SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

July 2024, Vol. 67, Iss. 7, 172205:1–172205:14
https://doi.org/10.1007/s11432-023-4009-4

Distributed generalized Nash equilibrium seeking: event-triggered coding-decoding-based secure communication

Shaofu YANG[1], Wenying XU[2*], Wangli HE[3] & Jinde CAO[2]

[1]*School of Computer Science and Engineering, Southeast University, Nanjing* 211189*, China;*
[2]*School of Mathematics, Southeast University, Nanjing* 211189*, China;*
[3]*Key Laboratory of Advanced Control and Optimization for Chemical Processes, East China University of Science and Technology, Shanghai* 200237*, China*

Received 22 August 2023/Revised 23 December 2023/Accepted 29 January 2024/Published online 27 June 2024

Abstract In this paper, we consider the distributed generalized Nash equilibrium (GNE) seeking problem in strongly monotone games. The transmission among players is implemented through a digital communication network with limited bandwidth. For improving communication efficiency or/and security, an event-triggered coding-decoding-based communication is first proposed, where the data (decision variable) are first mapped to a series of finite-level codewords and, only when an event condition is satisfied, then sent to the neighboring agents. Moreover, a distributed communication-efficient GNE seeking algorithm is constructed accordingly, and the overrelaxation scheme is further taken into consideration. Through primal-dual analysis, the proposed algorithm is proven to converge to a variational GNE with fixed step-sizes by recasting it as an inexact forward-backward iteration. Finally, numerical simulations illustrate the benefit of the proposed algorithms in terms of saving communication resources.

Keywords noncooperative games, event-triggered communication, coding-decoding, quantization

1 Introduction

Noncooperative games arise from the practical need for modeling the decision-making process of selfish agents, where each agent aims at minimizing its individual but inter-dependent cost function. Typical examples can be found in spectrum access in cognitive radio networks [1], traffic networks [2], smart-grid management [3, 4], demand response in competitive markets [5, 6], and opinion dynamics in social networks [7]. Coupling constraints often occur because a group of agents may compete for the limited and shared network resources, resulting in coupling restrictions on the feasible decision sets of agents. Therefore, the computation problem of generalized Nash equilibrium (GNE) has stirred extensive research attention (see [8–12], and the references therein).

For tackling the GNE problem, an elegant, operator-theoretic approach has recently been proposed in [13,14], based on which the GNE seeking problem is reformulated as that of finding zeros of a monotone operator. For instance, based on the operator splitting approach, a distributed algorithm has been introduced in [13] to seek an NE with full-decision information, where all decisions of competitors are implicitly assumed to be available to agents. However, in large-scale network scenarios, each agent is subject to limited communication capacities and possibly has access to partial decisions of few neighboring agents. This game setting is called partial-decision information setting [8, 9]. In such a setting, in order to make up for the lack of decisions (concerning non-neighboring competitors), a great deal of auxiliary variables are introduced for each agent to estimate and reconstruct decisions of all the other agents [8, 15–17]. At each iteration, each agent should send the augmented state that combines the real decision with the estimates of decisions of all other agents. Note that the number of estimate variables is proportional to that of agents, and thus a significant communication burden could be inevitable when a

* Corresponding author (email: wyxu@seu.edu.cn)

lot of agents participate in the game. Therefore, it is of significant importance to develop communication-efficient algorithms for seeking a GNE.

Over the past few years, as an efficient way of communication, event-triggered communication (ETC) has attracted great attention [18–22]. Different from traditional time-triggered communication, the key idea of ETC is that information transmission is activated only when a pre-designed event condition is satisfied (rather than when time lapses). An event condition is generally related to the change of system variables and is designed to greatly reduce the communication frequency (and also the number of communication rounds) while maintaining system performance almost unchanged. Remarkably, there have been a great deal of research effort made toward distributed event-triggered scheme (ETS) for control/cooperation problems. Nevertheless, the distributed GNE problem has not received adequate attention yet in terms of communication-efficient solution methods.

On the other hand, in a communication network with limited bandwidth, only a limited number of bits of data are allowed to be transmitted at a certain time [23–28]. As such, how to take advantage of the limited network resources for guaranteeing the desired algorithm performance becomes a key issue in the GNE seeking problem, especially in the partial-decision information setting. In this scenario, data-coding schemes appear as a kind of resource-saving scheme, which map the original data to specific codewords with fewer bits [29]. In addition, the information exchange among players through communication networks poses potential vulnerabilities to attackers, which may lead to unintended consequences. In this respect, the transmission based on data coding provides an effective encryption-like way to reduce security risks. Therefore, the coding-decoding (CD)-based scheme can be considered an efficient and secure communication scheme in multi-agent decision problems.

Inspired by the above points, the objective of this paper is to investigate a distributed communication-efficient algorithm for tackling the GNE seeking problem subject to shared affine coupling constraints. The main novelties and contributions of the paper are highlighted as follows: (1) for the efficient communication and security purposes, we establish a novel event-triggered, coding-decoding-based (ETCD-based) algorithm, where the compressed codeword is transmitted to the neighboring agents only when a local event condition is triggered; (2) an inexact forward-backward iteration is developed to prove that the proposed algorithm converges to a variational GNE with fixed step sizes under some mild technical assumptions; (3) we introduce an overrelaxation scheme to the algorithm design; and numerical simulations illustrate that the proposed overrelaxed, ETCD-based algorithm has the potential of further improving the convergence performance and reducing the communication burden as well as improving communication security.

Notation. $\mathbb{R}^n$ ($\mathbb{R}^n_+$) and $\mathbb{R}^{n\times m}$ mean the n-dimensional (nonnegative) Euclidean space and $n \times m$ real matrices. Let $\mathbf{1}_n$ or $\mathbf{0}_n$ denote n-dimensional column vectors with all ones or zeros. $I_{n\times m}$ and $\mathbf{0}_{n\times m}$ represent, respectively, the identity matrix or zero matrix of $n\times m$ dimensions. For $x, y \in \mathbb{R}^n$, $x^{\mathrm{T}}y = \langle x, y\rangle$ means the inner product of x and y. $\|x\| = \sqrt{x^{\mathrm{T}}x}$ denotes the norm induced by the inner product $\langle \cdot, \cdot\rangle$. Given n vectors $x_1, \ldots, x_n$, define $\mathrm{col}(x_1, \ldots, x_n) = [x_1^{\mathrm{T}}, \cdots, x_n^{\mathrm{T}}]^{\mathrm{T}}$. For a matrix $A = [a_{i,j}] \in \mathbb{R}^{n\times m}$, let $a_{i,j}$ (or $[A]_{i,j}$) be the matrix entry in the i-th row and j-th column. Let $\|A\|_\infty = \max_{1\leqslant i\leqslant n}\sum_{j=1}^m |a_{i,j}|$ be the induced matrix, and let $A \succ 0$ ($A \succeq 0$) denote a symmetric positive definite (semidefinite) matrix. For $A \succ 0$, $\langle x, y\rangle_A := \langle Ax, y\rangle$ and $\|x\|_A := \sqrt{\langle Ax, x\rangle}$ represent the A-induced inner product and A-induced norm, respectively, $q_1(A)$, $q_2(A)$, and $q_{\max}(A)$ mean the smallest, the second smallest, and maximum eigenvalues of A. $\mathrm{diag}\{\cdots\}$ stands for a block-diagonal matrix. $A \otimes B$ denotes the Kronecker product of two matrices A and B. $\Pi_{i=1}^n \Omega_i$ means the Cartesian product of the sets Ω_i, $i = 1, \ldots, n$.

2 Problem formulation

2.1 Game formulation

Consider a noncooperative game involving a group of agents $\mathcal{V} := \{1, 2, \ldots, N\}$ that subject to coupling constraints. Let x_i mean the local decision (strategy or action) of agent i and $\Omega_i \subseteq \mathbb{R}^{n_i}$ stands for its private decision set. $\Omega := \Pi_{i=1}^N \Omega_i \subseteq \mathbb{R}^n$ represents the overall action space and $n := \sum_{i=1}^N n_i$. Let $\boldsymbol{x} := \mathrm{col}(x_1, \ldots, x_N) \in \Omega$ denote the stacked decision vector of all agents. The objective function of agent $i \in \mathcal{V}$ is defined as $J_i(x_i, \boldsymbol{x}_{-i})$, where $\boldsymbol{x}_{-i} := \mathrm{col}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N) \in \mathbb{R}^{n-n_i}$ mean the decision vector of all agents except i. In this paper, the agents compete for shared and limited resources, which can be reflected by the following affine coupling constraints:

$$\boldsymbol{X} := \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \sum_{i=1}^{N} A_i x_i \leqslant \sum_{i=1}^{N} b_i \right\} \cap \Omega, \tag{1}$$

where $b_i \in \mathbb{R}^m$ and $A_i \in \mathbb{R}^{m \times n_i}$ are local data that are only available to agent i. Such coupling constraints (1) arise in some practical applications, such as demand response in competitive markets [6]. A game can be further represented by

$$\forall i \in \mathcal{V}: \quad \min_{x_i \in \mathbb{R}^{n_i}} J_i(x_i, \boldsymbol{x}_{-i}) \quad \text{s.t.} \quad x_i \in \mathcal{X}_i(\boldsymbol{x}_{-i}), \tag{2}$$

where $\mathcal{X}_i(\boldsymbol{x}_{-i})$ is the feasible set of agent i, which is denoted by $\mathcal{X}_i(\boldsymbol{x}_{-i}) := \{x_i \in \Omega_i \mid A_i x_i \leqslant \sum_{j \in \mathcal{V} \setminus \{i\}} (b_j - A_j x_j)\}$.

Definition 1 (GNE). $\boldsymbol{x}^*$ is a GNE if

$$J_i(x_i^*, \boldsymbol{x}_{-i}^*) \leqslant \inf\{J_i(y_i, \boldsymbol{x}_{-i}^*) \mid (y_i, \boldsymbol{x}_{-i}^*) \in \boldsymbol{X}\} \tag{3}$$

for each player $i \in \mathcal{V}$.

Assumption 1. The non-empty set $\boldsymbol{X}$ satisfies Slater's constraint qualification. For each $i \in \mathcal{V}$, the non-empty set Ω_i is closed and convex. The function J_i is continuous, and $J_i(\cdot, \boldsymbol{x}_{-i})$ is convex and continuously differentiable for every $\boldsymbol{x}_{-i}$.

This paper discusses variational GNE (v-GNE) [30], in which Lagrangian multipliers are the same (i.e., $\lambda_1^* = \lambda_2^* = \cdots = \lambda_N^*$) for all the agents.

Let $F(\boldsymbol{x})$ represent the pseudo-gradient mapping with the form of $F(\boldsymbol{x}) := \mathrm{col}(\{\nabla_{x_i} J_i(x_i, \boldsymbol{x}_{-i})\}_{i \in \mathcal{V}})$. This paper is concerned with a v-GNE [30], and $\boldsymbol{x}^* \in \boldsymbol{X}$ is a v-GNE of game (2) if it is the solution of

$$\mathrm{VI}(F, \boldsymbol{X}): \quad \langle F(\boldsymbol{x}^*), \boldsymbol{x} - \boldsymbol{x}^* \rangle \geqslant 0, \ \boldsymbol{x} \in \boldsymbol{X}. \tag{4}$$

In addition, $\boldsymbol{x}^*$ solves $\mathrm{VI}(F, \boldsymbol{X})$ iff $\exists$ a $\lambda^* \in \mathbb{R}^m$ satisfying the following Karush-Kuhn-Tucker (KKT) conditions:

$$(\forall i \in \mathcal{V}): \begin{cases} \mathbf{0}_{n_i} \in \nabla_{x_i} J_i(x_i^*, \boldsymbol{x}_{-i}^*) + A_i^{\mathrm{T}} \lambda^* + N_{\Omega_i}(x_i^*), \\ \mathbf{0}_m \in -(A\boldsymbol{x}^* - b) + N_{\mathbb{R}_+^m}(\lambda^*), \end{cases} \tag{5}$$

where $A := [A_1, A_2, \ldots, A_N]$ and $b := \sum_{i=1}^N b_i$. By [31, Corollary 2.2.6], under Assumption 1, there exists a solution for $\mathrm{VI}(F, \boldsymbol{X})$ (4); and by [30, Theorem 4.8], $\boldsymbol{x}^*$ is the v-GNE of the game (2) if ($\boldsymbol{x}^*$, λ^*) satisfies the conditions (5).

Assumption 2. Suppose that F is strongly monotone, i.e., for $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathrm{dom}(F)$, $\langle \boldsymbol{x} - \boldsymbol{y}, F(\boldsymbol{x}) - F(\boldsymbol{y}) \rangle \geqslant \mu \|\boldsymbol{x} - \boldsymbol{y}\|^2$ for some $\mu > 0$, and is Lipschitz continuous, i.e. for $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{dom}(F)$, $\|F(\boldsymbol{x}) - F(\boldsymbol{y})\| \leqslant \mu_\ell \|\boldsymbol{x} - \boldsymbol{y}\|$ for some $\mu_\ell > 0$.

The above assumptions are commonly used. By [32, Theorem 2.3.3], under Assumptions 1 and 2, the game (2) has a unique v-GNE.

2.2 Communication network

Let an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ represent information exchange among agents, in which the set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ involves M edges. If agent i can send information to j, then let $(i, j) \in \mathcal{E}$ and j is called a neighbor of i. The neighbor set of agent i is defined as $\mathcal{N}_i := \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$. Graph $\mathcal{G}$ is undirected if $(i, j) \in \mathcal{E} \Leftrightarrow (j, i) \in \mathcal{E}$. Let a path from agent i to j be a sequence of edges $(i, p_1), (p_1, p_2), \ldots, (p_m, j)$ in $\mathcal{G}$ with agents p_s $(s = 1, 2, \ldots, m)$. The undirected graph $\mathcal{G}$ is connected iff there exists a path between any two agents.

For graph $\mathcal{G}$, let $W = [w_{i,j}] \in \mathbb{R}^{N \times N}$ and $L = [l_{i,j}] \in \mathbb{R}^{N \times N}$ be weighted adjacency and Laplacian matrices, respectively. To be specific, define $W = [w_{i,j}]$ with $w_{i,j} > 0$ if $j \in \mathcal{N}_i$ and $w_{i,j} = 0$ otherwise. Let $W = W^{\mathrm{T}}$. Define $L = [l_{i,j}]$ with $l_{i,j} = -w_{i,j}$ if $i \neq j$ and $l_{i,i} = \sum_{j=1}^N w_{i,j}$. Next, the edges in $\mathcal{E}$ are labeled as e_l for $l \in \{1, 2, \ldots, M\}$ and an arbitrary but fixed orientation is assigned for each edge. Then, let $D = [d_{l,i}] \in \mathbb{R}^{M \times N}$ mean the weighted incidence matrix, where its (l, i)-entry is 0 if edge e_l and agent i are not incident; otherwise, it is $\sqrt{w_{i,j}}$ or $-\sqrt{w_{i,j}}$. In this regard, $L = D^{\mathrm{T}} D$.

Assumption 3. Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is assumed to be undirected and connected.

It is well known that, in the partial-decision information setting, it is hard for each agent to access to the full decisions $\boldsymbol{x}_{-i}$ of the others. In this scenario, agent i estimates agent j's decision that is defined as $\boldsymbol{x}_j^i$, and let $\boldsymbol{x}_{-i}^i := \mathrm{col}(\boldsymbol{x}_1^i, \cdots \boldsymbol{x}_{i-1}^i, \boldsymbol{x}_{i+1}^i, \ldots, \boldsymbol{x}_N^i)$ and $\boldsymbol{x}^i = \mathrm{col}(\boldsymbol{x}_j^i)_{j\in\mathcal{V}} \in \mathbb{R}^n$. Let $\lambda_i \in \mathbb{R}_+^m$ be the local estimate for λ^* in (5). Furthermore, the variables x_i, $\boldsymbol{x}_j^i$, $\boldsymbol{x}_{-i}^i$, $\boldsymbol{x}^i$, and λ_i at iteration (time) k are represented as $x_{i,k}$, $\boldsymbol{x}_{j,k}^i$, $\boldsymbol{x}_{-i,k}^i$, $\boldsymbol{x}_k^i$, and $\lambda_{i,k}$.

Traditional GNE seeking algorithms developed in partial-decision information setting [9] usually require that each agent i should transmit its full current state $\varpi_k^i := (\boldsymbol{x}_k^i, \lambda_{i,k})$ to its neighbors at every periodic sampling instant $k \in \mathbb{N}$. Such a kind of communication mode is likely to lead to a lot of unnecessary information passing through the communication network, thereby inevitably consuming a great deal of resources and increasing communication burden. As such, this paper aims to develop a communication-efficient algorithm that combines the advantages of ETC and CD-based schemes so as to mitigate the communication burden, while guaranteeing the convergence to the v-GNE.

3 Distributed ETCD-based algorithm

For the efficient communication and security purposes, an ETCD-based communication scheme is proposed in this section, which takes full advantages of the CD-based scheme and ETC while guaranteeing asymptotic convergence to the NE.

3.1 Design CD-based scheme

A coding algorithm of agent i is given as follows:

$$C_i : \begin{cases} \chi_k^i = Q\left(\frac{1}{s_{k-1}}(\boldsymbol{x}_k^i - \boldsymbol{\zeta}_{k-1}^i)\right), \\ \varphi_{i,k} = Q\left(\frac{1}{s_{k-1}}(\lambda_{i,k} - \eta_{i,k-1})\right), \\ \boldsymbol{\zeta}_k^i = s_{k-1}\chi_k^i + \boldsymbol{\zeta}_{k-1}^i,\ \boldsymbol{\zeta}_0^i = 0, \\ \eta_{i,k} = s_{k-1}\varphi_{i,k} + \eta_{i,k-1},\ \eta_{i,0} = 0, \end{cases} \tag{6}$$

where $\boldsymbol{\zeta}_k^i := \mathrm{col}(\boldsymbol{\zeta}_{1,k}^i, \ldots, \boldsymbol{\zeta}_{N,k}^i)$ and $\eta_{i,k}$ are auxiliary variables introduced to obtain χ_k^i and $\varphi_{i,k}$. χ_k^i and $\varphi_{i,k}$ are the codeword to be sent to agent $j \in \mathcal{N}_i$. For $\gamma = [\gamma_1, \ldots, \gamma_n]^{\mathrm{T}}$, the uniform quantization function is defined as $Q(\gamma) := [q(\gamma_1), \ldots, q(\gamma_n)]^{\mathrm{T}}$ with

$$q(\gamma_i) = \begin{cases} d, & \frac{2d-1}{2} < \gamma_i \leqslant \frac{2d+1}{2}, \\ -q(-\gamma_i), & \gamma_i < -\frac{1}{2}, \end{cases} \tag{7}$$

where $d = 0, 1, \ldots$. For such a quantizer, $\|\gamma - Q(\gamma)\|_\infty \leqslant \frac{1}{2}$. Here, s_k is a scaling function that is used to adjust the quantization error, which will be determined subsequently.

A decoding algorithm of agent i is designed as follows:

$$D_{i\to j} : \begin{cases} \bar{\boldsymbol{x}}_k^{ji} = s_{k-1}\chi_k^j + \bar{\boldsymbol{x}}_{k-1}^{ji}, & \bar{\boldsymbol{x}}_0^{ji} = 0, \\ \bar{\lambda}_{j,k}^i = s_{k-1}\varphi_{j,k} + \bar{\lambda}_{j,k-1}^i,\ \bar{\lambda}_{j,0}^i = 0, \end{cases} \tag{8}$$

where $\bar{\boldsymbol{x}}_k^{ji} := \mathrm{col}(\bar{\boldsymbol{x}}_{1,k}^{ji}, \ldots, \bar{\boldsymbol{x}}_{N,k}^{ji})$ and $\bar{\lambda}_{j,k}^i$ are the states obtained after decryption.

3.2 Design ETCD scheme

An ETCD scheme is developed in this subsection to schedule when agent i's data are coded as $\bar{\varpi}_k^i := (\chi_k^i, \varphi_{i,k})$ and transmitted to its neighbors. After being received, $(\chi_k^i, \varphi_{i,k})$ are decoded by agent $j \in \mathcal{N}_i$ to obtain the data $(\bar{\boldsymbol{x}}_k^{ij}, \bar{\lambda}_{i,k}^j)$. Under the ETCD scheme, if the coding-decoding and communication process is triggered at instant k for agent i, then k is called its triggering instant.

For agent $i \in \mathcal{V}$, its l-th triggering instant is defined as k_l^i ($l \in \mathbb{N}$) and let the set $\mathfrak{N}^i := \{k_0^i, k_1^i, \ldots\}$ include all of its triggering instants. Without loss of generality, we suppose that each agent's data are

Algorithm 1 Distributed ETCD scheme

Initialization: Set $x_{i,0} \in \Omega_i$, $\boldsymbol{x}^i_{-i,0} \in \mathbb{R}^{n-n_i}$, $\boldsymbol{\zeta}^i_0 = \bar{\boldsymbol{x}}^{ij}_0 = \mathbf{0}_n$, $\eta_{i,0} = \bar{\lambda}^j_{i,0} = z_{i,0} = \mathbf{0}_m$, $\lambda_{i,0} \in \mathbb{R}^m_+$, $\forall i \in \mathcal{V}$, $j \in \mathcal{N}_i$.
Iteration:
while $k \geqslant 1$ **do**
  if $k \in \mathfrak{N}^i$, **then**
    Coding:
    $\chi^i_k = Q(\frac{1}{s_{k-1}}(\boldsymbol{x}^i_k - \boldsymbol{\zeta}^i_{k-1}))$, $\varphi_{i,k} = Q(\frac{1}{s_{k-1}}(\lambda_{i,k} - \eta_{i,k-1}))$, $\boldsymbol{\zeta}^i_k = s_{k-1}\chi^i_k + \boldsymbol{\zeta}^i_{k-1}$, $\eta_{i,k} = s_{k-1}\varphi_{i,k} + \eta_{i,k-1}$;
    Send χ^i_k and $\varphi_{i,k}$ to neighboring agents $j \in \mathcal{N}_i$;
    Decoding: for agents $j \in \mathcal{N}_i$, $\bar{\boldsymbol{x}}^{ij}_k = s_{k-1}\chi^i_k + \bar{\boldsymbol{x}}^{ij}_{k-1}$, $\bar{\lambda}^j_{i,k} = s_{k-1}\varphi_{i,k} + \bar{\lambda}^j_{i,k-1}$;
  else
    $\boldsymbol{\zeta}^i_k = \boldsymbol{\zeta}^i_{k-1}$, $\eta_{i,k} = \eta_{i,k-1}$, $\bar{\boldsymbol{x}}^{ij}_k = \bar{\boldsymbol{x}}^{ij}_{k-1}$, $\bar{\lambda}^j_{i,k} = \bar{\lambda}^j_{i,k-1}$;
  end if
end while

Algorithm 2 Distributed ETCD-based v-GNE seeking algorithm

Initialization: Set $x_{i,0} \in \Omega_i$, $\boldsymbol{x}^i_{-i,0} \in \mathbb{R}^{n-n_i}$, $\boldsymbol{\zeta}^i_0 = \bar{\boldsymbol{x}}^{ij}_0 = \mathbf{0}_n$, $\eta_{i,0} = \bar{\lambda}^j_{i,0} = z_{i,0} = \mathbf{0}_m$, $\lambda_{i,0} \in \mathbb{R}^m_+$, $\forall i \in \mathcal{V}$, $j \in \mathcal{N}_i$;
Iteration:
while $k \geqslant 1$ **do**
  Receive $\boldsymbol{\zeta}^i_k$, $\eta_{i,k}$, $\bar{\boldsymbol{x}}^{ji}_k$, and $\bar{\lambda}^i_{j,k}$ from Algorithm 1;
  Strategy update:
$$x_{i,k+1} = \mathrm{proj}_{\Omega_i}(x_{i,k} - \tau_i(\nabla J_i(x_{i,k}, \boldsymbol{x}^i_{-i,k}) + A_i^{\mathrm{T}}\lambda_{i,k} + c\textstyle\sum_{j\in\mathcal{N}_i} w_{i,j}(\boldsymbol{\zeta}^i_{i,k} - \bar{\boldsymbol{x}}^{ji}_{i,k})));$$
$$\boldsymbol{x}^i_{-i,k+1} = \boldsymbol{x}^i_{-i,k} - \tau_i c\textstyle\sum_{j\in\mathcal{N}_i} w_{i,j}(\boldsymbol{\zeta}^i_{-i,k} - \bar{\boldsymbol{x}}^{ji}_{-i,k});$$
$$z_{i,k+1} = z_{i,k} + \delta\textstyle\sum_{j\in\mathcal{N}_i} w_{i,j}(\eta_{i,k} - \bar{\lambda}^i_{j,k});$$
$$\lambda_{i,k+1} = \mathrm{proj}_{\mathbb{R}^m_+}(\lambda_{i,k} + \sigma_i(A_i(2x_{i,k+1} - x_{i,k}) - b_i - (2z_{i,k+1} - z_{i,k})));$$
  $k \leftarrow k+1$;
end while
Here $\boldsymbol{\zeta}^i_{-i,k} := \mathrm{col}(\boldsymbol{\zeta}^i_{1,k}, \ldots, \boldsymbol{\zeta}^i_{i-1,k}, \boldsymbol{\zeta}^i_{i+1,k}, \ldots, \boldsymbol{\zeta}^i_{N,k})$ and $\bar{\boldsymbol{x}}^{ji}_{-i,k} := \mathrm{col}(\bar{\boldsymbol{x}}^{ji}_{1,k}, \ldots, \bar{\boldsymbol{x}}^{ji}_{i-1,k}, \bar{\boldsymbol{x}}^{ji}_{i+1,k}, \ldots, \bar{\boldsymbol{x}}^{ji}_{N,k})$.

coded and sent to its neighbors at $k = 0$, that is, $k^i_0 = 0$ for all $i \in \mathcal{V}$. Then, the instant k is the triggering one for agent i (i.e., $k \in \mathfrak{N}^i$) only if

$$\|\boldsymbol{x}^i_k - \boldsymbol{\zeta}^i_{k-1}\|^2_\infty + \|\lambda_{i,k} - \eta_{i,k-1}\|^2_\infty \geqslant (B^i_k)^2, \tag{9}$$

where $B^i_k > 0$ is a threshold to be determined later. The ETCD-based scheme is presented in Algorithm 1.

It is worthwhile to mention that the introduction of the ETCD scheme combined with the quantization process adds uncertainties into the released data, thereby enhancing the security of the released data and meanwhile releasing the communication burden.

Assumption 4. For all $i \in \mathcal{V}$, $B^i_k \leqslant B_k$ ($k \in \mathbb{N}$) and $\sum_{k=0}^{\infty} B_k < \infty$.

3.3 Design ETCD-based v-GNE seeking algorithm

In the following, a distributed ETCD-based algorithm is developed in Algorithm 2 for seeking the v-GNE.

In Algorithm 2, $c > 0$ is design parameters, τ_i, δ, $\sigma_i > 0$ are fixed step-sizes of agent i, and $W = [w_{i,j}]$ is the weighted adjacency matrix of graph $\mathcal{G}$.

Similar to [9, Eqs. (13) and (14)], we define $\mathcal{M} := \mathrm{diag}((\mathcal{M}_i)_{i\in\mathcal{V}})$ and $\mathcal{T} := \mathrm{diag}((\mathcal{T}_i)_{i\in\mathcal{V}})$ with

$$\mathcal{M}_i := \begin{bmatrix} \mathbf{0}_{n_i\times n_{<i}} & I_{n_i} & \mathbf{0}_{n_i\times n_{>i}} \end{bmatrix}, \tag{10}$$

$$\mathcal{T}_i := \begin{bmatrix} I_{n_{<i}} & \mathbf{0}_{n_{<i}\times n_i} & \mathbf{0}_{n_{<i}\times n_{>i}} \\ \mathbf{0}_{n_{>i}\times n_{<i}} & \mathbf{0}_{n_{>i}\times n_i} & I_{n_{>i}} \end{bmatrix}, \tag{11}$$

where $n_{<i} := \sum_{j<i,j\in\mathcal{V}} n_j$ and $n_{>i} := \sum_{j>i,j\in\mathcal{V}} n_j$. Let $\boldsymbol{x} := \mathrm{col}(x_i)_{i\in\mathcal{V}} \in \mathbb{R}^n$ and $\boldsymbol{x} := \mathrm{col}(\boldsymbol{x}^i)_{i\in\mathcal{V}} \in \mathbb{R}^{nN}$, then $\boldsymbol{x} = \mathcal{M}^{\mathrm{T}}\boldsymbol{x} + \mathcal{T}^{\mathrm{T}}\mathcal{T}\boldsymbol{x}$.

Lemma 1. Under the ETCD-based scheme (9), the following equality always holds:

$$\|\boldsymbol{x}^i_k - \boldsymbol{\zeta}^i_k\|^2_\infty + \|\lambda_{i,k} - \eta_{i,k}\|^2_\infty \leqslant \max\left\{(B^i_k)^2, \frac{s^2_{k-1}}{2}\right\}.$$

Proof. If k is the triggering instant, i.e., $k \in \mathfrak{N}^i$, then the codeword $\bar{\varpi}_k^i = (\chi_k^i, \varphi_{i,k})$ to the neighboring agents, and $\boldsymbol{\zeta}_k^i$ and $\eta_{i,k}$ are updated with (6). It follows from (6) that

$$\|\boldsymbol{\zeta}_k^i - \boldsymbol{x}_k^i\|_\infty = \|s_{k-1}\chi_k^i + \boldsymbol{\zeta}_{k-1}^i - \boldsymbol{x}_k^i\|_\infty = \left\| s_{k-1}\left(\chi_k^i - \frac{1}{s_{k-1}}(\boldsymbol{x}_k^i - \boldsymbol{\zeta}_{k-1}^i)\right)\right\|_\infty \leqslant \frac{s_{k-1}}{2}$$

and similarly, $\|\eta_{i,k} - \lambda_{i,k}\|_\infty \leqslant \frac{s_{k-1}}{2}$. If k is not the triggering instant, then $\boldsymbol{\zeta}_k^i = \boldsymbol{\zeta}_{k-1}^i$ and $\eta_{i,k} = \eta_{i,k-1}$, which implies that

$$\|\boldsymbol{x}_k^i - \boldsymbol{\zeta}_k^i\|_\infty^2 + \|\lambda_{i,k} - \eta_{i,k}\|_\infty^2 \leqslant (B_k^i)^2. \tag{12}$$

For $\forall i \in \mathcal{V}$, let $\bar{\boldsymbol{\epsilon}}_k^x := \mathrm{col}(\bar{e}_{i,k}^x)_{i\in\mathcal{V}}$ and $\bar{\boldsymbol{\epsilon}}_k^\lambda := \mathrm{col}(\bar{e}_{i,k}^\lambda)_{i\in\mathcal{V}}$ with

$$\bar{e}_{i,k}^x := \boldsymbol{\zeta}_k^i - \boldsymbol{x}_k^i, \quad \bar{\boldsymbol{e}}_{i,k}^x := \sum_{j\in\mathcal{N}_i} w_{i,j}\left(\bar{e}_{i,k}^x - \bar{e}_{j,k}^x\right), \quad \text{and} \quad \bar{e}_{i,k}^\lambda := \eta_{i,k} - \lambda_{i,k}.$$

Let $\boldsymbol{x}_k := \mathrm{col}(x_{i,k})_{i\in\mathcal{V}}$, $\boldsymbol{x}_k := \mathrm{col}(\boldsymbol{x}_k^i)_{i\in\mathcal{V}}$, $\boldsymbol{\lambda}_k := \mathrm{col}(\lambda_{i,k})_{i\in\mathcal{V}}$, $\boldsymbol{z}_k := \mathrm{col}(z_{i,k})_{i\in\mathcal{V}}$, $\bar{\boldsymbol{\epsilon}}_k^x := \mathrm{col}(\bar{\boldsymbol{e}}_{i,k}^x)_{i\in\mathcal{V}}$, and $\bar{\boldsymbol{\epsilon}}_k^\lambda := \mathrm{col}(\bar{e}_{i,k}^\lambda)_{i\in\mathcal{V}}$. Then, Algorithm 2 can be written as follows:

$$\boldsymbol{x}_{k+1} = \mathrm{proj}_\Omega\left(\boldsymbol{x}_k - \boldsymbol{\tau}_x\left(\boldsymbol{F}(\boldsymbol{x}_k) + \boldsymbol{A}^\mathrm{T}\boldsymbol{\lambda}_k + c\mathcal{M}\boldsymbol{L}_x\boldsymbol{x}_k + c\mathcal{M}\bar{\boldsymbol{\epsilon}}_k^x\right)\right), \tag{13}$$

$$\mathcal{T}\boldsymbol{x}_{k+1} = \mathcal{T}\boldsymbol{x}_k - \boldsymbol{\tau}_s c\left(\mathcal{T}\boldsymbol{L}_x\boldsymbol{x}_k + \mathcal{T}\bar{\boldsymbol{\epsilon}}_k^x\right), \tag{14}$$

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \delta\left(\boldsymbol{L}_\lambda\boldsymbol{\lambda}_k + \boldsymbol{L}_\lambda\bar{\boldsymbol{\epsilon}}_k^\lambda\right), \tag{15}$$

$$\boldsymbol{\lambda}_{k+1} = \mathrm{proj}_{\mathbb{R}_+^{mN}}\left(\boldsymbol{\lambda}_k - \boldsymbol{\sigma}\left(\boldsymbol{A}(2\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) - \boldsymbol{b} - (2\boldsymbol{z}_{k+1} - \boldsymbol{z}_k)\right)\right), \tag{16}$$

where $\boldsymbol{A} = \mathrm{diag}(A_i)_{i\in\mathcal{V}}$, $\boldsymbol{b} = \mathrm{diag}(b_i)_{i\in\mathcal{V}}$, $\boldsymbol{L}_x = L \otimes I_n$, $\boldsymbol{L}_\lambda = L \otimes I_m$, $\boldsymbol{\tau}_x = \mathrm{diag}(\tau_i I_{n_i})_{i\in\mathcal{V}}$, $\boldsymbol{\tau}_s = \mathrm{diag}(\tau_i I_{n-n_i})_{i\in\mathcal{V}}$, and $\boldsymbol{\sigma} = \mathrm{diag}(\sigma_i I_m)_{i\in\mathcal{V}}$. Let $\boldsymbol{F}(\boldsymbol{x}) = \mathrm{col}(\nabla_{x_i} J_i(x_i, \boldsymbol{x}_{-i}^i))_{i\in\mathcal{V}}$ be the extend pseudo-gradient mapping, and obviously, $\boldsymbol{F}(1_N \otimes \boldsymbol{x}) = F(\boldsymbol{x})$.

Assumption 5. Suppose that the mapping $\boldsymbol{F}$ is Lipschitz continuous, i.e., $\exists\, \mu_0 > 0$, $\|\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}')\| \leqslant \mu_0\|\boldsymbol{x} - \boldsymbol{x}'\|$ for all pairs $\boldsymbol{x}$ and $\boldsymbol{x}'$.

3.4 Convergence analysis

In this subsection, we introduce a forward-backward iteration to rewrite Algorithm 2 and then find zeros of the sum of two operators.

For edges of graph $\mathcal{G}$, define the auxiliary variables $\boldsymbol{y} := \mathrm{col}(y_l)_{l\in\mathcal{E}'}$ with $y_l \in \mathbb{R}^m$ and $\mathcal{E}' := \{1, \ldots, M\}$. By defining $\boldsymbol{\omega} = \mathrm{col}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) \in \boldsymbol{\Omega}$ with $\boldsymbol{\Omega} := \mathbb{R}^{nN} \times \mathbb{R}^{mM} \times \mathbb{R}_+^{mN}$, we introduce two operators $\mathfrak{E}$ and $\mathfrak{F}$ as

$$\mathfrak{E}: \boldsymbol{\omega} \rightarrow \begin{bmatrix} \mathcal{M}^\mathrm{T}\boldsymbol{F}(\boldsymbol{x}) + c\boldsymbol{L}_x\boldsymbol{x} \\ \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix},$$

$$\mathfrak{F}: \boldsymbol{\omega} \rightarrow \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \mathcal{M}^\mathrm{T}\boldsymbol{A}^\mathrm{T} \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{D}_\lambda \\ -\boldsymbol{A}\mathcal{M} & \boldsymbol{D}_\lambda^\mathrm{T} & \boldsymbol{0} \end{bmatrix}\boldsymbol{\omega} + \begin{bmatrix} \mathcal{M}^\mathrm{T}N_\Omega(\mathcal{M}\boldsymbol{x}) \\ \boldsymbol{0} \\ N_{\mathbb{R}_+^{mN}}(\boldsymbol{\lambda}) \end{bmatrix}, \tag{17}$$

where $N_\Omega(\mathcal{M}\boldsymbol{x}) = N_\Omega(\boldsymbol{x}) = \Pi_{i=1}^N N_{\Omega_i}(x_i)$ and $N_{\mathbb{R}_+^{mN}}(\boldsymbol{\lambda}) = \Pi_{i=1}^N N_{\mathbb{R}_+^m}(\lambda_i)$. Recall the incidence matrix D that satisfies $L = D^\mathrm{T}D$, and then define $\boldsymbol{D}_\lambda = D \otimes I_m$ and $\boldsymbol{L}_\lambda = \boldsymbol{D}_\lambda^\mathrm{T}\boldsymbol{D}_\lambda$. Then, we define Ψ as

$$\Psi = \begin{bmatrix} \boldsymbol{\tau}^{-1} & \boldsymbol{0} & -\mathcal{M}^\mathrm{T}\boldsymbol{A}^\mathrm{T} \\ \boldsymbol{0} & \boldsymbol{\delta}^{-1} & \boldsymbol{D}_\lambda \\ -\boldsymbol{A}\mathcal{M} & \boldsymbol{D}_\lambda^\mathrm{T} & \boldsymbol{\sigma}^{-1} \end{bmatrix}, \tag{18}$$

where $\boldsymbol{\tau} = \mathrm{diag}(\tau_i I_n)_{i\in\mathcal{V}}$, $\boldsymbol{\tau}^{-1} = \mathrm{diag}(\tau_i^{-1} I_n)_{i\in\mathcal{V}}$, $\boldsymbol{\delta}^{-1} = \mathrm{diag}(\delta^{-1} I_m)_{i\in E}$, and $\boldsymbol{\sigma}^{-1} = \mathrm{diag}(\sigma_i^{-1} I_m)_{i\in\mathcal{V}}$.

Lemma 2. If Assumptions 1–3 hold, then the following statements hold:

(i) $\mathrm{zer}(\mathfrak{E}+\mathfrak{F})\neq\emptyset$.

(ii) If $\mathrm{col}(\boldsymbol{x}^*,\boldsymbol{y}^*,\boldsymbol{\lambda}^*)\in\mathrm{zero}(\mathfrak{E}+\mathfrak{F})$, then $\boldsymbol{x}^*=1_N\otimes\boldsymbol{x}^*$ and $\boldsymbol{\lambda}^*=1_N\otimes\lambda^*$, and $(\boldsymbol{x}^*,\lambda^*)$ satisfies the KKT conditions in (5). Therefore, $\boldsymbol{x}^*$ is the v-GNE of the game (2).

Proof. The proof is analogous to [9, Theorem 1], and is thus omitted here.

Lemma 3. Let $\boldsymbol{\omega}_k=\mathrm{col}(\boldsymbol{x}_k,\boldsymbol{y}_k,\boldsymbol{\lambda}_k)$ with $\mathfrak{E}$, $\mathfrak{F}$, Ψ, given in (17) and (18). Suppose $\Psi\succ 0$ and $\Psi^{-1}\mathfrak{F}$ are maximally monotone. Then, the sequence $(\boldsymbol{x}_k,\boldsymbol{z}_k,\boldsymbol{\lambda}_k)_{k\in\mathbb{N}}$ generated by Algorithm 2 with any initial condition $\mathrm{col}(\boldsymbol{x}_0,\boldsymbol{z}_0=\mathbf{0}_{mN},\boldsymbol{\lambda}_0)$ can be rewritten as

$$\boldsymbol{\omega}_{k+1}=T_2\circ(T_1\boldsymbol{\omega}_k-\Psi^{-1}\boldsymbol{\epsilon}_k^q),\ \forall k\in\mathbb{N} \tag{19}$$

with $\boldsymbol{\omega}_0=\mathrm{col}(\boldsymbol{x}_0,\boldsymbol{y}_0=\mathbf{0}_{mM},\boldsymbol{\lambda}_0)$. Here $\boldsymbol{\epsilon}_k^q=\mathrm{col}(c\bar{\boldsymbol{\epsilon}}_k^x,-\boldsymbol{D}_\lambda\bar{\boldsymbol{\epsilon}}_k^\lambda,\mathbf{0})$, $T_1:=\mathrm{Id}-\Psi^{-1}\mathfrak{E}$ and $T_2:=(\mathrm{Id}+\Psi^{-1}\mathfrak{F})^{-1}$ with $\boldsymbol{D}_\lambda$ given in (17), and $\bar{\boldsymbol{\epsilon}}_k^x$ and $\bar{\boldsymbol{\epsilon}}_k^\lambda$ given in (13)–(15).

Proof. See Appendix B.

Notice that the iteration (19) involves an unavoidable error $\Psi^{-1}\boldsymbol{\epsilon}_k^q$ that is called inexact forward-backward iteration.

Next, Lemma 4 can be obtained by applying Gershgorin circular theorem [33].

Lemma 4. For any $\vartheta>0$, $\Psi-\vartheta I_{Nn+Mm+Nm}\succ 0$, if we choose τ_i, δ, and σ_i in Algorithm 2 with $0<\tau_i<(\|A_i^{\mathrm{T}}\|_\infty+\vartheta)^{-1}$, $0<\delta<(2+\vartheta)^{-1}$, and $0<\sigma_i<(\|A_i^{\mathrm{T}}\|_\infty+2+\vartheta)^{-1}$.

Proof. It follows from Gershgorin circular theorem and (18) that each eigenvalue λ_p of $\Psi-\vartheta I_{Nn+Mm+Nm}$ lies within at least one of the Gershgorin discs, i.e., $|\lambda_p-{\tau_i}^{-1}-\vartheta|\leqslant\|A_i^{\mathrm{T}}\|_\infty$, $|\lambda_p-\delta^{-1}-\vartheta|\leqslant 2$, and $|\lambda_p-{\sigma_i}^{-1}-\vartheta|\leqslant\|A_i^{\mathrm{T}}\|_\infty+2$ $(\forall i\in\mathcal{V})$. Obviously, $\lambda_p>0$.

Let $c_0:=\frac{1}{q_2(L)}(\frac{(\mu_\ell+\mu_0)^2}{4\mu}+\mu_0)$ and

$$\Theta:=\begin{bmatrix}\frac{\mu}{N} & -\frac{\mu_\ell+\mu_0}{2\sqrt{N}}\\ -\frac{\mu_\ell+\mu_0}{2\sqrt{N}} & cq_2(L)-\mu_0\end{bmatrix}; \tag{20}$$

then $\Theta\succ 0$ for any $c>c_0$. The minimum eigenvalue of Θ is defined as $\mu_c:=q_1(\Theta)$.

From (6) and (8), one has $\boldsymbol{\zeta}_k^i=\bar{\boldsymbol{x}}_k^{ij}$ and $\eta_{i,k}=\bar{\lambda}_{i,k}^j$ for $j\in\mathcal{N}_i$ with $k=0,1,\ldots$. Then, we can rewrite the iteration (19) as the form of the Krasnosel'skii-Mann fixed-point iteration with errors [34]:

$$\boldsymbol{\omega}_{k+1}=\boldsymbol{\omega}_k+(T_2\circ T_1\boldsymbol{\omega}_k-\boldsymbol{\omega}_k)+\tilde{\boldsymbol{\epsilon}}_k^q, \tag{21}$$

where $\tilde{\boldsymbol{\epsilon}}_k^q:=T_2\circ(T_1\boldsymbol{\omega}_k-\Psi^{-1}\boldsymbol{\epsilon}_k^q)-T_2\circ T_1\boldsymbol{\omega}_k$ and $\boldsymbol{\epsilon}_k^q$ is given in (19).

Lemma 5. The step size τ_i, δ, σ_i are chosen as in Lemma 4. Let $c>c_0$ and $\vartheta>\frac{1}{2\beta}$, where c_0,μ_c in (20) and $\beta\in(0,\mu_c/(\mu_0)^2]$. Under ETCD scheme given in Algorithm 1, if Assumptions 1–5 hold, then $\sum_{k=1}^{\infty}\|\tilde{\boldsymbol{\epsilon}}_k^q\|<\infty$, or equivalently, $\lim_{k\to\infty}\|\tilde{\boldsymbol{\epsilon}}_k^q\|=0$.

Proof. See Appendix C.

Theorem 1. Let $c>c_0$ and $\vartheta>\frac{1}{2\beta}$, where c_0,μ_c are given in (20) and $\beta\in(0,\mu_c/(\mu_0)^2]$, and set the step size τ_i, δ, and σ_i as in Lemma 4. Suppose that Assumptions 1–5 hold. Then, under Algorithm 2 with ETCD scheme given in Algorithm 1 with $\sum_{k=0}^{\infty}s_k<\infty$, the pair $(\boldsymbol{x}_k,\boldsymbol{\lambda}_k)$ converges to a solution of the KKT system (5) and in turn $\boldsymbol{x}_k$ converges to the v-GNE of the game (2).

Proof. See Appendix D.

Remark 1. Algorithm 2 takes full advantages of ETC and CD-based schemes, thereby exhibiting the merit of efficient and secure communication. In Theorem 1, a sufficient condition is established to guarantee the asymptotic convergence of Algorithm 2 to the v-GNE without any error induced by the ETCD-based scheme.

4 Distributed ETCD-based algorithm with overrelaxation

In this section, an overrelaxed version of Algorithm 2 is further proposed to achieve a possible convergence acceleration. The convergence of the algorithm is then discussed.

A fully distributed overrelaxed algorithm under the ETCD-based communication (9) is established for seeking a v-GNE as Algorithm 3.

Let $\widehat{x}_k:=\mathrm{col}(\widehat{x}_{i,k})_{i\in\mathcal{V}}$, $\widehat{\boldsymbol{x}}_k^i=\mathrm{col}(\widehat{\boldsymbol{x}}_{j,k}^i)_{j\in\mathcal{V}}$, $\widehat{\boldsymbol{x}}_k:=\mathrm{col}(\widehat{\boldsymbol{x}}_k^i)_{i\in\mathcal{V}}$, $\widehat{\boldsymbol{\lambda}}_k:=\mathrm{col}(\widehat{\lambda}_{i,k})_{i\in\mathcal{V}}$, and $\widehat{\boldsymbol{z}}_k:=\mathrm{col}(\widehat{z}_{i,k})_{i\in\mathcal{V}}$.

Algorithm 3 Distributed ETCD-based algorithm with overrelaxation

Iteration:

while $k \geqslant 1$ **do**

Receive $\boldsymbol{\zeta}_k^i$, $\eta_{i,k}$, $\bar{\boldsymbol{x}}_k^{ji}$, and $\bar{\lambda}_{j,k}^i$ from the ETCD scheme in Algorithm 1;

(1) Strategy update:

$\widehat{x}_{i,k+1} = \mathrm{proj}_{\Omega_i}(x_{i,k} - \tau_i(\nabla J_i(x_{i,k}, \boldsymbol{x}_{-i,k}^i) + A_i^{\mathrm{T}}\lambda_{i,k} + c\sum_{j\in\mathcal{N}_i} w_{i,j}(\boldsymbol{\zeta}_{i,k}^i - \bar{\boldsymbol{x}}_{i,k}^{ji})))$;

$\widehat{\boldsymbol{x}}_{-i,k+1}^i = \boldsymbol{x}_{-i,k}^i - \tau_i c\sum_{j\in\mathcal{N}_i} w_{i,j}(\boldsymbol{\zeta}_{-i,k}^i - \bar{\boldsymbol{x}}_{-i,k}^{ji})$;

$\widehat{z}_{i,k+1} = z_{i,k} + \delta\sum_{j\in\mathcal{N}_i} w_{i,j}(\eta_{i,k} - \bar{\lambda}_{j,k}^i)$;

$\widehat{\lambda}_{i,k+1} = \mathrm{proj}_{\mathbb{R}_+^m}(\lambda_{i,k} + \sigma_i(A_i(2\widehat{x}_{i,k+1} - x_{i,k}) - b_i - (2\widehat{z}_{i,k+1} - z_{i,k})))$;

(2) Acceleration:

$x_{i,k+1} = x_{i,k} + \alpha(\widehat{x}_{i,k+1} - x_{i,k})$, $\boldsymbol{x}_{-i,k+1}^i = \boldsymbol{x}_{-i,k}^i + \alpha(\widehat{\boldsymbol{x}}_{-i,k+1}^i - \boldsymbol{x}_{-i,k}^i)$;

$z_{i,k+1} = z_{i,k} + \alpha(\widehat{z}_{i,k+1} - z_{i,k})$, $\lambda_{i,k+1} = \lambda_{i,k} + \alpha(\widehat{\lambda}_{i,k+1} - \lambda_{i,k})$;

$k \leftarrow k+1$;

end while

Similar to Lemma 3, Algorithm 3 can be recast (with the change of variables $\boldsymbol{z} = \boldsymbol{D}_\lambda^{\mathrm{T}}\boldsymbol{y}$ and $\widehat{\boldsymbol{z}} = \boldsymbol{D}_\lambda^{\mathrm{T}}\widehat{\boldsymbol{y}}$) in the following form:

$$\begin{cases} \widehat{\boldsymbol{\omega}}_{k+1} = T_2 \circ T_1\boldsymbol{\omega}_k + \tilde{\boldsymbol{\epsilon}}_k^q, \\ \boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k + \alpha(\widehat{\boldsymbol{\omega}}_{k+1} - \boldsymbol{\omega}_k), \end{cases} \tag{22}$$

where $\widehat{\boldsymbol{\omega}}_k := \mathrm{col}(\widehat{\boldsymbol{x}}_k, \widehat{\boldsymbol{y}}_k, \widehat{\boldsymbol{\lambda}}_k)$, $\boldsymbol{\omega}_k = \mathrm{col}(\boldsymbol{x}_k, \boldsymbol{y}_k, \boldsymbol{\lambda}_k)$, $\alpha \in [1,2)$ is a relaxation parameter, and T_1, T_2 are given in (21). That is,

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k + \alpha(T_2 \circ T_1\boldsymbol{\omega}_k + \tilde{\boldsymbol{\epsilon}}_k^q - \boldsymbol{\omega}_k). \tag{23}$$

It follows from Lemma 6 that $T_2 \circ T_1$ is υ-restricted averaged with $\upsilon = \frac{2\beta\vartheta}{4\beta\vartheta - 1}$. Let

$$R := \left(1 - \frac{1}{\upsilon}\right)\mathrm{Id} + \frac{1}{\upsilon}T, \tag{24}$$

and $\bar{\alpha} = \upsilon\alpha$, then by [35, Prop. 4.25], $\mathrm{fix}(R) = \mathrm{fix}(T_2 \circ T_1)$ and R is restricted nonexpansive, i.e., $\|R\boldsymbol{\omega} - R\boldsymbol{\omega}'\|_\Psi^2 \leqslant \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\Psi^2$ for any $\boldsymbol{\omega}$ and for any $\boldsymbol{\omega}' \in V_x$, The iteration (23) can be written as

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k + \widehat{\boldsymbol{\epsilon}}_k), \tag{25}$$

where $\widehat{\boldsymbol{\epsilon}}_k = \frac{1}{\upsilon}\tilde{\boldsymbol{\epsilon}}_k^q$.

Theorem 2. Let $c > c_0$ and $\vartheta > \frac{1}{2\beta}$, where c_0, μ_c are given in (20) and $\beta \in (0, \mu_c/(\mu_0)^2]$, and $1 < \alpha < 2$ in Algorithm 3 satisfying $\frac{2}{\alpha}(1 - \frac{1}{4\beta\vartheta}) > 1$, and set the step size τ_i, δ, and σ_i as in Lemma 4. Suppose that Assumptions 1–5 hold. Then, under Algorithm 3 with the ETCD scheme given in Algorithm 1, the pair $(\boldsymbol{x}_k, \boldsymbol{\lambda}_k)$ converges to a solution of the KKT system (5) and in turn $\boldsymbol{x}_k$ converges to the v-GNE of the game (2).

Proof. Consider any $\boldsymbol{\omega}^* \in \mathrm{zer}(\mathfrak{E} + \mathfrak{F}) = \mathrm{fix}(T_2 \circ T_1) = \mathrm{fix}(R)$ (i.e., $\boldsymbol{\omega}^* \in V_x$ by Lemma 2). From (25), it follows that

$$\begin{aligned} \|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi &= \|\boldsymbol{\omega}_k + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k + \widehat{\boldsymbol{\epsilon}}_k) - \boldsymbol{\omega}^*\|_\Psi \\ &\leqslant \|(1-\bar{\alpha})(\boldsymbol{\omega}_k - \boldsymbol{\omega}^*) + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}^* + \widehat{\boldsymbol{\epsilon}}_k)\|_\Psi \\ &\leqslant (1-\bar{\alpha})\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\| + \bar{\alpha}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi + \bar{\alpha}\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi. \end{aligned} \tag{26}$$

Since R is restricted nonexpansive,

$$\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi \leqslant \|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\| + \bar{\alpha}\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi. \tag{27}$$

Then, $\{\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi\}$ is bounded (i.e., $\sup_{l\in\mathbb{N}}\|\boldsymbol{\omega}_l - \boldsymbol{\omega}^*\| < \infty$), since $\sum_{k=0}^\infty \widehat{\boldsymbol{\epsilon}}_k = \frac{1}{\upsilon}\sum_{k=0}^\infty \tilde{\boldsymbol{\epsilon}}_k^q < \infty$ by Theorem 1 and (23). Similar to Theorem 1, we can further conclude that $\{\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi\}$ converges for any $\boldsymbol{\omega}^* \in \mathrm{fix}(R)$. Next, we consider

$$\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi^2 \leqslant \|(1-\bar{\alpha})(\boldsymbol{\omega}_k - \boldsymbol{\omega}^*) + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}^*)\|_\Psi^2 + \bar{\alpha}^2\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi^2$$

$$
\begin{aligned}
&+ 2\bar{\alpha}\|(1-\bar{\alpha})(\boldsymbol{\omega}_k - \boldsymbol{\omega}^*) + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}^*)\|_\Psi \|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi \\
\leqslant &\|(1-\bar{\alpha})(\boldsymbol{\omega}_k - \boldsymbol{\omega}^*) + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}^*)\|_\Psi^2 + \hat{\epsilon}'_k,
\end{aligned} \tag{28}
$$

where $\hat{\epsilon}'_k := 2\bar{\alpha}\sup_{l\in\mathbb{N}}\|\boldsymbol{\omega}_l - \boldsymbol{\omega}^*\|\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi + \bar{\alpha}^2\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi^2$. In addition,

$$
\begin{aligned}
\|(1-\bar{\alpha})(\boldsymbol{\omega}_k - \boldsymbol{\omega}^*) + \bar{\alpha}(R\boldsymbol{\omega}_k - \boldsymbol{\omega}^*)\|_\Psi^2 =& (1-\bar{\alpha})\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi^2 + \bar{\alpha}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi^2 - (1-\bar{\alpha})\bar{\alpha}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k\|_\Psi^2 \\
\leqslant& \|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi^2 - (1-\bar{\alpha})\bar{\alpha}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k\|_\Psi^2.
\end{aligned}
$$

Substituting the above inequality into (28) yields

$$
\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi^2 \leqslant \|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi^2 + \hat{\epsilon}'_k - (1-\bar{\alpha})\bar{\alpha}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k\|_\Psi^2,
$$

which implies that

$$
(1-\bar{\alpha})\bar{\alpha}\sum_{k=0}^{\infty}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k\|_\Psi^2 \leqslant \|\boldsymbol{\omega}_0 - \boldsymbol{\omega}^*\|_\Psi^2 + \sum_{k=0}^{\infty}\hat{\epsilon}'_k.
$$

Since $\sum_{k=0}^{\infty}\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi \leqslant \sqrt{q_{\max}(\Psi)}\sum_{k=0}^{\infty}\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi < \infty$, then $\sum_{k=0}^{\infty}\|\widehat{\boldsymbol{\epsilon}}_k\|_\Psi^2 < \infty$. In addition, $\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|$ is bounded, hence $\sum_{k=0}^{\infty}\hat{\epsilon}'_k < \infty$. Due to $(1-\bar{\alpha})\bar{\alpha} > 0$, $\sum_{k=1}^{\infty}\|R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k\|_\Psi$ converges and $\lim_{k\to\infty} R\boldsymbol{\omega}_k - \boldsymbol{\omega}_k = 0$ since $\Psi \succ 0$. The following analysis is similar to that in Theorem 1, and is thus omitted.

Remark 2. We notice that there have been some initial and excellent results on continuous-time distributed NE seeking algorithms [15, 36, 37], where the limitation of communication resources is not taken into consideration. Compared with [15,36,37], this paper is concerned with discrete-time distributed algorithms through a digital communication network with limited bandwidth. As such, the existing continuous-time distributed algorithms through perfect communication networks are no longer suitable. This motivates us to develop a new communication-efficient discrete-time algorithm for seeking the NE.

Remark 3. Comparing with existing literature [8–10,13], the novelties of this paper can be summarized: (i) the addressed research problem is new that makes the first attempt to discuss the distributed ETCD-based v-GNE seeking problem; (ii) an inexact forward-backward iteration is constructed to reveal the uncertainties and errors jointly induced by ETC and CD-based scheme; and (iii) the communication-efficient algorithms are developed to significantly reduce communication rounds and improve data security, while guaranteeing the asymptotic convergence to the v-GNE.

5 Numerical simulation

5.1 Nash-cournot game

Consider N firms (agents) that produce a commodity and compete over m markets (i.e., $M_1,\ldots,M_m$) with market capacity constraints see [9]. The production limitation of the firm $i\in\mathcal{V}$ is $x_i\in\Omega_i\subset\mathbb{R}^{n_i}$ and the maximal capacity of the market M_k is r_k $(k=1,\ldots,m)$. Let $r=\mathrm{col}(r_k)_{k=1,\ldots,m}$ and $\boldsymbol{x}=\mathrm{col}(x_i)_{i\in\mathcal{V}}$, and then define the shared affine constraint $A\boldsymbol{x}<r$, where $A=[A_1,\ldots,A_N]\in\mathbb{R}^{m\times n}$ and $n=\sum_{i=1}^{N}n_i$. The matrix $A_i\in\mathbb{R}^{m\times n_i}$ means which markets the firm i joins in. For the p-th column of A_i (i.e., $[A_i]_{:p}$), its k-th element is 1 if the firm i delivers $[x_i]_p$ amount of production to the market M_k; and is 0 otherwise.

Each firm $i\in\mathcal{V}$ aims at minimizing the function $J_i(x_i,\boldsymbol{x}_{-i})$ with the form of

$$
J_i(x_i,\boldsymbol{x}_{-i}) = x_i^{\mathrm{T}}Q_ix_i + q_i^{\mathrm{T}}x_i - P(A\boldsymbol{x})^{\mathrm{T}}A_ix_i, \tag{29}
$$

where $q_i\in\mathbb{R}^{n_i}$ and $Q_i\succ 0\in\mathbb{R}^{n_i\times n_i}$. Let the map $P=\bar{P}-\Xi A\boldsymbol{x}$ be the total supply of each market to the price of the corresponding market, where $\Xi=\mathrm{diag}(\xi_k)_{k=1,\ldots,m}\in\mathbb{R}^{m\times m}$ and $\bar{P}=\mathrm{col}(\bar{P}_k)_{k=1,\ldots,m}\in\mathbb{R}^m$.

5.2 Numerical results

Let $m=3$ and $N=6$ (see Figure 1), and firms are connected in a ring topology. Let the nonzero elements of $W=[w_{i,j}]$ be 1. The local constraint of firm i is defined as $0<x_i<X_i$ with each component of X_i randomly selected from $[1,1.5]$. Let $b_i=\frac{r}{N}$ with r_k randomly drawn from $[0.5,1]$. Q_i is a diagonal matrix and its entries are randomly generated from $[0.1,0.8]$. The elements of q_i, $\bar{P}_k$, and ξ_k are randomly chosen

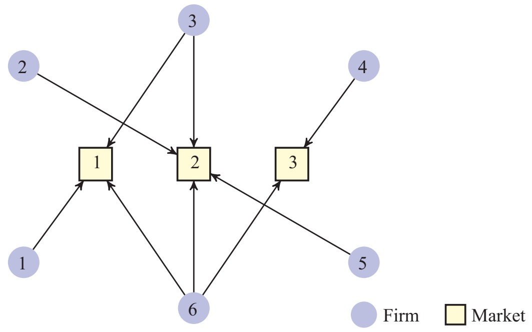

Figure 1 (Color online) Network Nash-Cournot game, where an edge from firm i to market j means that firm i participates in market j.

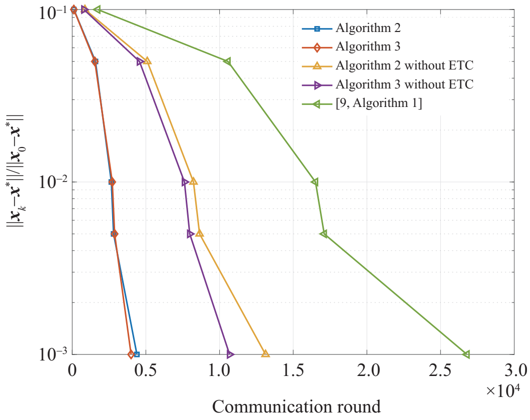

Figure 2 (Color online) Average communication rounds of all firms for reaching $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|/\|\boldsymbol{x}_0-\boldsymbol{x}^*\| < 10^{-1}, 5\times10^{-1}, 10^{-2}, 5\times10^{-2}$, and 10^{-3} under different algorithms with/without ETC.

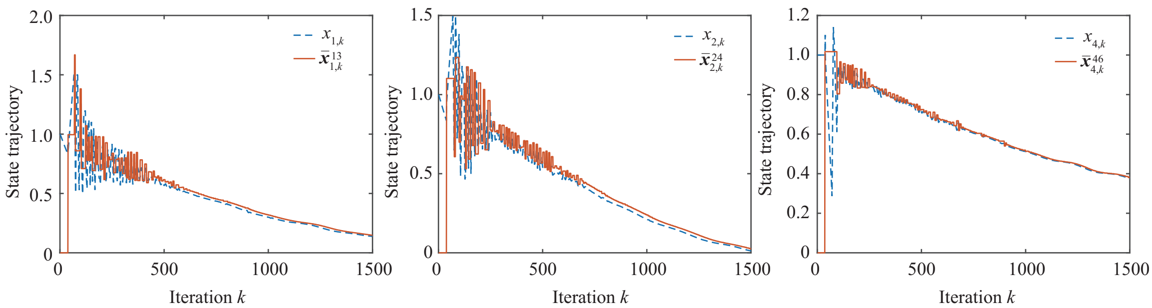

Figure 3 (Color online) State trajectories $x_{1,k}$, $x_{2,k}$, and $x_{4,k}$ and their decoded values $\bar{\boldsymbol{x}}_{1,k}^{13}$, $\bar{\boldsymbol{x}}_{2,k}^{24}$, and $\bar{\boldsymbol{x}}_{4,k}^{46}$, respectively.

from $[0.1, 0.6]$, $[2, 4]$, and $[0.5, 1]$, respectively. We choose $c = 100$, $\tau_i = 0.001$, $\delta = 0.015$, $\sigma_i = 0.016$, $s_k = \frac{2}{k^{1.1}}$, and $B_k^i = B_k = \frac{200}{k^{1.1}}$ for all $i \in \mathcal{V}$. Set $\alpha = 1.1$ in Algorithm 3.

First, we analyze the average communication rounds of all firms to reach targeted convergence precisions (i.e., $\|\boldsymbol{x}_k - \boldsymbol{x}^*\|/\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| < 10^{-1}$, 5×10^{-1}, 10^{-2}, 5×10^{-2}, and 10^{-3}). In this scenario, we compare the performance of Algorithms 2 and 3 with the corresponding cases without using ETC. It can be seen from Figure 2 that Algorithms 2 and 3 (that adopt ETC) can reach a convergence precision in fewer communication rounds than [9, Algorithm 1] and Algorithms 2 and 3 without using ETC (i.e., $B_k^i = 0$). Figure 2 illustrates that the utilization of ETC is capable of significantly saving communication resources to achieve a desired convergence precision: to achieve a convergence precision of around 10^{-1}, Algorithms 2 and 3 need 15 times less communication rounds than [9, Algorithm 1], and around 8 times less than Algorithms 2 and 3 without ETC.

In addition, Figure 3 presents the decision states and their decoded values for all firms under the distributed ETCD-based algorithm (see Algorithm 2), which illustrates that the decoding value can track the actual decision state eventually. Moreover, we discuss the convergence rate of Algorithms 2 and 3 by comparing the termination time (that is, the iteration k when $\|\boldsymbol{x}_k - \boldsymbol{x}^*\|/\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| < 10^{-1}$, 10^{-2}, and 10^{-3} hold for the first time). It can be seen from Figure 4 that Algorithm 3 has a faster convergence than Algorithm 2, and the advantage of Algorithm 3 comes from the introduction of the overrelaxation scheme.

6 Conclusion

This paper has investigated the distributed GNE computation problem in strongly monotone games. For communication efficiency/security purposes, an ETCD-based communication scheme has been developed with taking full advantages of CD-based scheme and ETC. Based on this, distributed communication-

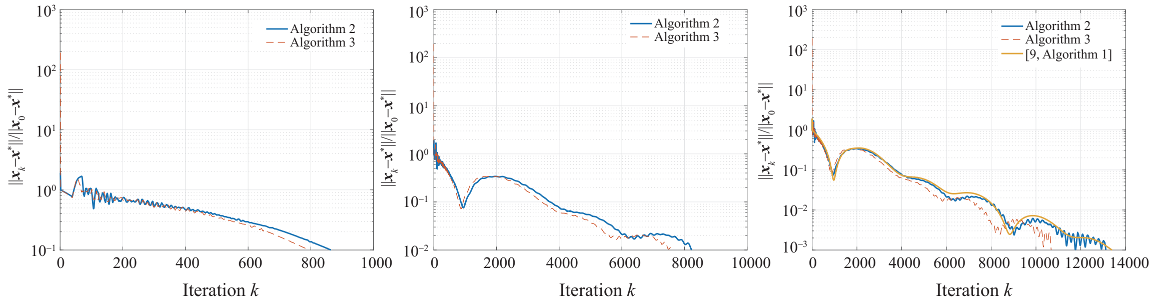

Figure 4 (Color online) Termination time.

efficient GNE seeking algorithms have been accordingly constructed (with overrelaxation), and further analyzed by resorting to an inexact forward-backward iteration. Finally, several numerical simulations have been given to illustrate that the proposed algorithms have capacity of significantly saving communication resources over the state-of-the-art while guaranteeing similar performance. The extension of our results to generalized games under directed communication networks and the convergence rate analysis are left as future research.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62173087, 62176056, 12326311), Young Elite Scientists Sponsorship Program by the China Association for Science and Technology (Grant No. 2021QNRC001), Joint Fund of Ministry of Education for Equipment Pre-research (Grant Nos. 8091B022234, 8091B030723), Fundamental Research Funds for the Central Universities, and Alexander von Humboldt Foundation of Germany. This work was also supported by ZhiShan Youth Scholar Program from Southeast University. The authors would like to thank the anonymous reviewers for the comments that greatly improve the paper. They also appreciate the discussion and suggestions from Dr. Sergio Grammatico.

References

1 Cheng N, Zhang N, Lu N, et al. Opportunistic spectrum access for CR-VANETs: a game-theoretic approach. IEEE Trans Veh Technol, 2013, 63: 237–251
2 Xiong Y H, Gan J R, An B, et al. Optimal electric vehicle fast charging station placement based on game theoretical framework. IEEE Trans Intell Transp Syst, 2017, 19: 2493–2504
3 Maharjan S, Zhu Q Y, Zhang Y, et al. Dependable demand response management in the smart grid: a Stackelberg game approach. IEEE Trans Smart Grid, 2013, 4: 120–132
4 Wang D, Wang Z, Wu Z J, et al. Distributed convex optimization for nonlinear multi-agent systems disturbed by a second-order stationary process over a digraph. Sci China Inf Sci, 2022, 65: 132201
5 Jia L Y, Tong L. Dynamic pricing and distributed energy management for demand response. IEEE Trans Smart Grid, 2016, 7: 1128–1136
6 Li N, Chen L J, Dahleh M A. Demand response using linear supply function bidding. IEEE Trans Smart Grid, 2015, 6: 1827–1838
7 Proskurnikov A V, Matveev A S, Cao M. Opinion dynamics in social networks with hostile camps: consensus vs. polarization. IEEE Trans Automat Contr, 2016, 61: 1524–1536
8 Belgioioso G, Nedic A, Grammatico S. Distributed generalized Nash equilibrium seeking in aggregative games on time-varying networks. IEEE Trans Automat Contr, 2021, 66: 2061–2075
9 Pavel L. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach. IEEE Trans Automat Contr, 2020, 65: 1584–1597
10 Bianchi M, Belgioioso G, Grammatico S. Fast generalized Nash equilibrium seeking under partial-decision information. Automatica, 2022, 136: 110080
11 Xiong M H, Zhang B Y, Yuan D M, et al. Distributed quantized mirror descent for strongly convex optimization over time-varying directed graph. Sci China Inf Sci, 2022, 65: 202202
12 Ye M J, Han Q L, Ding L, et al. Distributed Nash equilibrium seeking in games with partial decision information: a survey. Proc IEEE, 2023, 111: 140–157
13 Yi P, Pavel L. An operator splitting approach for distributed generalized Nash equilibria computation. Automatica, 2019, 102: 111–121
14 Belgioioso G, Grammatico S. Semi-decentralized Nash equilibrium seeking in aggregative games with separable coupling constraints and non-differentiable cost functions. IEEE Control Syst Lett, 2017, 1: 400–405
15 Ye M J, Hu G Q. Distributed Nash equilibrium seeking by a consensus based approach. IEEE Trans Automat Contr, 2017, 62: 4811–4818
16 Gadjov D, Pavel L. A passivity-based approach to Nash equilibrium seeking over networks. IEEE Trans Automat Contr, 2019, 64: 1077–1092
17 de Persis C, Grammatico S. Distributed averaging integral Nash equilibrium seeking on networks. Automatica, 2019, 110: 108548
18 Dimarogonas D V, Frazzoli E, Johansson K H. Distributed event-triggered control for multi-agent systems. IEEE Trans Automat Contr, 2012, 57: 1291–1297
19 Xu W Y, Hu G Q, Ho D W C, et al. Distributed secure cooperative control under denial-of-service attacks from multiple adversaries. IEEE Trans Cybern, 2020, 50: 3458–3467
20 Singh N, Data D, George J, et al. SPARQ-SGD: event-triggered and compressed communication in decentralized optimization. IEEE Trans Automat Contr, 2023, 68: 721–736

21 Xu W Y, He W L, Ho D W C, et al. Fully distributed observer-based consensus protocol: adaptive dynamic event-triggered schemes. Automatica, 2022, 139: 110188
22 Xu W Y, Yang S F, Grammatico S, et al. An event-triggered distributed generalized Nash equilibrium seeking algorithm. In: Proceedings of the 60th IEEE Conference on Decision and Control (CDC), 2021. 4301–4306
23 Gao C, Wang Z D, He X, et al. Sampled-data-based fault-tolerant consensus control for multi-agent systems: a data privacy preserving scheme. Automatica, 2021, 133: 109847
24 Yi P, Hong Y G. Quantized subgradient algorithm and data-rate analysis for distributed optimization. IEEE Trans Control Netw Syst, 2014, 1: 380–392
25 Lei J L, Yi P, Shi G D, et al. Distributed algorithms with finite data rates that solve linear equations. SIAM J Optim, 2020, 30: 1191–1222
26 Li T, Fu M Y, Xie L H, et al. Distributed consensus with limited communication data rate. IEEE Trans Automat Contr, 2010, 56: 279–292
27 Liu S, Li T, Xie L H. Distributed consensus for multiagent systems with communication delays and limited data rate. SIAM J Control Optim, 2011, 49: 2239–2262
28 Zhang Q, Zhang J F. Quantized data–based distributed consensus under directed time-varying communication topology. SIAM J Control Optim, 2013, 51: 332–352
29 Wang L C, Wang Z D, Zhao D, et al. Event-based state estimation under constrained bit rate: an encoding-decoding approach. Automatica, 2022, 143: 110421
30 Facchinei F, Kanzow C. Generalized Nash equilibrium problems. Ann Oper Res, 2010, 175: 177–211
31 Clark L. Nonlinear systems analysis. IEEE Trans Automa Control, 1980, 25: 613–614
32 Facchinei F, Pang J S. Finite-Dimensional Variational Inequalities and Complementarity Problems. Berlin: Springer, 2007
33 Horn R A, Johnson C R. Matrix Analysis. Cambridge: Cambridge University Press, 2012
34 Combettes P L. Quasi-fejérian analysis of some optimization algorithms. In: Proceedings of Inherently Parallel Algorithms in Feasibility and Optimization and their Applications, 2001. 115–152
35 Baischke H H, Combettes P L. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Berlin: Springer, 2010
36 Lu K H, Zhu Q X. Nonsmooth continuous-time distributed algorithms for seeking generalized Nash equilibria of noncooperative games via digraphs. IEEE Trans Cybern, 2022, 52: 6196–6206
37 Lu K H, Jing G S, Wang L. Distributed algorithms for searching generalized Nash equilibrium of noncooperative games. IEEE Trans Cybern, 2019, 49: 2362–2371

Appendix A Monotone operators

The following concepts come from [35]. Let $\mathfrak{E}:\mathbb{R}^n\to 2^{\mathbb{R}^n}$ be a set-valued operator and Id be the identity operator. Let $\mathrm{gra}(\mathfrak{E})=\{(x,u)\in\mathbb{R}^n\times\mathbb{R}^n|u\in\mathfrak{E}(x)\}$, $\mathrm{gra}(\mathfrak{E}^{-1})=\{(u,x)|(x,u)\in\mathrm{gra}(\mathfrak{E})\}$, and $\mathrm{zer}(\mathfrak{E})=\{x|\mathbf{0}\in\mathfrak{E}(x)\}$. $\mathfrak{E}$ is monotone if $\langle x-y,u-v\rangle\geqslant 0$ for $\forall(x,u)$, $\forall(y,v)\in\mathrm{gra}(\mathfrak{E})$, and is maximally monotone if $\mathrm{gra}(\mathfrak{E})$ is not strictly contained in the graph of any other monotone operator. Let $\mathcal{J}_{\mathfrak{E}}=(\mathrm{Id}+\mathfrak{E})^{-1}$, and $\mathrm{fix}(T)$ means the set of fixed points of T. If $\mathfrak{F}$ is single-valued, then $\mathrm{zer}(\mathfrak{E}+\mathfrak{F})=\mathrm{fix}(\mathcal{J}_{\mathfrak{E}}\circ(\mathrm{Id}-\mathfrak{F}))$. Let ∂f denote sub-differential operator and prox_f mean the proximal operator of f. Define the indicator function of Ω as $\iota_\Omega(x)=0$ if $x\in\Omega$ and $\iota_\Omega(x)=\infty$ if $x\notin\Omega$. For a closed convex set Ω, $\mathrm{bd}(\Omega)$ and $\mathrm{int}(\Omega)$ mean the boundary and interior of Ω. $\partial\iota_\Omega$ is the normal cone operator of Ω, i.e., $\partial\iota_\Omega=N_\Omega(x)$, where $N_\Omega(x)=\{v|\langle v,y-x\rangle\leqslant 0,\forall y\in\Omega\}$ if $x\in\mathrm{bd}(\Omega)$, and $N_\Omega(x)=0$ if $x\in\mathrm{int}(\Omega)$, where $\mathrm{dom}(N_\Omega)=\Omega$. $\mathrm{proj}_\Omega(x)=\arg\min_{y\in\Omega}\|x-y\|^2$, and $\mathrm{prox}_{\iota_\Omega}(x)=\mathcal{J}_{N_\Omega}(x)=\mathrm{proj}_\Omega(x)$.

The operator T is α-averaged ($\alpha\in(0,1)$), if $\exists$ a nonexpansive operator T' satisfying $T=(1-\alpha)\mathrm{Id}+\alpha T'$. Let $\mathcal{A}(\alpha)$ mean the set of α-averaged operators. T is called firmly nonexpansive (β-cocoercive) if $T\in\mathcal{A}(\frac{1}{2})$ ($\beta T\in\mathcal{A}(\frac{1}{2})$ for $\beta>0$). If f is convex differentiable with θ-Lipschitz gradient ∇f, then ∇f is $\frac{1}{\theta}$-cocoercive. $\mathfrak{E}$ is maximally monotone $\Leftrightarrow$ $\mathcal{J}_{\mathfrak{E}}=(\mathrm{Id}+\mathfrak{E})^{-1}$ is firmly nonexpansive. If Ω is closed and convex, proj_Ω is firmly nonexpansive.

Appendix B Proof of Lemma 3

Proof. Eq. (19) can be rewritten as

$$-\mathfrak{E}(\boldsymbol{\omega}_k)-\boldsymbol{\epsilon}_k^q\in\mathfrak{F}(\boldsymbol{\omega}_{k+1})+\Psi(\boldsymbol{\omega}_{k+1}-\boldsymbol{\omega}_k). \tag{B1}$$

Using $\mathfrak{E}$ and $\mathfrak{F}$, the update for $\boldsymbol{x}_k$ in (B1) is

$$-\mathcal{M}^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{x}_k)-c\boldsymbol{L}_x\boldsymbol{x}_k-c\bar{\boldsymbol{\epsilon}}_k^x\in\mathcal{M}^{\mathrm{T}}N_\Omega(\mathcal{M}\boldsymbol{x}_{k+1})+\mathcal{M}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}_{k+1}+\boldsymbol{\tau}^{-1}(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k)-\mathcal{M}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\lambda}_{k+1}-\boldsymbol{\lambda}_k),$$

which implies that for some $\nu\in N_\Omega(x_{k+1})$,

$$-\boldsymbol{\tau}\mathcal{M}^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{x}_k)-c\boldsymbol{\tau}\boldsymbol{L}_x\boldsymbol{x}_k-c\boldsymbol{\tau}\bar{\boldsymbol{\epsilon}}_k^x=\boldsymbol{\tau}\mathcal{M}^{\mathrm{T}}(\nu+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}_{k+1})+(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k)-\boldsymbol{\tau}\mathcal{M}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\lambda}_{k+1}-\boldsymbol{\lambda}_k). \tag{B2}$$

Recall $\boldsymbol{\tau}_x=\mathrm{diag}((\tau_i I_{n_i})_{i\in\mathcal{V}})$ and $\boldsymbol{\tau}=\mathrm{diag}((\tau_i I_n)_{i\in\mathcal{V}})$. It can be shown that $\boldsymbol{\tau}_x\mathcal{M}=\mathcal{M}\boldsymbol{\tau}$ using $\mathcal{M}=\mathrm{diag}((\mathcal{M}_i)_{i\in\mathcal{N}})$. Similarly, $\boldsymbol{\tau}_s\mathcal{T}=\mathcal{T}\boldsymbol{\tau}$. With $\mathcal{M}\mathcal{M}^{\mathrm{T}}=I_n$ and $\mathcal{M}^{\mathrm{T}}\mathcal{M}+\mathcal{T}^{\mathrm{T}}\mathcal{T}=I_{Nn}$, we have $\boldsymbol{\tau}_x=\mathcal{M}\boldsymbol{\tau}\mathcal{M}^{\mathrm{T}}$ and $\mathcal{M}^{\mathrm{T}}\boldsymbol{\tau}_x\mathcal{M}+\mathcal{T}^{\mathrm{T}}\boldsymbol{\tau}_s\mathcal{T}=\boldsymbol{\tau}$. Then, premultiplying (B2) by $\mathcal{M}$ and $\mathcal{T}$, respectively, result in

$$\begin{aligned}&x_k-\boldsymbol{\tau}_x\left(\boldsymbol{F}(\boldsymbol{x}_k)+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}_k+c\mathcal{M}\boldsymbol{L}_x\boldsymbol{x}_k+c\mathcal{M}\bar{\boldsymbol{\epsilon}}_k^x\right)\in x_{k+1}+\boldsymbol{\tau}_x N_\Omega(x_{k+1});\\&\mathcal{T}\boldsymbol{x}_{k+1}=\mathcal{T}\boldsymbol{x}_k-c\mathcal{T}\boldsymbol{\tau}\boldsymbol{L}_x\boldsymbol{x}_k-c\mathcal{T}\boldsymbol{\tau}\bar{\boldsymbol{\epsilon}}_k^x.\end{aligned} \tag{B3}$$

Since N_Ω is a cone and $\tau_i>0$, $\boldsymbol{\tau}_x N_\Omega(x)=\Pi_{i=1}^N\tau_i N_{\Omega_i}(x_i)=N_\Omega(x)$. With $\mathrm{proj}_\Omega(x)=(\mathrm{Id}+N_\Omega)^{-1}(x)$ and $\mathcal{T}\boldsymbol{\tau}=\boldsymbol{\tau}_s\mathcal{T}$, Eq. (B3) is equivalent to (13) and (14). The $\boldsymbol{y}_k$ update in (B1) is

$$\boldsymbol{y}_{k+1}=\boldsymbol{y}_k+\delta\boldsymbol{D}_\lambda\boldsymbol{\lambda}_k+\delta\boldsymbol{D}_\lambda\bar{\boldsymbol{\epsilon}}_k^\lambda. \tag{B4}$$

Recall $\boldsymbol{L}_\lambda=\boldsymbol{D}_\lambda^{\mathrm{T}}\boldsymbol{D}_\lambda$. Note that $\boldsymbol{z}_0=\boldsymbol{D}_\lambda\boldsymbol{y}_0$ due to $\boldsymbol{z}_0=\mathbf{0}_{mN}$ and $\boldsymbol{y}_0=\mathbf{0}_{mM}$. Hence, $\boldsymbol{D}_\lambda^{\mathrm{T}}\boldsymbol{y}_k$ in (B4) is identical to $\boldsymbol{z}_k$ in (16), i.e., $\boldsymbol{z}_k=\boldsymbol{D}_\lambda\boldsymbol{y}_k$, $\forall k\in\mathbb{N}$. Using $\boldsymbol{\sigma}N_{\mathbb{R}_+^{mN}}(\boldsymbol{\lambda})=N_{\mathbb{R}^{mN}}(\boldsymbol{\lambda})$ and $P_{\mathbb{R}_+^{mN}}(\boldsymbol{\lambda})=(\mathrm{Id}+N_{\mathbb{R}_+^{mN}})^{-1}$, the $\boldsymbol{\lambda}_k$ update in (B1) is equivalent to (16).

Lemma 6. Let $c > c_0$ and $\vartheta > \frac{1}{2\beta}$, where c_0, μ_c in (20) and $\beta \in (0, \mu_c/(\mu_0)^2]$, and let τ_i, δ, and σ_i satisfy the requirements in Lemma 4. If Assumptions 1–3 and 5 hold, then the following hold under $\|\cdot\|_\Psi$ (the Ψ-induced norm).

(i) $\Psi^{-1}\mathfrak{F}$ is maximally monotone and $T_2 \in \mathcal{A}(\frac{1}{2})$;

(ii) $\Psi^{-1}\mathfrak{E}$ is $\beta\vartheta$-restricted cocoercive and T_1 is restricted nonexpansive, such that

$$\|T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}'\|_\Psi^2 \leqslant \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\Psi^2 - (2\beta\vartheta - 1)\|\boldsymbol{\omega} - \boldsymbol{\omega}' - (T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}')\|_\Psi^2$$

holds for $\forall\boldsymbol{\omega}$ and for $\forall\boldsymbol{\omega}' \in V_x$, where $\Omega_V = V_x \times \mathbb{R}^{mM} \times \mathbb{R}_+^{Nm}$ and $V_x = \{1_N \otimes \boldsymbol{x} | \boldsymbol{x} \in \mathbb{R}^n\}$;

(iii) $T_2 \circ T_1$ is υ-restricted averaged with $\upsilon := \frac{2\beta\vartheta}{4\beta\vartheta - 1} \in (\frac{1}{2}, 1)$, i.e., for $\forall\boldsymbol{\omega}$ and for $\forall\boldsymbol{\omega}' \in V_x$,

$$\|T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}'\|_\Psi^2 \leqslant \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\Psi^2 - \frac{1-\upsilon}{\upsilon}\|(\boldsymbol{\omega} - \boldsymbol{\omega}') - (T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}')\|_\Psi^2. \tag{B5}$$

Proof. (i) The operator $\mathfrak{F}$ in (17) can be defined as $\mathfrak{F}\boldsymbol{\omega} = \mathcal{M}^{\mathrm{T}} N_\Omega \mathcal{M}(\boldsymbol{x}) \times \mathbf{0}_{mM} \times N_{\mathbb{R}_+^{mN}}(\boldsymbol{\lambda}) + U\boldsymbol{\omega}$, where

$$U = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathcal{M}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{D}_\lambda \\ -\boldsymbol{A}\mathcal{M} & \boldsymbol{D}_\lambda^{\mathrm{T}} & \mathbf{0} \end{bmatrix}. \tag{B6}$$

Since U is a skew-symmetric matrix, i.e., $U^{\mathrm{T}} = -U$, U is maximally monotone [35, Example 20.30]. By [35, Ex. 20.40], N_Ω and $N_{\mathbb{R}_+^{mN}}$ are maximally monotone, as is $\mathbf{0}_{mM}$. By [32, Proposition 12.5.5], $\mathcal{M}^{\mathrm{T}} N_\Omega \mathcal{M}$ is maximally monotone due to the full row rank $\mathcal{M}$. Since the direct sum of maximally monotone operators [35, Proposition 12.23], $\mathcal{M}^{\mathrm{T}} N_\Omega \mathcal{M} \times \mathbf{0}_{mM} \times N_{\mathbb{R}_+^{mN}}$ is maximally monotone. By [35, Corollary 24.4], $\mathfrak{F}$ is maximally monotone since $\mathrm{dom}(U\boldsymbol{\omega}) = \mathbb{R}^{Nn+mM+Nm}$.

Ψ is positive definite and nonsingular. For any $(x, u) \in \Psi^{-1}\mathfrak{F}$ and $(y, v) \in \Psi^{-1}\mathfrak{F}$, $\Psi u \in \Psi\Psi^{-1}\mathfrak{F}x \in \mathfrak{F}x$ and $\Psi v \in \Psi\Psi^{-1}\mathfrak{F}y \in \mathfrak{F}y$. Then, $\langle x - y, u - v\rangle_\Psi = \langle x - y, \Psi u - \Psi v\rangle \geqslant 0$ since $\mathfrak{F}$ is maximally monotone. Thus, $\Psi^{-1}\mathfrak{F}$ is monotone under the Ψ-induced norm $\|\cdot\|_\Psi$. Choose $(y, v) \in \boldsymbol{\Omega} \times \mathbb{R}^{n+Mm+Nm}$ satisfying $\langle x - y, u - v\rangle_\Psi \geqslant 0$ for any $(x, u) \in \mathrm{gra}\Psi^{-1}\mathfrak{F}$. Next, we prove $(y, v) \in \mathrm{gra}\Psi^{-1}\mathfrak{F}$. Choose $(x, u') \in \mathrm{gra}\mathfrak{F}$, then $(x, \Psi^{-1}u') \in \mathrm{gra}\Psi^{-1}\mathfrak{F}$. This implies that $\langle x - y, \Psi^{-1}u' - v\rangle_\Psi \geqslant 0$, which is equivalent to $\langle x - y, \Psi\Psi^{-1}u' - \Psi v\rangle \geqslant 0$. Since $\mathfrak{F}$ is maximally monotone, $(y, \Psi v) \in \mathrm{gra}\mathfrak{F}$, that is, $(y, v) \in \mathrm{gra}\Psi^{-1}\mathfrak{F}$. Thus, $\Psi^{-1}\mathfrak{F}$ is maximally monotone under $\|\cdot\|_\Psi$. By [35, Prop. 23.7], $T_2 = (\mathrm{Id} + \Psi^{-1}\mathfrak{F})^{-1}$ is firmly nonexpansive under $\|\cdot\|_\Psi$, and thus $T_2 \in \mathcal{A}(\frac{1}{2})$.

(ii) The space $\mathbb{R}^{Nn}$ can be decomposed as $\mathbb{R}^{Nn} = V_x \oplus V_x^\perp$ into $V_x = \{1_N \otimes \boldsymbol{x} | \boldsymbol{x} \in \mathbb{R}^n\}$ and its orthogonal complement. For $\boldsymbol{x} \in \mathbb{R}^{Nn}$, let $\boldsymbol{x} = \boldsymbol{x}^\parallel + \boldsymbol{x}^\perp$ with $\boldsymbol{x}^\parallel \in V_x$, $\boldsymbol{x}^\perp \in V_x^\perp$, and $(\boldsymbol{x}^\parallel)^{\mathrm{T}}\boldsymbol{x}^\perp = \mathbf{0}$. Here $\boldsymbol{x}^\parallel = 1_N \otimes \boldsymbol{x}$ for some $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{L}_x\boldsymbol{x}^\parallel = \mathbf{0}$, $(\boldsymbol{x}^\perp)^{\mathrm{T}}\boldsymbol{L}_x\boldsymbol{x}^\perp \geqslant q_2(L)\|\boldsymbol{x}^\perp\|^2$, and $\boldsymbol{F}(\boldsymbol{x}^\parallel) = F(\boldsymbol{x})$. For any $\boldsymbol{x}' \in V_x$, $\boldsymbol{x}' = 1_N \otimes \boldsymbol{x}'$ for some $\boldsymbol{x}' \in \mathbb{R}^n$, and $\boldsymbol{F}(\boldsymbol{x}') = F(\boldsymbol{x}')$.

Let $\boldsymbol{\omega} = \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}\}$, and $\boldsymbol{\omega}' = \{\boldsymbol{x}', \boldsymbol{y}', \boldsymbol{\lambda}'\} \in \Omega_V$ with $\boldsymbol{x}' \in V_x$. Recall $\mathcal{M}\boldsymbol{x}^\parallel = \boldsymbol{x}$, $\mathcal{M}\boldsymbol{x}' = \boldsymbol{x}'$, $\boldsymbol{L}_x\boldsymbol{x}^\parallel = \mathbf{0}$, and $\boldsymbol{L}_x\boldsymbol{x}' = \mathbf{0}$. Then, for $\mathfrak{E}$ in (18), $\langle\boldsymbol{\omega} - \boldsymbol{\omega}', \mathfrak{E}(\boldsymbol{\omega}) - \mathfrak{E}(\boldsymbol{\omega}')\rangle = (\boldsymbol{x} - \boldsymbol{x}')^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}^\parallel) + F(\boldsymbol{x}) - F(\boldsymbol{x}')) + c(\boldsymbol{x}^\perp)^{\mathrm{T}}\boldsymbol{L}_x\boldsymbol{x}^\perp + (\boldsymbol{x}^\perp)^{\mathrm{T}}(\mathcal{M}^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}^\parallel) + F(\boldsymbol{x}) - F(\boldsymbol{x}')))$. In addition, $\|\mathcal{M}^{\mathrm{T}}\boldsymbol{x}^\perp\| \leqslant \|\mathcal{M}^{\mathrm{T}}\|\|\boldsymbol{x}^\perp\| = \|\boldsymbol{x}^\perp\|$ due to $\|\mathcal{M}^{\mathrm{T}}\| = 1$. Using the Lipschitz properties of F and $\boldsymbol{F}$, and the strongly monotone property of F,

$$\langle\boldsymbol{\omega} - \boldsymbol{\omega}', \mathfrak{E}(\boldsymbol{\omega}) - \mathfrak{E}(\boldsymbol{\omega}')\rangle \geqslant -(\mu_0 + \mu_\ell)\|\boldsymbol{x} - \boldsymbol{x}'\|\|\boldsymbol{x}^\perp\| + (cq_2(L) - \mu_0)\|\boldsymbol{x}^\perp\|^2 + \mu\|\boldsymbol{x} - \boldsymbol{x}'\|^2.$$

It follows from $\|\boldsymbol{x} - \boldsymbol{x}'\| = \frac{1}{\sqrt{N}}\|\boldsymbol{x}^\parallel - \boldsymbol{x}'\|$ that $\langle\boldsymbol{\omega} - \boldsymbol{\omega}', \mathfrak{E}(\boldsymbol{\omega}) - \mathfrak{E}(\boldsymbol{\omega}')\rangle \geqslant \widetilde{\boldsymbol{x}}^{\mathrm{T}}\Theta\widetilde{\boldsymbol{x}} \geqslant \mu_c\|\widetilde{\boldsymbol{x}}\|^2$ with $\widetilde{\boldsymbol{x}} = [\|\boldsymbol{x}^\parallel - \boldsymbol{x}'\|, \|\boldsymbol{x}^\perp\|]^{\mathrm{T}}$. Using $\|\boldsymbol{x} - \boldsymbol{x}'\|^2 = \|\widetilde{\boldsymbol{x}}\|^2$ and the Lipschitz property of $\boldsymbol{F}$, $\langle\boldsymbol{\omega} - \boldsymbol{\omega}', \mathfrak{E}(\boldsymbol{\omega}) - \mathfrak{E}(\boldsymbol{\omega}')\rangle \geqslant \beta\|\mathfrak{E}\boldsymbol{\omega} - \mathfrak{E}\boldsymbol{\omega}'\|^2$ for any $\boldsymbol{\omega}$ and for any $\boldsymbol{\omega}' \in \Omega_V$, due to $0 < \beta \leqslant \mu_c/(\mu_0)^2$. With $\frac{1}{q_{\max}(\Psi^{-1})} = q_1(\Psi)$, $\|\mathfrak{E}\boldsymbol{\omega} - \mathfrak{E}\boldsymbol{\omega}'\|^2 \geqslant q_1(\Psi)\|\Psi^{-1}\mathfrak{E}\boldsymbol{\omega} - \Psi^{-1}\mathfrak{E}\boldsymbol{\omega}'\|_\Psi^2$. From Lemma 4, it follows that $q_1(\Psi) \geqslant \vartheta$. Thus, for any $\boldsymbol{\omega}$ and for any $\boldsymbol{\omega}' \in \Omega_V$, $\langle\boldsymbol{\omega} - \boldsymbol{\omega}', \Psi^{-1}\mathfrak{E}\boldsymbol{\omega} - \Psi^{-1}\mathfrak{E}\boldsymbol{\omega}'\rangle_\Psi \geqslant \beta\vartheta\|\Psi^{-1}\mathfrak{E}\boldsymbol{\omega} - \Psi^{-1}\mathfrak{E}\boldsymbol{\omega}'\|_\Psi^2$, and $\|T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}'\|_\Psi^2 \leqslant \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\Psi^2 - (2\beta\vartheta - 1)\|\boldsymbol{\omega} - \boldsymbol{\omega}' - (T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}')\|_\Psi^2$, due to $T_1 = \mathrm{Id} - \Psi^{-1}\mathfrak{E}$. This implies that T_1 is restricted nonexpansive, since $2\beta\vartheta > 1$ by assumptions.

(iii) Since $T_2 \in \mathcal{A}(\frac{1}{2})$, for any $\boldsymbol{\omega}$ and for any $\boldsymbol{\omega}' \in V_x$,

$$\|T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}'\|_\Psi^2 \leqslant \|T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}'\|_\Psi^2 - \|T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}' - (T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}')\|_\Psi^2.$$

Let $s_0 = \frac{1}{2\beta\vartheta} \in (0, 1)$. By (ii) of Lemma 6,

$$\begin{aligned} \|T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}'\|_\Psi^2 \leqslant{}& \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\Psi^2 - \|T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}' - (T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}')\|_\Psi^2 \\ & - \frac{1-s_0}{s_0}\|\boldsymbol{\omega} - \boldsymbol{\omega}' - (T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}')\|_\Psi^2. \end{aligned} \tag{B7}$$

Next, we use $\|\lambda x + (1-\lambda)y\|^2 = \lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \lambda(1-\lambda)\|x - y\|^2$ by [35, Cor. 2.14], and then

$$\begin{aligned} &\|T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}' - (T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}')\|_\Psi^2 + \frac{1-s_0}{s_0}\|\boldsymbol{\omega} - \boldsymbol{\omega}' - (T_1\boldsymbol{\omega} - T_1\boldsymbol{\omega}')\|_\Psi^2 \\ &\quad \geqslant (1-s_0)\|(\boldsymbol{\omega} - \boldsymbol{\omega}') - (T_2 \circ T_1\boldsymbol{\omega} - T_2 \circ T_1\boldsymbol{\omega}')\|_\Psi^2. \end{aligned}$$

Substituting the above inequality into (B7) yields (B5).

Appendix C Proof of Lemma 5

Proof. It follows from Lemma 6 that T_2 is nonexpansive, and hence

$$\|\tilde{\boldsymbol{\epsilon}}_k^q\| = \|T_2 \circ (T_1\boldsymbol{\omega}_k - \Psi^{-1}\boldsymbol{\epsilon}_k^q) - T_2 \circ T_1\boldsymbol{\omega}_k\| \leqslant \|(T_1\boldsymbol{\omega}_k - \Psi^{-1}\boldsymbol{\epsilon}_k^q) - T_1\boldsymbol{\omega}_k\| = \| - \Psi^{-1}\boldsymbol{\epsilon}_k^q\|.$$

Since $q_{\max}(\Psi^{-1}) = 1/q_1(\Psi) \leqslant \frac{1}{\vartheta}$, $\|\tilde{\boldsymbol{\epsilon}}_k\| \leqslant \frac{1}{\vartheta}\|\boldsymbol{\epsilon}_k^q\|$. Recall $\boldsymbol{\epsilon}_k^q = \mathrm{col}(c\bar{\boldsymbol{\epsilon}}_k^x, -\boldsymbol{D}_\lambda\bar{\boldsymbol{\epsilon}}_k^\lambda, \boldsymbol{0})$; then

$$\|\tilde{\boldsymbol{\epsilon}}_k^q\| \leqslant \frac{1}{\vartheta}(\|c\bar{\boldsymbol{\epsilon}}_k^x\|^2 + \|\boldsymbol{D}_\lambda\bar{\boldsymbol{\epsilon}}_k^\lambda\|^2)^{\frac{1}{2}} \leqslant \frac{c_1}{\vartheta}(\|\bar{\boldsymbol{\epsilon}}_k^x\|^2 + \|\bar{\boldsymbol{\epsilon}}_k^\lambda\|^2)^{\frac{1}{2}},$$

where $c_1 = \max\{c, \sqrt{q_{\max}(L)}\}$. From the definition of $\bar{\boldsymbol{\epsilon}}_k^x$ and the convexity of the Euclidean norm function, it follows that

$$\|\bar{\boldsymbol{\epsilon}}_k^x\|^2 = \sum_{i=1}^N \|\bar{\boldsymbol{\epsilon}}_{i,k}^x\|^2 = \sum_{i=1}^N \left\| \sum_{j\in\mathcal{N}_i} w_{i,j}\left(\bar{e}_{i,k}^x - \bar{e}_{j,k}^x\right)\right\|^2 \leqslant \sum_{i=1}^N l_{i,i} \sum_{j\in\mathcal{N}_i} w_{i,j}\|\bar{e}_{i,k}^x - \bar{e}_{j,k}^x\|^2.$$

With $w_{i,j} = w_{j,i}$, one has $\sum_{i=1}^N\sum_{j=1}^N w_{i,j}\|\bar{e}_{j,k}^x\|^2 = \sum_{j=1}^N\sum_{i=1}^N w_{j,i}\|\bar{e}_{i,k}^x\|^2 = \sum_{i=1}^N\sum_{j=1}^N w_{i,j}\|\bar{e}_{i,k}^x\|^2$. Thus,

$$\|\bar{\boldsymbol{\epsilon}}_k^x\|^2 \leqslant \sum_{i=1}^N l_{i,i}\sum_{j=1}^N w_{i,j}(\|\bar{e}_{i,k}^x\|^2 + \|\bar{e}_{j,k}^x\|^2) \leqslant 2w_*^2\sum_{i=1}^N \|\bar{e}_{i,k}^x\|^2$$

with $w_* = \max_{i\in\mathcal{V}}\{\sum_{j=1}^N w_{i,j}\}$. According to Lemma 1, under the ETCD scheme (9), $\|\bar{e}_{i,k}^x\|^2 + \|\bar{e}_{i,k}^\lambda\|^2 \leqslant (B_k^i)^2$ for $i\in\mathcal{V}$. It follows from Assumption 4 that $\|\bar{\boldsymbol{\epsilon}}_k^x\|^2 + \|\bar{\boldsymbol{\epsilon}}_k^\lambda\|^2 \leqslant c_2\sum_{i=1}^N(\|\bar{e}_{i,k}^x\|^2 + \|\bar{e}_{i,k}^\lambda\|^2) \leqslant c_2 N B_k^2$ and hence $\|\tilde{\boldsymbol{\epsilon}}_k^q\| \leqslant \bar{c}B_k$, where $c_2 = \max\{2w_*^2, 1\}$ and $\bar{c} = \frac{c_1\sqrt{c_2 N}}{\vartheta}$. In addition, $\sum_{k=1}^\infty \|\tilde{\boldsymbol{\epsilon}}_k^q\| < \infty$ due to $\sum_{i=1}^N B_k < \infty$, which implies that $\lim_{k\to\infty}\|\tilde{\boldsymbol{\epsilon}}_k^q\| = 0$.

Appendix D Proof of Theorem 1

Proof. It follows from [35, Prop. 25.1] and Lemma 2 that $\mathrm{fix}(T_2\circ T_1) = \mathrm{zer}(\Psi^{-1}\mathfrak{E} + \Psi^{-1}\mathfrak{F}) = \mathrm{zer}(\mathfrak{E}+\mathfrak{F}) \neq \emptyset$ since $\Psi \succ 0$, where T_1, T_2 are given in (19).

Consider any $\boldsymbol{\omega}^* \in \mathrm{zer}(\mathfrak{E}+\mathfrak{F}) = \mathrm{fix}(T_2\circ T_1)$, (i.e., $\boldsymbol{\omega}^*\in V_x$ by Lemma 2). From (B5), it follows that

$$\begin{aligned}\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi =& \|T_2\circ T_1\boldsymbol{\omega}_k + \tilde{\boldsymbol{\epsilon}}_k^q - T_2\circ T_1\boldsymbol{\omega}^*\|_\Psi \\ \leqslant& \|T_2\circ T_1\boldsymbol{\omega}_k - T_2\circ T_1\boldsymbol{\omega}^*\|_\Psi + \|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi \\ \leqslant& \|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi + \|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi. \end{aligned} \tag{D1}$$

By Lemma 5, $\sum_{k=1}^\infty \|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi \leqslant \sqrt{q_{\max}(\Psi)}\sum_{k=1}^\infty\|\tilde{\boldsymbol{\epsilon}}_k^q\| < \infty$, which implies that $\sum_{k=1}^\infty\|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi^2 < \infty$ and $\{\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi\}$ is bounded, i.e., $\sup_{l\in\mathbb{N}}\|\boldsymbol{\omega}_l - \boldsymbol{\omega}^*\|_\Psi < \infty$.

Let $\alpha_c := \underline{\lim}_{k\to\infty}\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi$; then there exists $\{n_k\}_{k\in\mathbb{N}}$ satisfying $\lim_{k\to\infty}\|\boldsymbol{\omega}_{n_k} - \boldsymbol{\omega}^*\|_\Psi = \alpha_c$. For $\forall\varepsilon > 0$, $\exists m_0\in\mathbb{N}$ such that $\|\boldsymbol{\omega}_{n_k} - \boldsymbol{\omega}^*\|_\Psi - \alpha_c < \frac{\varepsilon}{2}$ (for $k > m_0$) and $\sum_{k=n_{m_0}}^\infty \|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi < \frac{\varepsilon}{2}$. By (D1),

$$\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi \leqslant \|\boldsymbol{\omega}_{n_{m_0}} - \boldsymbol{\omega}^*\|_\Psi + \sum_{s=n_{m_0}}^\infty \|\tilde{\boldsymbol{\epsilon}}_s^q\|_\Psi \leqslant \alpha_c + \varepsilon,$$

which implies that $\overline{\lim}_{k\to\infty}\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi \leqslant \underline{\lim}_{k\to\infty}\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi + \varepsilon$. Thus, $\{\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi\}$ converges for any $\boldsymbol{\omega}^*\in\mathrm{zer}(\mathfrak{E}+\mathfrak{F})$, since $\varepsilon > 0$ is arbitrarily small. From (19), it follows that

$$\begin{aligned}\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi^2 =& \|T_2\circ T_1\boldsymbol{\omega}_k - T_2\circ T_1\boldsymbol{\omega}^* + \tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi^2 \\ \leqslant& \|T_2\circ T_1\boldsymbol{\omega}_k - T_2\circ T_1\boldsymbol{\omega}^*\|_\Psi^2 + \epsilon_k', \end{aligned} \tag{D2}$$

where $\epsilon_k' := 2\sup_{l\in\mathbb{N}}\|\boldsymbol{\omega}_l - \boldsymbol{\omega}^*\|\|\tilde{\boldsymbol{\epsilon}}_k\|_\Psi + \|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi^2$. Since $T_2\circ T_1$ is υ-restricted averaged, using (B5), Eq. (D2) can be written as

$$\|\boldsymbol{\omega}_{k+1} - \boldsymbol{\omega}^*\|_\Psi^2 \leqslant \|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi^2 + \epsilon_k' - \frac{1-\upsilon}{\upsilon}\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^* - (T_2\circ T_1\boldsymbol{\omega}_k - T_2\circ T_1\boldsymbol{\omega}^*)\|_\Psi^2,$$

which is equivalent to $\frac{1-\upsilon}{\upsilon}\sum_{k=0}^\infty\|\boldsymbol{\omega}_k - T_2T_1\boldsymbol{\omega}_k\|_\Psi \leqslant \|\boldsymbol{\omega}_0 - \boldsymbol{\omega}^*\|_\Psi + \epsilon'$. Here, $\epsilon' := \sum_{k=0}^\infty \epsilon_k' < \infty$ since $\sum_{k=1}^\infty\|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi < \infty$, $\sum_{k=1}^\infty\|\tilde{\boldsymbol{\epsilon}}_k^q\|_\Psi^2 < \infty$, and $\sup_{l\in\mathbb{N}}\|\boldsymbol{\omega}_l - \boldsymbol{\omega}^*\|_\Psi < \infty$. Due to $\frac{1-\upsilon}{\upsilon} > 0$, $\sum_{k=1}^\infty\|\boldsymbol{\omega}_k - T_2T_1\boldsymbol{\omega}_k\|_\Psi$ converges, and hence $\lim_{k\to\infty}\boldsymbol{\omega}_k - T_2T_1\boldsymbol{\omega}_k = 0$ since $\Psi\succ 0$.

Since $\{\|\boldsymbol{\omega}_k - \boldsymbol{\omega}^*\|_\Psi\}$ is bounded, $\{\boldsymbol{\omega}_k\}$ is bounded. Then, there exists a subsequence $\{\boldsymbol{\omega}_{m_k}\}$ such that $\lim_{k\to\infty}\boldsymbol{\omega}_{m_k} = \bar{\boldsymbol{\omega}}^*$ for some $\bar{\boldsymbol{\omega}}^*$. Note that $T_2\circ T_1$ is continuous and single-valued, $\lim_{k\to\infty}\boldsymbol{\omega}_{m_k} - T_2\circ T_1\boldsymbol{\omega}_{m_k} = 0$ implies that $\bar{\boldsymbol{\omega}}^* = T_2\circ T_1\bar{\boldsymbol{\omega}}^*$, i.e., $\bar{\boldsymbol{\omega}}^*\in\mathrm{fix}(T_2\circ T_1)$, namely, $\bar{\boldsymbol{\omega}}^*\in\mathrm{zer}(\mathfrak{E}+\mathfrak{F})$. Thus, $\{\|\boldsymbol{\omega}_k - \bar{\boldsymbol{\omega}}^*\|\}$ converges, and $\lim_{k\to\infty}\|\boldsymbol{\omega}_k - \bar{\boldsymbol{\omega}}^*\| = 0$ since $\lim_{k\to\infty}\|\boldsymbol{\omega}_{m_k} - \bar{\boldsymbol{\omega}}^*\| = 0$. Therefore, $\lim_{k\to\infty}\boldsymbol{\omega}_k = \bar{\boldsymbol{\omega}}^*$. By Lemmas 2 and 3, the conclusion can be obtained.