Article

Alternating Inertial and Overrelaxed Algorithms for Distributed Generalized Nash Equilibrium Seeking in Multi-Player Games

Zhangcheng Feng [1], Wenying Xu [1] and Jinde Cao [1,2,*]

1 Jiangsu Provincial Key Laboratory of Networked Collective Intelligence, School of Mathematics, Southeast University, Nanjing 211189, China; 213172860@seu.edu.cn (Z.F.); wyxu@seu.edu.cn (W.X.)
2 Yonsei Frontier Lab, Yonsei University, Seoul 03722, Korea
* Correspondence: jdcao@seu.edu.cn; Tel.: +86-025-52090588

Abstract: This paper investigates the distributed computation issue of generalized Nash equilibrium (GNE) in a multi-player game with shared coupling constraints. Two kinds of relatively fast distributed algorithms are constructed with alternating inertia and overrelaxation in the partial-decision information setting. We prove their convergence to GNE with fixed step-sizes by resorting to the operator splitting technique under the assumptions of Lipschitz continuity of the extended pseudo-gradient mappings. Finally, one numerical simulation is given to illustrate the efficiency and performance of the algorithm.

Keywords: generalized Nash equilibrium; distributed algorithm; partial decision; networked games

1. Introduction

Game theory is the study of mathematical models for describing competition and cooperation interaction among intelligent rational decision-makers [1]. In the past few years, networked games have received increasing attention due to their wide applications in different areas such as competitive economy [2], power allocation in interference channel models [3,4], environmental pollution control [5], cloud computing [6], wireless communication [7–9], and adversarial classification [10,11].

The Nash equilibrium (NE) is a set of strategies where each player's choice is its best response to the choices of the other players of the game [12]. An NE in games with shared coupling constraints is referred to as generalized Nash equilibrium (GNE) [13]. In order to compute the GNE, a great number of algorithms have been proposed [14–18], most of which depend on *full-decision* information, i.e., each player is assumed to have full access to all of the other players' actions. However, such an assumption could be impractical in large-scale distributed networks [19,20]. To overcome this shortcoming, fully distributed algorithms under the *partial-decision* information setting have recently become a research topic that attracts recurring interest.

Under the partial-decision information setting, each player can communicate only with its neighbors (instead of all its opponents) via a certain communication graph. In this case, the player has no direct access to some necessary decision information involving its cost function. In order to make up for the missing information, the player estimates other players' actions and exchanges its estimates with neighbors. Such an estimate would tend to be the real actions of players by designing an appropriate consensus protocol [21]. So far, some efforts have been devoted to the GNE seeking problem with partial-decision information. For example, an adaption of a fictitious play algorithm for large-scale games is introduced in [22], and information exchange techniques for aggregative games are studied in [23]. An operator-theoretic approach has been introduced to analyze GNE problems [16,21], under which the problem is cast as finding a zero of a sum of monotone

operators through primal-dual analysis and show its convergence by reformulating it as a forward–backward fixed-point iteration.

Compared with the existing distributed algorithms for diminishing steps [24], the algorithm for fixed steps has the potential to exhibit a faster convergence [16]. Very recently, some distributed proximal algorithms and project-gradient algorithms have been proposed for seeking the GNE with fixed steps [16,25–28]. It is worth noting that most of the existing algorithms, under the partial-decision information setting, require that the extended pseudo-gradient mapping in the augmented space of actions and estimates is strictly/strongly monotone. Such an assumption seems strong and how to relax it becomes a technical difficulty. In this paper, we would like to investigate the GNE seeking algorithm under a mild assumption of the extended pseudo-gradient mapping, like [21].

In addition, some refined GNE seeking algorithms with inertia and relaxation have been proposed in ([16], [Alg. 6.1]), ([29], [Alg. 2]) and ([25], [Alg. 3]) to accelerate the convergence to GNE. Although the fast convergence of the mentioned algorithms has been validated numerically, more computation resources are inevitably required at each iteration. Note that the computation resources could be limited and expensive in many situations. Inspired by the above discussion, in this paper, we combine a projection based algorithm via a doubly augmented operator splitting from the work [21] with the inertia/overrelaxation idea from the paper [25]. Specifically, we design distributed GNE seeking algorithms to balance the convergence rate and computation consumption in games with shared coupling constraints under a partial-decision information setting. Two kinds of fully distributed algorithms, i.e., alternating inertial algorithms and alternating overrelaxed algorithms, are proposed with fixed step-sizes. Their convergence to the GNE are guaranteed under a mild assumption on the extended pseudo-gradient mappings, compared to [26], by using the Karush–Kuhn–Tucker (KKT) conditions of an optimization problem and variational inequality. Finally, a numerical example is provided to show the effectiveness of our algorithms that are validated numerically to have a relatively fast convergence rate.

The remainder of the paper is organized as follows. In Section 2, we introduce some notations and backgroud theory. Section 3 describes the problem that we are interested in, formulates it into mathematical model, and rewrites the game into a problem of finding the solution of the stochastic variational inequality (SVI). In Section 4, we propose two alternating fully distributed GNE seeking algorithms under a partial-decision information setting and assumptions to guarantee convergence; the convergence analysis is also presented in this section. We present numerical results in Section 5 and finally conclude in Section 6.

Notations: Let $\mathbf{R}^m(\mathbf{R}^m_+)$ represent an m-dimensional (non-negative) Euclidean space. $\mathbf{0}_n \in \mathbf{R}^n$ is an $n-$dimensional vector with all elements equal to 0, and $I_m \in \mathbf{R}^{m \times m}$ is the identity matrix with $m \times m$ dimension. $\mathbf{1}_N$ denotes the N-dimension column vector with all elements equal to 1. We denote $\Omega_1 \times \cdots \times \Omega_n$ or $\prod_{i=1}^n \Omega_i$ as the Cartesian product of the sets $\Omega_i, i = 1, \cdots, n$. For given n column vectors $x_1, \cdots, x_n$, $\mathrm{col}(x_1, \cdots, x_n) = [x_1^\top, \cdots, x_n^\top]^\top$. Let $[x]_k$ denote the k-th element in column vector x, let $\langle x, y \rangle = x^\top y$ denote the inner product of x, y, and $\|x\| = \sqrt{x^\top x}$ denotes the norm induced by the inner product $\langle \cdot, \cdot \rangle$. $\Phi \succ 0$ stands for a symmetric positive definite matrix. Similarly, the Φ-induced product is $\langle x, y \rangle_\Phi = \langle \Phi x, y \rangle$, and the Φ-induced norm is $\|x\|_\Phi = \sqrt{\langle \Phi x, x \rangle}$. $\otimes$ is the Kronecker product, and $\mathrm{diag}(A_1, \cdots, A_n)$ denotes the block diagonal matrix with $A_1, \cdots, A_N$ on its diagonal. Suppose $A \in \mathbf{R}^{m \times n}$, then $\|A\|_\infty = \max\{\sum_{k=1}^n |[A_i]_{1k}|, \cdots, \sum_{k=1}^n |[A_i]_{mk}|\}$, where $[A_i]_{jk}$ denotes the element of A_i in the j-th row and k-th column.

2. Preliminary

2.1. Operator Theory

The following concepts are reviewed from [30]. Let $\mathfrak{A} : \mathbf{R}^m \to 2^{\mathbf{R}^m}$ be a set-valued operator. Denote Id as the identity operator, i.e., $Id(x) = x$. The graph of $\mathfrak{A}$ is $gra\mathfrak{A} = \{(x, u) \in \mathbf{R}^m \times \mathbf{R}^m | u \in \mathfrak{A}x\}$. The zero set of operator $\mathfrak{A}$ is $zer\mathfrak{A} = \{x \in \mathbf{R}^m | \mathbf{0} \in \mathfrak{A}x\}$. Define the resolvent of operator $\mathfrak{A}$ as $R_{\mathfrak{A}} = (Id + \mathfrak{A})^{-1}$. An operator $\mathfrak{A}$ is called monotone if $\forall(x, u), \forall(y, v) \in gra\mathfrak{A}$, we have $\langle x - y, u - v \rangle \geq 0$. Moreover, it is maximally monotone

if $gra\mathfrak{A}$ is not strictly contained in the graph of any other monotone operator, i.e., for every $(x, u) \in \mathbf{R}^m \times \mathbf{R}^m, (x, u) \in gra\mathfrak{A} \Leftrightarrow \forall (y, v) \in gra\mathfrak{A}, \langle x - y, u - v \rangle \geq 0$. $\mathfrak{A}$ is nonexpansive if it is Lipschitz continuous with constant 1, i.e., $\forall x, y \in \mathbf{R}^m, \|\mathfrak{A}x - \mathfrak{A}y\| \leq \|x - y\|$, and is firmly nonexpansive if $\|\mathfrak{A}x - \mathfrak{A}y\|^2 + \|(Id - \mathfrak{A})x - (Id - \mathfrak{A})y\|^2 \leq \|x - y\|^2$. The operator $\mathfrak{A}$ is α-averaged with the constant $\alpha \in [0, 1]$, denoted by $\mathfrak{A} \in \mathcal{A}(\alpha)$, if $\forall x, y \in \mathbf{R}^m, \|\mathfrak{A}x - \mathfrak{A}y\|^2 \leq \|x - y\|^2 - (1 - \alpha)/\alpha \|(Id - \mathfrak{A})x - (Id - \mathfrak{A})y\|^2$. We can easily derive that if $\mathfrak{A}$ is averaged then it is nonexpansive, and $\mathfrak{A}$ is firmly nonexpansive if and only if it is $1/2$-averaged. $\mathfrak{A}$ is β-cocoercive for $\beta > 0$, if $\forall x, y \in \mathbf{R}^m, \beta \|\mathfrak{A}x - \mathfrak{A}y\|^2 \leq \langle x - y, \mathfrak{A}x - \mathfrak{A}y \rangle$. The normal cone operator of the set Ω is defined as

$$N_\Omega(x) = \begin{cases} \varnothing & x \notin \Omega \\ \{v | \langle v, y - x \rangle \leq 0, \forall y \in \Omega\} & x \in bd(\Omega) \\ \mathbf{0} & x \in int(\Omega). \end{cases} \quad (1)$$

Let the projection of x onto Ω be $P_\Omega(x) = \arg\min_{y \in \Omega} \|x - y\|$, and $P_\Omega(x) = R_{N_\Omega}(x) = (Id + N_\Omega)^{-1}(x)$.

2.2. Graph Theory

Let the graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ describe the information exchanged among agents, where $\mathcal{N} := \{1, \cdots, N\}$ is the set of players and $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ is the edge set. If player i can obtain information from player j, then $(i, j) \in \mathcal{E}$ and j belong to player i's neighbor set $\mathcal{N}_i := \{j | (i, j) \in \mathcal{E}\}$. $\mathcal{G}$ is said to be undirected when $(i, j) \in \mathcal{E}$ if and only if $(j, i) \in \mathcal{E}$. Let $W := [w_{ij}] \in \mathbf{R}^{N \times N}$ be the weighted adjacency matrix of $\mathcal{G}$ with $w_{ij} > 0$ if $j \in \mathcal{N}_i$ and $w_{ij} = 0$ otherwise. Assume that $W = W^\top$. The degree matrix is defined as $Deg := \mathrm{diag}[d_1, \cdots, d_N] = \mathrm{diag}[\sum_{j=1}^N w_{1j}, \cdots, \sum_{j=1}^N w_{Nj}]$, and the weighted Laplacian of graph $\mathcal{G}$ is $L := Deg - W$. If $\mathcal{G}$ is connected and undirected, then 0 is an eigenvalue of L, and the eigenvalues of L are $0 < s_2(L) \leq \cdots \leq s_N(L)$ in ascending order.

3. Game Formulation

In this section, we build a mathematical setup about the problem considered.

Consider a set of players $\mathcal{N} = \{1, \cdots, N\}$, where every player $i \in \mathcal{N}$ controls its local decision variable $x_i \in \Omega_i \subseteq \mathbf{R}^{n_i}$ and Ω_i is the private decision set of player i. Denote $n := \sum_{i=1}^N n_i$ and $\Omega := \Omega_1 \times \cdots \times \Omega_N \in \mathbf{R}^n$, then the stacked vector of all the players' decisions $x := \mathrm{col}(x_i)_{i \in \mathcal{N}} \in \mathbf{R}^n$ is called the decision profile. We also write $x = (x_i, x_{-i})$, where $x_{-i} := \mathrm{col}(x_j)_{j \in \mathcal{N}/\{i\}} = \mathrm{col}(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N)$ denotes all of the decisions except player i's.

The local objective function of each player $i \in \mathcal{N}$ is denoted by $J_i(x_i, x_{-i})$, and the affine coupling constrained set is defined as

$$K := \prod_{i=1}^N \Omega_i \cap \{x \in \mathbf{R}^n | Ax \leq b\} \quad (2)$$

where $A := [A_1, \cdots, A_N] \in \mathbf{R}^{m \times n}, A_i \in \mathbf{R}^{m \times n_i}$ and $b := \sum_{i=1}^N b_i \in \mathbf{R}^m$. Here, A_i and b_i are the local data only accessible to player i. Define the feasible set of player i as $K_i(x_{-i}) := \{x_i \in \mathbf{R}^{n_i} | (x_i, x_{-i}) \in K\}$, which implies that the feasible set of each player depends on the action of the other players. Every player aims to optimize its objective function, and the game can be represented by the inter-dependent optimization problems

$$\forall i \in \mathcal{N} : \min_{x_i \in \mathbf{R}^{n_i}} J_i(x_i, x_{-i}) \qquad \text{s.t. } x_i \in K_i(x_{-i}). \quad (3)$$

Definition 1. *A GNE of game* (3) *is a collective strategy* $x^* = \mathrm{col}(x_i^*)_{i \in \mathcal{N}}$ *such that for all* $i \in \mathcal{N}$

$$x_i^* \in \arg\min J_i(x_i, x_{-i}^*) \qquad \text{s.t. } x_i \in K_i(x_{-i}). \quad (4)$$

In order to deal with the coupling constraints and solve the problems, we define the Lagrange function of each player $i \in \mathcal{N}$:

$$\mathcal{L}_i(x_i, \lambda_i; x_{-i}) = J_i(x_i, x_{-i}) + \lambda_i^\top (Ax - b) \tag{5}$$

where $\lambda_i \in \mathbf{R}_+^m$ is a dual variable. According to optimization theory, if x_i^* is an optimal solution to (3), then there exists $\lambda_i^* \in \mathbf{R}^{n_i}$ such that the following KKT conditions are satisfied:

$$\begin{cases} \nabla_{x_i} \mathcal{L}_i(x_i^*, \lambda_i^*; x_{-i}^*) = \mathbf{0}_{n_i} \\ \langle \lambda_i^*, Ax^* - b \rangle = 0 \\ -(Ax^* - b) \geq 0 \\ \lambda_i^* \geq \mathbf{0}. \end{cases} \tag{6}$$

By using the normal cone operator, the KKT conditions (6) are equivalent to

$$\begin{cases} \mathbf{0}_{n_i} \in \nabla_{x_i} J_i(x_i^*, x_{-i}^*) + A_i^\top \lambda_i^* + N_{\Omega_i}(x_i^*) \\ \mathbf{0}_m \in -(Ax^* - b) + N_{\mathbf{R}_+^m}(\lambda_i^*). \end{cases} \tag{7}$$

Note that by the definition of a normal cone, one has $N_{\mathbf{R}_+^m}(\lambda_i^*) = \varnothing$ when $\lambda_i^* \notin \mathbf{R}_+^m$, which implies $\lambda_i^* \in \mathbf{R}_+^m$ (equivalently $[\lambda_i^*]_k \geq 0$) when (7) holds. Furthermore, $N_{\mathbf{R}_+^m} = \prod_{k=1}^m N_{\mathbf{R}_+}$, that is, if $[\lambda_i^*]_k = 0$, then $N_{\mathbf{R}_+}([\lambda_i^*]_k) = -\mathbf{R}_+$, and thus $[Ax^* - b]_k \leq 0$; if $[\lambda_i^*]_k > 0$, then $N_{\mathbf{R}_+}([\lambda_i^*]_k) = 0$, and hence $[Ax^* - b]_k = 0$. This result implies that $Ax^* - b \leq 0$ and $\langle \lambda_i^*, Ax^* - b \rangle = 0$.

We consider the GNE with the same Lagrangian multipliers for every player, i.e., $\lambda_1^* = \lambda_2^* = \cdots = \lambda_N^* = \lambda^*$, which is called *variational* GNE (v-GNE). The v-GNE x^* is a solution of the following inequality $VI(F, K)$:

$$\langle F(x^*), x - x^* \rangle \geq 0, \forall x \in K \tag{8}$$

where F is the *pseudo-gradient* mapping of the game with the following form:

$$F(x) := \text{col}(\nabla_{x_i} J_i(x_i, x_{-i}))_{i \in \mathcal{N}}. \tag{9}$$

Assumption 1. *Given x_{-i}, $J_i(x_i, x_{-i})$ is continuously differentiable and convex in x_i, and Ω_i is nonempty, compact and convex for each player i, then K is nonempty and satisfies Slater's constraint qualification.*

Assumption 2. *F is μ-monotone and θ_0-Lipschitz continuous, i.e., for any point x and x', $\langle x - x', F(x) - F(x') \rangle \geq \mu \|x - x'\|^2$ and $\|F(x) - F(x')\| \leq \theta_0 \|x - x'\|$.*

It follows from ([15], [Theorem 4.8]) that x^* solves $VI(F, K)$ (8) if and only if there exists a $\lambda^* \in \mathbf{R}^m$ such that the KKT conditions are satisfied:

$$\begin{cases} \mathbf{0}_n \in F(x^*) + A^\top \lambda^* + N_\Omega(x^*) \\ \mathbf{0}_m \in -(Ax^* - b) + N_{\mathbf{R}_+^m}(\lambda^*) \end{cases} \tag{10}$$

where $N_\Omega(x^*) = \prod_{i=1}^N N_{\Omega_i}(x_i^*)$.

Assumption 1 guarantees the existence of the v-GNE for game (3) by ([31], [Corollary 2.2.5]). The goal of this paper is to design distributed algorithms for seeking the v-GNE under a *partial-decision* information setting, where both the computational cost and convergence rate are taken into consideration.

4. Alternating Distributed v-GNE Algorithms

In this section, we propose two kinds of distributed algorithms for seeking the v-GNE of game (3) with partial-decision information, where each player controls its own actions and exchanges information with its neighbors via the communication graph.

Remark 1. *Some GNE seeking algorithms with inertia and overrelaxation have been proposed* [28,29]. *Although the fast convergence of these algorithms has been validated numerically, more computation resources are inevitably required at each iteration. Note that the computation resources could be limited and expensive in many situations. Inspired by the above discussion, in this section we design distributed GNE seeking algorithms with alternated inertia and alternated overrelaxation, where both fast convergence rate and low computation consumption are taken into consideration.*

Suppose that player $i \in \mathcal{N}$ controls its local decision $x_i \in \mathbf{R}^{n_i}$ and $\lambda_i \in \mathbf{R}^m_+$ (i.e., the estimation of λ^* in (10)). In order to make up for the lack of non-neighbors' information, we introduce an auxiliary variable $\mathbf{x}_i$ for each player i that provides the estimation of the other players' decisions. To be specific, $\mathbf{x}_i = col(\mathbf{x}_i^j)_{j \in \mathcal{N}}$ where $\mathbf{x}_i^j$ denotes the player i's estimation of player j's decision and $\mathbf{x}_i^i = x_i$. We can also rewrite $\mathbf{x}_i = (x_i, \mathbf{x}_i^{-i})$, where $\mathbf{x}_i^{-i}$ represents player i's estimation vector except its own decisions. In addition, an auxiliary variable $z_i \in \mathbf{R}^m_+$ is introduced for each player $i \in \mathcal{N}$. We assume that each player exchanges its local variable $\{\mathbf{x}_i, \lambda_i, z_i\}$ with its neighbor through the communication graph $\mathcal{G}$.

Assumption 3. *The communication graph $\mathcal{G}$ is undirected and connected.*

4.1. Alternating Inertial Distributed v-GNE Seeking Algorithm

In this subsection, we propose an alternating inertial distributed algorithm for seeking the v-GNE, where the inertia is adopted intermittently (see Algorithm 1). Here, $x_{i,k}, \mathbf{x}_{i,k}^{-i}$ and $z_{i,k}, \lambda_{i,k}$ denote $x_i, \mathbf{x}_i^{-i}, z_i, \lambda_i$ at iteration k, respectively, and $\tilde{x}_{i,k}, \tilde{\mathbf{x}}_{i,k}^{-i}, \tilde{z}_{i,k}, \tilde{\lambda}_{i,k}$ denote $\tilde{x}_i, \tilde{\mathbf{x}}_i^{-i}, \tilde{z}_i, \tilde{\lambda}_i$ at iteration k, respectively. ρ is the inertial parameter, c is the coupling parameter, and τ_i, ν_i, σ_i are the fixed step-sizes of player i in the update step. P_{Ω_i} is the projection operator on to the set Ω_i.

Let $\mathbf{x} := col(\mathbf{x}_i)_{i \in \mathcal{N}}$, $\mathbf{z} := col(z_i)_{i \in \mathcal{N}}$ and $\boldsymbol{\lambda} := col(\lambda_i)_{i \in \mathcal{N}}$. Let $\tilde{\mathbf{x}} := col(\tilde{\mathbf{x}}_i)_{i \in \mathcal{N}}$ with $\tilde{\mathbf{x}}_i := (\tilde{x}_i, \tilde{\mathbf{x}}_i^{-i})$, $\tilde{\mathbf{z}} := col(\tilde{z}_i)_{i \in \mathcal{N}}$ and $\tilde{\boldsymbol{\lambda}} := col(\tilde{\lambda}_i)_{i \in \mathcal{N}}$. In addition, $A := \text{diag}((A_i)_{i \in \mathcal{N}})$. $\mathbf{L}_\lambda := L \otimes I_m$, $\mathbf{L}_x := L \otimes I_n$, $\boldsymbol{b} := col(b_i)_{i \in \mathcal{N}}$, $\boldsymbol{\tau}^{-1} := \text{diag}((\tau_i^{-1} I_n)_{i \in \mathcal{N}})$, $\boldsymbol{\nu}^{-1} := \text{diag}((\nu_i^{-1} I_n)_{i \in \mathcal{N}})$ and $\boldsymbol{\sigma}^{-1} := \text{diag}((\sigma_i^{-1} I_n)_{i \in \mathcal{N}})$.

The *extended pseudo-gradient* mapping **F** is defined as

$$\mathbf{F}(\mathbf{x}) := col(\nabla_{x_i} J_i(x_i, \mathbf{x}_i^{-i}))_{i \in \mathcal{N}}. \tag{11}$$

Let $\varpi := col(\mathbf{x}, \mathbf{z}, \boldsymbol{\lambda}) \in \mathbf{\Omega}$, where $\mathbf{\Omega} := \mathbf{R}^{Nn} \times \mathbf{R}^{Nm} \times \mathbf{R}^{Nm}_+$, and we define operators $\mathfrak{A}, \mathfrak{B}$ and matrix Φ as follows:

$$\mathfrak{A} : \varpi \mapsto \begin{bmatrix} \mathcal{R}^\top N_\Omega(\mathcal{R}\mathbf{x}) \\ \mathbf{0} \\ N_{\mathbf{R}^{Nm}_+}(\boldsymbol{\lambda}) \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathcal{R}^\top A^\top \\ \mathbf{0} & \mathbf{0} & -\mathbf{L}_\lambda \\ -A\mathcal{R} & \mathbf{L}_\lambda & \mathbf{0} \end{bmatrix} \varpi$$

$$\mathfrak{B} : \varpi \mapsto \begin{bmatrix} \mathcal{R}^\top \mathbf{F}(\mathbf{x}) + c\mathbf{L}_x\mathbf{x} \\ \mathbf{0} \\ \mathbf{L}_\lambda\boldsymbol{\lambda} + \boldsymbol{b} \end{bmatrix} \tag{12}$$

$$\Phi := \begin{bmatrix} \boldsymbol{\tau}^{-1} & \mathbf{0} & -\mathcal{R}^\top A^\top \\ \mathbf{0} & \boldsymbol{\nu}^{-1} & \mathbf{L}_\lambda \\ -A\mathcal{R} & \mathbf{L}_\lambda & \boldsymbol{\sigma}^{-1} \end{bmatrix} \tag{13}$$

where $\mathcal{R} := \text{diag}((\mathcal{R}_i)_{i \in \mathcal{N}})$ with

$$\mathcal{R}_i := \begin{bmatrix} \mathbf{0}_{n_i \times n_{<i}} & I_{n_i} & \mathbf{0}_{n_i \times n_{>i}} \end{bmatrix}, \tag{14}$$

$n_{<i} := \sum_{j<i} n_j$ and $n_{>i} := \sum_{j>i} n_j$.

Algorithm 1 Distributed alternating inertial v-GNE seeking.

Initialization: $x_{i,0} \in \Omega_i, \mathbf{x}_{i,0}^{-i} \in \mathbf{R}^{n-n_i}, \lambda_{i,0} \in \mathbf{R}_+^m, z_{i,0} \in \mathbf{R}^m$
Acceleration: Set $\rho_k = 0$ if k is even, $\rho_k = \rho$ if k is odd.

$$\begin{aligned}
\tilde{x}_{i,k} &= x_{i,k} + \rho_k(x_{i,k} - x_{i,k-1}) \\
\tilde{\mathbf{x}}_{i,k}^{-i} &= \mathbf{x}_{i,k}^{-i} + \rho_k(\mathbf{x}_{i,k}^{-i} - \mathbf{x}_{i,k-1}^{-i}) \\
\tilde{z}_{i,k} &= z_{i,k} + \rho_k(z_{i,k} - z_{i,k-1}) \\
\tilde{\lambda}_{i,k} &= \lambda_{i,k} + \rho_k(\lambda_{i,k} - \lambda_{i,k-1})
\end{aligned}$$

Update:

$$\begin{aligned}
x_{i,k+1} &= P_{\Omega_i}(\tilde{x}_{i,k} - \tau_i(\nabla_{x_i} J_i(\tilde{x}_{i,k}, \tilde{\mathbf{x}}_{i,k}^{-i}) + A_i^\top \tilde{\lambda}_{i,k} \\
&\qquad + c \sum_{j \in \mathcal{N}_i} w_{ij}(\tilde{x}_{i,k} - \tilde{\mathbf{x}}_{j,k}^i))) \\
\mathbf{x}_{i,k+1}^{-i} &= \tilde{\mathbf{x}}_{i,k}^{-i} - \tau_i c \sum_{j \in \mathcal{N}_i} w_{ij}(\tilde{\mathbf{x}}_{i,k}^{-i} - \tilde{\mathbf{x}}_{j,k}^{-i}) \\
z_{i,k+1} &= \tilde{z}_{i,k} + \nu_i \sum_{j \in \mathcal{N}_i} w_{ij}(\tilde{\lambda}_{i,k} - \tilde{\lambda}_{j,k}) \\
\lambda_{i,k+1} &= P_{\mathbf{R}_+^m}(\tilde{\lambda}_{i,k} + \sigma_i(A_i(2x_{i,k+1} - \tilde{x}_{i,k}) - b_i \\
&\qquad - \sum_{j \in \mathcal{N}_i} w_{ij}(2(z_{i,k+1} - z_{j,k+1}) - (\tilde{z}_{i,k} - \tilde{z}_{j,k})) \\
&\qquad - \sum_{j \in \mathcal{N}_i} w_{ij}(\tilde{\lambda}_{i,k} - \tilde{\lambda}_{j,k})))
\end{aligned}$$

Let $\varpi_k := \text{col}(\mathbf{x}_k, z_k, \lambda_k), \tilde{\varpi}_k := \text{col}(\tilde{\mathbf{x}}_k, \tilde{z}_k, \tilde{\lambda}_k)$, where $\mathbf{x}_k, z_k, \lambda_k, \tilde{\mathbf{x}}_k, \tilde{z}_k, \tilde{\lambda}_k$ denote $\mathbf{x}, z, \lambda$, $\tilde{\mathbf{x}}, \tilde{z}, \tilde{\lambda}$ at iteration k, respectively. Suppose that $\Phi \succ 0$ and $\Phi^{-1}\mathfrak{A}$ is maximally monotone, then Algorithm 1 is equivalent to

$$\begin{cases} \varpi_{k+1} = T(\varpi_k), & \text{if } k \text{ is even} \\ \varpi_{k+1} = T(\varpi_k + \rho(\varpi_k - \varpi_{k-1})), & \text{if } k \text{ is odd} \end{cases} \tag{15}$$

where $\Phi, \mathfrak{A}, \mathfrak{B}$ in (12)–(13), $T := T_2 \circ T_1, T_1 := Id - \Phi^{-1}\mathfrak{B}$, and $T_2 := (Id + \Phi^{-1}\mathfrak{A})^{-1}$.

Lemma 1. *Suppose* $\Phi \succ 0$ *and* $\Phi^{-1}\mathfrak{A}$ *is maximally monotone, then any limit point* $\bar{\varpi} = \text{col}(\bar{\mathbf{x}}, \bar{z}, \bar{\lambda})$ *of Algorithm 1 is a zero of* $\mathfrak{A} + \mathfrak{B}$ *and a fixed point of* $T_2 \circ T_1$.

Proof. By the continuity of the right hand of (15), $\bar{\varpi} = T(\bar{\varpi})$. Since Φ is positive definite,

$$\begin{aligned}
\bar{\varpi} = T_2 \circ T_1(\bar{\varpi}) &:= (Id + \Phi^{-1}\mathfrak{A})^{-1} \circ (Id - \Phi^{-1}\mathfrak{B})(\bar{\varpi}) \\
&\Leftrightarrow (Id + \Phi^{-1}\mathfrak{A})(\bar{\varpi}) \in (Id - \Phi^{-1}\mathfrak{B})(\bar{\varpi}) \\
&\Leftrightarrow \mathbf{0} \in \Phi^{-1}(\mathfrak{A} + \mathfrak{B})(\bar{\varpi}) \\
&\Leftrightarrow \mathbf{0} \in (\mathfrak{A} + \mathfrak{B})(\bar{\varpi}).
\end{aligned}$$

$\square$

In order to show the convergence of the algorithm, the following assumptions are introduced.

Assumption 4. *The extended pseudo-gradient mapping* **F** *in* (11) *is θ-Lipschitz continuous, i.e., there exists θ > 0 such that for any* **x** *and* **x**′, $\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}')\| \le \theta \|\mathbf{x} - \mathbf{x}'\|$.

Let $c_{\min} := \frac{1}{s_2(L)}\left(\frac{(\theta+\theta_0)^2}{4\mu} + \theta\right)$ with μ, θ_0 in Assumption 2, and θ in Assumption 4. Let $E_x := \{\mathbf{x} \in \mathbf{R}^{Nn} | \mathbf{x} = \mathbf{1}_N \otimes x, x \in \mathbf{R}^n\}$. It follows from ([21], [Lemma 4]) that if c is selected such that $c > c_{\min}$, then $\mathfrak{A}$ is maximally monotone and $\mathfrak{B}$ is β-restricted cocoercive, i.e., for any ϖ and any $\varpi' \in \Omega_E$, where $\Omega_E := E_x \times \mathbf{R}^{Nm} \times \mathbf{R}_+^{Nm}$,

$$\langle \varpi - \varpi', \mathfrak{B}\varpi - \mathfrak{B}\varpi' \rangle \ge \beta \|\mathfrak{B}\varpi - \mathfrak{B}\varpi'\|^2, \tag{16}$$

where $0 < \beta \le \min\{\frac{1}{2d^*}, \frac{\mu}{\theta^2}\}$, and d^* is the maximal weighted degree of $\mathcal{G}$, i.e., $d^* = \max\{\sum_{j=1}^N w_{1j}, \cdots, \sum_{j=1}^N w_{Nj}\}$.

Similar to [21], a mild assumption (Assumption 4) on the pseudo-gradient mapping **F** is required only, while the requirement of strong monotonicity is relaxed for **F**.

Theorem 1. *Suppose Assumptions 1-4 hold. Choose $c > c_{min}$, $\delta > \frac{1}{2\beta}$, and the step sizes $\tau_i \le \frac{1}{\|A_i^\top\|_\infty + \delta}$, $\nu_i \le \frac{1}{2d_i + \delta}$, and $\sigma_i \le \frac{1}{\|A_i\|_\infty + 2d_i + \delta}$. Then for any $\rho \in [0, \frac{1}{2})$, the sequence $\{\mathbf{x}_k, \mathbf{z}_k, \boldsymbol{\lambda}_k\}_{k\in\mathbb{N}}$ generated by Algorithm 1 converges to the equilibrium $(\mathbf{x}^*, \mathbf{z}^*, \boldsymbol{\lambda}^*)$, where $\mathbf{x}^* = \mathbf{1}_N \otimes x^*$ and x^* is a v-GNE of the game* (3).

Proof. It follows from the Gershgorin's circle theorem ([32], [§6.8 Theorem 1]) that, given any $\delta > 0$, Φ is positive definite and $\Phi - \delta I_{n+2mN}$ is positive semi-definite if the step sizes $\tau_i \le \frac{1}{\|A_i^\top\|_\infty + \delta}$ and $\nu_i \le \frac{1}{2d_i + \delta}$.

Next, we first show the convergence of $\{\varpi_{2k}\}$ and then show the convergence of $\{\varpi_k\}$.

By ([21], [Lemma 6]), we have $T_2 \in \mathcal{A}(\frac{1}{2})$ and T_1 is $\frac{1}{2\beta\delta}$-*restricted* averaged, i.e., for any ϖ and any $\varpi' \in \Omega_E$,

$$\|T_1\varpi - T_1\varpi'\|_\Phi^2 \le \|\varpi - \varpi'\|_\Phi^2 \\ - (2\beta\delta - 1)\|\varpi - \varpi' - (T_1\varpi - T_1\varpi')\|_\Phi^2. \tag{17}$$

It follows from ([30], [Proposition 4.32]) that $T = T_2 \circ T_1$ is α-*restricted* averaged, with $\alpha = \frac{2}{3}$ when $\delta > \frac{1}{\beta}$. Let ϖ^* be a fixed point of T, then $\varpi^* \in \Omega_E$ according to ([21], [Theorem 1]).

(i) For the subsequence $\{\varpi_{2k}\}$, by (15), we have $\varpi_{2(k+1)} = T(T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}))$. Then, by T is $\alpha - restricted$ averaged, we obtain

$$\|\varpi_{2k+2} - \varpi^*\|_\Phi^2 \\ = \|T(T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k})) - \varpi^*\|_\Phi^2 \\ \le \|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi^*\|_\Phi^2 \\ - \frac{1-\alpha}{\alpha}\|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi_{2k+2}\|_\Phi^2. \tag{18}$$

By resorting to $\|\alpha x + (1-\alpha)y\|^2 + \alpha(1-\alpha)\|x-y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2$,

$$\|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi^*\|_\Phi^2 \\ = (1+\rho)\|T(\varpi_{2k}) - \varpi^*\|_\Phi^2 - \rho\|\varpi_{2k} - \varpi^*\|_\Phi^2 \\ + (1+\rho)\rho\|T(\varpi_{2k}) - \varpi_{2k}\|_\Phi^2, \tag{19}$$

and by using (17) again, (18) can be rewritten as

$$
\begin{aligned}
&\|\varpi_{2k+2} - \varpi^*\|_\Phi^2 \\
&\leq (1+\rho)\|\varpi_{2k} - \varpi^*\|_\Phi^2 - \rho\|\varpi_{2k} - \varpi^*\|_\Phi^2 \\
&\quad - (1+\rho)\frac{1-\alpha}{\alpha}\|T(\varpi_{2k}) - \varpi_{2k}\|_\Phi^2 \\
&\quad + (1+\rho)\rho\|T(\varpi_{2k}) - \varpi_{2k}\|_\Phi^2 \\
&\quad - \frac{1-\alpha}{\alpha}\|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi_{2k+2}\|_\Phi^2 \\
&= \|\varpi_{2k} - \varpi^*\|_\Phi^2 - (1+\rho)\left(\frac{1-\alpha}{\alpha} - \rho\right)\|T(\varpi_{2k}) - \varpi_k\|_\Phi^2 \\
&\quad - \frac{1-\alpha}{\alpha}\|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi_{2k+2}\|_\Phi^2.
\end{aligned}
\tag{20}
$$

Choose $\rho \leq \frac{1-\alpha}{\alpha} = \frac{1}{2}$, then

$$
\begin{aligned}
\|\varpi_{2k+2} - \varpi^*\|_\Phi^2 &\leq \|\varpi_{2k} - \varpi^*\|_\Phi^2 \\
&\quad - \frac{1-\alpha}{\alpha}\|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi_{2k+2}\|_\Phi^2.
\end{aligned}
\tag{21}
$$

This result implies that the sequence $\{\|\varpi_{2k} - \varpi^*\|_\Phi^2\}$ is decreasing and non-negative, and thus converges. Moreover, we have

$$
\sum_{k=0}^\infty \|T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi_{2(k+1)}\|_\Phi^2 < \infty
$$

and $T(\varpi_{2k}) + \rho(T(\varpi_{2k}) - \varpi_{2k}) - \varpi_{2(k+1)} \to 0$. Note that since $\{\varpi_{2k}\}$ is bounded, then there exists a convergent subsequence $\{\varpi_{2n_k}\}$ that converges to $\tilde{\varpi}$. Obviously,

$$
\varpi_{2(n_k+1)} = T(T(\varpi_{2n_k}) + \rho(T(\varpi_{2n_k}) - \varpi_{2n_k})).
$$

Let $k \to \infty$, we have $\tilde{\varpi} = T\tilde{\varpi}$, which implies that $\tilde{\varpi}$ is a fixed point of T and thus $\{\|\varpi_{2k} - \tilde{\varpi}\|_\Phi^2\}$ converges. Since $\{\|\varpi_{2n_k} - \tilde{\varpi}\|_\Phi^2\}$ converges to 0, $\{\varpi_{2k}\}$ converges to $\tilde{\varpi}$.

(ii) T is *restricted* nonexpansive since it is $\frac{2}{3}$-*restricted* averaged, and then one obtains

$$
\|\varpi_{2k+1} - \tilde{\varpi}\| = \|T\varpi_{2k} - T\tilde{\varpi}\| \leq \|\varpi_{2k} - \tilde{\varpi}\|,
\tag{22}
$$

which implies that the odd subsequence $\{\varpi_{2k+1}\}$ also converges to $\tilde{\varpi}$, and thus $\{\varpi_k\}$ converges to $\tilde{\varpi}$. Note that $\Phi \succ 0$ and $\Phi^{-1}\mathfrak{A}$ is maximally monotone. $\tilde{\varpi}$ is a fixed point of T, and hence is a zero of $\mathfrak{A} + \mathfrak{B}$ by Lemma 1. It follows from ([21], [Theorem 1]) that given any $\tilde{\varpi} := \mathrm{col}(\mathbf{x}^*, \mathbf{z}^*, \boldsymbol{\lambda}^*) \in zer(\mathfrak{A} + \mathfrak{B})$, then $\mathbf{x}^* = \mathbf{1}_N \otimes x^*$, and x^* solves $VI(F, K)$ (8), that is, x^* is a v-GNE of game (3). $\square$

4.2. Alternating Overrelaxed Distributed v-GNE Seeking Algorithm

In this subsection, an alternating overrelaxed distributed algorithm is constructed for seeking the v-GNE, presented in Algorithm 2, and also that η is an overrelaxed parameter. Here the partial-decision information setting is considered.

Algorithm 2 Distributed alternating overrelaxed v-GNE seeking.

Initialization: $x_{i,0} \in \Omega_i, \mathbf{x}_{i,0}^{-i} \in \mathbf{R}^{n-n_i}, \lambda_{i,0} \in \mathbf{R}_+^m, z_{i,0} \in \mathbf{R}^m$

Update:

$$\tilde{x}_{i,k} = P_{\Omega_i}(x_{i,k} - \tau_i(\nabla_{x_i} J_i(x_{i,k}, \mathbf{x}_{i,k}^{-i}) + A_i^\top \lambda_{i,k}$$
$$+ c \sum_{j \in \mathcal{N}_i} w_{ij}(x_{i,k} - \mathbf{x}_{j,k}^i)))$$

$$\tilde{\mathbf{x}}_{i,k}^{-i} = \mathbf{x}_{i,k}^{-i} - \tau_i c \sum_{j \in \mathcal{N}_i} w_{ij}(\mathbf{x}_{i,k}^{-i} - \mathbf{x}_{j,k}^{-i})$$

$$\tilde{z}_{i,k} = z_{i,k} + \nu_i \sum_{j \in \mathcal{N}_i} w_{ij}(\lambda_{i,k} - \lambda_{j,k})$$

$$\tilde{\lambda}_{i,k} = P_{\mathbf{R}_+^m}(\lambda_{i,k} + \sigma_i(A_i(2\tilde{x}_{i,k} - x_{i,k}) - b_i$$
$$- \sum_{j \in \mathcal{N}_i} w_{ij}(2(\tilde{z}_{i,k} - \tilde{z}_{j,k}) - (z_{i,k} - z_{j,k}))$$
$$- \sum_{j \in \mathcal{N}_i} w_{ij}(\lambda_{i,k} - \lambda_{j,k})))$$

Acceleration: Set $\eta_k = 1$ if k is even, $\eta_k = \eta$ if k is odd.

$$x_{i,k+1} = \tilde{x}_{i,k} + (\eta_k - 1)(\tilde{x}_{i,k} - x_{i,k})$$
$$\mathbf{x}_{i,k+1}^{-i} = \tilde{\mathbf{x}}_{i,k}^{-i} + (\eta_k - 1)(\tilde{\mathbf{x}}_{i,k}^{-i} - \mathbf{x}_{i,k}^{-i})$$
$$z_{i,k+1} = \tilde{z}_{i,k} + (\eta_k - 1)(\tilde{z}_{i,k} - z_{i,k})$$
$$\lambda_{i,k+1} = \tilde{\lambda}_{i,k} + (\eta_k - 1)(\tilde{\lambda}_{i,k} - \lambda_{i,k})$$

Similar to (15), we suppose that $\Phi \succ 0$ and $\Phi^{-1}\mathfrak{A}$ is maximally monotone, then Algorithm 2 is equivalent to

$$\begin{cases} \varpi_{k+1} = T(\varpi_k) & \text{if } k \text{ is even} \\ \varpi_{k+1} = T(\varpi_k) + (\eta - 1)(T(\varpi_k) - \varpi_k) & \text{if } k \text{ is odd.} \end{cases} \quad (23)$$

where $\varpi_k = \text{col}(\mathbf{x}_k, \mathbf{z}_k, \boldsymbol{\lambda}_k)$ and T is given in (15).

Next, we prove the convergence of Algorithm 2 to a v-GNE.

Theorem 2. *Suppose Assumptions 1–4 hold. Take any $c > c_{min}$, $\delta > \frac{1}{2\beta}$, and the step sizes $\tau_i \leq \frac{1}{\|A_i^\top\|_\infty + \delta}$, $\nu_i \leq \frac{1}{2d_i + \delta}$, and $\sigma_i \leq \frac{1}{\|A_i\|_\infty + 2d_i + \delta}$. Then, for any $\eta \in [1, \frac{3}{2})$, the sequence $\{\mathbf{x}_k, \mathbf{z}_k, \boldsymbol{\lambda}_k\}_{k \in \mathbb{N}}$ generated by Algorithm 2 converges to the equilibrium $(\mathbf{x}^*, \mathbf{z}^*, \boldsymbol{\lambda}^*)$, where $\mathbf{x}^* = \mathbf{1}_N \otimes x^*$ and x^* is a v-GNE of the game (3).*

Proof. Similar to Theorem 1, we first show the convergence of $\{\varpi_{2k}\}$, and then prove the convergence of $\{\varpi_k\}$. Note that $T = T_2 \circ T_1$ is α-restricted averaged with $\alpha = \frac{2}{3}$ when $\delta > \frac{1}{\beta}$. Let ϖ^* be any fixed point of T.

First, we consider the subsequence $\{\varpi_{2k}\}$, and according to (23) and (17), one has

$$\|\varpi_{k+2} - \varpi^*\|_\Phi^2$$

$$\leq \eta[\|T(\varpi_k) - \varpi^*\|_\Phi^2 - \frac{1-\alpha}{\alpha}\|T(\varpi_k) - T(T(\varpi_k))\|_\Phi^2]$$

$$+ (1-\eta)[\|\varpi_k - \varpi^*\|_\Phi^2 - \frac{1-\alpha}{\alpha}\|\varpi_k - T(\varpi_k)\|_\Phi^2]$$

$$- \eta(1-\eta)\|T(T(\varpi_k)) - T(\varpi_k)\|_\Phi^2$$

$$= (1-\eta)\|\varpi_k - \varpi^*\|_\Phi^2 + \eta\|T(\varpi_k) - \varpi^*\|_\Phi^2$$

$$+ [-\eta(\frac{1}{\alpha} - \eta)]\|T(\varpi_k) - T(T(\varpi_k))\|_\Phi^2$$

$$- (1-\eta)\frac{1-\alpha}{\alpha}\|T(\varpi_k) - \varpi_k\|_\Phi^2$$

$$\leq \|\varpi_k - \varpi^*\|_\Phi^2 - \frac{1-\alpha}{\alpha}\|T(\varpi_k) - \varpi_k\|_\Phi^2$$

$$+ [-\eta(\frac{1}{\alpha} - \eta)]\|T(\varpi_k) - T(T(\varpi_k))\|_\Phi^2$$

where the first equality holds due to $\|\alpha x + (1-\alpha)y\|^2 + \alpha(1-\alpha)\|x - y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2$. By choosing $\eta \leq \frac{1}{\alpha}$, we have

$$\|\varpi_{k+2} - \varpi^*\|_\Phi^2 \leq \|\varpi_k - \varpi^*\|_\Phi^2 - \frac{1-\alpha}{\alpha}\|T(\varpi_k) - \varpi_k\|_\Phi^2 \tag{24}$$

which implies that $\{\|\varpi_{k+2} - \varpi^*\|_\Phi^2\}$ is monotonically decreasing and bounded, and is thus convergent. Furthermore,

$$\sum_{k=0}^{\infty}(\|\varpi_{2(k+1)} - \varpi^*\|_\Phi^2 - \|\varpi_{2k} - \varpi^*\|_\Phi^2)$$

$$\leq -\frac{1-\alpha}{\alpha}\sum_{k=0}^{\infty}\|T(\varpi_{2k}) - \varpi_{2k}\|_\Phi^2, \tag{25}$$

that is, $\sum_{k=0}^{\infty}\|T(\varpi_{2k}) - \varpi_{2k}\|_\Phi^2 \leq \|\varpi_0 - \varpi^*\|_\Phi^2 < \infty$, and hence

$$\|T(\varpi_{2k}) - \varpi_{2k}\| \to 0. \tag{26}$$

Note that if $\{\varpi_{2k}\}$ is bounded, there exists a convergent subsequence $\{\varpi_{2n_k}\} \to \tilde{\varpi}$ for some limit $\tilde{\varpi}$.

Let $k \to \infty$ in (26), we have $T(\tilde{\varpi}) \to \tilde{\varpi}$ which implies $\tilde{\varpi}$ is a fixed point of T, thus $\{\|\varpi_{2k} - \tilde{\varpi}\|_\Phi^2\}$ converges. Since $\{\|\varpi_{2n_k} - \tilde{\varpi}\|_\Phi^2\} \to 0$, $\{\varpi_{2k}\}$ converges to $\tilde{\varpi}$.

(ii) If T is *restricted* nonexpansive since it is $\frac{2}{3}$-*restricted* averaged, then one obtains

$$\|\varpi_{2k+1} - \tilde{\varpi}\| = \|T(\varpi_{2k}) - T(\tilde{\varpi})\| \leq \|\varpi_{2k} - \tilde{\varpi}\|, \tag{27}$$

which implies that the sequence $\{\varpi_{2k+1}\}$ converges to the same limit of $\{\varpi_{2k}\}$, and thus $\{\varpi_k\}$ converges to $\tilde{\varpi}$. $\quad\square$

5. Numerical Simulation

In this section, we consider a classic Nash–Cournot game over a network as [21], where there are N firms and each firm $i \in \{1, \cdots, N\}$ produces commodities to participate in the competition over m markets (see Figure 1). Each market (denoted by $M_1, \cdots, M_m$) has limited capacity. Here, the partial-decision information setting is considered where each firm has limited access to its neighboring firms' information over the communication graph as in Figure 2.
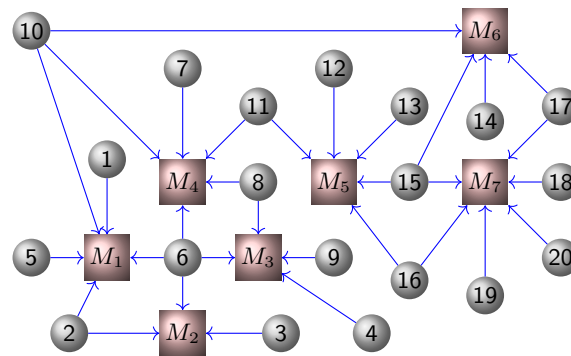
Figure 1. Network of Nash–Gournot game, an arrow from *i* to M_k means firm *i* participates in market M_k's competition.
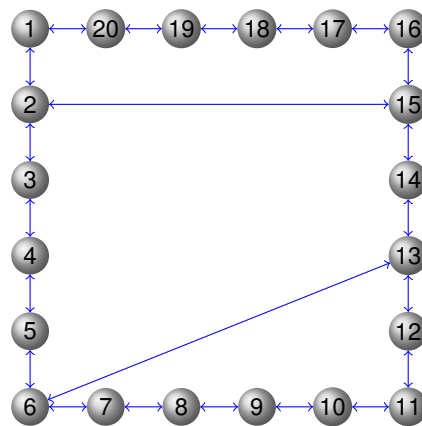

Figure 2. Communication graph among all firms; an edge from *i* to *j* means firm *i* and *j* can exchange information through the graph.

We assume that firm *i* participates in n_i markets by producing $x_i \in \mathbf{R}^{n_i}$ amount of commodities and its production is limited by the set $\Omega_i \in \mathbf{R}^{n_i}$. The local matrix $A_i \in \mathbf{R}^{m \times n_i}$ for firm *i* represents which markets it participates in. Specifically, for the *j*-th column of A_i, its *k*-th element is 1 if and only if firm *i* delivers $[x_i]_j$ amount of production to market M_k; all other elements are 0. Each market M_k has a maximal capacity of $r_k > 0$, that is, $Ax \leq r$, where $A = [A_1, \cdots, A_N]$, $x = \text{col}(x_i)_{i \in \mathcal{N}} \in \mathbf{R}^n$, $n = \sum_{i=1}^{N} n_i$ and $r = \text{col}(r_k)_{k=1,\cdots,m} \in \mathbf{R}^m$. Suppose that each firm *i* has the production cost $c_i(x_i) : \Omega_i \to \mathbf{R}$, and the price function $P : \mathbf{R}^n \to \mathbf{R}^m$ maps the total supply of each market to the market's price vector. The local objective function of firm *i* is $J_i(x_i, x_{-i}) = c_i(x_i) - (P(Ax))^\top A_i x_i$.

Suppose $N = 20$ and $m = 7$. Let $\Omega_i = \{x_i \in \mathbf{R}^{n_i} | 0 \leq x_i \leq \Theta_i\}$, where each component of Θ_i is randomly drawn from $(5, 10)$. r_k is randomly drawn from $(1, 2)$. The local cost function of firm *i* is $c_i(x_i) = x_i^\top H_i x_i + h_i^\top x_i$, where H_i is a diagonal matrix with the elements randomly drawn from $(1, 8)$ and h_i is randomly drawn from $(1, 2)$. The price function is taken as the linear function $P = \bar{P} - DAx$ with $\bar{P} = \text{col}(\bar{P}_k)_{k=1,\cdots,m} \in \mathbf{R}^m$ and $D = \text{diag}(d_k)_{k=1,\cdots,m} \in \mathbf{R}^{m \times m}$, where $\bar{P}_k$ and d_k are randomly drawn from $(10, 20)$ and $(1, 3)$, respectively. Set the step-sizes as $c = 100$, $\tau_i = 0.003$, $\nu_i = 0.02$, and $\sigma_i = 0.003$.

First, Figure 3 shows that the convergence to the v-GNE can be guaranteed under Algorithms 1 and 2, and the trajectories of the local decision $x_{i,k}$ of firms $1, 6, 10, 11$ are displayed in Figure 4. It can be seen in Figure 5 that the estimates on the firms 1 and 3 asymptotically tend to their real actions by using the proposed algorithms.

Then, it can be seen from Figure 6, that both of the proposed Algorithms 1 and 2 converge to the GNE x^* with a faster convergence as compared with ([21], [Alg. 1]), where Algorithm 1 has the fastest convergent rate. From Figure 7 we can see that the proposed Algorithm 1 also has a faster convergence rate than ([25], [Alg. 3]). We set $\alpha = 4.3 \times 10^{-3}$ in ([25], [Alg. 3]), the same step-sizes τ_i, ν_i and σ_i and the same other

parameters as Algorithms 1 and 2 in ([21], [Alg. 1]) and ([25], [Alg. 3]). On the other hand, as compared with the algorithm with inertia, Algorithm 1 with alternating inertia requires less computation resources. Thus, Algorithm 1 could be the best choice when both fast convergence rate and low computation cost are taken into consideration.
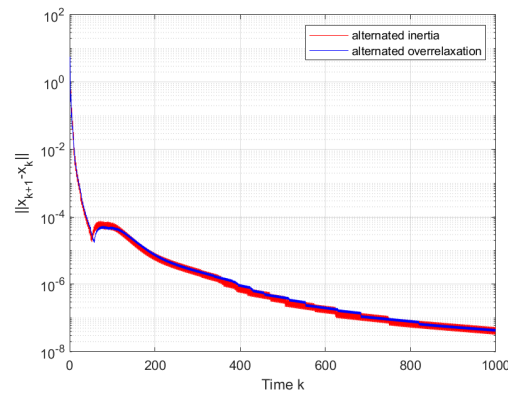

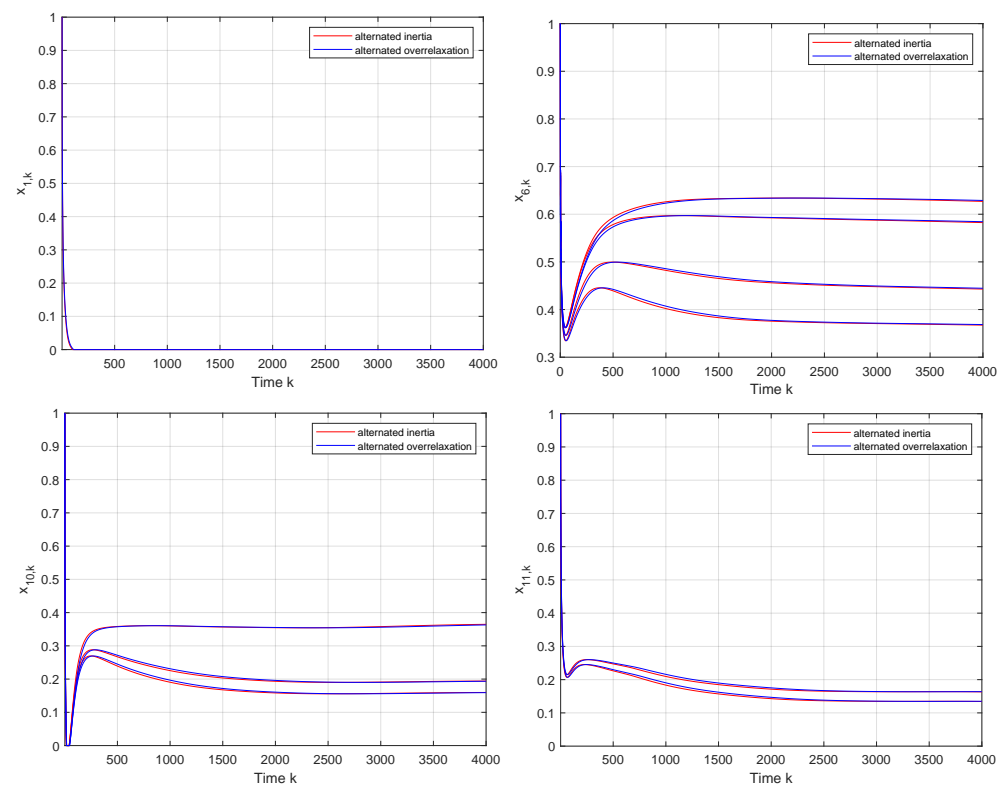**Figure 3.** The trajectories of $\|x_{k+1} - x_k\|$ generated by Algorithms 1 and 2.

Figure 4. The trajectories of local decisions $x_{i,k}$ of firms $1, 6, 10$ and 11 by Algorithms 1 and 2, respectively.
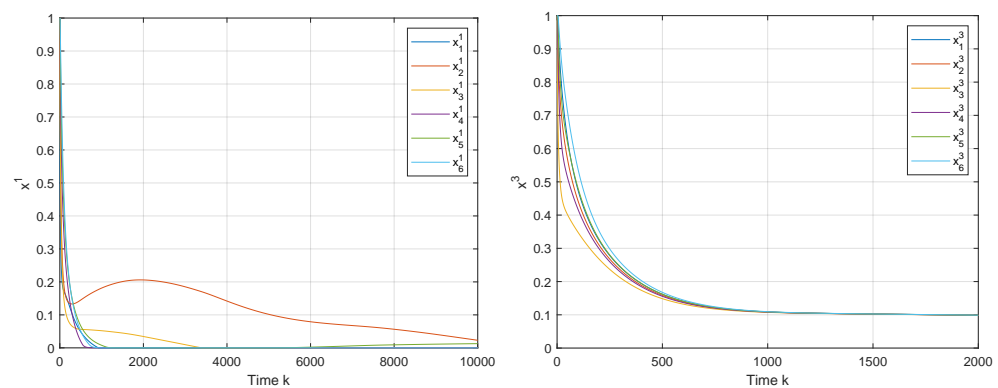
Figure 5. The trajectories of the estimate variable $\mathbf{x}_j^1$ from firms 1–6 generated by Algorithm 1 (left); and the trajectories of the estimate variable $\mathbf{x}_j^3$ from firm 1–6 generated by Algorithm 2 (right).
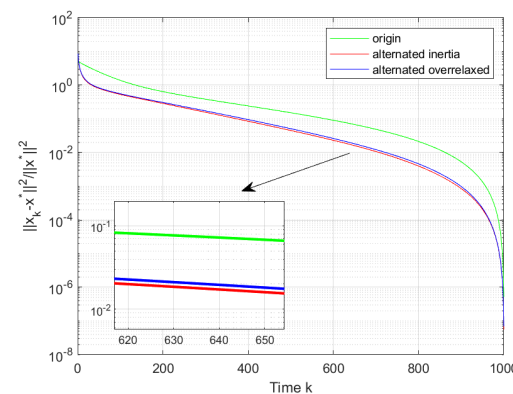

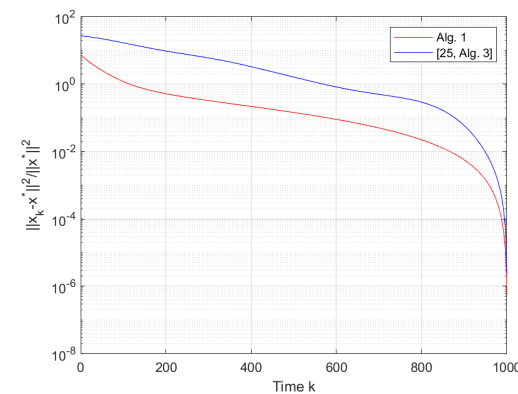**Figure 6.** Relative error $\|x_k - x^*\|^2 / \|x^*\|^2$ generated by ([21], [Algorithm 1]), Algorithms 1 and 2.

Figure 7. Relative error $\|x_k - x^*\|^2 / \|x^*\|^2$ generated by Algorithm 1 and ([25], [Alg. 3]).

Remark 2. *It is worthwhile to note that the introduction of the inertia and overrelaxation steps has the potential of accelerating the convergence rate. As such, in this paper, the inertial and overrelaxed distributed algorithms are developed based on the pseudo-gradient method for seeking generalized Nash equilibrium in multi-player games. The similar inertia idea has been considered in the proximal-point algorithm (see ([25], [Alg. 3])). However, the proximal-point algorithm generally needs to solve the optimization problem at each step k, which may be time-consuming and possibly costs a great amount of computation resources in many situations. As such, pseudo-gradient algorithms with inertia and overrelaxation were constructed in this paper, which successfully guarantees the convergence to v-GNE with a fast convergence rate. Moreover, we note that the introduction of the inertia and overrelaxation steps increases the computation burden, and thus two alternating inertial and overrelaxed algorithms are established in Algorithms 1 and 2 to balance the convergence rate and computation burden. In order to better display the effectiveness of our*

algorithms, we have added the comparison with ([25], [Alg. 3]) in the simulation part (see Figure 7). From Figure 7, it can be seen that Algorithm 1 in this paper outperforms the ([25], [Alg. 3]) in terms of the convergence rate.

6. Conclusions

This paper has studied the GNE computation issue in multi-player games with shared coupling constraints under the partial-decision information setting. Two distributed algorithms with alternating inertia and alternating overrelaxation have been developed, respectively, with fixed step-sizes. Both algorithms have guaranteed the convergence to the GNE under a mild assumption, which have the potential of improving the convergence rate and saving computation cost. Finally, one simulation example has been provided to show the effectiveness of the proposed algorithms. Further research topics can be focused on stochastic NE seeking problems subject to time-varying topologies with and without event-triggered communication protocols.

Author Contributions: Z.F.: Conceptualization, writing, editing and methodology; W.X.: Conceptualization, writing, editing and methodology; J.C.: Review, editing and supervision. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Myerson, R.B. *Game Theory*; Harvard University Press: Cambridge, MA, USA, 2013.
2. Arrow, K.J.; Debreu, G. Existence of an equilibrium for a competitive economy. *Econometrica* **1954**, *22*, 265–290. [CrossRef]
3. Pang, J.-S.; Scutari, G.; Facchinei, F.; Wang, C.X. Distributed power allocation with rate constraints in Gaussian frequency-selective interference channels. *IEEE Trans. Inf. Theory* **2007**, *54*, 3471–3489. [CrossRef]
4. Liu, Y.; Guo, L.; Lu, C.; Chai, Y.; Gao, S.; Xu, B. A fully distributed voltage optimization method for distribution networks considering integer constraints of step voltage regulators. *IEEE Access* **2019**, *7*, 60055–60066. [CrossRef]
5. Breton, M.; Zaccour, G.; Zahaf, M. A game-theoretic formulation of joint implementation of environmental projects. *Eur. J. Oper. Res.* **2006**, *168*, 221–239. [CrossRef]
6. Ardagna, D.; Panicucci, B.; Passacantando, M. A game theoretic formulation of the service provisioning problem in cloud systems. In Proceedings of the 20th International Conference on World Wide Web, Hyderabad India, 28 March–1 April 2011; pp. 177–186.
7. Pang, J.-S.; Scutari, G.; Palomar, D.P.; Facchinei, F. Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach. *IEEE Trans. Signal Process.* **2010**, *58*, 3251–3271. [CrossRef]
8. Tang, Y.; Gao, H.; Zhang, W.; Kurths, J. Leader-following consensus of a class of stochastic delayed multi-agent systems with partial mixed impulses. *Automatica* **2015**, *53*, 346–354. [CrossRef]
9. Zhao, C.; Xiang, B.; Chen, B. Fusion estimation for discrete time-varying systems with bounded nonlinearities. *IEEE Access* **2019**, *7*, 27097–27105. [CrossRef]
10. Brückner, M.; Kanzow, C.; Scheffer, T. Static prediction games for adversarial learning problems. *J. Mach. Learn. Res.* **2012**, *13*, 2617–2654.
11. Xu, M.; Zhang, Y.; Zhang, D.; Chen, B. Distributed robust dimensionality reduction fusion estimation under dos attacks and uncertain covariances. *IEEE Access* **2021**, *9*, 10328–10337. [CrossRef]
12. Holt, C.A.; Roth, A.E. The nash equilibrium: A perspective. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3999–4002. [CrossRef]
13. Debreu, G. A social equilibrium existence theorem. *Proc. Natl. Acad. Sci. USA* **1952**, *38*, 886–893. [CrossRef] [PubMed]
14. Dutang, C. *A Survey of GNE Computation Methods: Theory and Algorithms*; HAL: Lyon, France, 2013.
15. Facchinei, F.; Kanzow, C. Generalized Nash equilibrium problems. *Ann. Oper. Res.* **2010**, *175*, 177–211. [CrossRef]

16. Yi, P.; Pavel, L. A distributed primal-dual algorithm for computation of generalized Nash equilibria via operator splitting methods. In Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control, Melbourne, VIC, Australia, 12–15 December 2017; pp. 3841–3846.
17. Carlson, F.; Krupar, K.R. Real and complex monotone communication games. *Engl. J.* **2014**, *64*, 65. [CrossRef]
18. Zhu, Y.; Yu, W.; Wen, G.; Chen, G.; Ren, W. Continuous-time distributed subgradient algorithm for convex optimization with general constraints. *IEEE Trans. Autom. Control* **2018**, *64*, 1694–1701. [CrossRef]
19. Frihauf, P.; Krstic, M.; Basar, T. Nash equilibrium seeking in noncooperative games. *IEEE Trans. Autom. Control* **2011**, *57*, 1192–1207. [CrossRef]
20. Bimpikis, K.; Ehsani, S.; Ilkılıç, R. Cournot competition in networked markets. *Manag. Sci.* **2019**, *65*, 2467–2481. [CrossRef]
21. Pavel, L. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach. *IEEE Trans. Autom. Control* **2020**, *65*, 1584–1597. [CrossRef]
22. Swenson, B.; Kar, S.; Xavier, J. Empirical centroid fictitious play: An approach for distributed learning in multi-agent games. *IEEE Trans. Signal Process.* **2015**, *63*, 3888–3901. [CrossRef]
23. Koshal, J.; Nediæ, A.; Shanbhag, U.V. Distributed algorithms for aggregative games on graphs. *Oper. Res.* **2016**, *64*, 680–704. [CrossRef]
24. Zhu, M.; Frazzoli, E. Distributed robust adaptive equilibrium computation for generalized convex games. *Automatica* **2016**, *63*, 82–91. [CrossRef]
25. Bianchi, M.; Belgioioso, G.; Grammatico, S. Fast generalized Nash equilibrium seeking under partial-decision information. *arXiv* **2020**, arXiv:2003.09335.
26. Pavel, L. A doubly-augmented operator splitting approach for distributed gne seeking over networks. In Proceedings of the 2018 IEEE Conference on Decision and Control, Miami, FL, USA, 17–19 December 2018; pp. 3529–3534.
27. Ye, M.; Hu, G. Distributed nash equilibrium seeking by a consensus based approach. *IEEE Trans. Autom. Control* **2017**, *62*, 4811–4818. [CrossRef]
28. Zhu, Y.; Yu, W.; Wen, G.; Chen, G. Distributed nash equilibrium seeking in an aggregative game on a directed graph. *IEEE Trans. Autom. Control* **2020**, *66*, 2746–2753. [CrossRef]
29. Yi, P.; Pavel, L. An operator splitting approach for distributed generalized Nash equilibria computation. *Automatica* **2019**, *102*, 111–121. [CrossRef]
30. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*; Springer: New York, NY, USA, 2011.
31. Facchinei, F.; Pang, J.-S. *Finite-Dimensional Variational Inequalities and Complementarity Problems*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007.
32. Franklin, J.N. *Matrix Theory*; Courier Corporation: North Chelmsford, MA, USA, 2012.