Ensuring both Almost Sure Convergence and Differential Privacy in Nash Equilibrium Seeking on Directed Graphs

Yongqiang Wang, Tamer Başar

***Abstract*—We propose a distributed Nash equilibrium seeking approach that can achieve both almost sure convergence and rigorous differential privacy with finite cumulative privacy budget, which is in sharp contrast to existing differential-privacy solutions for networked games that have to trade provable convergence for differential privacy. The approach is applicable when the communication graph is directed and unbalanced. Numerical comparison results with existing counterparts confirm the effectiveness of the proposed approach.**

I. Introduction

Nash equilibrium (NE) seeking in game theory addresses the problem where multiple players compete to minimize their individual cost functions [1]. In many application scenarios, individual players only have access to the decisions of their local neighbors, which is usually termed as games in the partial-decision information setting [2], [3]. In contrast to the classical full-decision information setting where a player knows the past actions of all other players, in the partial-decision information setting, individual players cannot evaluate their cost functions or gradients due to lack of necessary information. Consequently, players have to exchange action information with their local neighbors for NE seeking.

Significant inroads have been made in fully distributed NE seeking (see, e.g., [4], [5], [6], [7]). However, all of these distributed algorithms require players to share explicit (estimated) decisions in every iteration, which is problematic when sensitive information is involved. In fact, given that in noncooperative games the players are not fully cooperative, it is important for individual players to protect their private information, which, otherwise, might be exploited by others. Recently, several results have been reported on privacy protection in NE seeking (see, e.g., [8], [9]). However, most of these results assume the presence of a coordinator. In the fully distributed case, the authors in [10] exploit spatially-correlated noise to protect the privacy of players. However, their approach is only effective when the communication graph satisfies certain properties. Recently, the authors of [11] have used a constant uncertain parameter to obfuscate individual players' pseudo-gradients. However, the privacy strength enabled by such a constant scalar is weak in the sense that only the exact value of the cost function is prevented from being uniquely identifiable. As differential privacy has emerged as the de facto standard for privacy protection due to its strong resilience against arbitrary post-processing [12], recent results in [13] and [14] propose differential-privacy mechanisms for aggregative games, which, however, have to sacrifice provable convergence to the exact Nash equilibrium.

In this paper, we introduce a distributed NE seeking approach on directed graphs that can ensure both almost sure convergence and rigorous ϵ-differential privacy. We propose to gradually weaken the inter-player interaction to attenuate the effect of differential-privacy noise in shared messages on NE seeking. Note that inter-player interaction is necessary for all players' convergence to the NE (which exists and is unique in our case), and thus we judiciously design the weakening factor sequence and the stepsize sequence, under which we prove that our approach can ensure provable convergence to the exact unique NE even in the presence of differential-privacy noise. We prove that the algorithm is ϵ-differentially private with a finite cumulative privacy budget, even when the number of iterations tends to infinity. It is worth noting that compared with our recent results on differentially-private distributed optimization [15], [16], the results here for NE seeking are fundamentally different: agents in distributed optimization are cooperative in computing a common objective function, whereas players in games are competitive and only mind their own individual cost functions. Moreover, different from our recent results in [17] which address aggregative games on symmetric communication graphs, this paper addresses general networked games (that are not necessarily aggregative) on directed communication graphs that could be unbalanced.

Notations: We use $\mathbb{R}^d$ to denote the Euclidean space of dimension d. We write I_d for the identity matrix of dimension d, and $\mathbf{1}_d$ for the d-dimensional column vector with all entries equal to 1. For a vector x, $[x]_i$ denotes its ith element. We write $x > 0$ (resp. $x \geq 0$) if all elements of x are positive (resp. non-negative). We use $\langle \cdot, \cdot \rangle$ to denote the inner product and $\|x\|_2$ for the standard Euclidean norm of a vector x. We use $\|x\|_1$ to represent the ℓ_1 norm of a vector x. We write $\|A\|$ for the matrix norm induced by a vector norm $\|\cdot\|$. For two vectors u and v with the same dimension, we write $u \leq v$ to mean that each entry of u is no larger than the corresponding entry of v. We use *a.s.* to denote *almost surely* or *almost sure*, depending on the syntax context.

II. Problem Formulation and Preliminaries

A. On Networked Games

We consider a networked game among a set of m players, i.e., $[m] = \{1, 2, \ldots, m\}$. Player i is characterized by a feasible action set $\Omega_i \subseteq \mathbb{R}^{d_i}$ and a cost function $f_i(x_i, x_{-i})$

The work of the first author was supported in part by the National Science Foundation under Grants ECCS-1912702, CCF-2106293, CCF-2215088, and CNS-2219487. Research of the second author was supported in part by the ONR MURI Grant N00014-16-1-2710 and in part by the Army Research Laboratory, under Cooperative Agreement Number W911NF-17-2-0196.

Yongqiang Wang is with the Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634, USA yongqiw@clemson.edu

Tamer Başar is with the Coordinated Science Lab, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA basar1@illinois.edu

where $x_i \in \Omega_i$ is the decision of player i and $x_{-i} \triangleq [x_1^T, \cdots, x_{i-1}^T, x_{i+1}^T, \cdots, x_m^T]^T$ denotes the joint decisions of all players except player i. Note that we allow different x_i to have different dimensions d_i.

Traditionally when a mediator/coordinator exists, every player i can access all other players' decisions x_{-i}. Then, the game that player i faces can be formulated as:

$$\min f_i(x_i, x_{-i}) \quad \text{s.t.} \quad x_i \in \Omega_i \text{ and } x_{-i} \in \Omega_{-i}. \quad (1)$$

The function $f_i(\cdot)$ is assumed to be known to player i only.

At the NE $x^* = [(x_1^*)^T, \ldots, (x_m^*)^T]^T \in \mathbb{R}^D$ with $D = \sum_{i=1}^m d_i$, each player has $f_i(x_i^*, x_{-i}^*) \leq f_i(x_i, x_{-i}^*), \forall x_i \in \Omega_i$. Namely, at the NE, no player can unilaterally reduce its cost by changing its own decision.

We consider a scenario where no mediator/coordinator exists, and players share decisions locally among neighbors, which is commonly referred to as the partial-decision information scenario [2]. We use a directed graph $\mathcal{G} = ([m], \mathcal{E})$ to denote the communication pattern where $[m] = \{1, 2, \ldots, m\}$ is the set of nodes (players) and $\mathcal{E} \subseteq [m] \times [m]$ is the edge set of ordered node pairs describing the interactions among players. We also use the notion of directed graph induced by a weight matrix $L = \{L_{ij}\} \in \mathbb{R}^{m\times m}$, denoted as $\mathcal{G}_L = ([m], \mathcal{E}_L)$. More specifically, in $\mathcal{G}_L = ([m], \mathcal{E}_L)$, a directed edge (i, j) from agent j to agent i exists, i.e., $(i, j) \in \mathcal{E}_L$ if and only if $L_{ij} > 0$. For a player $i \in [m]$, its in-neighbor set $\mathbb{N}_i^{\text{in}}$ is defined as the collection of players j such that $L_{ij} > 0$; similarly, the out-neighbor set $\mathbb{N}_i^{\text{out}}$ of player i is the collection of players j such that $L_{ji} > 0$.

Assumption 1. *For all $i \in [m]$, $f_i(x_i, x_{-i})$ is convex and differentiable in x_i over $\mathbb{R}^{d_i}$ under each given x_{-i}.*

To characterize the NE of the game (1), we also define

$$\phi(x) \triangleq \left(F_1^T(x_1, x_{-1}), \cdots, F_m^T(x_m, x_{-m})\right)^T, \quad (2)$$

where $F_i(x_i, x_{-i}) \triangleq \nabla_{x_i} f_i(x_i, x_{-i})$.

Assumption 2. *$\phi(x)$ is strongly monotone over $\Omega \triangleq \Omega_1 \times \cdots \times \Omega_m$, i.e., for all $x \neq x'$ in Ω, $(\phi(x) - \phi(x'))^T (x - x') \geq \mu\|x - x'\|$ holds for some $\mu > 0$. Each mapping $F_i(x_i, x_{-i})$ is Lipschitz continuous in both of its arguments, x_i and x_{-i}. Namely, for all $x_i, y_i \in \mathbb{R}^{d_i}$ and $x_{-i}, y_{-i} \in \mathbb{R}^{D-d_i}$, where $D = \sum_{i=1}^m d_i$, $\|F_i(x_i, x_{-i}) - F_i(y_i, x_{-i})\|_2 \leq K_1\|x_i - y_i\|_2$ and $\|F_i(x_i, x_{-i}) - F_i(x_i, y_{-i})\|_2 \leq K_2\|x_{-i} - y_{-i}\|_2$ hold for all $i \in [m]$, where K_1 , K_2 are some constants.*

Assumption 2 ensures that (1) has a unique NE x^* [18].

Assumption 3. *The off-diagonal entries of the matrix $L = \{L_{ij}\} \in \mathbb{R}^{m\times m}$ are non-negative and its diagonal entries $L_{ii} = -\sum_{j=1}^m L_{ij}$ satisfy $L_{ii} > -1$ for all $i \in [m]$. Moreover, the digraph $\mathcal{G}_L$ is strongly connected.*

In the analysis of our approach, we use the following results:

Lemma 1. *(Lemma 2 of [15]) Let $\{v^k\}$,$\{\alpha^k\}$, and $\{p^k\}$ be random nonnegative scalar sequences, and $\{q^k\}$ be a deterministic nonnegative scalar sequence satisfying $\sum_{k=0}^\infty \alpha^k < \infty$ a.s., $\sum_{k=0}^\infty q^k = \infty$, $\sum_{k=0}^\infty p^k < \infty$ a.s., and*

$$\mathbb{E}\left[v^{k+1}|\mathcal{F}^k\right] \leq (1 + \alpha^k - q^k)v^k + p^k, \quad \forall k \geq 0 \quad a.s.$$

where $\mathcal{F}^k = \{v^\ell, \alpha^\ell, p^\ell; 0 \leq \ell \leq k\}$. Then, $\sum_{k=0}^\infty q^k v^k < \infty$ and $\lim_{k\to\infty} v^k = 0$ hold almost surely.

Lemma 2. *(Lemma 5 of [15]) Let $\{\mathbf{v}^k\} \subset \mathbb{R}^d$ and $\{\mathbf{u}^k\} \subset \mathbb{R}^p$ be random nonnegative vector sequences, and $\{a^k\}$ and $\{b^k\}$ be random nonnegative scalar sequences such that*

$$\mathbb{E}\left[\mathbf{v}^{k+1}|\mathcal{F}^k\right] \leq (V^k + a^k\mathbf{1}\mathbf{1}^T)\mathbf{v}^k + b^k\mathbf{1} - H^k\mathbf{u}^k, \quad \forall k \geq 0$$

holds a.s., where $\{V^k\}$ and $\{H^k\}$ are random sequences of nonnegative matrices and $\mathbb{E}\left[\mathbf{v}^{k+1}|\mathcal{F}^k\right]$ denotes the conditional expectation given $\mathbf{v}^\ell, \mathbf{u}^\ell, a^\ell, b^\ell, V^\ell, H^\ell$ for $\ell = 0, 1, \ldots, k$. Assume that $\{a^k\}$ and $\{b^k\}$ satisfy $\sum_{k=0}^\infty a^k < \infty$ and $\sum_{k=0}^\infty b^k < \infty$ a.s., and that there exists a (deterministic) vector $\pi > 0$ such that $\pi^T V^k \leq \pi^T$ and $\pi^T H^k \geq 0$ hold a.s. for all $k \geq 0$. Then, we have 1) $\{\pi^T\mathbf{v}^k\}$ converges to some random variable $\pi^T\mathbf{v} \geq 0$ a.s.; 2) $\{\mathbf{v}^k\}$ is bounded a.s.; and 3) $\sum_{k=0}^\infty \pi^T H^k\mathbf{u}^k < \infty$ holds almost surely.

Lemma 3. *[15] Let $\{v^k\}$ be a nonnegative sequence, and $\{\alpha^k\}$ and $\{\beta^k\}$ be positive sequences satisfying $\sum_{k=0}^\infty \alpha^k = \infty$ and $\lim_{k\to\infty} \alpha^k = 0$, and $\frac{\beta^k}{\alpha^k}$ converges to 0 with a polynomial decay rate. If there exists a $K \geq 0$ such that $v^{k+1} \leq (1 - \alpha^k)v^k + \beta^k$ holds for all $k \geq K$, then we always have $v^k \leq C\frac{\beta^k}{\alpha^k}$ for all k, where C is some constant.*

B. On Differential Privacy

We adopt the notion of ϵ-differential privacy (DP) for continuous bit streams [19], which has recently been applied to distributed optimization (see, e.g., [20] as well as our work [15]). To enable DP, we inject Laplace noise $\text{Lap}(\nu)$ to all shared messages, where $\nu > 0$ is a constant parameter of the probability density function $\frac{1}{2\nu}e^{-\frac{|x|}{\nu}}$. One can verify that $\text{Lap}(\nu)$ has mean zero and variance $2\nu^2$. We represent the networked game $\mathcal{P}$ in (1) by three parameters $(\Omega, \mathbb{F}, \mathcal{G}_L)$, where $\Omega \triangleq \Omega_1 \times \cdots \times \Omega_m$ is the domain of decision variables, $\mathbb{F} \triangleq \{f_1, \cdots, f_m\}$, and $\mathcal{G}_L$ denotes the communication graph. We define "adjacency" between two games as follows:

Definition 1. *Two networked games $\mathcal{P} \triangleq (\Omega, \mathbb{F}, \mathcal{G}_L)$ and $\mathcal{P}' \triangleq (\Omega', \mathbb{F}', \mathcal{G}'_L)$ are adjacent if the following conditions hold:*

- *$\Omega = \Omega'$ and $\mathcal{G}_L = \mathcal{G}'_L$;*
- *there exists an $i \in [m]$ such that $f_i \neq f_i'$ but $f_j = f_j'$ for all $j \in [m]$, $j \neq i$;*
- *f_i and f_i', which are not the same, have similar behaviors around x^*, the NE of $\mathcal{P}$. More specifically, there exits some $\delta > 0$ such that for all v and v' in $B_\delta(x^*) \triangleq \{u : u \in \mathbb{R}^D, \|u - x^*\| < \delta\}$, we have $f_i(v) = f_i'(v')$.*

Remark 1. *Note that in Definition 1, for the sake of notational simplicity, we use $f_i(v)$ to represent $f_i(v_i, v_{-i})$. Moreover, in Definition 1, since the difference between f_i and f_i' can be arbitrary, additional restrictions have to be imposed to ensure rigorous DP. Different from [13], [20] which restrict all gradients to be uniformly bounded, we add the third condition, which, together with the proposed noise-robust algorithm, allows us to ensure rigorous DP while maintaining accurate convergence.*

Given a distributed algorithm, we represent an execution of this algorithm as $\mathcal{A}$, which is an infinite sequence of the iteration variable ϑ, i.e., $\mathcal{A} = \{\vartheta^0, \vartheta^1, \cdots\}$. We consider adversaries that can observe all communicated messages. Thus, the observation part of an execution is the infinite sequence of shared messages, which is represented by $\mathcal{O}$. We define the mapping from execution sequence to observation sequence by $\mathcal{R}_{\mathcal{P},\vartheta^0}(\mathcal{A}) \triangleq \mathcal{O}$, where ϑ^0 denotes the initial condition. Given a networked game $\mathcal{P}$, an initial condition ϑ^0, and observation sequence $\mathcal{O}$, $\mathcal{R}^{-1}_{\mathcal{P},\vartheta^0}(\mathcal{O})$ is the set of executions $\mathcal{A}$ that can generate the observation $\mathcal{O}$.

Definition 2. *(ϵ-DP, [20]). For a given $\epsilon > 0$, an algorithm $\mathcal{A}$ is ϵ-differentially private if for any two adjacent networked games $\mathcal{P}$ and $\mathcal{P}'$, any set of observation sequences $\mathcal{O}_s \subseteq \mathbb{O}$ ($\mathbb{O}$ is the set of all possible observation sequences), and any initial state ϑ^0, we always have $\mathbb{P}[\mathcal{R}_{\mathcal{P},\vartheta^0}(\mathcal{A}) \in \mathcal{O}_s] \leq e^{\epsilon}\mathbb{P}[\mathcal{R}_{\mathcal{P}',\vartheta^0}(\mathcal{A}) \in \mathcal{O}_s,]$, where the probability $\mathbb{P}$ is taken over the randomness over iteration processes.*

III. A Differentially-Private NE Seeking Algorithm

We present in this section a fully distributed NE seeking algorithm (Algorithm 1 below).

Algorithm 1: Distributed NE seeking with provable convergence and differential privacy

---

Parameters: Stepsize sequence $\{\lambda^k > 0\}$ and decreasing sequence $\{\gamma^k > 0\}$.

Every player i maintains one decision variable $x^k_{(i)i}$, and $m-1$ estimates $x^k_{(i)-i} \triangleq [(x^k_{(i)1})^T, \cdots, (x^k_{(i)i-1})^T, (x^k_{(i)i+1})^T, \cdots, (x^k_{(i)m})^T]^T$ of other players' decision variables. Player i sets $x^0_{(i)\ell}$ randomly in $\mathbb{R}^{d_\ell}$ for all $\ell \in [m]$.

for $k = 1, 2, \ldots$ **do**

a) For both its decision variable $x^k_{(j)j}$ and estimate variables $x^k_{(j)1}, \cdots, x^k_{(j)j-1}, x^k_{(j)j+1}, \cdots, x^k_{(j)m}$, every player j adds respective persistent DP noise $\zeta^k_{(j)1}, \cdots, \zeta^k_{(j)m}$, and then sends the obscured values $x^k_{(j)1} + \zeta^k_{(j)1}, \cdots, x^k_{(j)m} + \zeta^k_{(j)m}$ to all players $i \in \mathbb{N}^{\text{out}}_j$.

b) After receiving $x^k_{(j)1} + \zeta^k_{(j)1}, \cdots, x^k_{(j)m} + \zeta^k_{(j)m}$ from all $j \in \mathbb{N}^{\text{in}}_i$, player i updates its decision and estimate variables:

$$x^{k+1}_{(i)i} = x^k_{(i)i} + \gamma^k \sum_{j\in\mathbb{N}^{\text{in}}_i} L_{ij}(x^k_{(j)i} + \zeta^k_{(j)i} - x^k_{(i)i}) - \lambda^k F_i(x^k_{(i)i}, x^k_{(i)-i}),$$
$$x^{k+1}_{(i)\ell} = x^k_{(i)\ell} + \gamma^k \sum_{j\in\mathbb{N}^{\text{in}}_i} L_{ij}(x^k_{(j)\ell} + \zeta^k_{(j)\ell} - x^k_{(i)\ell}), \ \forall \ell \neq i. \tag{3}$$

c) **end**

Remark 2. *The sequences $\{\gamma^k\}$ and $\{\lambda^k\}$, and the DP noise parameter are hard-coded into players' programs and need no adjustment/coordination in implementation. This enables our algorithm to be implementable in a fully distributed manner.*

IV. A General Convergence Result

We first have to establish some general results necessary for the convergence analysis of Algorithm 1.

Lemma 4. *Under Assumption 3, we have the following properties when $\gamma^k > 0$ in Algorithm 1 is sufficiently small:*

1) *the eigenvectors of the matrix $I + \gamma^k L$ are time-invariant;*
2) *$I + \gamma^k L$ has a unique positive left eigenvector u^T (associated with eigenvalue 1) satisfying $u^T\mathbf{1} = m$;*
3) *the spectral radius of $I + \gamma^k L - \frac{\mathbf{1}u^T}{m}$ is upper-bounded by $1 - \alpha\gamma^k$, where $0 < \alpha < 1$;*
4) *there exists an L-dependent matrix norm $\|\cdot\|_L$ such that $\|I + \gamma^k L - \frac{\mathbf{1}u^T}{m}\|_L \leq 1 - \alpha\gamma^k$ for $0 < \alpha < 1$ when γ^k is small enough. Moreover, this norm has an associated inner product $\langle\cdot,\cdot\rangle_L$, i.e., $\|x\|^2_L = \langle x, x\rangle_L$.*

Proof. 1) Representing the eigenvalues and associated eigenvectors of L as $\{\varrho_1, \cdots, \varrho_m\}$ and $\{v_1, \cdots, v_m\}$, respectively, we can verify that the eigenvalues and associated eigenvectors of $I + \gamma^k L$ are given by $\{1 + \gamma^k\varrho_1, \cdots, 1 + \gamma^k\varrho_m\}$ and $\{v_1, \cdots, v_m\}$, respectively.

2) One can obtain from [21] (or Lemma 1 in [22]) that $I + L$ has a unique positive left eigenvector u^T (associated with eigenvalue 1) satisfying $u^T\mathbf{1} = m$. Hence, using statement 1), $I + \gamma^k L$ has a unique positive left eigenvector u^T (associated with eigenvalue 1) satisfying $u^T\mathbf{1} = m$.

3) Representing the eigenvalues of L by $\{\varrho_1, \cdots, \varrho_m\}$, the eigenvalues of $I + L$ can be expressed as $\{1 + \varrho_1, \cdots, 1 + \varrho_m\}$. Under Assumption 3, $I + L$ is irreducible. Using the Perron–Frobenius theorem, one can obtain that $I + L$ has one unique eigenvalue equal to one and all its other eigenvalues strictly less than one in absolute value, implying that one and only one of ϱ_i is zero. Represent this eigenvalue of L as $\varrho_m = 0$ without loss of generality. Then we have $|1 + \varrho_i| < 1$ for all $1 \leq i \leq m-1$. One can verify that the eigenvalues of $I + \gamma^k L - \frac{\mathbf{1}u^T}{m}$ are given by $\{1 + \gamma^k\varrho_1, \cdots, 1 + \gamma^k\varrho_{m-1}, 0\}$. Next, we prove the third statement by showing that there exists an α satisfying $|1 + \gamma^k\varrho_i| < 1 - \alpha\gamma^k$ for every $i = 1, 2, \cdots, m-1$.

We represent ϱ_i as $\varrho_i = a_i + \mathrm{i}b_i$, where a_i and b_i are real numbers, and i is the imaginary unit. Because $|1 + \varrho_i| < 1$ holds for all $i = 1, 2, \cdots, m-1$, we have $a_i < 0$ for $i = 1, 2, \cdots, m-1$. Under the new representation of ϱ_i, $|1 + \gamma_k\varrho_i|$ becomes $\sqrt{(1 - |a_i|\gamma^k)^2 + (b_i\gamma^k)^2}$. So we only have to prove $\sqrt{(1 - |a_i|\gamma^k)^2 + (b_i\gamma^k)^2} < 1 - \gamma^k\alpha$ for some $0 < \alpha < 1$ when γ^k is small enough. Squaring both sides of the inequality yields $\alpha^2(\gamma^k)^2 - 2\alpha\gamma^k > (\gamma^k)^2(a_i^2 + b_i^2) - 2\gamma^k|a_i|$, i.e.,

$$\alpha^2 - \frac{2}{\gamma^k}\alpha > (a_i^2 + b_i^2) - \frac{2|a_i|}{\gamma^k}. \tag{4}$$

When γ^k is less than $\frac{2|a_i|}{a_i^2 + b_i^2}$ (note $a_i < 0$ for $i = 1, \cdots, m-1$), the right hand side of (4) is negative whereas the left hand side is a quadratic function of α with two x-intercepts given by $\alpha = 0$ and $\alpha = \frac{2}{\gamma^k}$. So there always exits an α in the interval (0, 1) making the left hand side of (4) larger than its right hand side, and hence making (4) hold. Therefore, there always exists an $0 < \alpha < 1$ making $|1 + \gamma_k\varrho_i| < 1 - \alpha\gamma^k$ hold when $\gamma^k > 0$ is less than $\frac{2|a_i|}{a_i^2 + b_i^2}$.

Given that the above derivation is independent of i, we have $|1 + \gamma_k\varrho_i| < 1 - \alpha\gamma^k$ for some $0 < \alpha < 1$ and all $i = 1, \cdots, m-1$, and hence the third statement in the Lemma.

4) According to [23], there exits a matrix norm $\|\cdot\|_L$, which only depends on the eigenvectors (or unitary Schur-decomposition matrix in the more general complex-matrix

case) of $I+\gamma^k L-\frac{\mathbf{1}u^T}{m}$, such that the norm of $I+\gamma^k L-\frac{\mathbf{1}u^T}{m}$ is arbitrarily close to its spectral radius. Statement 2) proves that the eigenvectors of $I+\gamma^k L-\frac{\mathbf{1}u^T}{m}$ are time-invariant and independent of γ^k. Hence, the matrix norm $\|\cdot\|_L$ is independent of γ^k. Using Statement 3) yields that $\|I+\gamma^k L-\frac{\mathbf{1}u^T}{m}\|_L<1-\alpha\gamma^k$ holds for some $0<\alpha<1$ when γ^k is smaller than $\min\{\frac{2|a_1|}{a_1^2+b_1^2},\cdots,\frac{2|a_{m-1}|}{a_{m-1}^2+b_{m-1}^2}\}$. Moreover, from [23], $\|\cdot\|_L$ can be expressed as $\|x\|_L=\|\hat{L}x\|_2$ for some $\hat{L}$ determined by L. Thus, the norm $\|\cdot\|_L$ satisfies the Parallelogram Law and, hence, has an associated inner product $\langle\cdot,\rangle_L$. We refer the reader to Sec. II.B of [24] for an instantiation of this norm and inner product (which only depends on the left eigenvector u^T) for directed graphs. ■

Using the time-invariant positive left eigenvector $u\triangleq[u_1,\cdots,u_m]^T$ of $I+\gamma^k L$ from Lemma 3, we define a weighted average $\bar{x}_i^k\triangleq\frac{1}{m}\sum_{\ell=1}^m u_\ell x_{(\ell)i}^k$ of player i's decision variable $x_{(i)i}^k$ and other players' estimates $x_{(j)i}^k$ ($j\neq i$) of this decision variable. We also define the assembly of the ith decision variable $\mathbf{x}_i^k$ as well as the assembly of the weighted average $\bar{\mathbf{x}}_i^k$ as

$$\mathbf{x}_i^k=\begin{bmatrix}(x_{(1)i}^k)^T\\ \vdots\\ (x_{(m)i}^k)^T\end{bmatrix}\in\mathbb{R}^{m\times d_i},\ \bar{\mathbf{x}}_i^k=\begin{bmatrix}(\bar{x}_i^k)^T\\ \vdots\\ (\bar{x}_i^k)^T\end{bmatrix}\in\mathbb{R}^{m\times d_i}.\quad(5)$$

To measure the distance between matrix variables $\mathbf{x}_i^k$ and $\bar{\mathbf{x}}_i^k$, we define a matrix norm for an arbitrary vector norm $\|\cdot\|_L$. Specifically, for a matrix $X\in\mathbb{R}^{m\times d_i}$, we define $\|X\|_L\triangleq\big\|\big[\|X_{(1)}\|_L,\cdots,\|X_{(d_i)}\|_L\big]\big\|_2$, where $X_{(i)}$ denotes the ith column of X. $\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L$ measures the distance between all players' $x_{(j)i}^k$ for $j\in[m]$ and their average $\bar{x}_i^k$.

Based on the inner product $\langle\cdot,\cdot\rangle_L$ for vectors, we also define an inner product for matrices consistent with the $\|\cdot\|_L$ norm for matrices. More specifically, for a matrix X in $\mathbb{R}^{m\times d_i}$, we define $\langle X,X\rangle_L$ as: $\langle X,X\rangle_L=\sum_{i=1}^{d_i}\langle X_{(i)},X_{(i)}\rangle_L$, where $X_{(i)}$ denotes the ith column of X. Since for any column $X_{(i)}$ of a matrix X, we have $\|X_{(i)}\|_L^2=\langle X_{(i)},X_{(i)}\rangle_L$, one can verify that $\|X\|_L^2=\langle X,X\rangle_L$ holds for any matrix $X\in\mathbb{R}^{m\times d_i}$. In addition, we have the following result:

Lemma 5. *For any norm $\|\cdot\|_L$, $X\in\mathbb{R}^{m\times d_i}$, and $W\in\mathbb{R}^{m\times m}$, we always have $\|WX\|_L\le\|W\|_L\|X\|_L$. Furthermore, there exist constants $\delta_{L,2}$ and $\delta_{2,L}$ such that $\|X\|_L\le\delta_{L,2}\|X\|_2$ and $\|X\|_2\le\delta_{2,L}\|X\|_L$ hold for any $X\in\mathbb{R}^{m\times d_i}$.*

Proof. The proof follows from the line of reasoning in Lemma 5 and Lemma 6 in [22], and hence is not included here. ■

Based on the above results, we have the following convergence result for general distributed algorithms for problem (1):

Proposition 1. *Under Assumptions 1 and 2, let $x^*=[(x_1^*)^T,\ldots,(x_m^*)^T]^T$ be the unique NE of (1). If, under interaction matrix L, a distributed NE-seeking algorithm generates sequences $\{\mathbf{x}_i^k\}$ for all $i\in[m]$ such that there exists some $T\ge0$ to make the following relations hold a.s. for all $k\ge T$:*

$$\begin{bmatrix}\mathbb{E}\left[\sum_{i=1}^m\|\bar{x}_i^{k+1}-x_i^*\|_2^2|\mathcal{F}^k\right]\\ \mathbb{E}\left[\sum_{i=1}^m\|\mathbf{x}_i^{k+1}-\bar{\mathbf{x}}_i^{k+1}\|_L^2|\mathcal{F}^k\right]\end{bmatrix}\le\left(\begin{bmatrix}1&\kappa_1\gamma^k\\0&1-\kappa_2\gamma^k\end{bmatrix}+a^k\mathbf{1}\mathbf{1}^T\right)\begin{bmatrix}\sum_{i=1}^m\|\bar{x}_i^k-x_i^*\|_2^2\\ \sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2\end{bmatrix}+b^k\mathbf{1}-c^k\begin{bmatrix}\left(\phi(\bar{x}^k)-\phi(x^*)\right)^T(\bar{x}^k-x^*)\\0\end{bmatrix},\quad(6)$$

where

- *$\|\cdot\|_L$ is an L-dependent norm, $\mathcal{F}^k=\{\mathbf{x}_i^\ell,\ i\in[m],\ 0\le\ell\le k\}$, $\bar{x}^k=\left[(\bar{x}_1^k)^T,\cdots,(\bar{x}_m^k)^T\right]^T$;*
- *the random nonnegative scalar sequences $\{a^k\}$, $\{b^k\}$ satisfy $\sum_{k=0}^\infty a^k<\infty$ and $\sum_{k=0}^\infty b^k<\infty$, respectively, a.s., and the deterministic nonnegative sequences $\{c^k\}$ and $\{\gamma^k\}$ satisfy $\sum_{k=0}^\infty c^k=\infty$ and $\sum_{k=0}^\infty\gamma^k=\infty$;*
- *the scalars κ_1 and κ_2 satisfy $\kappa_1>0$ and $0<\kappa_2\gamma^k<1$, respectively, for all $k\ge0$.*

Then, $\lim_{k\to\infty}\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L=0$ and $\lim_{k\to\infty}\|\bar{x}_i^k-x_i^*\|=0$ a.s. for all i, implying $\lim_{k\to\infty}\|x_{(i)i}^k-x_i^*\|=0$ a.s. for all i.

Proof. Since we have $\left(\phi(\bar{x}^k)-\phi(x^*)\right)^T(\bar{x}^k-x^*)>0$ for all k from Assumption 2, by letting $\mathbf{v}^k=\left[\sum_{i=1}^m\|\bar{x}_i^k-x_i^*\|_2^2,\ \sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2\right]^T$, we can arrive at the following relationship from (6) *a.s.* for all $k\ge T$:

$$\mathbb{E}\left[\mathbf{v}^{k+1}|\mathcal{F}^k\right]\le\left(\begin{bmatrix}1&\kappa_1\gamma^k\\0&1-\kappa_2\gamma^k\end{bmatrix}+a^k\mathbf{1}\mathbf{1}^T\right)\mathbf{v}^k+b^k\mathbf{1}.\quad(7)$$

By setting $\pi=[1,\frac{\kappa_1}{\kappa_2}]^T$, we have $\pi^T\begin{bmatrix}1&\kappa_1\gamma^k\\0&1-\kappa_2\gamma^k\end{bmatrix}=\pi^T$.

Since the results of Lemma 1 and Lemma 2 are asymptotic, they remain valid when the starting index is shifted from $k=0$ to $k=T$, for an arbitrary $T\ge0$. Thus, relation (7) implies that $\lim_{k\to\infty}\pi^T\mathbf{v}^k$ exists *a.s.*, and that $\{\sum_{i=1}^m\|\bar{x}_i^k-x_i^*\|_2^2\}$ and $\{\sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2\}$ are bounded almost surely.

Consider the second element of $\mathbf{v}^k$ in (7), which should satisfy $\mathbb{E}\left[\sum_{i=1}^m\|\mathbf{x}_i^{k+1}-\bar{\mathbf{x}}_i^{k+1}\|_L^2|\mathcal{F}^k\right]\le(1-\kappa_2\gamma^k)\sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2+\beta^k\quad\forall k\ge0$, *a.s.*, where $\beta^k\triangleq a^k\left(\sum_{i=1}^m\left(\|\bar{x}_i^k-x_i^*\|_2^2+\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2\right)\right)+b^k$. Using the assumption that $\sum_{k=0}^\infty a^k<\infty$ holds *a.s.*, and the proven results that $\{\sum_{i=1}^m\|\bar{x}_i^k-x_i^*\|_2^2\}$ and $\{\sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2\}$ are bounded *a.s.*, one obtains $\sum_{k=0}^\infty\beta^k<\infty$ *a.s.*Thus, under the assumption of the proposition, $\sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2$ satisfies the conditions of Lemma 1 with $q^k=\kappa_2\gamma^k$ and $p^k=\beta^k$. Therefore, we have the following relationship *a.s.*:

$$\sum_{k=0}^\infty\kappa_2\gamma^k\sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2<\infty,\ \lim_{k\to\infty}\sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2=0.\quad(8)$$

We next proceed to prove $\sum_{i=1}^m\|\bar{x}_i^k-x_i^*\|^2\to0$ almost surely. One can verify that under $\sum_{k=0}^\infty a^k<\infty$ and $\sum_{k=0}^\infty b^k<\infty$, the inequality in (6) satisfies the relationship in Lemma 2 with $\mathbf{v}^k=\left[\sum_{i=1}^m\|\bar{x}_i^k-x_i^*\|_2^2,\ \sum_{i=1}^m\|\mathbf{x}_i^k-\bar{\mathbf{x}}_i^k\|_L^2\right]^T$, $V^k=\begin{bmatrix}1&\kappa_1\gamma^k\\0&1-\kappa_2\gamma^k\end{bmatrix}$, $H^k=\begin{bmatrix}c^k&0\\0&0\end{bmatrix}$, and $\pi^T=[1,\frac{\kappa_1}{\kappa_2}]^T$. Therefore, from Lemma 2, we arrive at the conclusion that $\pi^T \mathbf{v}^k$ converges *a.s.*, i.e., $\sum_{i=1}^m \|\bar{x}_i^k - x_i^*\|_2^2 + \frac{\kappa_1}{\kappa_2}\sum_{i=1}^m \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2$ converges almost surely. Since $\sum_{i=1}^m \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2$ has been proven to converge *a.s.* (see (8)), we know that $\sum_{i=1}^m \|\bar{x}_i^k - x_i^*\|_2^2$ (equivalent to $\|\bar{x}^k - x^*\|_2^2$) converges almost surely. Lemma 2 also implies $\sum_{k=0}^\infty \pi^T H^k \mathbf{u}^k < \infty$ *a.s.*, i.e., $\sum_{k=0}^\infty \left[1, \frac{\kappa_1}{\kappa_2}\right]^T \begin{bmatrix} c^k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \left(\phi(\bar{x}^k) - \phi(x^*)\right)^T (\bar{x}^k - x^*) \\ 0 \end{bmatrix} < \infty$, or

$$\sum_{k=0}^\infty c^k \left(\phi(\bar{x}^k) - \phi(x^*)\right)^T (\bar{x}^k - x^*) < \infty. \tag{9}$$

Next, using (9) and the proven *a.s.* convergence of $\|\bar{x}^k - x^*\|^2$, we prove that $\bar{x}^k$ converges *a.s.* to x^*. The condition $\sum_{k=0}^\infty c^k = \infty$, the property $\left(\phi(\bar{x}^k) - \phi(x^*)\right)^T (\bar{x} - x^*) > 0$ (see Assumption 2), and (9) imply that there exists a subsequence of $\{\bar{x}^k\}$, say $\{\bar{x}^{k_\ell}\}$, along which $\left(\phi(\bar{x}^k) - \phi(x^*)\right)^T (\bar{x}^k - x^*)$ converges *a.s.* to zero. The strongly monotone condition on $\phi(\cdot)$ in Assumption 2 implies that $\{\bar{x}^{k_\ell}\}$ must converge *a.s.* to x^*. This and the fact that $\|\bar{x}^k - x^*\|^2$ converges *a.s.* imply that $\bar{x}^k$ converges *a.s.* to x^*. Further note that $\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2$ converging to zero implies $x_{(\ell)i}^k$ converging to $\bar{x}_i^k$ for all $\ell \in [m]$. Therefore, we have $x_{(i)i}^k$ converging to x_i^* *a.s.* for all $i \in [m]$. ∎

V. Convergence Analysis for Algorithm 1

In this section, based on Proposition 1, we establish the convergence of Algorithm 1 to the unique NE under persistent DP noise satisfying the following assumption:

Assumption 4. *For every $i, \ell \in [m]$ and every k, conditional on the state $x_{(i)\ell}^k$, the DP noise $\zeta_{(i)\ell}^k$ that player i adds to its shared decision (or estimates of other players' decisions) satisfies $\mathbb{E}\left[\zeta_{(i)\ell}^k \mid x_{(i)\ell}^k\right] = 0$, $\mathbb{E}\left[\|\zeta_{(i)\ell}^k\|^2 \mid x_{(i)\ell}^k\right] = (\sigma_i^k)^2$, and $\sum_{k=0}^\infty (\gamma^k)^2 \max_{i\in[m]} (\sigma_i^k)^2 < \infty$, where $\{\gamma^k\}$ is from Algorithm 1. Furthermore, $\mathbb{E}\left[\|x_{(i)\ell}^0\|^2\right] < \infty$, $\forall i, \ell \in [m]$.*

Remark 3. *Since γ^k decreases with time, Assumption 4 even allows the sequence $\{\sigma_i^k\}$ to increase with time. For example, for $\gamma^k = \mathcal{O}(\frac{1}{k^{0.9}})$, if $\{\sigma_i^k\}$ increases with time with a rate no larger than $\mathcal{O}(k^{0.3})$, the summable condition still holds. Allowing $\{\sigma_i^k\}$ to increase with time is key to enabling strong ϵ-DP, which will be detailed later in Theorem 2.*

Theorem 1. *Under Assumptions 1-4, if there exists some $T \geq 0$ such that the sequences $\{\gamma^k\}$ and $\{\lambda^k\}$ satisfy*

$$\sum_{k=T}^\infty \gamma^k = \infty, \sum_{k=T}^\infty \lambda^k = \infty, \sum_{k=T}^\infty (\gamma^k)^2 < \infty, \sum_{k=T}^\infty \frac{(\lambda^k)^2}{\gamma^k} < \infty,$$

Algorithm 1 converges a.s. to the unique NE of problem (1).

Proof. The basic idea is to prove that $\sum_{i=1}^m \|\bar{x}_i^k - x_i^*\|^2$ and $\sum_{i=1}^m \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2$ satisfy the conditions in Proposition 1.

Part I: The evolution of $\sum_{i=1}^m \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2$.

From Algorithm 1, one can obtain the dynamics of $\mathbf{x}_i^k$:

$$\mathbf{x}_i^{k+1} = (I + \gamma^k L)\mathbf{x}_i^k + \gamma^k L_o \boldsymbol{\zeta}_i^k - \lambda^k e_i F_i^T(x_{(i)i}^k, x_{(i)-i}^k), \tag{10}$$

where $e_i \in \mathbb{R}^m$ is a unitary vector with the ith element equal to 1 and all the other elements equal to zero, $L_o \in \mathbb{R}^{m\times m}$ is the matrix obtained by replacing all diagonal entries of matrix L with zero, and $\boldsymbol{\zeta}_i^k = [\zeta_{(1)i}^k, \cdots \zeta_{(m)i}^k]^T \in \mathbb{R}^{m\times d_i}$.

One can obtain that $\bar{\mathbf{x}}_i^k = \frac{\mathbf{1}u^T \mathbf{x}_i^k}{m}$ always holds, which, in combination with (10), yields

$$\bar{\mathbf{x}}_i^{k+1} = \bar{\mathbf{x}}_i^k + \gamma^k \frac{\mathbf{1}u^T L_o}{m} \boldsymbol{\zeta}_i^k - \frac{\lambda^k \mathbf{1} u_i F_i^T(x_{(i)i}^k, x_{(i)-i}^k)}{m}, \tag{11}$$

where u_i is the ith entry of u. Note that in the last equality we used the property $u^T(I + \gamma^k L) = u^T$ from Lemma 4.

Combining (10) with (11) yields

$$\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1} = W^k \mathbf{x}_i^k + \gamma^k \Pi_{L_o} \boldsymbol{\zeta}_i^k - \lambda^k \Pi_{e_i} F_i^T(x_{(i)i}^k, x_{(i)-i}^k), \tag{12}$$

where we have defined $W^k \triangleq I + \gamma^k L - \frac{\mathbf{1}u^T}{m}$, $\Pi_{L_o} \triangleq L_o - \frac{\mathbf{1}u^T L_o}{m}$, and $\Pi_{e_i} \triangleq e_i - \frac{u_i \mathbf{1}}{m}$.

The second statement of Lemma 4 implies $W^k \mathbf{1} = 0$ and further $W^k \bar{\mathbf{x}}_i^k = 0$. Hence, we can subtract $W^k \bar{\mathbf{x}}_i^k = 0$ from the right hand side of (12) to obtain

$$\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1} = W^k(\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k) + \gamma^k \Pi_{L_o} \boldsymbol{\zeta}_i^k - \lambda^k \Pi_{e_i} F_i^T(x_{(i)i}^k, x_{(i)-i}^k). \tag{13}$$

Taking the $\|\cdot\|_L$ norm on both sides leads to

$$\begin{aligned} \|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 &\leq \|W^k(\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k) - \lambda^k \Pi_{e_i} F_i^T(x_{(i)i}^k, x_{(i)-i}^k)\|_L^2 \\ &+ (\gamma^k)^2 \|\Pi_{L_o}\|_L^2 \|\boldsymbol{\zeta}_i^k\|_L^2 \\ &+ 2\left\langle W^k(\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k) - \lambda^k \Pi_{e_i} F_i^T(x_{(i)i}^k, x_{(i)-i}^k), \gamma^k \Pi_{L_o} \boldsymbol{\zeta}_i^k \right\rangle_L. \end{aligned} \tag{14}$$

Taking the conditional expectation on both sides, with respect to $\mathcal{F}_k = \{\mathbf{x}_i^\ell; 0 \leq \ell \leq k, i \in [m]\}$, leads to

$$\begin{aligned} \mathbb{E}\left[\|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 | \mathcal{F}^k\right] &\leq (\gamma^k)^2 \|\Pi_{L_o}\|_L^2 \delta_{L,2}^2 m (\sigma_i^k)^2 \\ &+ \|W^k(\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k) - \lambda^k \Pi_{e_i} F_i^T(x_{(i)i}^k, x_{(i)-i}^k)\|_L^2, \end{aligned} \tag{15}$$

where we have used Assumption 4 and the property $\|\boldsymbol{\zeta}_i^k\|_L^2 \leq \delta_{L,2}^2 \|\boldsymbol{\zeta}_i^k\|_2^2$ (see Lemma 5).

Further using Lemma 5, we can simplify (15) as

$$\begin{aligned} \mathbb{E}\left[\|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 | \mathcal{F}^k\right] &\leq (\gamma^k)^2 \|\Pi_{L_o}\|_L^2 \delta_{L,2}^2 m (\sigma_i^k)^2 \\ &+ \left(\|W^k\|_L \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L + \lambda^k \|\Pi_{e_i}\|_L \|F_i^T(x_{(i)i}^k, x_{(i)-i}^k)\|_L\right)^2. \end{aligned} \tag{16}$$

According to Lemma 4, we have $\|W^k\|_L \leq 1 - \alpha\gamma^k$ for some $0 < \alpha < 1$ when γ^k is sufficiently small. Given that $\{\gamma^k\}$ is square summable, we have $\|W^k\|_L \leq 1 - \alpha\gamma^k$ for some $0 < \alpha < 1$ when k is larger than some T. Therefore, (16) means that there always exists a $T \geq 0$ such that we have

$$\begin{aligned} \mathbb{E}\left[\|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 | \mathcal{F}^k\right] &\leq (\gamma^k)^2 \|\Pi_{L_o}\|_L^2 \delta_{L,2}^2 m (\sigma_i^k)^2 \\ &+ \left((1 - \alpha\gamma^k)\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L + \lambda^k \|\Pi_{e_i}\|_L \|F_i^T(x_{(i)i}^k, x_{(i)-i}^k)\|_L\right)^2 \end{aligned} \tag{17}$$

for $k \geq T$.

Applying to the second term on the right hand side of (17) the inequality $(a+b)^2 \leq (1+\epsilon)a^2 + (1+\epsilon^{-1})b^2$, valid for any scalars a, b, and $\epsilon > 0$ [25], we can obtain the following relationship by setting ϵ as $\frac{\gamma^k \alpha}{1 - \gamma^k \alpha}$ (which further results in $1 + \epsilon = \frac{1}{1 - \gamma^k \alpha}$ and $1 + \epsilon^{-1} = \frac{1}{\gamma^k \alpha}$):

$$\begin{aligned} \mathbb{E}\left[\|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 | \mathcal{F}^k\right] &\leq (\gamma^k)^2 \|\Pi_{L_o}\|_L^2 \delta_{L,2}^2 m (\sigma_i^k)^2 \\ &+ (1 - \alpha\gamma^k)\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2 + \frac{(\lambda^k)^2 \|\Pi_{e_i}\|_L^2}{\gamma^k \alpha} \|F_i^T(x_{(i)i}^k, x_{(i)-i}^k)\|_L^2. \end{aligned} \tag{18}$$

Next, we use Assumption 2 to bound $\|F_i^T(x_{(i)i}^k, x_{(i)-i}^k)\|_L^2$.

At the NE point x^*, we always have $F_i(x_i^*, x_{-i}^*) = 0$ for all $i \in [m]$, which implies

$$\begin{aligned} &\|F_i(x_{(i)i}^k, x_{(i)-i}^k)\|_L^2 \\ &\le \delta_{L,2}^2 \|F_i(x_{(i)i}^k, x_{(i)-i}^k) - F_i(x_{(i)i}^k, x_{-i}^*) + F_i(x_{(i)i}^k, x_{-i}^*) \\ &\quad - F_i(x_i^*, x_{-i}^*)\|_2^2 \\ &\le 2K_2^2\delta_{L,2}^2 \|x_{(i)-i}^k - x_{-i}^*\|_2^2 + 2K_1^2\delta_{L,2}^2\|x_{(i)i}^k - x_i^*\|_2^2, \end{aligned} \tag{19}$$

where in the last inequality we used Assumption 2.

Using inequalities $\|x_{(i)-i}^k - x_{-i}^*\|_2^2 \le 2\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 + 2\|\bar{x}_{-i}^k - x_{-i}^*\|_2^2$ and $\|x_{(i)i}^k - x_i^*\|_2^2 \le 2\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + 2\|\bar{x}_i^k - x_i^*\|_2^2$, where $\bar{x}_{-i}^k \triangleq [(\bar{x}_1^k)^T, \cdots, (\bar{x}_{i-1}^k)^T, (\bar{x}_{i+1}^k)^T, \cdots, (\bar{x}_m^k)^T]^T$, we can obtain

$$\begin{aligned} &\|F_i(x_{(i)i}^k, x_{(i)-i}^k)\|_L^2 \le 4K_1^2\delta_{L,2}^2(\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + \|\bar{x}_i^k - x_i^*\|_2^2) \\ &+ 4K_2^2\delta_{L,2}^2(\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 + \|\bar{x}_{-i}^k - x_{-i}^*\|_2^2). \end{aligned} \tag{20}$$

Plugging (20) into (18) leads to

$$\begin{aligned} &\mathbb{E}\left[\|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 | \mathcal{F}^k\right] \\ &\le (1-\alpha\gamma^k)\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2 + (\gamma^k)^2\|\Pi_{L_o}\|_L^2\delta_{L,2}^2 m(\sigma_i^k)^2 \\ &+ \tfrac{4(\lambda^k)^2\|\Pi_{e_i}\|_L^2 K_2^2\delta_{L,2}^2}{\gamma^k\alpha}(\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 + \|\bar{x}_{-i}^k - x_{-i}^*\|_2^2) \\ &+ \tfrac{4(\lambda^k)^2\|\Pi_{e_i}\|_L^2 K_1^2\delta_{L,2}^2}{\gamma^k\alpha}(\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + \|\bar{x}_i^k - x_i^*\|_2^2). \end{aligned} \tag{21}$$

Summing (21) from $i=1$ to $i=m$ and noting $\sum_{i=1}^m \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_2^2 = \sum_{i=1}^m(\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + \|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2)$ and $\sum_{i=1}^m \|\bar{x}_{-i}^k - x_{-i}^*\|_2^2 = (m-1)\sum_{i=1}^m \|\bar{x}_i^k - x_i^*\|_2^2$, we obtain

$$\begin{aligned} &\mathbb{E}\left[\textstyle\sum_{i=1}^m \|\mathbf{x}_i^{k+1} - \bar{\mathbf{x}}_i^{k+1}\|_L^2 | \mathcal{F}^k\right] \le (1-\alpha\gamma^k)\textstyle\sum_{i=1}^m \|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2 \\ &+ (\gamma^k)^2\|\Pi_{L_o}\|_L^2\delta_{L,2}^2 m\textstyle\sum_{i=1}^m(\sigma_i^k)^2 \\ &+ \tfrac{4m(\lambda^k)^2\|\Pi_{e_i}\|_L^2\tilde{K}^2\delta_{L,2}^2}{\gamma^k\alpha}\textstyle\sum_{i=1}^m\|\bar{x}_i^k - x_i^*\|_2^2 \\ &+ \tfrac{4(\lambda^k)^2\|\Pi_{e_i}\|_L^2\tilde{K}^2\delta_{L,2}^2\delta_{2,L}^2}{\gamma^k\alpha}\textstyle\sum_{i=1}^m\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2, \end{aligned} \tag{22}$$

where we have defined $\tilde{K} \triangleq \max\{K_1, K_2\}$.

Part II: The evolution of $\sum_{i=1}^m \|\bar{x}_i^k - x_i^*\|^2$.

Subtracting x_i^* from both sides of (11) yields

$$\begin{aligned} &\|\bar{x}_i^{k+1} - x_i^*\|_2^2 \le \|\bar{x}_i^k - x_i^*\|^2 + 2(\gamma^k)^2\tfrac{\|u^T L_o\|_2^2}{m^2}\|\boldsymbol{\zeta}_i^k\|_2^2 \\ &\quad + \tfrac{2(\lambda^k)^2u_i^2}{m^2}\|F_i(x_{(i)i}^k, x_{(i)-i}^k)\|_2^2 \\ &\quad + 2\left\langle \bar{x}_i^k - x_i^*, \gamma^k\tfrac{(u^T L_o\boldsymbol{\zeta}_i^k)^T}{m} - \tfrac{1}{m}\lambda^k u_i F_i(x_{(i)i}^k, x_{(i)-i}^k)\right\rangle. \end{aligned} \tag{23}$$

Taking the conditional expectation on both sides, with respect to $\mathcal{F}_k = \{\mathbf{x}_i^\ell; 0 \le \ell \le k,\ i \in [m]\}$, leads to

$$\begin{aligned} &\mathbb{E}\left[\|\bar{x}_i^{k+1} - x_i^*\|_2^2|\mathcal{F}^k\right] \le \|\bar{x}_i^k - x_i^*\|_2^2 + \tfrac{2(\gamma^k)^2\|u^T L_o\|_2^2(\sigma_i^k)^2}{m} \\ &+ \tfrac{2(\lambda^k)^2u_i^2\|F_i(x_{(i)i}^k, x_{(i)-i}^k)\|_2^2}{m^2} - \tfrac{2u_i\lambda^k\langle\bar{x}_i^k - x_i^*, F_i(x_{(i)i}^k, x_{(i)-i}^k)\rangle}{m}. \end{aligned} \tag{24}$$

Next we bound the last two terms in (24). For the second last term, we can bound it similarly to (20):

$$\begin{aligned} &\|F_i(x_{(i)i}^k, x_{(i)-i}^k)\|_2^2 \le 4K_2^2\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 \\ &+ 4K_2^2\|\bar{x}_{-i}^k - x_{-i}^*\|_2^2 + 4K_1^2\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + 4K_1^2\|\bar{x}_i^k - x_i^*\|_2^2. \end{aligned} \tag{25}$$

For the inner-product term in (24), we bound it using $F_i(x_i^*, x_{-i}^*) = 0$ and split it as follows:

$$\begin{aligned} &2\lambda^k\left\langle \bar{x}_i^k - x_i^*, F_i(x_{(i)i}^k, x_{(i)-i}^k) - F_i(\bar{x}_i^k, \bar{x}_{-i}^k)\right\rangle \\ &+ 2\lambda^k\left\langle \bar{x}_i^k - x_i^*, F_i(\bar{x}_i^k, \bar{x}_{-i}^k) - F_i(x_i^*, x_{-i}^*)\right\rangle. \end{aligned} \tag{26}$$

For the first inner-product term on the right hand side of (26), using the Cauchy-Schwarz inequality yields

$$\begin{aligned} &2\lambda^k\left\langle \bar{x}_i^k - x_i^*, F_i(x_{(i)i}^k, x_{(i)-i}^k) - F_i(\bar{x}_i^k, \bar{x}_{-i}^k)\right\rangle \\ &\ge -\tfrac{(\lambda^k)^2\|\bar{x}_i^k - x_i^*\|_2^2}{\gamma^k} - \gamma^k\|F_i(x_{(i)i}^k, x_{(i)-i}^k) - F_i(\bar{x}_i^k, \bar{x}_{-i}^k)\|_2^2. \end{aligned} \tag{27}$$

The Lipschitz assumption in Assumption 2 implies

$$\begin{aligned} &\|F_i(x_{(i)i}^k, x_{(i)-i}^k) - F_i(\bar{x}_i^k, \bar{x}_{-i}^k)\|_2^2 \\ &\le 2\|F_i(x_{(i)i}^k, x_{(i)-i}^k) - F_i(\bar{x}_i^k, x_{(i)-i}^k)\|_2^2 \\ &\quad + 2\|F_i(\bar{x}_i^k, x_{(i)-i}^k) - F_i(\bar{x}_i^k, \bar{x}_{-i}^k)\|_2^2 \\ &\le 2K_1^2\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + 2K_2^2\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2. \end{aligned} \tag{28}$$

Combining (26), (27), and (28) leads to

$$\begin{aligned} &2\lambda^k\left\langle \bar{x}_i^k - x_i^*, F_i(x_{(i)i}^k, x_{(i)-i}^k)\right\rangle \\ &\ge -\tfrac{(\lambda^k)^2\|\bar{x}_i^k - x_i^*\|_2^2}{\gamma^k} - 2K_1^2\gamma^k\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 \\ &\quad - 2K_2^2\gamma^k\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 \\ &\quad + 2\lambda^k\left\langle \bar{x}_i^k - x_i^*, F_i(\bar{x}_i^k, \bar{x}_{-i}^k) - F_i(x_i^*, x_{-i}^*)\right\rangle. \end{aligned} \tag{29}$$

Further substituting (25) and (29) into (24) yields

$$\begin{aligned} &\mathbb{E}\left[\|\bar{x}_i^{k+1} - x_i^*\|_2^2|\mathcal{F}^k\right] \le \|\bar{x}_i^k - x_i^*\|_2^2 + 2(\gamma^k)^2\tfrac{\|u^T L_o\|_2^2}{m}(\sigma_i^k)^2 \\ &+ \tfrac{8(\lambda^k)^2u_i^2K_2^2}{m^2}(\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 + \|\bar{x}_{-i}^k - x_{-i}^*\|_2^2) \\ &+ \tfrac{8(\lambda^k)^2u_i^2K_1^2}{m^2}(\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + \|\bar{x}_i^k - x_i^*\|_2^2) \\ &+ \tfrac{u_i(\lambda^k)^2\|\bar{x}_i^k - x_i^*\|_2^2}{m\gamma^k} + \tfrac{2u_iK_1^2\gamma^k}{m}\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 \\ &+ \tfrac{2u_iK_2^2\gamma^k}{m}\|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2 \\ &- \tfrac{2u_i\lambda^k}{m}\left\langle \bar{x}_i^k - x_i^*, F_i(\bar{x}_i^k, \bar{x}_{-i}^k) - F_i(x_i^*, x_{-i}^*)\right\rangle. \end{aligned} \tag{30}$$

Summing (30) from $i=1$ to $i=m$, and using the relationship $\sum_{i=1}^m\|\bar{x}_{-i}^k - x_{-i}^*\|_2^2 = (m-1)\sum_{i=1}^m\|\bar{x}_i^k - x_i^*\|_2^2$ and $\sum_{i=1}^m\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_2^2 = \sum_{i=1}^m\left(\|x_{(i)i}^k - \bar{x}_i^k\|_2^2 + \|x_{(i)-i}^k - \bar{x}_{-i}^k\|_2^2\right)$ lead to

$$\begin{aligned} &\mathbb{E}\left[\textstyle\sum_{i=1}^m\|\bar{x}_i^{k+1} - x_i^*\|_2^2|\mathcal{F}^k\right] \\ &\le \textstyle\sum_{i=1}^m\|\bar{x}_i^k - x_i^*\|_2^2 + 2(\gamma^k)^2\tfrac{\|u^T L_o\|_2^2}{m}\textstyle\sum_{i=1}^m(\sigma_i^k)^2 \\ &+ \tfrac{8(\lambda^k)^2u_i^2\tilde{K}^2}{m^2}\textstyle\sum_{i=1}^m\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_2^2 + \tfrac{8(\lambda^k)^2u_i^2\tilde{K}^2}{m}\textstyle\sum_{i=1}^m\|\bar{x}_i^k - x_i^*\|_2^2 \\ &+ \tfrac{u_i(\lambda^k)^2\sum_{i=1}^m\|\bar{x}_i^k - x_i^*\|_2^2}{m\gamma^k} + \tfrac{2u_i\tilde{K}^2\gamma^k}{m}\textstyle\sum_{i=1}^m\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_2^2 \\ &- \tfrac{2u_i\lambda^k}{m}\left(\phi(\bar{x}^k) - \phi(x^*)\right)^T(\bar{x}^k - x^*), \end{aligned} \tag{31}$$

where $\bar{x}^k = \left[(\bar{x}_1^k)^T, \cdots, (\bar{x}_m^k)^T\right]^T$ and $\tilde{K} \triangleq \max\{K_1, K_2\}$.

Part III: Combination of Step I and Step II.

By combining (22) and (31), and using Assumption 4, we have $\sum_{i=1}^m\|\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k\|_L^2$ and $\sum_{i=1}^m\|\bar{x}_i^k - x_i^*\|_2^2$ satisfying the conditions in Proposition 1 with $\kappa_1 = \frac{2u_i\tilde{K}^2\delta_{L,2}^2}{m}$, $\kappa_2 = \alpha$, $a^k = \max\{a_1^k, a_2^k, a_3^k, a_4^k, a_5^k\}$, $a_1^k \triangleq$

$\frac{4(\lambda^k)^2\|\Pi_{e_i}\|_L^2\tilde{K}^2\delta_{L,2}^2\delta_{2,L}^2}{\gamma^k\alpha}$, $a_2^k \triangleq \frac{4m(\lambda^k)^2\|\Pi_{e_i}\|_L^2\tilde{K}^2\delta_{L,2}^2}{\gamma^k\alpha}$, $a_3^k \triangleq \frac{8(\lambda^k)^2u_i^2\tilde{K}^2\delta_{L,2}^2}{m^2}$, $a_4^k \triangleq \frac{8(\lambda^k)^2u_i^2\tilde{K}^2}{m}$, $a_5^k \triangleq \frac{u_i(\lambda^k)^2}{m\gamma^k}$, $b^k = \max\{b_1^k, b_2^k\}$, $b_1^k \triangleq (\gamma^k)^2\|\Pi_{L_o}\|_L^2\delta_{L,2}^2 m\sum_{i=1}^m(\sigma_i^k)^2$, $b_2^k \triangleq 2(\gamma^k)^2\frac{\|u^T L_o\|_2^2}{m}\sum_{i=1}^m(\sigma_i^k)^2$, and $c^k = \frac{2u_i\lambda^k}{m}$. ■

Remark 4. *The requirement on γ^k and λ^k in the statement of Theorem 1 can be satisfied, for example, by setting $\gamma^k = \mathcal{O}(\frac{1}{k^a})$ and $\lambda^k = \mathcal{O}(\frac{1}{k^b})$ with $a, b \in \mathbb{R}$ satisfying $0.5 < a < b \leq 1$ and $2b - a > 1$.*

VI. Privacy Analysis of Algorithm 1

Definition 3. *For any initial state ϑ^0 and any adjacent networked games $\mathcal{P}$ and $\mathcal{P}'$, the sensitivity of an NE seeking algorithm at iteration k is*

$$\Delta^k \triangleq \sup_{\mathcal{O}\in\mathbb{O}}\left\{\sup_{\vartheta\in\mathcal{R}^{-1}_{\mathcal{P},\vartheta^0}(\mathcal{O}),\, \vartheta'\in\mathcal{R}^{-1}_{\mathcal{P}',\vartheta^0}(\mathcal{O})}\|\vartheta^k - \vartheta'^k\|_1\right\}. \quad (32)$$

Based on this definition, we obtain the following result:

Lemma 6. *At each iteration k, if each player in Algorithm 1 adds a vector noise $\zeta_{(i)\ell}^k \in \mathbb{R}^{d_i}$ (consisting of d_i independent Laplace noises with parameter ν^k) to each of its shared message $x_{(i)\ell}^k$ such that $\sum_{k=1}^{T_0}\frac{\Delta^k}{\nu^k} \leq \bar{\epsilon}$, then Algorithm 1 is ϵ-differentially private with the cumulative privacy budget from iterations $k = 0$ to $k = T_0$ less than $\bar{\epsilon}$.*

Proof. The result can be obtained following the derivation of Lemma 2 in [20] (see also Theorem 3 in [13]). ■

Before giving the main results, we first use Definition 1 and the guaranteed convergence in Theorem 1 to confine the sensitivity. Note that when the conditions in the statement of Theorem 1 are satisfied, our algorithm ensures convergence of both $\mathcal{P}$ and $\mathcal{P}'$ to their respective NEs, which are the same under the third requirement in Definition 1. This means that $\left\|F_i(x_{(i)i}^k, x_{(i)-i}^k) - F'_i(x'^k_{(i)i}, x'^k_{(i)-i})\right\|_1 = 0$ will hold when k is sufficiently large (for the iterates in both $\mathcal{P}$ and $\mathcal{P}'$ to enter the neighborhood B_δ in Definition 1, upon which the evolution in $\mathcal{P}$ and $\mathcal{P}'$ will be identical). Furthermore, the ensured convergence also means that $F_i(x_{(i)i}^k, x_{(i)-i}^k)$ and $F'_i(x'^k_{(i)i}, x'^k_{(i)-i})$ are always bounded. Hence, there always exists some constant C such that the following relation holds for all $k \geq 0$ under the conditions of Theorem 1:

$$\left\|F_i(x_{(i)i}^k, x_{(i)-i}^k) - F'_i(x'^k_{(i)i}, x'^k_{(i)-i})\right\|_1 \leq C\gamma^k \quad (33)$$

Theorem 2. *Under the conditions in Theorem 1, if all elements of $\zeta_{(i)1}^k, \cdots, \zeta_{(i)m}^k$ follow Laplace distribution $\mathrm{Lap}(\nu^k)$ with $(\sigma_i^k)^2 = 2(\nu^k)^2$ satisfying Assumption 4, then Algorithm 1 is ϵ-differentially private with the cumulative privacy budget from $k = 0$ to $k = T_0$ bounded by $\epsilon \leq \sum_{k=1}^{T_0}\frac{C\varsigma^k}{\nu^k}$, where C is given in (33) and $\varsigma^k \triangleq \sum_{p=1}^{k-1}(\Pi_{q=p}^{k-1}(1-\bar{L}\gamma^q))\gamma^{p-1}\lambda^{p-1} + \gamma^{k-1}\lambda^{k-1}$ with $\bar{L} \triangleq \min_i\{|L_{ii}|\}$. The cumulative privacy budget is finite even as $T_0 \to \infty$ when $\{\frac{\lambda^k}{\nu^k}\}$ is summable.*

Proof. According to Definition 3, the sensitivity at iteration k is determined by $\|\vartheta^k - \vartheta'^k\|_1$. Since $\mathcal{P}$ and $\mathcal{P}'$ are adjacent, only one of their cost functions is different. Pick this different cost function as the ith one, i.e., $f_i(\cdot)$, without loss of generality. Given that the observations under $\mathcal{P}$ and $\mathcal{P}'$ are identical, we have $x_{(i)\ell}^k = x'^k_{(i)\ell}$ for all $k \geq 0$ and $\ell \neq i$.

By defining $x_{(j):}^k \triangleq [(x_{(j)1}^k)^T, \cdots, (x_{(j)m}^k)^T]^T$, we have

$$\|\vartheta^k - \vartheta'^k\|_1 =$$
$$\left\|\left[(x_{(1):}^k)^T, \cdots, (x_{(m):}^k)^T\right]^T - \left[(x'^k_{(1):})^T, \cdots, (x'^k_{(m):})^T\right]^T\right\|_1$$
$$= \left\|\left[\ x_{(i):}^k - x'^k_{(i):}\ \right]\right\|_1 = \left\|\left[\ x_{(i)i}^k - x'^k_{(i)i}\ \right]\right\|_1,$$

where in the second-to-last equality we used the fact that only the ith cost function is different, and in the last equality, we used the fact that $x_{(i)\ell}^k$ and $x'^k_{(i)\ell}$ for $\ell \neq i$ are updated independently of $F_i(\cdot,\cdot)$ and $F'_i(\cdot,\cdot)$, and hence are the same when observations are identical in $\mathcal{P}$ and $\mathcal{P}'$.

Representing $F_i^k = F_i(x_{(i)i}^k, x_{(i)-i}^k)$, we have the following relationship from (3):

$$\begin{aligned} x_{(i)i}^{k+1} - x'^{k+1}_{(i)i} &= x_{(i)i}^k + \gamma^k\textstyle\sum_{j\in\mathbb{N}_i^{\text{in}}}L_{ij}(x_{(j)i}^k + \zeta_{(j)i}^k - x_{(i)i}^k) - \lambda^k F_i^k \\ &\quad - x'^k_{(i)i} - \gamma^k\textstyle\sum_{j\in\mathbb{N}_i^{\text{in}}}L_{ij}(x'^k_{(j)i} + \zeta'^k_{(j)i} - x'^k_{(i)i}) + \lambda^k F'^k_i \\ &= (1 - \gamma^k|L_{ii}|)(x_{(i)i}^k - x'^k_{(i)i}) - \lambda^k(F_i^k - F'^k_i), \end{aligned}$$

where we have used the fact that the shared messages $x_{(j)i}^k + \zeta_{(j)i}^k$ and $x'^k_{(j)i} + \zeta'^k_{(j)i}$ are the same. Since all conditions of Theorem 1 are satisfied, Theorem 1 ensures convergence in both $\mathcal{P}$ and $\mathcal{P}'$, implying that the sensitivity Δ^k satisfies

$$\Delta^{k+1} \leq (1 - |L_{ii}|\gamma^k)\Delta^k + C\gamma^k\lambda^k \quad (34)$$

where C is from (33). Hence, we can arrive at the first privacy statement by iteration.

For the infinity horizon result, we exploit Lemma 3. More specially, Lemma 3 implies that (34) guarantees $\Delta^k < \bar{C}\lambda^k$ for some $\bar{C}$. Hence, according to Lemma 6, we know that the privacy budget is always finite when the sequence $\{\frac{\lambda^k}{\nu^k}\}$ is summable. ■

VII. Numerical Simulations

We use the networked Nash-Cournot game in [2], [4], [7] to evaluate our approach. Due to space limitations, we suppress the application details of this game and only provide the mathematical representation. More specifically, we consider 20 players with each player having a cost function $f_i(x_i, x) = x_i^T Q_i x_i + q_i^T x_i - (\bar{P} - \Xi Bx)^T B_i^T x_i$, where $x_i \in \mathbb{R}^{d_i}$ with $1 \leq d_i \leq N$. $Q_i \in \mathbb{R}^{d_i\times d_i}$ is a randomly generated positive definite matrix and $q_i \in \mathbb{R}^{d_i}$. $\bar{P}$ is a positive vector and Ξ is a diagonal matrix with positive diagonal entries, both of which are randomly chosen in the numerical simulation. B is constructed as $B \triangleq [B_1, \cdots, B_N]$, where $B_i \in \mathbb{R}^{N\times d_i}$ is chosen following [7]. The communication graph is generated randomly but is assured to be strongly connected.

To evaluate the proposed approach, we inject vector noise $\zeta_{(i)\ell}^k$ ($1 \leq \ell \leq 20$) in every message $x_{(i)\ell}^k$ that player i shares in every iteration. Each element of the noise vector $\zeta_{(i)\ell}^k$ follows Laplace distribution with parameter $\nu^k = 1 + 0.1k^{0.2}$. We set $\lambda^k = \frac{0.1}{1+0.1k}$ and $\gamma^k = \frac{1}{1+0.1k^{0.9}}$, respectively, which satisfy the conditions in Theorems 1 and 2. We ran our Algorithm

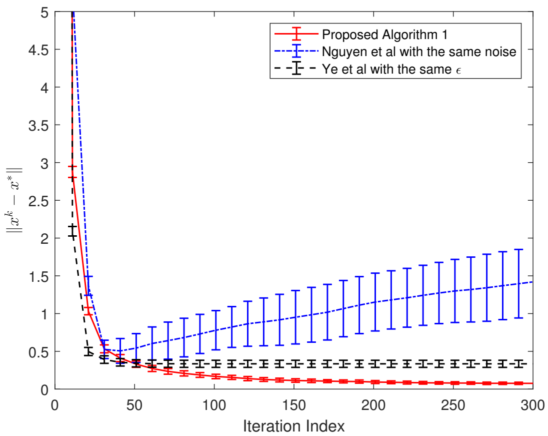

Fig. 1. Comparison with the existing distributed NE seeking algorithm in [7] (under the same noise level) and the differential-privacy approach for aggregative games in [13] (under the same privacy budget ϵ).

1 for 100 times and calculated the average of the distance $\|x^k - x^*\|$ as a function of k. We also calculated the variance of the distance of the 100 runs as a function of k. The trajectories of the average and variance are given by the red curve and error bars in Fig. 1. For comparison, we also ran the distributed NE seeking algorithm proposed by Nguyen et al. in [7] under the same noise level, and the DP approach for networked games proposed by Ye et al. in [13] under the same privacy budget ϵ. Note that [13] addresses undirected graphs but its DP strategy, i.e., geometrically decreasing stepsizes for a finite privacy budget, can be adapted to the directed-graph scenario. The average errors/variances of the two approaches are given by the blue and black curve/error bars in Fig. 1. The comparison clearly shows that our approach has a better accuracy.

VIII. Conclusions

This paper has introduced a distributed NE seeking approach that can ensure both almost sure convergence and rigorous ϵ-DP, even when the number of iterations tends to infinity. The simultaneous achievement of both goals is in sharp contrast to existing DP solutions for aggregative games that trade provable convergence for privacy, and to our knowledge, has not been achieved before for general networked games. The approach is applicable to general directed graphs. Numerical results confirm effectiveness of the proposed approach.

References

[1] P. Frihauf, M. Krstic, and T. Başar, "Nash equilibrium seeking in noncooperative games," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1192–1207, 2011.

[2] L. Pavel, "Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach," *IEEE Trans. Autom. Control*, vol. 65, no. 4, pp. 1584–1597, 2020.

[3] G. Belgioioso, A. Nedić, and S. Grammatico, "Distributed generalized Nash equilibrium seeking in aggregative games under partial-decision information via dynamic tracking," in *Proc. IEEE Conf. Decis. Control*. IEEE, 2019, pp. 5948–5954.

[4] J. Koshal, A. Nedić, and U. V. Shanbhag, "Distributed algorithms for aggregative games on graphs," *Oper. Res.*, vol. 64, no. 3, pp. 680–704, 2016.

[5] F. Salehisadaghiani and L. Pavel, "Distributed Nash equilibrium seeking: A gossip-based algorithm," *Automatica*, vol. 72, pp. 209–216, 2016.

[6] M. Bianchi, G. Belgioioso, and S. Grammatico, "Fast generalized Nash equilibrium seeking under partial-decision information," *Automatica*, vol. 136, p. 110080, 2022.

[7] D. T. A. Nguyen, D. T. Nguyen, and A. Nedić, "Distributed Nash equilibrium seeking over time-varying directed communication networks," *https://arxiv.org/abs/2201.02323*, 2022.

[8] Y. Lu and M. Zhu, "Game-theoretic distributed control with information-theoretic security guarantees," *IFAC-PapersOnLine*, vol. 48, no. 22, pp. 264–269, 2015.

[9] I. Shilov, H. Le Cadre, and A. Busic, "Privacy impact on generalized Nash equilibrium in peer-to-peer electricity market," *Oper. Res. Lett.*, vol. 49, no. 5, pp. 759–766, 2021.

[10] S. Gade, A. Winnicki, and S. Bose, "On privatizing equilibrium computation in aggregate games over networks," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 3272–3277, 2020.

[11] M. Shakarami, C. De Persis, and N. Monshizadeh, "Distributed dynamics for aggregative games: Robustness and privacy guarantees," *Int. J. Robust Nonlinear Control*, vol. 32, no. 9, pp. 5048–5069, 2022.

[12] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy." *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3-4, pp. 211–407, 2014.

[13] M. Ye, G. Hu, L. Xie, and S. Xu, "Differentially private distributed Nash equilibrium seeking for aggregative games," *IEEE Trans. Autom. Control*, vol. 67, no. 5, pp. 2451–2458, 2021.

[14] J. Wang, J.-F. Zhang, and X. He, "Differentially private distributed algorithms for stochastic aggregative games," *Automatica*, vol. 142, p. 110440, 2022.

[15] Y. Wang and A. Nedić, "Tailoring gradient methods for differentially-private distributed optimization," *IEEE Trans. Autom. Control*, 2023(early access 4 May 2023).

[16] Y. Wang and T. Başar, "Quantization enabled privacy protection in decentralized stochastic optimization," *IEEE Trans. Autom. Control*, vol. 68, no. 7, pp. 4038–4052, 2023.

[17] Y. Wang and A. Nedić, "Differentially-private distributed algorithms for aggregative games with guaranteed convergence," *IEEE Trans. Autom. Control*, 2022.

[18] G. Scutari, F. Facchinei, J.-S. Pang, and D. P. Palomar, "Real and complex monotone communication games," *IEEE Trans. Inf. Theory.*, vol. 60, no. 7, pp. 4197–4231, 2014.

[19] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum, "Differential privacy under continual observation," in *Proc. Annu. ACM Symp. Theory Comput.*, 2010, pp. 715–724.

[20] Z. Huang, S. Mitra, and N. Vaidya, "Differentially private distributed optimization," in *Proc. Int. Conf. Distrib. Comput. Netw.*, New York, NY, USA, 2015.

[21] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge university press, 1990.

[22] S. Pu, W. Shi, J. Xu, and A. Nedić, "Push-pull gradient methods for distributed optimization in networks," *IEEE Trans. Autom. Control*, vol. 66, no. 1, pp. 1–16, 2021.

[23] Y. Wang, "The spectral radius of a square matrix can be approximated by its "weighted" spectral norm," *https://arxiv.org/abs/2304.10421*, 2023.

[24] R. Xin, A. K. Sahu, U. A. Khan, and S. Kar, "Distributed stochastic optimization with gradient tracking over strongly-connected networks," in *Proc. IEEE Conf. Decis. Control*. IEEE, 2019, pp. 8353–8358.

[25] Y. Wang and A. Nedić, "Decentralized gradient methods with time-varying uncoordinated stepsizes: Convergence analysis and privacy design," *IEEE Transactions on Automatic Control*, 2023.