad \left.\left. F\left(z_{\sigma_{s_0}(1)}, z_{\sigma_{s_0}(2)}, \ldots, z_{\sigma_{s_0}(n)}\right)\right)\right) \\
&\leq \psi\left(\max_{1\leq i\leq n} p\left(gz_{\sigma_{j_0}(i)}, gz_{\sigma_{s_0}(i)}\right)\right) \\
&\quad - \varphi\left(\max_{1\leq i\leq n} p\left(gz_{\sigma_{j_0}(i)}, gz_{\sigma_{s_0}(i)}\right)\right) \\
&= \psi(\Lambda) - \varphi(\Lambda) \leq \psi(M) - \varphi(\Lambda).
\end{aligned} \tag{63}$$

If $\Lambda > 0$, then $\varphi(\Lambda) > 0$, so $\psi(M) \leq \psi(M) - \varphi(\Lambda) < \psi(M)$, which is impossible. Then, $\Lambda = 0$, and (63) implies that $\psi(M) \leq \psi(\Lambda) - \varphi(\Lambda) = \psi(0) - \varphi(0) = 0$, so $\psi(M) = 0$. Therefore, $p(z_i, z_j) = 0$ for all i and j. □

Example 33. Let $X = \mathbb{R}$ provided with its usual partial order $\leq$ and the partial metric $p(x, y) = \max(|x|, |y|)$. Let $n \in \mathbb{N}$, and let $a_1, a_2, \ldots, a_n \in \mathbb{R} \setminus \{0\}$ real numbers such that there exist $i_0, j_0 \in \{1, 2, \ldots, n\}$ verifying $a_{i_0} < 0 < a_{j_0}$. Let $N > |a_1| + |a_2| + \cdots + |a_n|$, and consider $F(x_1, x_2, \ldots, x_n) = (a_1x_1 + a_2x_2 + \cdots + a_nx_n)/N$ and $gx = x$, for all $x, x_1, x_2, \ldots, x_n \in X$. Then, F is monotone nondecreasing in those arguments for which $a_i > 0$ and monotone nonincreasing in those arguments for which $a_i < 0$. Furthermore, taking $k = (|a_1| + |a_2| + \cdots + |a_n|)/N \in (0, 1)$, it follows that

$$\begin{aligned}
&\left|F\left(x_1, x_2, \ldots, x_n\right)\right| \\
&\quad \leq \frac{|a_1|\,|x_1| + |a_2|\,|x_2| + \cdots + |a_n|\,|x_n|}{N} \\
&\quad \leq \frac{|a_1| + |a_2| + \cdots + |a_n|}{N} \max\left(|x_1|, |x_2|, \ldots, |x_n|\right) \\
&\quad = k \max\left(|x_1|, |x_2|, \ldots, |x_n|\right).
\end{aligned} \tag{64}$$

Therefore,

$$\begin{aligned}
&p\left(F\left(x_1, x_2, \ldots, x_n\right), F\left(y_1, y_2, \ldots, y_n\right)\right) \\
&\quad = \max\left(\left|F\left(x_1, x_2, \ldots, x_n\right)\right|, \left|F\left(y_1, y_2, \ldots, y_n\right)\right|\right) \\
&\quad \leq \max\left(k \max\left(|x_1|, |x_2|, \ldots, |x_n|\right),\right. \\
&\qquad\qquad \left. k \max\left(|y_1|, |y_2|, \ldots, |y_n|\right)\right) \\
&\quad = k \max_i \left(p\left(x_i, y_i\right)\right).
\end{aligned} \tag{65}$$

If $\psi(t) = t$ and $\varphi(t) = (1 - k)t$, all conditions of Theorems 24 and 30 (and Corollaries 31 and 32) are satisfied. Indeed, it is clear that $(0, 0, \ldots, 0)$ is the unique fixed point of F.

The following example is based on Examples 1.9 and 2.2 in [29].

Example 34. Let $X = \{0, 1, 2, 3, 4\}$, and let p be the partial metric on X given by $p(x, y) = \max(x, y)$ for all $x, y \in X$. Then, (X, p) is complete, and p generates the discrete topology on X (indeed, d_p is the Euclidean metric on X). Consider on X the following partial order:

$$x, y \in X, \qquad x \leq y \Longleftrightarrow x = y \quad \text{or} \quad (x, y) = (0, 2). \tag{66}$$

Consider $F : X^n \to X$ and $g : X \to X$ defined by

$$\begin{aligned}
F\left(x_1, x_2, \ldots, x_n\right) &= \begin{cases} 0, & \text{if } x_1, x_2, \ldots, x_n \in \{0, 1, 2\}, \\ 1, & \text{otherwise,} \end{cases} \\
gx &= \begin{cases} 0, & \text{if } x = 0, \\ 2, & \text{if } x \in \{0.5, 1\}, \\ 3, & \text{if } x \in \{1.5, 2\}. \end{cases}
\end{aligned} \tag{67}$$

It is not difficult to prove the following statements.

(1) F and g are α_p-continuous mappings (since d_p generates the discrete topology on X).

(2) F and g are commuting.

(3) If $y, z \in X$ verify $gy \leq gz$, then either $y, z \in \{0, 1, 2\}$ or $y, z \in \{3, 4\}$. Then, F has the mixed $(g, \leq)$-monotone property on X.

(4) If $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in X$ verify $gx_i \leq_i gy_i$ for all i, then $F(x_1, x_2, \ldots, x_n) = F(y_1, y_2, \ldots, y_n)$. In particular, (17) holds (whatever ψ and φ; for instance, $\psi(t) = 2t$ and $\varphi(t) = \log(t + 1)$ for all $t \geq 0$).

For simplicity, henceforth, suppose that n is even, and let A (resp., B) be the set of all odd (resp., even) numbers in $\{1, 2, \ldots, n\}$.

(5) For a mapping $\sigma : \Lambda_n \to \Lambda_n$, we use the notation $\sigma \equiv (\sigma(1), \sigma(2), \ldots, \sigma(n))$ and consider

$$\sigma_i \equiv (i, i + 1, \ldots, n - 1, n, 1, 2, \ldots, i - 1) \quad \forall i. \tag{68}$$

Then, $\sigma_i \in \Omega_{A,B}$ if i is odd, and $\sigma_i \in \Omega'_{A,B}$ if i is even. Let $\Upsilon = (\sigma_1, \sigma_2, \ldots, \sigma_n)$.

(6) Take $x_0^i = 0$ if i is odd and $x_0^i = 2$ if i is even. Then, $gx_0^i \leq_i F(x_0^{\sigma_i(1)}, x_0^{\sigma_i(2)}, \ldots, x_0^{\sigma_i(n)})$ for all i.

(7) $(X, d_p, \leq)$ has the sequential g-monotone property.

Therefore, we can apply Theorems 24 and 30, and Corollaries 31 and 32, to conclude that F and g have a unique Υ-coincidence point, which is $(0, 0, \ldots, 0)$.

Acknowledgments

This work has been partially supported by Junta de Andaluca, by projects FQM-268, FQM-178 and FQM-235 of the Andalusian CICYE.

References

[1] D. J. Guo and V. Lakshmikantham, "Coupled fixed points of nonlinear operators with applications," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 11, no. 5, pp. 623–632, 1987.

[2] T. Gnana Bhaskar and V. Lakshmikantham, "Fixed point theorems in partially ordered metric spaces and applications," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 65, no. 7, pp. 1379–1393, 2006.

[3] J.-X. Fang, "Common fixed point theorems of compatible and weakly compatible maps in Menger spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 5-6, pp. 1833–1843, 2009.

[4] X.-Q. Hu, "Common coupled fixed point theorems for contractive mappings in fuzzy metric spaces," *Fixed Point Theory and Applications*, Article ID 363716, 14 pages, 2011.

[5] V. Lakshmikantham and L. Ćirić, "Coupled fixed point theorems for nonlinear contractions in partially ordered metric spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 70, no. 12, pp. 4341–4349, 2009.

[6] S. Shakeri, L. J. B. Ćirić, and R. Saadati, "Common fixed point theorem in partially ordered L-fuzzy metric spaces," *Fixed Point Theory and Applications*, vol. 2010, Article ID 125082, 13 pages, 2010.

[7] V. Berinde and M. Borcut, "Tripled fixed point theorems for contractive type mappings in partially ordered metric spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 15, pp. 4889–4897, 2011.

[8] V. Berinde, "Approximating common fixed points of noncommuting almost contractions in metric spaces," *Fixed Point Theory*, vol. 11, no. 2, pp. 179–188, 2010.

[9] M. Borcut and V. Berinde, "Tripled coincidence theorems for contractive type mappings in partially ordered metric spaces," *Applied Mathematics and Computation*, vol. 218, no. 10, pp. 5929–5936, 2012.

[10] M. Turinici, "Product fixed points in ordered metric spaces," http://arxiv.org/abs/1110.3079 .

[11] A. Roldán, J. Martínez-Moreno, and C. Roldán, "Multidimensional fixed point theorems in partially ordered complete metric spaces," *Journal of Mathematical Analysis and Applications*, vol. 396, no. 2, pp. 536–545, 2012.

[12] M. Berzig and B. Samet, "An extension of coupled fixed point's concept in higher dimension and applications," *Computers & Mathematics with Applications*, vol. 63, no. 8, pp. 1319–1334, 2012.

[13] E. Karapınar and V. Berinde, "Quadruple fixed point theorems for nonlinear contractions in partially ordered metric spaces," *Banach Journal of Mathematical Analysis*, vol. 6, no. 1, pp. 74–89, 2012.

[14] S. G. Matthews, "Partial metric topology, general topology and its applications," in *Proceedings of the 8th Summer Conference, Queen's College*, vol. 728, pp. 183–197, Annals of the New York Academy of Sciences, 1994.

[15] İ. M. Erhan, E. Karapınar, and A. Öztürk, "Fixed point theorems on quasi-partial metric spaces," *Mathematical and Computer Modelling*, vol. 57, no. 9-10, pp. 2442–2448, 2013.

[16] W. Shatanawi, B. Samet, and M. Abbas, "Coupled fixed point theorems for mixed monotone mappings in ordered partial metric spaces," *Mathematical and Computer Modelling*, vol. 55, no. 3-4, pp. 680–687, 2012.

[17] H. Aydi, E. Karapınar, and W. Shatanawi, "Coupled fixed point results for (ψ, φ)-weakly contractive condition in ordered partial metric spaces," *Computers & Mathematics with Applications*, vol. 62, no. 12, pp. 4449–4460, 2011.

[18] B. S. Choudhury, N. Metiya, and A. Kundu, "Coupled coincidence point theorems in ordered metric spaces," *Annali dell'Universitá di Ferrara*, vol. 57, no. 1, pp. 1–16, 2011.

[19] E. Karapınar and N. V. Luong, "Quadruple fixed point theorems for nonlinear contractions," *Computers & Mathematics with Applications*, vol. 64, no. 6, pp. 1839–1848, 2012.

[20] N. V. Luong and N. X. Thuan, "Coupled fixed points in partially ordered metric spaces and application," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 3, pp. 983–992, 2011.

[21] R. H. Haghi, Sh. Rezapour, and N. Shahzad, "Be careful on partial metric fixed point results," *Topology and its Applications*, vol. 160, no. 3, pp. 450–454, 2013.

[22] L. Ćirić, M. Abbas, B. Damjanović, and R. Saadati, "Common fuzzy fixed point theorems in ordered metric spaces," *Mathematical and Computer Modelling*, vol. 53, no. 9-10, pp. 1737–1741, 2011.

[23] N. Shobkolaei, S. M. Vaezpour, and S. Sedghi, "A common fixed point theorem on ordered partial metric spaces," *Journal of Applied Sciences Research*, vol. 1, pp. 3433–3439, 2011.

[24] E. Karapınar, N. Shobkolaei, S. Sedghi, and S. M. Vaezpour, "A common fixed point theorem for cyclic operators on partial metric spaces," *Filomat*, vol. 26, pp. 407–414, 2012.

[25] E. Karapınar and İ. M. Erhan, "Fixed point theorems for operators on partial metric spaces," *Applied Mathematics Letters*, vol. 24, no. 11, pp. 1894–1899, 2011.

[26] T. Abdeljawad, E. Karapınar, and K. Taş, "Existence and uniqueness of a common fixed point on partial metric spaces," *Applied Mathematics Letters*, vol. 24, no. 11, pp. 1900–1904, 2011.

[27] M. S. Khan, M. Swaleh, and S. Sessa, "Fixed point theorems by altering distances between the points," *Bulletin of the Australian Mathematical Society*, vol. 30, no. 1, pp. 1–9, 1984.

[28] A. Roldán, J. Martínez-Moreno, C. Roldán, and E. Karapınar, "Meir-Keeler type multidimensional fixed point theorems in partially ordered metric spaces," *Abstract and Applied Analysis*, vol. 2013, Article ID 406026, 9 pages, 2013.

[29] N. M. Hung, E. Karapinar, and N. V. Luong, "Coupled coincidence point theorem in partially ordered metric spaces via implicit relation," *Abstract and Applied Analysis*, vol. 2012, Article ID 796964, 14 pages, 2012.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 319241, 7 pages
http://dx.doi.org/10.1155/2013/319241

Research Article

Convergence of a New Modified Ishikawa Type Iteration for Common Fixed Points of Total Asymptotically Strict Pseudocontractive Semigroups

Yuanheng Wang and Chunjie Wang

Department of Mathematics, Zhejiang Normal University, Jinhua, Zhejiang 321004, China

Correspondence should be addressed to Yuanheng Wang; wangyuanheng@yahoo.com.cn

Received 28 April 2013; Revised 3 June 2013; Accepted 3 June 2013

Academic Editor: Abdellah Bnouhachem

The purpose of this paper is to give a new modified Ishikawa type iteration algorithm for common fixed points of total asymptotically strict pseudocontractive semigroups. Under the reduction of some conditions, both strong convergence and weak convergence of the iteration algorithm are proved in Banach spaces with new methods of proofs, respectively. The main results presented in this paper extend and improve the corresponding recent results of many others.

1. Introduction

Throughout this paper, we assume that E is a real Banach space with the norm $\|\cdot\|$, E^* the dual space of E, $\langle\cdot,\cdot\rangle$ the duality between E and E^*, and C a nonempty closed convex subset of E. $\mathbb{R}^+$ denotes the set of nonnegative real numbers and $\mathbb{N}$ the natural number set. The mapping $J : E \to 2^{E^*}$ with

$$J(x) = \left\{f^* \in E^* : \langle x, f^*\rangle = \|x\|^2, \|f^*\| = \|x\|\right\}, \quad x \in E, \tag{1}$$

is called the normalized duality mapping.

Let $T : C \to C$ be a nonlinear mapping. $F(T)$ denotes the set of the fixed points of T.

As we know, a mapping $T : C \to C$ is said to be pseudocontractive, if, for all $x, y \in C$, there exists $j(x-y) \in J(x-y)$, such that

$$\langle Tx - Ty, j(x-y)\rangle \le \|x-y\|^2. \tag{2}$$

Variational inequalities introduced by Stampacchia in the early sixties have had a great impact and influence on the development of almost all branches of pure and applied sciences and have witnessed an explosive growth in theoretical advances, algorithmic development, and so forth. Recently, some authors also studied the problem of finding the solution set of variational inequalities and the common element of the fixed point set for generalized nonexpansive mappings in the framework of real Hilbert spaces and Banach spaces. As is known to all, the variational inequality problem, nonlinear optimization problem, and fixed point problem are equivalent to each other under certain conditions.

In 2012, Chang et al. [1] introduced a more general class of pseudocontractive mappings and studied the methods for approximation of the split common fixed points.

Definition 1 (see [1]). (I) A mapping $T : C \to C$ is said to be $(\gamma, \mu_n, \xi_n, \phi)$-totally asymptotically strictly pseudocontractive, if there exist a constant $\gamma \in [0,1]$ and sequences $\{\mu_n\}$, $\{\xi_n\} \subset [0,\infty)$ with $\mu_n \to 0$ and $\xi_n \to 0$, such that, for all $x, y \in C$,

$$\begin{aligned}\|T^n x - T^n y\|^2 \le{}& \|x-y\|^2 + \gamma\|(I-T^n)x - (I-T^n)y\|^2 \\ &+ \mu_n\phi(\|x-y\|) + \xi_n, \quad \forall n \ge 1,\end{aligned} \tag{3}$$

where $\phi : [0,\infty) \to [0,\infty)$ is continuous and a strict increasing function with $\phi(0) = 0$.

(II) A mapping $T : C \to C$ is said to be (γ, k_n)-asymptotically strictly pseudocontractive, if there exist

a constant $\gamma \in [0,1)$ and a sequence $k_n \subset [1,\infty)$ with $k_n \to 1$, such that

$$\|T^n x - T^n y\|^2 \le k_n \|x-y\|^2 + \gamma \|(I-T^n)x - (I-T^n)y\|^2, \quad \forall x, y \in C. \tag{4}$$

Definition 2. (I) One-parameter family $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ is said to be a pseudocontractive semigroup on C, if the following conditions are satisfied:

(a) $T(0)x = x$ for each $x \in C$;

(b) $T(t+s)x = T(t)T(s)x$ for any $t, s \in \mathbb{R}^+$ and $x \in C$;

(c) the mapping $t \to T(t)x$ is continuous for any given $x \in C$;

(d) for any $t \ge 0$, $T(t)$ is pseudocontractive; that is, for any $x, y \in C$, there exists $j(x-y) \in J(x-y)$, such that

$$\langle T^n(t)x - T^n(t)y, j(x-y)\rangle \le \|x-y\|^2. \tag{5}$$

(II) One-parameter family $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ is said to be strict pseudocontractive semigroup on C, if the conditions (a)–(c) and the following condition (e) are satisfied.

(e) For any $x, y \in C$, there exist $j(x-y) \in J(x-y)$ and a bounded function $\eta : [0,\infty) \to [0,\infty)$, such that, for any $t \ge 0$,

$$\langle T(t)x - T(t)y, j(x-y)\rangle \le \|x-y\|^2 - \eta(t)\|(x-y) - (T(t)x - T(t)y)\|^2. \tag{6}$$

(III) $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ is said to be asymptotically strict pseudocontractive semigroup, if the conditions (a)–(c) and the following condition (f) are satisfied.

(f) There exist a bounded function $\eta : [0,\infty) \to [0,\infty)$ and a sequence $k_n \subset [1,\infty)$ with $k_n \to 1$ as $n \to \infty$. For any given $x, y \in C$, there exists $j(x-y) \in J(x-y)$ such that for, any $t \ge 0$,

$$\langle T^n(t)x - T^n(t)y, j(x-y)\rangle \le k_n\|x-y\|^2 - \eta(t)\|(x-y) - (T^n(t)x - T^n(t)y)\|^2. \tag{7}$$

Osilike and Akuchu [2] established an iterative scheme for approximation of common fixed points of a finite family of asymptotically pseudocontractive mappings. Miao et al. [3] introduced an implicit iteration process for a finite family of total asymptotically pseudocontractive maps. And in recent years, many researchers focused on the convergence of pseudocontractive and asymptotically strict pseudocontractive semigroups; see [4–8] and their references. In [9, 10] especially, the authors gave the modified Mann type iteration algorithm and studied its convergence.

Inspired and motivated by the above works, in this paper, we give a new modified Ishikawa type iteration algorithm for total asymptotically strict pseudocontractive semigroups. Under the reducation of some conditions, we prove both strong convergence and weak convergence of the iteration algorithm by using the method of the subsequence of a subsequence of the sequence $\{x_n\}$ in Banach spaces, respectively. The results presented in this paper extend and improve the corresponding recent results of many authors, such as [1, 7–10].

2. Preliminaries

This section contains some definitions, notations, and lemmas, which will be used in the proofs of our main results in the next section.

A Banach space E is said to be smooth if the limit $\lim_{t\to 0}((\|x+ty\| - \|x\|)/t)$ exists for each $x, y \in \{x \in E : \|x\| = 1\}$. It is well known that if E is reflexive and smooth, then the duality mapping J is single valued.

A Banach space E is said to have Opial condition if, for any sequence $\{x_n\} \subset E$ weakly convergent to $x_0 \in E$,

$$\liminf_{n\to\infty} \|x_n - x_0\| < \liminf_{n\to\infty} \|x_n - x\| \tag{8}$$

holds for any $x \ne x_0$.

A mapping T is said to be demiclosed, if, for any sequence $\{x_n\} \subset E$, $x_n \rightharpoonup y$ and $\|(I-T)x_n\| \to 0$ imply that $(I-T)y = 0$.

Definition 3 (see [9]). One-parameter $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ is said to be a $(\eta, \{\mu_n\}, \{\xi_n\}, \phi)$-total asymptotically strict pseudocontractive semigroup on C, if the conditions (a)–(c) in Definition 2 and the following condition (g) are satisfied.

(g) There exist a bounded function $\eta : [0,\infty) \to [0,\infty)$ and sequences $\{\mu_n\} \subset [0,\infty)$ and $\{\xi_n\} \subset [0,\infty)$ with $\mu_n \to 0$, $\xi_n \to 0$, as $n \to \infty$. For any given $x, y \in C$, there exists $j(x-y) \in J(x-y)$, such that

$$\langle T^n(t)x - T^n(t)y, j(x-y)\rangle \le \|x-y\|^2 - \eta(t)\|(x-y) - (T^n(t)x - T^n(t)y)\|^2 + \mu_n\phi(\|x-y\|) + \xi_n, \tag{9}$$

for any $t \ge 0$, for all $n \ge 1$, where $\phi : [0,\infty) \to [0,\infty)$ is continuous and strictly increasing function with $\phi(0) = 0$.

A $(\eta, \{\mu_n\}, \{\xi_n\}, \phi)$-total asymptotically strict pseudocontractive semigroup is said to be *uniformly Lipschitzian*, if there exists a bounded measurable function $L : [0,\infty) \to (0,\infty)$, such that

$$\|T^n(t)x - T^n(t)y\| \le L(t)\|x-y\|, \quad \forall x, y \in C, t \ge 0, n \in \mathbb{N}. \tag{10}$$

Remark 4. According to the definitions, it is obvious that a pseudocontractive semigroup is a strict pseudocontractive semigroup with $\eta(t) = 0$, and a strict pseudocontractive semigroup is an asymptotically strict pseudocontractive semigroup with $k_n = 1$. An asymptotically strict pseudocontractive semigroup is a $(\eta, \{\mu_n\}, \{\xi_n\}, \phi)$-total asymptotically strict pseudocontractive semigroup with $\phi(t) = t^2$, $\mu_n = k_n - 1$, and $\xi_n = 0$.

Definition 5 (see [11]). The normalized duality mapping J of a Banach space E is said to be weakly sequential continuous; if for all $\{x_n\} \subset E$, $x_n \rightharpoonup x$, then there exist $j(x_n) \in J(x_n)$, $j(x) \in J(x)$ such that $j(x_n) \overset{*}{\rightharpoonup} j(x)$, where weak convergence and weak star convergence are denoted by $\rightharpoonup$ and $\overset{*}{\rightharpoonup}$, respectively.

In order to prove the main results of this paper, the following lemmas should be used.

Lemma 6 (see [4]). *For any $x, y \in E$, one has*

$$\|x+y\|^2 \le \|x\|^2 + 2\langle y, j(x-y)\rangle, \quad \forall j(x-y) \in J(x-y). \tag{11}$$

Lemma 7 (see [12]). *Let $\{a_n\}$, $\{b_n\}$, and $\{\delta_n\}$ be the sequences of $\mathbb{R}^+$, which satisfy*

$$a_{n+1} \le (1+\delta_n) a_n + b_n, \quad \forall n \ge 1. \tag{12}$$

If $\sum_{n=1}^{\infty} \delta_n < \infty$, $\sum_{n=1}^{\infty} b_n < \infty$, then the limit $\lim_{n\to\infty} a_n$ exists.

3. Main Results

Theorem 8. *Let C be a nonempty closed convex subset of a real Banach space E, and let $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ be a uniformly Lipschitzian and $(\eta, \{\mu_n\}, \{\xi_n\}, \phi)$-total asymptotically strict pseudocontractive semigroup defined in Definition 3. Suppose that $F(\mathcal{T}) := \bigcap_{t\ge 0} F(T(t)) \ne \emptyset$ and there exists a compact subset K of E such that $\bigcap_{t\ge 0} T(t)(C) \subseteq K$. We assume that there exist positive constants M and M^*, such that $\phi(x) \le M^* x^2$ for all $x \ge M$. Let $\{x_n\}$ be the sequence defined by the modified Ishikawa type iteration algorithm:*

$$\begin{aligned} &x_1 \in C, \quad \text{chosen arbitrarily}, \\ &y_n = (1-\beta_n) x_n + \beta_n T^n(t) x_n, \\ &x_{n+1} = (1-\alpha_n) x_n + \alpha_n T^n(t) y_n. \end{aligned} \tag{13}$$

Then $\{x_n\}$ converges strongly to a common fixed point $x^* \in F(\mathcal{T})$ in C, if the following conditions are satisfied:

(i) $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$, $\sum_{n=1}^{\infty} \alpha_n = \infty$, $\sum_{n=1}^{\infty} \alpha_n \mu_n < \infty$, *and* $\sum_{n=1}^{\infty} \alpha_n \xi_n < \infty$;

(ii) $\beta_n \to 0$ *as* $n \to \infty$, $\sum_{n=1}^{\infty} \alpha_n \beta_n < \infty$;

(iii) $\eta = \inf_{t\ge 0} \eta(t) > 0$, $L = \sup_{t\ge 0} L(t) < +\infty$.

Proof. We divide the proof into four steps.

Step 1. Firstly, we prove that $\lim_{n\to\infty} \|x - p\|$ exists for any $p \in F(\mathcal{T})$.

By the definitions of $T(t)$ and $\{x_n\}$, we have

$$\begin{aligned} \|T^n(t) x_n - p\| &\le L \|x_n - p\|, \\ \|y_n - p\| &= \|(1-\beta_n) x_n + \beta_n T^n(t) x_n - p\| \\ &\le (1-\beta_n) \|x_n - p\| + \beta_n \|T^n(t) x_n - p\| \\ &\le (1-\beta_n + \beta_n L) \|x_n - p\| \\ &\le (1+L) \|x_n - p\|, \\ \|T^n(t) y_n - p\| &\le L \|y_n - p\| \le L(1+L) \|x_n + p\|. \end{aligned} \tag{14}$$

This follows from that

$$\begin{aligned} \|x_{n+1} - p\| &= \|(1-\alpha_n) x_n + \alpha_n T^n(t) y_n - p\| \\ &\le (1-\alpha_n) \|x_n - p\| + \alpha_n \|T^n(t) y_n - p\| \\ &\le (1-\alpha_n) \|x_n - p\| + \alpha_n L(1+L) \|x_n - p\| \\ &\le \left(1 + L + L^2\right) \|x_n - p\|, \\ \|y_n - x_n\| &= \|(1-\beta_n) x_n + \beta_n T^n(t) x_n - x_n\| \\ &= \beta_n \|x_n - T^n(t) x_n\| \\ &\le \beta_n (\|x_n - p\| + \|T^n(t) x_n - p\|) \\ &\le \beta_n (1+L) \|x_n - p\|, \\ \|T^n(t) y_n - x_n\| &\le \|T^n(t) y_n - p\| + \|x_n - p\| \\ &\le \left(1 + L + L^2\right) \|x_n - p\|, \\ \|x_{n+1} - x_n\| &= \|(1-\alpha_n) x_n + \alpha_n T^n(t) y_n - x_n\| \\ &\le \alpha_n \|T^n(t) y_n - x_n\| \\ &\le \alpha_n \left(1 + L + L^2\right) \|x_n - p\|. \end{aligned} \tag{15}$$

Since $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ is total asymptotically strict pseudocontractive semigroup, for any point $x_{n+1} \in C$ and $p \in F(\mathcal{T})$, by (9), we have

$$\begin{aligned} &\langle T^n(t) x_{n+1} - x_{n+1}, j(x_{n+1} - p)\rangle \\ &\quad \le -\eta(t) \|T^n(t) x_{n+1} - x_{n+1}\| + \mu_n \phi(\|x_{n+1} - p\|) + \xi_n. \end{aligned} \tag{16}$$

Since ϕ is an increasing function, it results in that $\phi(x) \le \phi(M)$, if $x \le M$; $\phi(x) \le M^* x^2$, if $x \ge M$. In either case, we can obtain that

$$\phi(x) \le \phi(M) + M^* x^2. \tag{17}$$

Hence, by Lemma 6, we have

$$\begin{aligned}
&\|x_{n+1}-p\|^2 \\
&\quad= \|x_n - p + \alpha_n (T^n(t) y_n - x_n)\|^2 \\
&\quad\le \|x_n-p\|^2 + 2\alpha_n \langle T^n(t) y_n - T^n(t) x_n, j(x_{n+1}-p)\rangle \\
&\qquad + 2\alpha_n \langle T^n(t) x_n - T^n(t) x_{n+1}, j(x_{n+1}-p)\rangle \\
&\qquad + 2\alpha_n \langle T^n(t) x_{n+1} - x_{n+1}, j(x_{n+1}-p)\rangle \\
&\qquad + 2\alpha_n \langle x_{n+1} - x_n, j(x_{n+1}-p)\rangle \\
&\quad\le \|x_n-p\|^2 + 2\alpha_n L \|y_n - x_n\| \cdot \|x_{n+1}-p\| \\
&\qquad + 2\alpha_n \|x_n - x_{n+1}\| \cdot \|x_{n+1}-p\| \\
&\qquad - 2\alpha_n \eta(t) \|T^n(t) x_{n+1} - x_{n+1}\|^2 \\
&\qquad + 2\alpha_n \|x_n - x_{n+1}\| \cdot \|x_{n+1}-p\| \\
&\qquad + 2\alpha_n \mu_n \phi(\|x_{n+1}-p\|) + 2\alpha_n \xi_n \\
&\quad\le \|x_n-p\|^2 + 2\alpha_n \beta_n L(1+L)\left(1+L+L^2\right)\|x_n-p\|^2 \\
&\qquad + 2\alpha_n^2\left(1+L+L^2\right)^2\|x_n-p\|^2 \\
&\qquad - 2\alpha_n \eta(t) \|T^n(t) x_{n+1} - x_{n+1}\|^2 \\
&\qquad + 2\alpha_n^2\left(1+L+L^2\right)^2\|x_n-p\|^2 \\
&\qquad + 2\alpha_n \mu_n \left[\phi(M) + M^*\left(1+L+L^2\right)\|x_n-p\|^2\right] \\
&\qquad + 2\alpha_n \xi_n \\
&\quad= (1+\delta_n)\|x_n-p\|^2 + b_n,
\end{aligned} \tag{18}$$

where $\delta_n = [2\alpha_n\beta_n L(1+L) + 4\alpha_n^2(1+L+L^2) + 2M^*\alpha_n\mu_n](1+L+L^2)$, $b_n = 2\alpha_n\mu_n\phi(M) + 2\alpha_n\xi_n$.

By the conditions (i) and (ii), we have $\sum_{n=1}^{\infty}\delta_n < \infty$, $\sum_{n=1}^{\infty} b_n < \infty$. Thus, by Lemma 7, we can obtain that $\lim_{n\to\infty}\|x_n - p\|$ exists.

Step 2. Now we prove that $\liminf_{n\to\infty}\|x_n - T^n(t)x_n\| = 0$.

From (18), we know that

$$\begin{aligned}
&2\alpha_n\eta(t)\|T^n(t)x_{n+1} - x_{n+1}\|^2 \\
&\quad\le \left(\|x_n-p\|^2 - \|x_{n+1}-p\|^2\right) + \delta_n\|x_n-p\|^2 + b_n.
\end{aligned} \tag{19}$$

As $\eta = \inf_{t\ge 0}\eta(t) > 0$, $A = \sup_n\|x_n - p\| < \infty$, we can have

$$\begin{aligned}
&\sum_{n=1}^{m} 2\alpha_n\eta\|T^n(t)x_{n+1} - x_{n+1}\|^2 \\
&\quad\le \sum_{n=1}^{m}\left[\left(\|x_n-p\|^2 - \|x_{n+1}-p\|^2\right) + \delta_n A^2 + b_n\right] \\
&\quad\le \|x_1-p\|^2 + A^2\sum_{n=1}^{m}\delta_n + \sum_{n=1}^{m} b_n.
\end{aligned} \tag{20}$$

Then,

$$\begin{aligned}
&\sum_{n=1}^{\infty} 2\alpha_n\eta\|T^n(t)x_{n+1} - x_{n+1}\|^2 \\
&\quad\le \sum_{n=1}^{\infty}\left[\left(\|x_n-p\|^2 - \|x_{n+1}-p\|^2\right) + \delta_n A^2 + b_n\right] \\
&\quad\le \|x_1-p\|^2 + A^2\sum_{n=1}^{\infty}\delta_n + \sum_{n=1}^{\infty} b_n < \infty.
\end{aligned} \tag{21}$$

Since $\sum_{n=1}^{\infty}\alpha_n = \infty$, then (21) implies that

$$\liminf_{n\to\infty}\|x_{n+1} - T^n(t)x_{n+1}\| = 0. \tag{22}$$

Otherwise, if $\liminf_{n\to\infty}\|x_{n+1} - T^n(t)x_{n+1}\| = c > 0$, then there exists an N, such that $\|x_n - T^n(t)x_n\| \ge c/2$, when $n \ge N$. So, we have

$$\begin{aligned}
&\sum_{n=1}^{\infty} 2\alpha_n\eta\|T^n(t)x_{n+1} - x_{n+1}\|^2 \\
&\quad= \sum_{n=1}^{N} 2\alpha_n\eta\|T^n(t)x_{n+1} - x_{n+1}\|^2 \\
&\qquad + \sum_{n=N}^{\infty} 2\alpha_n\eta\|T^n(t)x_{n+1} - x_{n+1}\|^2 \\
&\quad\ge \sum_{n=1}^{N} 2\alpha_n\eta\|T^n(t)x_{n+1} - x_{n+1}\|^2 + \frac{c^2}{2\eta}\sum_{n=N}^{\infty}\alpha_n = \infty.
\end{aligned} \tag{23}$$

This is in contradiction with (21).

Because

$$\begin{aligned}
&\|x_n - T^n(t)x_n\| \\
&\quad\le \|x_n - x_{n+1}\| + \|x_{n+1} - T^n(t)x_{n+1}\| \\
&\qquad + \|T^n(t)x_{n+1} - T^n(t)x_n\| \\
&\quad\le \|x_{n+1} - T^n(t)x_{n+1}\| + (1+L)\|x_n - x_{n+1}\| \\
&\quad\le \|x_{n+1} - T^n(t)x_{n+1}\| \\
&\qquad + \alpha_n(1+L)\left(1+L+L^2\right)A
\end{aligned} \tag{24}$$

and $\lim_{n\to\infty}\alpha_n = 0$, we have

$$\liminf_{n\to\infty}\|x_n - T^n(t)x_n\| = 0. \tag{25}$$

Step 3. Now we prove that $\liminf_{n\to\infty}\|x_n - T(t)x_n\| = 0$.

Consider

$$\begin{aligned}
&\|x_{n+1} - T(t)x_{n+1}\| \\
&\quad \le \|x_{n+1} - T^{n+1}(t)x_{n+1}\| \\
&\qquad + \|T^{n+1}(t)x_{n+1} - T(t)x_{n+1}\| \\
&\quad \le \|x_{n+1} - T^{n+1}(t)x_{n+1}\| \\
&\qquad + L\|T^{n}(t)x_{n+1} - x_{n+1}\| \\
&\quad \le \|x_{n+1} - T^{n+1}(t)x_{n+1}\| \\
&\qquad + L(\|T^{n}(t)x_{n} - T^{n}(t)x_{n}\| \\
&\qquad\quad + \|T^{n}(t)x_{n} - x_{n}\| + \|x_{n} - x_{n+1}\|) \\
&\quad \le \|x_{n+1} - T^{n+1}(t)x_{n+1}\| + L\|T^{n}(t)x_{n} - x_{n}\| \\
&\qquad + L(1+L)\|x_{n} - x_{n+1}\| \\
&\quad \le \|x_{n+1} - T^{n+1}(t)x_{n+1}\| + L\|T^{n}(t)x_{n} - x_{n}\| \\
&\qquad + \alpha_n L(1+L)\left(1+L+L^2\right)A.
\end{aligned} \tag{26}$$

Since $\lim_{n\to\infty}\alpha_n = 0$ and (25), we have

$$\liminf_{n\to\infty}\|x_n - T(t)x_n\| = 0. \tag{27}$$

Thus, there exists a subsequence $\{x_{n_k}\} \subseteq \{x_n\}$ such that

$$\lim_{n\to\infty}\|x_{n_k} - T(t)x_{n_k}\| = 0. \tag{28}$$

Step 4. Finally, we prove the sequence $\{x_n\}$ converges strongly to a common fixed point of the semigroup $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$.

Since K is a compact subset of E and $\bigcap_{t\ge0} T(t)(C) \subseteq K$, just as the proof in [9, 10], there exists a subsequence $\{x_{n_{k_i}}\} \subseteq \{x_{n_k}\} \subseteq \{x_n\} \subseteq C$, such that $T(t)x_{n_{k_i}} \to x^* \in K$. From (28), we have $\lim_{n_{k_i}\to\infty}\|T(t)x_{n_{k_i}} - x_{n_{k_i}}\| = 0$, and

$$\left\|x_{n_{k_i}} - x^*\right\| \le \left\|x_{n_{k_i}} - T(t)x_{n_{k_i}}\right\| + \left\|T(t)x_{n_{k_i}} - x^*\right\| \longrightarrow 0. \tag{29}$$

Hence we have that

$$\|T(t)x^* - x^*\| = \lim_{n_{k_i}\to\infty}\left\|x_{n_{k_i}} - T(t)x_{n_{k_i}}\right\| = 0. \tag{30}$$

That is, $x^* \in F(\mathcal{T})$.

Since, for any $p \in F(\mathcal{T})$, $\lim_{n\to\infty}\|x_n - p\|$ exists, $\lim_{n_{k_i}\to\infty}\|x_{n_{k_i}} - x^*\| = 0$, and $\{x_{n_{k_i}}\} \subseteq \{x_n\}$, so we have $\lim_{n\to\infty}\|x_n - p\| = 0$; that is, x_n converges strongly to an element $x^* = p$ of $F(\mathcal{T})$. □

Remark 9. (a) If we take $\beta_n = 0$ in the modified Ishikawa type iteration algorithm (13), then (13) is called the modified Mann type iteration algorithm in many articles, such as in [9, 10]. (b) In Theorem 8, because there is no limit to t of $T(t)$, so our result is stronger and the conditions here are less than in [9, 10]. For example, the conditions "for any bounded subset $D \subset C$,

$$\lim_{n\to\infty}\sup_{x\in D, s\in\mathbb{R}^+}\|T^n(s+t_n)x - T^n(t_n)x\| = 0" \tag{31}$$

in [9, 10] can be removed in Theorem 8. (c) The condition "there exists a compact subset K of E such that $\bigcap_{t\ge0} T(t)(C) \subseteq K$" does not look natural. But it easy to see that this condition is established naturally when we assume C is a compact subset of E. So, the result in Theorem 8 is still true if this condition is replaced by the condition "let C be a compact subset of E." If there is no compactness assumption, we can get the following weak convergence theorem.

Theorem 10. *Let E be a reflexive Banach space satisfying the opial condition and C be a nonempty bounded closed convex subset of E. Let $\mathcal{T} := \{T(t) : C \to C, t \ge 0\}$ be a uniformly Lipschitzian and $(\eta, \{\mu_n\}, \{\xi_n\}, \phi)$-total asymptotically strict pseudocontractive semigroup defined by Definition 3. Suppose that there exist positive constants M and M^*, such that $\phi(x) \le M^* x^2$, for all $x \ge M$, and $F(\mathcal{T}) := \bigcap_{t\ge0} F(T(t)) \ne \emptyset$. Let $\{x_n\}$ be the sequence defined by (13). Then $\{x_n\}$ converges weakly to a common fixed point $x^* \in F(\mathcal{T})$ in C, if the following conditions are satisfied.*

(i) $\sum_{n=1}^{\infty}\alpha_n^2 < \infty$, $\sum_{n=1}^{\infty}\alpha_n = \infty$, $\sum_{n=1}^{\infty}\alpha_n\mu_n < \infty$, $\sum_{n=1}^{\infty}\alpha_n\xi_n < \infty$.

(ii) $\beta_n \to 0$ *as* $n \to \infty$, *and* $\sum_{n=1}^{\infty}\alpha_n\beta_n < \infty$.

(iii) $\lambda = \inf_{t\ge0}\lambda(t) > 0$, $L = \sup_{t\ge0}L(t) < +\infty$.

Proof. It can be proved just like the proof in Theorem 8 that, for each $p \in F(\mathcal{T})$, $\lim_{n\to\infty}\|x_n - p\|$ exists, and, for all $t > 0$, $T(t)x_n$ is bounded, $\liminf\|T(t)x_n - x_n\| = 0$. Thus, there exists a subsequence $\{x_{n_k}\} \subseteq \{x_n\}$ such that $\lim_{n\to\infty}\|x_{n_k} - T(t)x_{n_k}\| = 0$.

Now we prove that $I - T(t)$ is demiclosed at zero (see [11]).

Since C is a closed and convex subset of a reflexive Banach space E, there exists a subsequence $\{x_{n_{k_i}}\} \subseteq \{x_{n_k}\} \subseteq \{x_n\}$, such that $x_{n_{k_i}} \rightharpoonup x^* \in C$. Without loss of generality, we can assume that $\{x_n\}$ replaces $\{x_{n_{k_i}}\}$ now.

In the following, we prove that $x^* = T(t)x^*$.

Firstly, we choose $\alpha \in (0, 1/(1+L))$ and $y_m = (1-\alpha)x + T^m(t)x$ for $m \ge 1$. Since $T(t)$ is uniformly Lipschitzian, we have

$$\begin{aligned}
\|x_n - T^m(t)x_n\| &\le \sum_{k=0}^{m-1}\left\|T^k(t)x_n - T^{k+1}(t)x_n\right\| \\
&\le \sum_{k=0}^{m-1} L\|x_n - T(t)x_n\| \\
&= mL\|x_n - T(t)x_n\| \longrightarrow 0, \qquad n \longrightarrow \infty.
\end{aligned} \tag{32}$$

Because $T(t)$ is totally asymptotically strictly pseudocontractive, we have

$$\begin{aligned}
&\langle (I-T^m(t))\,y_m, J(x-y_m)\rangle \\
&\quad = \langle (I-T^m(t))\,y_m, J(x-y_m)-J(x_n-y_m)\rangle \\
&\qquad + \langle (I-T^m(t))\,y_m, J(x_n-y_m)\rangle \\
&\quad = \langle (I-T^m(t))\,y_m, J(x-y_m)-J(x_n-y_m)\rangle \\
&\qquad + \langle (I-T^m(t))\,x_n, J(x_n-y_m)\rangle \\
&\qquad + \langle (I-T^m(t))\,y_m-(I-T^m(t))\,x_n, J(x_n-y_m)\rangle \\
&\quad = \langle (I-T^m(t))\,y_m, J(x-y_m)-J(x_n-y_m)\rangle \\
&\qquad + \langle (I-T^m(t))\,x_n, J(x_n-y_m)\rangle \\
&\qquad - \eta(t)\left\|(I-T^m(t))\,x_n-(I-T^m(t))\,y_m\right\|^2 \\
&\qquad + \mu_n\phi(\|x_n-y_n\|)+\xi_n \\
&\quad \le \langle (I-T^m(t))\,y_m, J(x-y_m)-J(x_n-y_m)\rangle \\
&\qquad + \langle (I-T^m(t))\,x_n, J(x_n-y_n)\rangle \\
&\qquad + \mu_n\left(M+M^*\|x_n-y_m\|^2\right)+\xi_n.
\end{aligned} \tag{33}$$

Since $x_n \rightharpoonup x^*$, $\lim_{n\to\infty}\|x_n-T(t)x_n\|=0$ (note the x_n of here instead of $x_{n_{k_i}}$), and J is weakly sequential continuous duality mapping, we have

$$\begin{aligned}
&\langle (I-T^m(t))\,x^*-(I-T^m(t))\,y_m, J(x^*-y_m)\rangle \\
&\quad \le (1+L)\left\|x^*-y_m\right\|^2 \\
&\quad \le (1+L)\,\alpha^2\|x^*-T^m(t)x^*\|^2.
\end{aligned} \tag{34}$$

Hence

$$\begin{aligned}
&\|x^*-T^m(t)x^*\|^2 \\
&\quad = \langle x^*-T^m(t)\,x^*, J(x^*-T^m(t)\,x^*)\rangle \\
&\quad = \frac{1}{\alpha}\langle x^*-T^m(t)\,x^*, J(x^*-y_m)\rangle \\
&\quad = \frac{1}{\alpha}\langle x^*-T^m(t)\,x^*-(y_m-T^m(t)\,y_m), J(x^*-y_m)\rangle \\
&\qquad + \frac{1}{\alpha}\langle y_m-T^m(t)\,y_m, J(x^*-y_m)\rangle \\
&\quad \le \alpha(1+L)\left\|x^*-T^m(t)\,x^*\right\|^2 \\
&\qquad + \frac{1}{\alpha}\left(\mu_m M+\mu_m M^*(\operatorname{diam} C)+\xi_n\right).
\end{aligned} \tag{35}$$

This implies that

$$\begin{aligned}
&\alpha(1-\alpha(1+L))\left\|x^*-T^m(t)\,x^*\right\|^2 \\
&\quad \le \mu_m M+\mu_m M^*(\operatorname{diam} C)+\xi_n, \quad \forall m\in\mathbb{N}.
\end{aligned} \tag{36}$$

Let $m\to\infty$; then we have $\|x^*-T^m(t)x^*\|\to 0$, as $m\to\infty$, for $\mu_n\to 0$, $\xi_n\to 0$. Hence, $T^m(t)x^*\to x^*$, as $m\to\infty$, and $T^{m+1}(t)x^*\to T(t)x^*$. By the continuity of $T(t)$, we have $T(t)x^*=x^*$.

Now, for the sequence $\{x_n\}$ generated by (13), we prove that $x_n\rightharpoonup x^*$.

Suppose the contrary; if there exists another subsequence $\{x_{n_j}\}\subset\{x_n\}$, such that $x_{n_j}\rightharpoonup y^*$ with $y^*\ne x^*$, then we have that $\lim_{n\to\infty}\|x_n-x^*\|$ and $\lim_{n\to\infty}\|x_n-y^*\|$ exist. Since E satisfies the Opial condition, we have

$$\begin{aligned}
&\liminf_{n_{k_i}\to\infty}\left\|x_{n_{k_i}}-x^*\right\| \\
&\quad < \lim_{n_{k_i}\to\infty}\left\|x_{n_{k_i}}-y^*\right\| = \lim_{n\to\infty}\|x_n-y^*\| \\
&\quad = \lim_{n_j\to\infty}\left\|x_{n_j}-y^*\right\| < \liminf_{n_j\to\infty}\left\|x_{n_j}-x^*\right\| \\
&\quad = \lim_{n\to\infty}\|x_n-x^*\| = \liminf_{n_{k_i}\to\infty}\left\|x_{n_{k_i}}-x^*\right\|.
\end{aligned} \tag{37}$$

This is a contraction, which shows $x^*=y^*$. Therefore, $x_n\rightharpoonup x^*\in F(\mathcal{T})$. This completes the proof. □

Remark 11. (a) Our results extend many other results that have been proved for this important class of general pseudocontractive mappings. For example, we extend the total asymptotically strict pseudocontractive mapping in [1] and Lipschitzian pseudo-contraction semigroup in [8] to the total asymptotically strict pseudocontractive semigroup. (b) In addition, we study the weak convergence of the total asymptotically strict pseudocontractive semigroup by using the demiclosedness of $I-T(t)$ which, in some way, extends the result in [11] in Banach spaces. (c) And the method by using the subsequence of a subsequence of the sequence $\{x_n\}$ in this paper is different from the previous references.

Acknowledgments

The authors would like to thank editors and referees for many useful comments and suggestions for the improvement of the paper. This work was partially supported by the Natural Science Foundation of Zhejiang Province (Y6100696) and the National Natural Science Foundation of China (11271330).

References

[1] S. S. Chang, L. Wang, Y. K. Tang, and L. Yang, "The split common fixed point problem for total asymptotically strictly pseudocontractive mappings," *Journal of Applied Mathematics*, vol. 2012, Article ID 385638, 13 pages, 2012.

[2] M. O. Osilike and B. G. Akuchu, "Common fixed points of a finite family of asymptotically pseudocontractive maps," *Fixed Point Theory and Applications*, vol. 2004, no. 2, pp. 81–88, 2004.

[3] Q. Miao, R. Chen, and H. Zhou, "Convergence of an implicit iteration process for a finite family of total asymptotically pseudocontractive maps," *International Journal of Mathematical Analysis*, vol. 2, no. 9–12, pp. 433–436, 2008.

[4] Y. H. Wang, "Strong convergence theorems for asymptotically weak G-pseudo-Ψ-contractive non-self-mappings with

the generalized projection in Banach spaces," *Abstract and Applied Analysis*, vol. 2012, Article ID 651304, 11 pages, 2012.

[5] X. Li, J. K. Kim, and N. Huang, "Viscosity approximation of common fixed points for L-Lipschitzian semigroup of pseudo-contractive mappings in Banach spaces," *Journal of Inequalities and Applications*, vol. 2009, Article ID 936121, 16 pages, 2009.

[6] W. Xu and Y. H. Wang, "Strong convergence of the iterative methods for hierarchical fixed point problems of an infinite family of strictly nonself pseudocontractions," *Abstract and Applied Analysis*, vol. 2012, Article ID 457024, 11 pages, 2012.

[7] S. Zhang, "Convergence theorem of common fixed points for Lipschitzian pseudo-contraction semi-groups in Banach spaces," *Applied Mathematics and Mechanics*, vol. 30, no. 2, pp. 145–152, 2009.

[8] J. Quan, S. Chang, and M. Liu, "Strong and weak convergence of an implicit iterative process for pseudocontractive semigroups in Banach space," *Fixed Point Theory and Applications*, vol. 2012, article 142, 2012.

[9] L. Yang and F. H. Zhao, "Large strong convergence theorems for total asymptotically strict pseudocontractive semigroup in Banach spaces," *Fixed Point Theory and Applications*, vol. 2012, article 24, 2012.

[10] S. Chang, Y. J. Cho, H. W. J. Lee, and C. K. Chan, "Strong convergence theorems for Lipschitzian demicontraction semigroups in Banach spaces," *Fixed Point Theory and Applications*, vol. 2011, Article ID 583423, 10 pages, 2011.

[11] Y.-H. Wang and Y.-H. Xia, "Strong convergence for asymptotically pseudocontractions with the demiclosedness principle in Banach spaces," *Fixed Point Theory and Applications*, vol. 2012, article 45, 2012.

[12] K. Aoyama, Y. Kimura, W. Takahashi, and M. Toyoda, "Approximation of common fixed points of a countable family of nonexpansive mappings in a Banach space," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 67, no. 8, pp. 2350–2360, 2007.

Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 912846, 6 pages
http://dx.doi.org/10.1155/2013/912846

Research Article

Convergence Analysis of the Relaxed Proximal Point Algorithm

Min Li[1] and Yanfei You[2]

[1] *School of Economics and Management, Southeast University, Nanjing 210096, China*
[2] *Department of Mathematics, Nanjing University, Nanjing 210093, China*

Correspondence should be addressed to Min Li; liminnju@yahoo.com

Received 3 May 2013; Accepted 9 June 2013

Academic Editor: Abdellah Bnouhachem

Recently, a worst-case $O(1/t)$ convergence rate was established for the Douglas-Rachford alternating direction method of multipliers (ADMM) in an ergodic sense. The relaxed proximal point algorithm (PPA) is a generalization of the original PPA which includes the Douglas-Rachford ADMM as a special case. In this paper, we provide a simple proof for the same convergence rate of the relaxed PPA in both ergodic and nonergodic senses.

1. Introduction

The finite-dimensional variational inequality (VI), denoted by VI(Ω, F), is to find a vector $w^* \in \Omega$ such that

$$\left(w - w^*\right)^T F\left(w^*\right) \geq 0, \quad \forall w \in \Omega, \tag{1}$$

where Ω is a nonempty closed convex set in $\Re^n$ and F is a monotone mapping from $\Re^n$ into itself. The solution set, denoted by Ω^* is assumed to be nonempty. We refer to [1–4] for the pivotal roles of VIs in various fields such as economics, transportation, and engineering.

As is well known, proximal point algorithm (PPA), which was presented originally in [5] and mainly developed in [6, 7], is a well-developed approach to solving VI(Ω, F). Let w^k be the current approximation of a solution of (1); then PPA generates the new iterate $w^{k+1} \in \Omega$ by solving the following auxiliary VI:

$$\left(w - w^{k+1}\right)^T \left[F\left(w^{k+1}\right) + \frac{1}{\beta}\left(w^{k+1} - w^k\right)\right] \geq 0, \tag{2}$$

where β is a positive constant. Compared to the monotone VI (1), (2) is easier to handle since it is a strongly monotone VI. In this paper, we focus on the relaxed proximal point algorithm (PPA) proposed by Gol'shtein and Tret'yakov in [8], which combines the PPA step (3a) with a relaxation step (3b) as follows:

$$\widetilde{w}^k \in \Omega, \quad \left(w - \widetilde{w}^k\right)^T \left[F\left(\widetilde{w}^k\right) + G\left(\widetilde{w}^k - w^k\right)\right] \geq 0, \quad \forall w \in \Omega, \tag{3a}$$

$$w^{k+1} := w^k - \gamma\left(w^k - \widetilde{w}^k\right), \tag{3b}$$

where $\gamma \in (0, 2)$ is a relaxation factor and G is a symmetric positive semidefinite matrix. In particular, γ is called an under-relaxation factor when $\gamma \in (0, 1)$ or an over-relaxation factor when $\gamma \in (1, 2)$, and the relaxed PPA reduces to the original PPA (2) when $\gamma = 1$ and $G = (1/\beta)I$. For convenience, we still use the notation $\|w\|_G^2$ to represent the nonnegative number $w^T G w$ in our analysis.

The Douglas-Rachford alternating direction methods of multipliers (ADMM) scheme proposed by Glowinski and Marrocco in [9] (see also [10]) is a commonplace tool to solve the convex minimization problem with linear constraints and a separable objective function as follows:

$$\min\left\{\theta_1(x) + \theta_2(y) \mid Ax + By = b,\ x \in \mathcal{X},\ y \in \mathcal{Y}\right\}, \tag{4}$$

where $A \in \Re^{m\times n_1}$, $B \in \Re^{m\times n_2}$, $b \in \Re^m$, $\mathcal{X} \subseteq \Re^{n_1}$, and $\mathcal{Y} \subseteq \Re^{n_2}$ are closed convex sets and θ_1: $\Re^{n_1} \to \Re$ and

θ_2: $\mathfrak{R}^{n_2} \rightarrow \mathfrak{R}$ are convex smooth functions. The iterative scheme of ADMM for solving (4) at the k-th iteration runs as

$$\begin{aligned} & x^{k+1} \in \mathscr{X}, \\ & \left(x - x^{k+1}\right)^T \left\{\nabla\theta_1\left(x^{k+1}\right) \right. \\ & \qquad \left. - A^T\left[\lambda^k - H\left(Ax^{k+1} + By^k - b\right)\right]\right\} \geq 0, \\ & \qquad \forall x \in \mathscr{X}, \end{aligned} \tag{5a}$$

$$\begin{aligned} & y^{k+1} \in \mathscr{Y}, \\ & \left(y - y^{k+1}\right)^T \left\{\nabla\theta_2\left(y^{k+1}\right) \right. \\ & \qquad \left. - B^T\left[\lambda^k - H\left(Ax^{k+1} + By^{k+1} - b\right)\right]\right\} \geq 0, \\ & \qquad \forall y \in \mathscr{Y}, \end{aligned} \tag{5b}$$

$$\lambda^{k+1} := \lambda^k - H\left(Ax^{k+1} + By^{k+1} - b\right), \tag{5c}$$

where $H := hI$ and h is a positive constant. As shown in [11], ADMM can be regarded as an application of the relaxed PPA with $\gamma = 1$ (i.e., the original PPA (2)) and

$$G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B^T HB & -B^T \\ 0 & -B & H^{-1} \end{pmatrix}. \tag{6}$$

Without further assumption on B, the matrix G defined previously can be guaranteed as a symmetric and positive semidefinite matrix. Recently, He and Yuan in [12] have shown a worst-case $O(1/t)$ convergence rate of the ADMM scheme (5a), (5b), and (5c) in an ergodic sense. You et al. in [13] have proved the same convergence rate of the Lagrangian PPA-based contraction methods with nonsymmetric linear proximal term in an ergodic sense. The purpose of this paper is to establish the $O(1/t)$ convergence rate of the relaxed PPA (3a) and (3b) in both ergodic and nonergodic senses.

2. Preliminaries

In this section, we review some preliminaries which are useful for further discussions. More specially, we recall a useful characterization on Ω^*, the variational reformulation of (4), the relationship of the ADMM in [9, 10], and the relaxed PPA in [8] for solving this variational reformulation.

First, we provide a useful characterization on Ω^* as Theorem 2.3.5 in [14] and Theorem 2.1 in [12].

Theorem 1. *The solution set of* $\mathrm{VI}(\Omega, F)$ *is convex, and it can be characterized as*

$$\Omega^* = \bigcap_{w\in\Omega} \left\{\widetilde{w} \in \Omega : (w - \widetilde{w})^T F(w) \geq 0\right\}. \tag{7}$$

Based on Theorem 1, $\widetilde{w} \in \Omega$ can be regarded as an ε-approximation solution of $\mathrm{VI}(\Omega, F)$ if it satisfies

$$\sup_{w\in\mathscr{D}} \left\{(\widetilde{w} - w)^T F(w)\right\} \leq \varepsilon, \tag{8}$$

where $\mathscr{D} \subseteq \Omega$ is some compact set. As Definition 1 in [15], we can take

$$\mathscr{D} = \mathscr{B}_\Omega(\widetilde{w}) := \{w \in \Omega \mid \|w - \widetilde{w}\| \leq 1\}. \tag{9}$$

In the following, we will give a variational reformulation of (4). It is easy to see that the model (4) can be characterized by a variational inequality problem: find $w^* = (x^*, y^*, \lambda^*) \in \Omega := \mathscr{X} \times \mathscr{Y} \times \mathfrak{R}^m$ such that

$$\mathrm{VI}(\Omega, F) : \left(w - w^*\right)^T F\left(w^*\right) \geq 0, \quad \forall w \in \Omega, \tag{10a}$$

where

$$w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad F(w) = \begin{pmatrix} \nabla\theta_1(x) - A^T\lambda \\ \nabla\theta_2(y) - B^T\lambda \\ Ax + By - b \end{pmatrix}. \tag{10b}$$

Note that the mapping F is monotone since θ_1 and θ_2 are convex. As shown in [11], the ADMM scheme (5a), (5b), and (5c) is identical with the following iterative scheme in a cyclical sense:

$$\begin{aligned} & \tilde{x}^k \in \mathscr{X}, \quad \left(x - \tilde{x}^k\right)^T \left\{\nabla\theta_1\left(\tilde{x}^k\right) \right. \\ & \qquad \left. - A^T\left[\lambda^k - H\left(A\tilde{x}^k + By^k - b\right)\right]\right\} \geq 0, \\ & \qquad \forall x \in \mathscr{X}, \end{aligned} \tag{11a}$$

$$\tilde{\lambda}^k := \lambda^k - H\left(A\tilde{x}^k + By^k - b\right), \tag{11b}$$

$$\begin{aligned} & \tilde{y}^k \in \mathscr{Y}, \quad \left(y - \tilde{y}^k\right)^T \left\{\nabla\theta_2\left(\tilde{y}^k\right) \right. \\ & \qquad \left. - B^T\left[\tilde{\lambda}^k - H\left(A\tilde{x}^k + B\tilde{y}^k - b\right)\right]\right\} \geq 0, \\ & \qquad \forall y \in \mathscr{Y}, \end{aligned} \tag{11c}$$

$$w^{k+1} = w^k - \left(w^k - \widetilde{w}^k\right). \tag{12}$$

Based on the definition (6) of the matrix G, we can rewrite (11a), (11b), (11c), and (12) as a special case of the relaxed PPA with $\gamma = 1$ immediately.

Lemma 2. *For given w^k, let $\widetilde{w}^k$ be generated by the ADMM scheme* (11a), (11b), *and* (11c). *Then, one has*

$$\begin{aligned} & \widetilde{w}^k \in \Omega, \quad \left(w - \widetilde{w}^k\right)^T \left\{F\left(\widetilde{w}^k\right) + G\left(\widetilde{w}^k - w^k\right)\right\} \geq 0, \\ & \qquad \forall w \in \Omega, \end{aligned} \tag{13}$$

where F and G are defined by (10b) *and* (6), *respectively.*

3. The Contraction of the Relaxed Proximal Point Algorithm

In this section, we prove the contraction of the relaxed PPA. First, we give an important lemma.

Lemma 3. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b)*, and let G be a symmetric positive semidefinite matrix. Then, one has*

$$\begin{aligned}\left(w-\widetilde{w}^k\right)^T F\left(\widetilde{w}^k\right) &\geq \frac{1}{2\gamma}\left(\left\|w-w^{k+1}\right\|_G^2-\left\|w-w^k\right\|_G^2\right)\\ &\quad+\left(1-\frac{\gamma}{2}\right)\left\|w^k-\widetilde{w}^k\right\|_G^2, \quad \forall w\in\Omega.\end{aligned} \tag{14}$$

Proof. First, using (3a), we have

$$\left(w-\widetilde{w}^k\right)^T F\left(\widetilde{w}^k\right)\geq\left(w-\widetilde{w}^k\right)^T G\left(w^k-\widetilde{w}^k\right), \quad \forall w\in\Omega. \tag{15}$$

Since $w^k-\widetilde{w}^k=(w^k-w^{k+1})/\gamma$ (see (3b)), we have

$$\left(w-\widetilde{w}^k\right)^T G\left(w^k-\widetilde{w}^k\right)=\frac{1}{\gamma}\left(w-\widetilde{w}^k\right)^T G\left(w^k-w^{k+1}\right). \tag{16}$$

Thus, it suffices to show that

$$\begin{aligned}\left(w-\widetilde{w}^k\right)^T G\left(w^k-w^{k+1}\right) &= \frac{1}{2}\left(\left\|w-w^{k+1}\right\|_G^2-\left\|w-w^k\right\|_G^2\right)\\ &\quad+\gamma\left(1-\frac{\gamma}{2}\right)\left\|w^k-\widetilde{w}^k\right\|_G^2.\end{aligned} \tag{17}$$

By setting $a=w$, $b=\widetilde{w}^k$, $c=w^k$, and $d=w^{k+1}$ in the identity

$$\begin{aligned}(a-b)^T G(c-d) &= \frac{1}{2}\left(\|a-d\|_G^2-\|a-c\|_G^2\right)\\ &\quad+\frac{1}{2}\left(\|c-b\|_G^2-\|d-b\|_G^2\right),\end{aligned} \tag{18}$$

we derive that

$$\begin{aligned}\left(w-\widetilde{w}^k\right)^T G\left(w^k-w^{k+1}\right) &= \frac{1}{2}\left(\left\|w-w^{k+1}\right\|_G^2-\left\|w-w^k\right\|_G^2\right)\\ &\quad+\frac{1}{2}\left(\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|w^{k+1}-\widetilde{w}^k\right\|_G^2\right).\end{aligned} \tag{19}$$

On the other hand, using (3b), we have

$$\begin{aligned}&\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|w^{k+1}-\widetilde{w}^k\right\|_G^2\\ &\quad=\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|\left(w^k-\widetilde{w}^k\right)-\left(w^k-w^{k+1}\right)\right\|_G^2\\ &\quad=\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|\left(w^k-\widetilde{w}^k\right)-\gamma\left(w^k-\widetilde{w}^k\right)\right\|_G^2\\ &\quad=\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2.\end{aligned} \tag{20}$$

Combining the last two equations, we obtain (17). The assertion (14) follows immediately. The proof is completed. □

With the proved lemma, we are now ready to show the contraction of the relaxed PPA (3a) and (3b).

Theorem 4. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b)*, and let G be a symmetric positive semidefinite matrix. Then, for any $k\geq 0$, one has*

$$\begin{aligned}&\left\|w^{k+1}-w^*\right\|_G^2\\ &\quad\leq\left\|w^k-w^*\right\|_G^2-\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2, \quad \forall w^*\in\Omega^*.\end{aligned} \tag{21}$$

Proof. Setting $w=w^*$ in (14), we get

$$\begin{aligned}2\gamma\left(w^*-\widetilde{w}^k\right)^T F\left(\widetilde{w}^k\right) &\geq\left\|w^*-w^{k+1}\right\|_G^2-\left\|w^*-w^k\right\|_G^2\\ &\quad+\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2.\end{aligned} \tag{22}$$

On the other hand, since F is monotone and $w^*\in\Omega^*$, we have

$$0\geq\left(w^*-\widetilde{w}^k\right)^T F\left(w^*\right)\geq\left(w^*-\widetilde{w}^k\right)^T F\left(\widetilde{w}^k\right). \tag{23}$$

It follows from the previous two inequalities that

$$\left\|w^{k+1}-w^*\right\|_G^2\leq\left\|w^k-w^*\right\|_G^2-\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2. \tag{24}$$

The proof is completed. □

4. Ergodic Worst-Case $O(1/t)$ Convergence Rate

In this section, we will establish an ergodic worst-case $O(1/t)$ convergence rate for the relaxed PPA in the sense that after t iterations of such an algorithm, we can find $\widetilde{w}\in\Omega$ such that

$$(\widetilde{w}-w)^T F(w)\leq\varepsilon, \quad \forall w\in\mathscr{B}_\Omega(\widetilde{w}), \tag{25}$$

with $\varepsilon=O(1/t)$ and $\mathscr{B}_\Omega(\widetilde{w}):=\{w\in\Omega\mid\|w-\widetilde{w}\|_G\leq 1\}$.

Theorem 5. *Let $\{w^k\}$ and $\{\widetilde{w}^k\}$ be the sequences generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. For any integer number $t > 0$, let*

$$\widetilde{w}_t := \frac{1}{t+1}\sum_{k=0}^{t}\widetilde{w}^k. \tag{26}$$

Then, one has $\widetilde{w}_t \in \Omega$ and

$$\left(\widetilde{w}_t - w\right)^T F(w) \le \frac{1}{2\gamma(t+1)}\left\|w^0 - w\right\|_G^2, \quad \forall w \in \Omega. \tag{27}$$

Proof. From (14), we have

$$\begin{aligned}&\left(w - \widetilde{w}^k\right)^T F\left(\widetilde{w}^k\right) + \frac{1}{2\gamma}\left\|w^k - w\right\|_G^2 \\ &\quad \ge \frac{1}{2\gamma}\left\|w^{k+1} - w\right\|_G^2, \quad \forall w \in \Omega.\end{aligned} \tag{28}$$

Since F is monotone, from the previous inequality, we have

$$\begin{aligned}&\left(w - \widetilde{w}^k\right)^T F(w) + \frac{1}{2\gamma}\left\|w^k - w\right\|_G^2 \\ &\quad \ge \frac{1}{2\gamma}\left\|w^{k+1} - w\right\|_G^2, \quad \forall w \in \Omega.\end{aligned} \tag{29}$$

Summing the inequality (29) over $k = 0, 1, \ldots, t$, we obtain

$$\begin{aligned}&\left[(t+1)w - \left(\sum_{k=0}^{t}\widetilde{w}^k\right)\right]^T F(w) + \frac{1}{2\gamma}\left\|w^0 - w\right\|_G^2 \\ &\quad \ge \frac{1}{2\gamma}\left\|w^{t+1} - w\right\|_G^2 \ge 0, \quad \forall w \in \Omega.\end{aligned} \tag{30}$$

Since $\sum_{k=0}^{t} 1/(t+1) = 1$, $\widetilde{w}_t$ is a convex combination of $\widetilde{w}^0, \widetilde{w}^1, \ldots, \widetilde{w}^t$ and thus $\widetilde{w}_t \in \Omega$. Using the notation of $\widetilde{w}_t$, we derive

$$\left(w - \widetilde{w}_t\right)^T F(w) + \frac{1}{2\gamma(t+1)}\left\|w^0 - w\right\|_G^2 \ge 0, \quad \forall w \in \Omega. \tag{31}$$

The assertion (27) follows from the previous inequality immediately. □

It follows from Theorem 4 that the sequence $\{\|w^k\|_G\}$ is bounded. According to (21), the sequence $\{\|\widetilde{w}^k\|_G\}$ is also bounded. Therefore, there exists a constant $D > 0$ such that

$$\left\|w^k\right\|_G \le D, \qquad \left\|\widetilde{w}^k\right\|_G \le D, \quad \forall k \ge 0. \tag{32}$$

Recall that $\widetilde{w}_t$ is the average of $\{\widetilde{w}^0, \widetilde{w}^1, \ldots, \widetilde{w}^t\}$. Thus, we have $\|\widetilde{w}_t\|_G \le D$. For any $w \in \mathscr{B}_\Omega(\widetilde{w}_t) := \{w \in \Omega \mid \|w - \widetilde{w}_t\|_G \le 1\}$, we get

$$\begin{aligned}&\left(\widetilde{w}_t - w\right)^T F(w) \\ &\quad \le \frac{1}{2\gamma(t+1)}\left\|w^0 - w\right\|_G^2 \\ &\quad \le \frac{1}{2\gamma(t+1)}\left(\left\|w^0 - \widetilde{w}_t\right\|_G + \left\|\widetilde{w}_t - w\right\|_G\right)^2 \\ &\quad \le \frac{1}{2\gamma(t+1)}\left(\left\|w^0\right\|_G + \left\|\widetilde{w}_t\right\|_G + \left\|\widetilde{w}_t - w\right\|_G\right)^2 \\ &\quad \le \frac{(2D+1)^2}{2\gamma(t+1)}.\end{aligned} \tag{33}$$

Thus, for any given $\varepsilon > 0$, after at most $t := \lceil((2D+1)^2/2\gamma\varepsilon) - 1\rceil$ iterations, we have

$$\left(\widetilde{w}_t - w\right)^T F(w) \le \varepsilon, \quad \forall w \in \mathscr{B}_\Omega\left(\widetilde{w}_t\right), \tag{34}$$

which means that $\widetilde{w}_t$ is an approximate solution of $\mathrm{VI}(\Omega, F)$ with an accuracy of $O(1/t)$. That is, a worst-case $O(1/t)$ convergence rate of the relaxed PPA in an ergodic sense is established.

Note that this convergence rate is in an ergodic sense and $\widetilde{w}_t$ is a convex combination of the previous vectors $\{\widetilde{w}^0, \widetilde{w}^1, \ldots, \widetilde{w}^t\}$ with equal weights. One may ask if we can establish the same convergence rate in a nonergodic sense directly for the sequence $\{w^k\}$ generated by the relaxed PPA (3a) and (3b), and this is the main purpose of the next section.

5. Nonergodic Worst-Case $O(1/t)$ Convergence Rate

This section shows that the relaxed PPA has a worst-case $O(1/t)$ convergence rate in a nonergodic sense. First, we establish two important inequalities in the following lemmas.

Lemma 6. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. Then, one has*

$$\left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)^T G\left[\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right] \ge 0. \tag{35}$$

Proof. Setting $w = \widetilde{w}^{k+1}$ in (3a), we have

$$\left(\widetilde{w}^{k+1} - \widetilde{w}^k\right)^T\left[F\left(\widetilde{w}^k\right) + G\left(\widetilde{w}^k - w^k\right)\right] \ge 0. \tag{36}$$

Note that (3a) is also true for $k := k+1$, and thus we have

$$\begin{aligned}&\left(w - \widetilde{w}^{k+1}\right)^T\left[F\left(\widetilde{w}^{k+1}\right) + G\left(\widetilde{w}^{k+1} - w^{k+1}\right)\right] \ge 0, \\ &\qquad \forall w \in \Omega.\end{aligned} \tag{37}$$

Setting $w = \widetilde{w}^k$ in the previous inequality, we obtain

$$\left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)^T \left[F\left(\widetilde{w}^{k+1}\right) + G\left(\widetilde{w}^{k+1} - w^{k+1}\right)\right] \geq 0. \quad (38)$$

Adding (36) and (38) and using the monotonicity of F, we get (35) immediately. □

Lemma 7. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. Then, one has*

$$\begin{aligned}&\left(w^k - \widetilde{w}^k\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\}\\ &\quad \geq \frac{1}{\gamma}\left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (39)$$

Proof. First, adding the term

$$\begin{aligned}&\left\{\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\}^T\\ &\quad \times G\left\{\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\}\end{aligned} \quad (40)$$

to the both sides of (35), we get

$$\begin{aligned}&\left(w^k - w^{k+1}\right)^T G\left\{\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\}\\ &\quad \geq \left\|\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (41)$$

Reordering $(w^k - w^{k+1}) - (\widetilde{w}^k - \widetilde{w}^{k+1})$ in the previous inequality to $(w^k - \widetilde{w}^k) - (w^{k+1} - \widetilde{w}^{k+1})$, we get

$$\begin{aligned}&\left(w^k - w^{k+1}\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\}\\ &\quad \geq \left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (42)$$

Substituting the term $w^k - w^{k+1} = \gamma(w^k - \widetilde{w}^k)$ (see (3b)) into the left-hand side of the last inequality, we obtain (39). The proof is completed. □

Next, we prove that $\{\|w^k - \widetilde{w}^k\|_G\}$ is monotonically nonincreasing.

Theorem 8. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. Then, one has*

$$\left\|w^{k+1} - \widetilde{w}^{k+1}\right\|_G \leq \left\|w^k - \widetilde{w}^k\right\|_G, \quad \forall k \geq 0. \quad (43)$$

Proof. Setting $a = w^k - \widetilde{w}^k$ and $b = w^{k+1} - \widetilde{w}^{k+1}$ in the identity

$$\|a\|_G^2 - \|b\|_G^2 = 2a^T G(a - b) - \|a - b\|_G^2, \quad (44)$$

we obtain

$$\begin{aligned}&\left\|w^k - \widetilde{w}^k\right\|_G^2 - \left\|w^{k+1} - \widetilde{w}^{k+1}\right\|_G^2\\ &\quad = 2\left(w^k - \widetilde{w}^k\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\}\\ &\qquad - \left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (45)$$

Inserting (39) into the first term of the right-hand side of the last equality and using $\gamma \in (0, 2)$, we obtain

$$\begin{aligned}&2\left(w^k - \widetilde{w}^k\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\}\\ &\qquad - \left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2\\ &\quad \geq \frac{2-\gamma}{\gamma}\left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2\\ &\quad \geq 0.\end{aligned} \quad (46)$$

The assertion (43) follows immediately. □

With Theorems 4 and 8, we can prove the worst-case $O(1/t)$ convergence rate in a nonergodic sense for the relaxed PPA.

Theorem 9. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. Then, for any integer $t \geq 0$, one has*

$$\left\|w^t - \widetilde{w}^t\right\|_G^2 \leq \frac{1}{\gamma(2-\gamma)(t+1)}\left\|w^0 - w^*\right\|_G^2, \quad \forall w^* \in \Omega^*. \quad (47)$$

Proof. Summing the inequality (21) over $k = 0, 1, \ldots, t$, we obtain

$$\begin{aligned}&\gamma(2-\gamma)\sum_{k=0}^{t}\left\|w^k - \widetilde{w}^k\right\|_G^2\\ &\quad \leq \left\|w^0 - w^*\right\|_G^2 - \left\|w^{t+1} - w^*\right\|_G^2\\ &\quad \leq \left\|w^0 - w^*\right\|_G^2, \quad \forall w^* \in \Omega^*.\end{aligned} \quad (48)$$

According to Theorem 8, the sequence $\{\|w^k - \widetilde{w}^k\|_G\}$ is monotonically nonincreasing. Therefore, we have

$$(t+1)\left\|w^t - \widetilde{w}^t\right\|_G^2 \leq \sum_{k=0}^{t}\left\|w^k - \widetilde{w}^k\right\|_G^2. \quad (49)$$

The assertion (47) follows from (48) and (49) immediately. □

Note that Ω^* is convex and closed (see Theorem 1). Let $d := \inf\{\|w^0 - w^*\|_G \mid w^* \in \Omega^*\}$. Then, for any given $\varepsilon > 0$, Theorem 9 shows that the relaxed PPA (3a) and (3b) needs at most $\lceil d^2/(\varepsilon\gamma(2-\gamma)) - 1\rceil$ iterations to ensure that $\|w^t - \widetilde{w}^t\|_G^2 \leq \varepsilon$. Recall that $\widetilde{w}^t$ is a solution of VI(Ω, F) if $\|w^t - \widetilde{w}^t\|_G^2 = 0$. In other words, if $\|w^t - \widetilde{w}^t\|_G^2 = 0$, we have $G(w^t - \widetilde{w}^t) = 0$ since G is a positive semidefinite matrix. And thus from (3a), it follows that

$$\left(w - \widetilde{w}^t\right)^T F\left(\widetilde{w}^t\right) \geq 0, \quad \forall w \in \Omega, \quad (50)$$

which means that $\widetilde{w}^t$ is a solution of VI(Ω, F) according to (1). A worst-case $O(1/t)$ convergence rate in a nonergodic sense for the relaxed PPA (3a) and (3b) is thus established from Theorem 9.

6. Concluding Remarks

This paper established the worst-case $O(1/t)$ convergence rate in both ergodic and nonergodic senses for the relaxed PPA. Recall that ADMM is a primal application of the relaxed PPA with $\gamma = 1$. And thus ADMM also has the same worst-case $O(1/t)$ convergence rate in both ergodic and nonergodic senses.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 11001053), Program for New Century Excellent Talents in University (Grant no. NCET-12-0111), and Natural Science Foundation of Jiangsu Province, China (Grant no. BK2012662).

References

[1] S. Dafermos and A. Nagurney, "Supply and demand equilibration algorithms for a class of market equilibrium problems," *Transportation Science*, vol. 23, no. 2, pp. 118–124, 1989.

[2] M. C. Ferris and J. S. Pang, "Engineering and economic applications of complementarity problems," *SIAM Review*, vol. 39, no. 4, pp. 669–713, 1997.

[3] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, vol. 88, Academic Press, New York, NY, USA, 1980.

[4] A. Nagurney, *Network Economics: A Variational Inequality Approach*, vol. 1, Kluwer Academic, Dordrecht, The Netherlands, 1993.

[5] B. Martinet, "Régularisation D'inéquations Variationnelles par Approximations Successives," *Revue Française dInformatique et de Recherche Opérationelle*, vol. 4, pp. 154–158, 1970.

[6] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[7] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 97–116, 1976.

[8] E. G. Gol'shtein and N. V. Tret'yakov, "Modified Lagrangians in convex programming and their generalizations," *Mathematical Programming Study*, no. 10, pp. 86–97, 1979.

[9] R. Glowinski and 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, no. R-2, pp. 41–76, 1975.

[10] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite-element approximations," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[11] X. J. Cai, G. Y. Gu, B. S. He, and X. M. Yuan, *A Proximal Point Algorithm Revisit on the Alternating Direction Method of Multipliers*, Science China Mathematics.

[12] 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.

[13] Y. F. You, X. L. Fu, and B. S. He, *Lagrangian PPA-Based Contraction Methods for Linearly Constrained Convex Optimization*, manuscript, 2013.

[14] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, Springer Series in Operations Research*, Springer, New York, NY, USA, 2003.

[15] Yu. Nesterov, "Gradient methods for minimizing composite functions," *Mathematical Programming*, 2012.