A Distributed Douglas-Rachford Based Algorithm for Stochastic GNE Seeking with Partial Information

Yuanhanqing Huang[1] and Jianghai Hu[1]

Abstract— **We consider the stochastic generalized Nash equilibrium problem (SGNEP) where a set of self-interested players, subject to certain global constraints, aim to optimize their local objectives that depend on their own decisions and the decisions of others and are influenced by some random factors. A distributed stochastic generalized Nash equilibrium seeking algorithm is proposed based on the Douglas-Rachford operator splitting scheme, which only requires local communications among neighbors. The proposed scheme significantly relaxes assumptions on co-coercivity and contractiveness in the existing literature, where the projected stochastic subgradient method is applied to provide approximate solutions to the augmented best-response subproblems for each player. Finally, we illustrate the validity of the proposed algorithm through a Nash-Cournot production game.**

I. INTRODUCTION

The Nash equilibrium problem (NEP), rooted in the seminal work [1], models a situation where a set of self-interested players aim to optimize their individual payoffs that depend not only on their own decisions but also on the decisions of others. The stochastic generalized Nash equilibrium problem (SGNEP) extends the NEP by considering additional global resource constraints that these players should collectively satisfy [2], [3], and by incorporating stochasticity in players' objectives. In SGNEPs, instead of deterministic objective functions, players optimize the expected values of uncertain objective functions that are dependent on some random variables. There has been a surge of interest in using SGNEPs to model problems arising in areas such as power markets [4], [5] and congestion control [6]. Nevertheless, it remains challenging to compute the Nash equilibria of SGNEPs, due to the absence of closed-form expressions of the objective functions. Fortunately, as has been shown in [7, Ch. 1.4], many SGNEPs can be formulated as stochastic variational inequalities (SVIs) and solved by leveraging existing results from this field, e.g. [8], [9], [10], [11], [12].

There is an enduring research interest in distributing the computation of Nash equilibria [13], [14], especially through the avenue of operator splitting technique [15], [16]. Significant attention and efforts have been devoted to designing algorithms to solve SGNEPs distributedly under the full-decision information setting where each player has access to all other players' decisions. The authors of [17] propose a solution based on the preconditioned forward-backward (FB) operator splitting with the expected-value pseudogradient assumed to be restricted co-coercive and approximated via the stochastic approximation (SA) scheme. To accelerate game dynamics and relax the co-coercivity assumption, [18] adopts a forward-backward-forward framework. The work in [19] provides an inexact generalization of the proximal best-response (BR) schemes to SNEPs whose corresponding proximal BR maps admit a contractive property. In addition to the distributed computation, in most cases, participants only have access to the local information and decisions of their neighbors, which constitutes a partial-decision information setting [20], [21], [22]. Far less has been studied when it comes to the distributed solution to SGNEPs under the partial-decision information setting. The only existing work to our best knowledge is [23], which also relies on the FB framework along with the SA method. The convergence of the proposed algorithms in [23] is analyzed under the assumption that the preconditioned forward operator is restricted co-coercive, which only allows comparatively small step sizes.

Our contributions can be summarized in the following respects. First, we propose a distributed algorithm to the SGNEP with merely partial information based on the Douglas-Rachford splitting and the proximal mapping. In the proposed algorithm, the involved players are asked to update their decision vectors in two separate steps: solving the augmented best-response subproblems, and projecting onto the local feasible sets after some linear transformations. The updates of their local estimates and dual variables only require some trivial linear transformations. This algorithm can deal with cases where the scenario-based objectives of players are nonsmooth, and relaxes some commonly-made assumptions such as the co-coercivity of the operators after the splitting. Second, we establish the convergence of the proposed algorithm without resorting to the contractive property. The proof in this paper is based on the Robbins-Siegmund theorem and extends the convergence results discussed in [19]. Drawing tools and techniques from stochastic approximation and convex analysis, we construct a feasible inexact solver based on the projected stochastic subgradient method and discuss the prescribed accuracy within which the inexact solver to the augmented best-response subproblems should achieve such that the convergence property is ensured. Complete proofs of the main statements and some intermediate results are omitted due to the space limit, which are available in [24].

Basic Notations: For a set of matrices $\{V_i\}_{i\in S}$, we let $\mathrm{blkd}(V_1, \ldots, V_{|S|})$ or $\mathrm{blkd}(V_i)_{i\in S}$ denote the diagonal concatenation of these matrices, $[V_1, \ldots, V_{|S|}]$ their horizontal

This work was supported by the National Science Foundation under Grant No. 2014816.

[1]The authors are with the Elmore Family School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 47907, USA {huan1282, jianghai}@purdue.edu

stack, and $[V_1; \cdots; V_{|S|}]$ their vertical stack. For a set of vectors $\{v_i\}_{i\in S}$, $[v_i]_{i\in S}$ or $[v_1; \cdots; v_{|S|}]$ denotes their vertical stack. For a matrix V and a pair of positive integers (i, j), $[V]_{(i,j)}$ denotes the entry on the ith row and the jth column of V. For a vector v and a positive integer i, $[v]_i$ denotes the ith entry of v. Denote $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$, $\mathbb{R}_+ := [0, +\infty)$, and $\mathbb{R}_{++} := (0, +\infty)$. $\mathbb{S}^n_+$ (resp. S^n_{++}) represents the set of all $n \times n$ symmetric positive semi-definite (resp. definite) matrices. $\iota_{\mathcal{S}}(x)$ is defined to be the indicator function of a set $\mathcal{S}$, i.e., if $x \in \mathcal{S}$, then $\iota_{\mathcal{S}}(x) = 0$; otherwise, $\iota_{\mathcal{S}}(x) = +\infty$. $N_S(x)$ denotes the normal cone to the set $S \subseteq \mathbb{R}^n$ at the point x: if $x \in S$, then $N_S(x) := \{u \in \mathbb{R}^n \mid \sup_{z\in S}\langle u, z - x\rangle \le 0\}$; otherwise, $N_S(x) := \varnothing$. If $S \in \mathbb{R}^n$ is a closed and convex set, the map $\mathrm{Pj}_S : \mathbb{R}^n \to S$ denotes the projection onto S, i.e., $\mathrm{Pj}_S(x) := \mathrm{argmin}_{v\in S}\|v - x\|_2$. We use $\rightrightarrows$ to indicate a point-to-set map. For an operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, $\mathrm{Zer}(T) := \{x \in \mathbb{R}^n \mid Tx \ni \mathbf{0}\}$ and $\mathrm{Fix}(T) := \{x \in \mathbb{R}^n \mid Tx \ni x\}$ denote its zero set and fixed point set, respectively. We denote $\mathrm{dom}(T)$ the domain of the operator T and $\mathrm{gra}(T)$ the graph of it. The resolvent and reflected resolvent of T are defined as $J_T := (I + A)^{-1}$ and $R_T := 2J_T - I$, respectively.

II. PROBLEM FORMULATION

A. Stochastic Game Formulation and SGNE

In this section, we formalize the stochastic generalized Nash equilibrium problem (SGNEP) on networks. There are N players participating in this game, indexed by $\mathcal{N} := \{1, \ldots, N\}$. Each player $i \in \mathcal{N}$ needs to determine its local decision vector $x_i \in \mathcal{X}_i$ to optimize its objective, where $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$ is the local feasible set/action space of player i. This Nash equilibrium seeking problem is generalized because, besides the local constraints $\{\mathcal{X}_i\}_{i\in\mathcal{N}}$, the decision vectors of these players should satisfy some global resource constraints, i.e., $\sum_{i\in\mathcal{N}} A_i x_i \le c$. Here, we have the matrix $A_i \in \mathbb{R}^{m\times n_i}$ with m denoting the number of the (global) affine coupling constraints, and the constant vector $c \in \mathbb{R}^m$ representing the quantities of available resources. Altogether, for each player i, the feasible set of the decision vector x_i is given by

$$\tilde{\mathcal{X}}_i(x_{-i}) := \mathcal{X}_i \cap \{x_i \in \mathbb{R}^{n_i} \mid A_i x_i + \textstyle\sum_{j\in\mathcal{N}_{-i}} A_j x_j \le c\}, \quad (1)$$

where $\mathcal{N}_{-i} := \mathcal{N}\backslash\{i\}$, and x_{-i} denotes the stack of decision vectors except that of player i. The feasible set of the collective decision vector $x := [x_1; \cdots; x_N]$ is given by

$$\tilde{\mathcal{X}} := \mathcal{X} \cap \{x \in \mathbb{R}^n \mid Ax - c \le \mathbf{0}\}, \quad (2)$$

where $\mathcal{X} := \prod_{i\in\mathcal{N}} \mathcal{X}_i$, $n := \sum_{i\in\mathcal{N}} n_i$, and $A := [A_1, \ldots, A_N]$.

To capture uncertainty in practical settings, we consider stochastic games where the objective function $\mathbb{J}_i(x_i; x_{-i})$ of each player i is the expected value of certain function J_i. Formally, given the decisions x_{-i} of the other players, each player i aims to solve the following local problem:

$$\begin{cases} \text{minimize}_{x_i\in\mathcal{X}_i} \ \mathbb{J}_i(x_i; x_{-i}) = \mathbb{E}_{\xi_i}[J_i(x_i; x_{-i}, \xi_i)] \\ \text{subject to} \qquad A_i x_i \le c - \sum_{j\in\mathcal{N}_{-i}} A_j x_j \end{cases}, \quad (3)$$

where $\xi_i : \Omega_i \to \mathbb{R}^{n_{\xi_i}}$ is a random variable in a well-defined probability space.

Given the above formulation of the SGNEP, we make the following assumptions that hold throughout the paper.

Assumption 1. *(Scenario-Based Objectives) For each player $i \in \mathcal{N}$, given any fixed sample $\omega_i \in \Omega_i$, the scenario-based objective $J_i(\cdot;\cdot, \xi_i(\omega_i))$ is proper and continuous. In addition, $J_i(x_i; x_{-i}, \xi_i(\omega_i))$ is a convex function w.r.t. x_i given any fixed x_{-i} and $\omega_i \in \Omega_i$.*

Assumption 2. *(Feasible Sets) Each local feasible set $\mathcal{X}_i$ is nonempty, compact, and convex. The collective feasible set $\tilde{\mathcal{X}}$ is nonempty, and the Mangasarian-Fromovitz constraint qualification (MFCQ) holds [7, Ch 3.2][25, Ch. 16.2.3].*

Collectively solving the problems in (3) gives rise to the stochastic generalized Nash equilibrium (SGNE), the formal definition of which is given as follows [17], [26]:

Definition 1. *The collective decision $x^* \in \tilde{\mathcal{X}}$ is a stochastic generalized Nash equilibrium (SGNE) if no player can benefit by unilaterally deviating from x^*. Specifically, for all $i \in \mathcal{N}$, $\mathbb{J}_i(x_i^*; x_{-i}^*) \le \mathbb{J}_i(x_i; x_{-i}^*)$ for any deviation $x_i \in \tilde{\mathcal{X}}_i(x_{-i}^*)$.*

We restrict our attention to a subset of these SGNEs where the players share the same coupled constraints, and hence all the Lagrangian multipliers are in consensus, i.e., $\lambda_1 = \cdots = \lambda_N$. This gives rise to a generalized variational inequality (GVI) problem. This subclass of the SGNEs, known as the variational stochastic generalized Nash equilibria (v-SGNEs) [3], [2], enforces the idea of economic fairness and enjoys better social stability/sensitivity [27]. We will focus on this subset since we can leverage a variety of tools that have been developed for solving (G)VIs [7, Ch. 10-12] and then design the modified best-response dynamics based on it.

Definition 2. *The collective decision $x^* \in \tilde{\mathcal{X}}$ is a variational stochastic generalized Nash equilibrium (v-SGNE) if x^* along with a suitable $g^* \in \prod_{i\in\mathcal{N}} \partial_{x_i}\mathbb{J}_i(x_i^*; x_{-i}^*)$ is a solution of the GVI$(\tilde{\mathcal{X}}, \prod_{i\in\mathcal{N}} \partial_{x_i}\mathbb{J}_i)$, i.e.,*

$$(x - x^*)^T g^* \ge 0, \forall x \in \tilde{\mathcal{X}}. \quad (4)$$

Under Assumptions 1 and 2, we can recast the problem in (3) into a set of inclusions by considering the Karush-Kuhn-Tucker (KKT) conditions of the above GVI:

$$\begin{aligned} &\mathbf{0} \in \partial_{x_i}\mathbb{J}_i(x_i^*; x_{-i}^*) + A_i^T\lambda + N_{\mathcal{X}_i}(x_i^*) \\ &\mathbf{0} \in -(Ax^* - c) + N_{\mathbb{R}^m_+}(\lambda), \end{aligned} \quad (5)$$

where λ is the Lagrangian multiplier for the global constraints in (3). Notice that the GVI in (4) is not completely equivalent to the initial SGNEP in (3), and it is possible that the game admits an SGNE while the GVI has no solution. We make the following assumption concerning the existence of v-SGNEs.

Assumption 3. *(Existence of v-SGNE) The SGNEP considered admits a nonempty set of v-SGNEs.*

Remark 1. *The existence and multiplicity of solutions of problems with continuously differentiable objectives have been extensively studied, and the related theories can be* *found in [7, Ch. 2&3]. For the problems with nonsmooth objectives, if the closed-form expressions of the objectives $\mathbb{J}_i(x_i; x_{-i})$ for all $i \in \mathcal{N}$ are available, we can check the existence of v-GNEs in this GNEP by [25, Prop. 12.11]. Otherwise, [26, Sect. 4] provides sufficient conditions to guarantee the existence of SGNEs based on the properties of scenario-based objectives.*

B. Network Game Formulation

In network games, there exists an underlying communication graph $\mathcal{G} = (\mathcal{N}_g, \mathcal{E}_g)$, where players can communicate with their neighbors through arbitrators on the edges. The node set $\mathcal{N}_g$ denotes the set of all players, and $\mathcal{E}_g \subseteq \mathcal{N}_g \times \mathcal{N}_g$ is the set of directed edges, the cardinalities of which are denoted by N_g and E_g, respectively. In this case, $\mathcal{N}_g = \mathcal{N}$ and $N_g = N$. We use (i, j) to denote a directed edge having node/player i as its tail and node/player j as its head. For notational brevity, let $\mathcal{N}_i$ denote the set of immediate neighbors of player i who can directly communicate with it, $\mathcal{N}_i^+ := \{j \in \mathcal{N} \mid (j, i) \in \mathcal{E}_g\}$ the set of in-neighbors of player i, and $\mathcal{N}_i^- := \{j \in \mathcal{N} \mid (i, j) \in \mathcal{E}_g\}$ the set of out-neighbors of player i. Note that although the multipliers we are going to introduce are defined in a directed fashion, we assume each node can send messages to both its in- and out-neighbors, and $\mathcal{G}$ should satisfy the following assumption.

Assumption 4. *(Communicability) The underlying communication graph $\mathcal{G} = (\mathcal{N}_g, \mathcal{E}_g)$ is undirected and connected. Besides, it has no self-loops.*

We next recast the SGNEP in (3) as the zero-finding problem of a certain operator that can be carried out distributedly over $\mathcal{G}$ via the network Lagrangian of this game and refer the interested reader to [28] for more details. Now for each player $i \in \mathcal{N}$, besides its local decision vector $y_i^i \in \mathcal{X}_i$, it keeps a local estimate $y_i^j \in \mathbb{R}^{n_j}$ of the player j's decision for all $j \in \mathcal{N}_{-i}$, which together constitutes its augmented decision vector y_i. We denote $y_i^{-i} := [y_i^j]_{j \in \mathcal{N}_{-i}}$ the vertical stack of $\{y_i^j\}_{j \in \mathcal{N}_{-i}}$ and $y_i := [y_i^j]_{j \in \mathcal{N}}$ the vertical stack of $\{y_i^j\}_{j \in \mathcal{N}}$, both in prespecified orders. Here, we use y_i^i to denote the local decision of each player i to distinguish from the case where only local decision x_i is maintained and considered. Denote $n_{<i} = \sum_{j \in \mathcal{N}, j<i} n_j$ and $n_{>i} = \sum_{j \in \mathcal{N}, j>i} n_j$. The extended feasible set $\hat{\mathcal{X}}$ is defined as $\hat{\mathcal{X}} := \hat{\mathcal{X}}_1 \times \hat{\mathcal{X}}_2 \times \cdots \times \hat{\mathcal{X}}_N$ with each one defined as $\hat{\mathcal{X}}_i := \mathbb{R}^{n_{<i}} \times \mathcal{X}_i \times \mathbb{R}^{n_{>i}}$. For brevity, we shall write $\{y_i\}$ in replacement of the more cumbersome notation $\{y_i\}_{i \in \mathcal{N}}$ and similarly for other variables on nodes and edges (e.g. the dual variables $\{\mu_{ji}\}_{(j,i) \in \mathcal{E}_g}$ will be denoted simply by $\{\mu_{ji}\}$), unless otherwise specified. For the variables involved in the reformulated zero-finding problem, we introduced a set of dual variables $\{\lambda_i\}$ to enforce the global resource constraints. Moreover, another two sets of dual variables $\{\mu_{ji}\}$ and $\{z_{ji}\}$ are introduced to guarantee the consensus of $\{y_i\}$ and $\{\lambda_i\}$. It is worth mentioning that $\{y_i\}$ and $\{\lambda_i\}$ are maintained by players while $\{\mu_{ji}\}$ and $\{z_{ji}\}$ are maintained by arbitrators on the edges.

We then give a brief introduction to two commonly used operators in the distributed solution of GNEP: the pseudogradient $\mathbb{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and the extended pseudogradient $\tilde{\mathbb{F}} : \mathbb{R}^{nN} \rightrightarrows \mathbb{R}^n$. The pseudogradient $\mathbb{F}$ is the vertical stack of the partial subgradients of the objective functions of all players, which is given as follows:

$$\mathbb{F} : x \mapsto [\partial_{x_i} \mathbb{J}_i(x_i; x_{-i})]_{i \in \mathcal{N}}. \tag{6}$$

In contrast, the extended pseudogradient $\tilde{\mathbb{F}}$ defined in (7) is a commonly used operator in the partial information setting, where each player keeps the local estimates of others' decisions and then uses these estimates as the parametric inputs.

$$\tilde{\mathbb{F}} : [y_i]_{i \in \mathcal{N}} \mapsto [\partial_{y_i^i} \mathbb{J}_i(y_i^i; y_i^{-i})]_{i \in \mathcal{N}}. \tag{7}$$

We introduce the following matrices, i.e., the individual selection matrices $\{\mathcal{R}_i\}_{i \in \mathcal{N}}$ and their diagonal concatenation $\mathcal{R} \in \mathbb{R}^{n \times nN}$, to incorporate the extended pseudogradient $\tilde{\mathbb{F}}$ into a fixed-point iteration:

$$\mathcal{R}_i = [\mathbf{0}_{n_i \times n_{<i}}, \mathbf{I}_{n_i}, \mathbf{0}_{n_i \times n_{>i}}], \mathcal{R} = \mathrm{blkd}(\mathcal{R}_1, \ldots, \mathcal{R}_N). \tag{8}$$

Finally, the set-valued operator $\mathbb{T}$ we are going to study is given below:

$$\mathbb{T} : \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{\lambda} \\ \boldsymbol{\mu} \\ \boldsymbol{z} \end{bmatrix} \mapsto \begin{bmatrix} \mathcal{R}^T(\tilde{\mathbb{F}}(\boldsymbol{y}) + \Lambda^T \boldsymbol{\lambda}) + B_n \boldsymbol{\mu} + \rho_\mu L_n \boldsymbol{y} + N_{\hat{\mathcal{X}}}(\boldsymbol{y}) \\ N_{\mathbb{R}_+^{mN}}(\boldsymbol{\lambda}) - \Lambda \mathcal{R} \boldsymbol{y} + \boldsymbol{c} + B_m \boldsymbol{z} + \rho_z L_m \boldsymbol{\lambda} \\ -B_n^T \cdot \boldsymbol{y} \\ -B_m^T \cdot \boldsymbol{\lambda} \end{bmatrix}, \tag{9}$$

where Λ is the diagonal concatenation of $\{A_i\}_{i \in \mathcal{N}}$, i.e., $\Lambda := \mathrm{blkd}(A_1, \ldots, A_N)$; $\boldsymbol{c}$ is the vertical stack of $\{c_i\}_{i \in \mathcal{N}}$ with $\sum_{i \in \mathcal{N}} c_i = c$; $B_n := (B \otimes I_n)$, $L_n := (L \otimes I_n)$, $B_m := (B \otimes I_m)$, $L_m := (L \otimes I_m)$, B and L are the incidence matrix and Laplacian matrix of the underlying communication graph, respectively, with $L = B \cdot B^T$; and $\boldsymbol{y}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, and $\boldsymbol{z}$ are the stack vectors of $\{y_i\}$, $\{\lambda_i\}$, $\{\mu_{ji}\}$, and $\{z_{ji}\}$, respectively; ψ denotes the stack of these variables, i.e., $\psi := [\boldsymbol{y}; \boldsymbol{\lambda}; \boldsymbol{\mu}; \boldsymbol{z}]$.

Theorem 1. *Suppose Assumptions 1 to 4 hold, and there exists $\psi^* := [\boldsymbol{y}^*; \boldsymbol{\lambda}^*; \boldsymbol{\mu}^*; \boldsymbol{z}^*] \in Zer(\mathbb{T})$. Then $\boldsymbol{y}^* = \mathbf{1}_N \otimes y^*$, $\boldsymbol{\lambda}^* = \mathbf{1}_N \otimes \lambda^*$, and (y^*, λ^*) satisfies the KKT conditions (5) for v-GNE with x^* replaced with y^*. Conversely, for a solution $(y^\dagger, \lambda^\dagger)$ of the KKT problem in (5), there exist $\boldsymbol{\mu}^\dagger$ and $\boldsymbol{z}^\dagger$ such that $\psi^\dagger := [\mathbf{1}_N \otimes y^\dagger; \mathbf{1}_N \otimes \lambda^\dagger; \boldsymbol{\mu}^\dagger; \boldsymbol{z}^\dagger] \in Zer(\mathbb{T})$.*

Proof: See the proof of Theorem 1 in [28].

Thus, finding a v-SGNE of the game in (3) is equivalent to solving for a zero point of the operator $\mathbb{T}$. To facilitate the convergence analysis of the algorithm to be proposed for the latter task, we make two parallel assumptions, either of which can guarantee the convergence to a v-GNE [28, Sect. 4].

Assumption 5. *(Convergence Condition) At least one of the following statements holds:*

(i) the operator $\mathcal{R}^T \tilde{\mathbb{F}} + \frac{\rho_\mu}{2} L_n$ is maximally monotone;

(ii) the pseudogradient $\mathbb{F}$ is strongly monotone and Lipschitz continuous, i.e., there exist $\eta > 0$ and $\theta_1 > 0$, such that $\forall x, x' \in \mathbb{R}^n$, $\langle x - x', \mathbb{F}(x) - \mathbb{F}(x') \rangle \geq \eta \|x - x'\|^2$ and $\|\mathbb{F}(x) - \mathbb{F}(x')\| \leq \theta_1 \|x - x'\|$. The operator $\mathcal{R}^T \tilde{\mathbb{F}}$ is Lipschitz continuous, i.e., there exists $\theta_2 > 0$, such that $\forall \boldsymbol{y}, \boldsymbol{y}' \in \mathbb{R}^{nN}$, $\|\tilde{\mathbb{F}}(\boldsymbol{y}) - \tilde{\mathbb{F}}(\boldsymbol{y}')\| \leq \theta_2 \|\boldsymbol{y} - \boldsymbol{y}'\|$.

III. AN AUGMENTED BEST-RESPONSE SCHEME

To compute the zeros of the operator $\mathbb{T}$ given in the previous section, we leverage the Douglas-Rachford (DR) splitting method which combines operator splitting and the Krasnosel'skill-Mann (K-M) schemes. Given a nonexpansive operator Q with a nonempty fixed point set $\mathrm{Fix}(Q)$, the K-M scheme [29, Sect. 5.2] suggests the following iteration:

$$\psi^{(k+1)} := \psi^{(k)} + \gamma^{(k)}(Q\psi^{(k)} - \psi^{(k)}), \tag{10}$$

where $(\gamma^{(k)})_{k\in\mathbb{N}}$ is a sequence such that $\gamma^{(k)} \in [0,1]$ for all $k \in \mathbb{N}$ and $\sum_{k\in\mathbb{N}} \gamma^{(k)}(1-\gamma^{(k)}) = \infty$. Here, we introduce a set of local bounded box constraints $\{\mathcal{X}_i^B\}$ which can be chosen manually as long as it satisfies $\mathcal{X}_i \subseteq \mathcal{X}_i^B$ for all $i \in \mathcal{N}$. We similarly define the extended box set $\hat{\mathcal{X}}^B := \hat{\mathcal{X}}_1^B \times \hat{\mathcal{X}}_2^B \times \cdots \hat{\mathcal{X}}_N^B$ where the extended box set of each player i is defined as $\hat{\mathcal{X}}_i^B := \mathbb{R}^{n_{<i}} \times \mathcal{X}_i^B \times \mathbb{R}^{n_{>i}}$. It is easy to see that the normal cones of $\hat{\mathcal{X}}^B$ and $\hat{\mathcal{X}}$ satisfy $N_{\hat{\mathcal{X}}^B} + N_{\hat{\mathcal{X}}} = N_{\hat{\mathcal{X}}}$. The motivation for introducing these box sets is to simplify the computation while maintaining boundedness for the convergence analysis as we will show later in this paper. We then split the operator $\mathbb{T}$ into the following operators $\mathbb{A}$ and $\mathcal{B}$ (for details, please refer to [28, Sect. 3]):

$$\mathbb{A} : \psi \mapsto (D + \mathbb{A}_y)\psi \text{ and } \mathcal{B} : \psi \mapsto (D + \mathcal{B}_y)\psi \tag{11}$$

with D, $\mathcal{A}_y$, and $\mathcal{B}_y$ defined by

$$D = \begin{bmatrix} \frac{\rho_\mu}{2} L_n & \frac{1}{2}(\Lambda\mathcal{R})^T & \frac{1}{2}B_n & 0 \\ -\frac{1}{2}\Lambda\mathcal{R} & \frac{\rho_z}{2} L_m & 0 & \frac{1}{2}B_m \\ -\frac{1}{2}B_n^T & 0 & 0 & 0 \\ 0 & -\frac{1}{2}B_m^T & 0 & 0 \end{bmatrix},$$

$$\mathbb{A}_y : \psi \mapsto \begin{bmatrix} \mathcal{R}^T\, \tilde{\mathbb{F}}(\boldsymbol{y}) + N_{\hat{\mathcal{X}}^{\mathcal{B}}}(\boldsymbol{y}) \\ \boldsymbol{c} \\ 0 \\ 0 \end{bmatrix}, \; \mathcal{B}_y : \psi \mapsto \begin{bmatrix} N_{\hat{\mathcal{X}}}(\boldsymbol{y}) \\ N_{\mathbb{R}_+^{mN}}(\lambda) \\ 0 \\ 0 \end{bmatrix}.$$

Furthermore, we introduce the following design matrix Φ for distributedly computing $J_{\Phi^{-1}\mathbb{A}}$ and $J_{\Phi^{-1}\mathcal{B}}$:

$$\Phi = \begin{bmatrix} \tau_1^{-1} - \frac{\rho_\mu}{2} L_n & -\frac{1}{2}(\Lambda\mathcal{R})^T & -\frac{1}{2}B_n & 0 \\ -\frac{1}{2}\Lambda\mathcal{R} & \tau_2^{-1} - \frac{\rho_z}{2} L_m & 0 & -\frac{1}{2}B_m \\ -\frac{1}{2}B_n^T & 0 & \tau_3^{-1} & 0 \\ 0 & -\frac{1}{2}B_m^T & 0 & \tau_4^{-1} \end{bmatrix}, \tag{12}$$

where $\boldsymbol{\tau}_1 := \mathrm{blkd}(\tau_{11}I_n, \ldots, \tau_{1N}I_n)$ with the scalars $\tau_{11} \in \mathbb{R}_{++}, \ldots, \tau_{1N} \in \mathbb{R}_{++}$; similarly for $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$ and $\boldsymbol{\tau}_4$. These step sizes $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_4$ should be small enough to guarantee that Φ is positive definite. Conservative upper bounds for these step sizes [28, Lemma 1] can be derived using the Gershgorin circle theorem [30].

Assumption 6. *The step sizes $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_4$ are chosen properly such that the design matrix Φ in (12) is positive definite.*

Notice that because of the incorporation of the design matrix Φ, now we are working in the inner product space $\mathcal{K}$ which is a real vector space endowed with the inner product $\langle\psi_1, \psi_2\rangle_{\mathcal{K}} = \psi_1^T \Phi \psi_2$. For brevity, let $\bar{\mathbb{A}} := \Phi^{-1}\mathbb{A}$ and $\bar{\mathcal{B}} := \Phi^{-1}\mathcal{B}$. In the DR splitting scheme, the general operator Q in (10) is given by $\mathscr{R}_* := R_{\bar{\mathcal{B}}} \circ R_{\bar{\mathbb{A}}}$ and it suggests the following exact iteration:

$$\tilde{\psi}^{(k+1)} := \mathscr{P}_*(\tilde{\psi}^{(k)}), \text{ with } \mathscr{P}_* = \mathrm{Id} + \gamma^{(k)}(\mathscr{R}_* - \mathrm{Id}). \tag{13}$$

Given a generic single-valued operator Q, we say that Q is restricted nonexpansive w.r.t. a set S if, for all $\psi \in \mathrm{dom}\, Q$ and $\psi^* \in S$, $\|Q\psi - Q\psi^*\| \le \|\psi - \psi^*\|$ [20]; if, in addition, $S = \mathrm{Fix}(Q)$, then Q is quasinonexpansive [29, Def. 4.1(v)]. From the main convergence results in [28, Thm. 2&3], if Assumptions 1 to 6 hold, even though $\mathscr{R}_*$ is not contractive, it possesses nonexpansiveness or just quasinonexpansiveness in the inner-product space $\mathcal{K}$, and hence the sequence $(y_i^{(k)})_{k\in\mathbb{N}}$ generated by the exact iteration above will converge to a v-SGNE of the original problem defined in (3). The detailed description of the exact iteration is given in [28, Algorithm 1].

However, unlike the problem setting in [28] where each player has a closed-form objective function, here the objective function is expected-value, and all too often its closed-form expression may be too complicated to analyze or not available at all. Consequently, the argmin operation in the first player loop of [28, Algorithm 1] can not be solved exactly. In this case, we need a desirable inexact solver such that, although at each iteration step, it can only get an approximate solution, the whole sequence can still eventually converge to a v-SGNE. We let $R_{\bar{\mathcal{A}}}$ denote the (scenario-based) approximate operator to the exact reflected resolvent $R_{\bar{\mathbb{A}}}$, and $\mathscr{R}$ denote the corresponding composite $R_{\bar{\mathcal{B}}} \circ R_{\bar{\mathcal{A}}}$. Substituting the operator $\mathscr{R}_*$ with $\mathscr{R}$ in [28, Algorithm 1] gives rise to the following approximate iteration:

$$\tilde{\psi}^{(k+1)} := \mathscr{P}(\tilde{\psi}^{(k)}), \text{ with } \mathscr{P} = \mathrm{Id} + \gamma^{(k)}(\mathscr{R} - \mathrm{Id}). \tag{14}$$

The details of (14) are presented in Algorithm 1. For brevity, let $\tilde{y}_{iL}^{-i(k)} := \sum_{j\in\mathcal{N}_i}(\tilde{y}_i^{-i(k)} - \tilde{y}_j^{-i(k)})$, and similarly for $\tilde{y}_{iL}^{i(k)}$, $\tilde{\lambda}_{iL}^{(k)}$, $\hat{y}_{iL}^{i(k+1)}$, and $\hat{\lambda}_{iL}^{(k+1)}$; let $\tilde{\mu}_{iB}^{-i(k)} := \sum_{j\in\mathcal{N}_i^+}\tilde{\mu}_{ji}^{-i(k)} - \sum_{j\in\mathcal{N}_i^-}\tilde{\mu}_{ij}^{-i(k)}$, and similarly for $\tilde{\mu}_{iB}^{i(k)}$, $\tilde{z}_{iB}^{(k)}$, $\hat{\mu}_{iB}^{(k+1)}$, and $\hat{z}_{iB}^{(k+1)}$; let $\hat{y}_{ji}^{(k+1)} := \hat{y}_i^{(k+1)} - \hat{y}_j^{(k+1)}$, and similarly for $\hat{\lambda}_{ji}^{(k+1)}$, $\bar{y}_{ji}^{(k+1)}$, and $\bar{\lambda}_{ji}^{(k+1)}$.

Depending on the inexact solver adopted, $R_{\bar{\mathcal{A}}}$ usually admits no explicit formulas. Yet, as will be shown later, we can still establish the convergence of Algorithm 1 based on the specific properties of $R_{\bar{\mathcal{A}}}$.

IV. CONVERGENCE ANALYSIS AND CONSTRUCTION OF INEXACT SOLVER

A. General Convergence Results with Approximate Solution

In this subsection, we investigate the sufficient conditions to guarantee the convergence of Algorithm 1 to a v-SGNE of the problem (3) through the Robbins-Siegmund theorem [31]. We start by defining the approximate error and its norm for each iteration k as $\epsilon^{(k)} := \mathscr{R}(\tilde{\psi}^{(k)}) - \mathscr{R}_*(\tilde{\psi}^{(k)})$ and $\varepsilon^{(k)} := \|\epsilon^{(k)}\|_{\mathcal{K}}$, where $\tilde{\psi}^{(k)} := [\tilde{\boldsymbol{y}}^{(k)}; \tilde{\boldsymbol{\lambda}}^{(k)}; \tilde{\boldsymbol{\mu}}^{(k)}; \tilde{\boldsymbol{z}}^{(k)}]$. We next introduce the residual function $\mathrm{res}(\tilde{\psi}) := \|\tilde{\psi} - \mathscr{R}_*(\tilde{\psi})\|_{\mathcal{K}}$ such that $\mathrm{res}(\tilde{\psi}^*) = 0$ is a necessary condition for $\tilde{\psi}^*$ to belong to the fixed-point set of $\mathscr{R}_*$. This relation can be easily checked by using [29, Prop. 26.1(iii)]. Let $\mathcal{F}_k$ denote the σ-field comprised of $\{\tilde{\psi}^{(0)}, \{\xi_i^{(0)}\}_{i\in\mathcal{N}}, \ldots, \{\xi_i^{(k-1)}\}_{i\in\mathcal{N}}\}$, where for each major iteration $k \in \mathbb{N}$, $\xi_i^{(k)} = \{\xi_{i,0}^{(k)}, \ldots, \xi_{i,T_i^{(k)}-1}^{(k)}\}$ and

Algorithm 1: Distributed v-SGNE Seeking under Partial-Decision Information

Initialize: $\{\tilde{y}_i^{(0)}\}, \{\tilde{\lambda}_i^{(0)}\}, \{\tilde{\mu}_{ji}^{(0)}\}, \{\tilde{z}_{ji}^{(0)}\}$;
Iterate until convergence:
for *player* $i \in \mathcal{N}$ **do**
- $y_i^{-i(k+1)} = \tilde{y}_i^{-i(k)} - \frac{\tau_{1i}}{2}(\rho_\mu \tilde{y}_{iL}^{-i(k)} + \tilde{\mu}_{iB}^{-i(k)})$;
- $y_i^{i(k+1)}$ inexactly solves: $\mathrm{argmin}_{v_i \in \mathcal{X}_i^B} [\mathbb{J}_i(v_i; y_i^{-i(k+1)}) + \frac{1}{2}((A_i^T \tilde{\lambda}_i^{(k)} + \rho_\mu \tilde{y}_{iL}^{i(k)} + \mu_{iB}^{i(k)})^T v_i + \frac{1}{\tau_{1i}} \|v_i - \tilde{y}_i^{i(k)}\|^2)]$;
- $\lambda_i^{(k+1)} = \tilde{\lambda}_i^{(k)} + \tau_{2i}(A_i(y_i^{i(k+1)} - \frac{1}{2}\tilde{y}_i^{i(k)}) - \frac{\rho_z}{2}\tilde{\lambda}_{iL}^{(k)} - \frac{1}{2}\tilde{z}_{iB}^{(k)} - c_i)$;
- $\hat{y}_i^{(k+1)} = 2y_i^{(k+1)} - \tilde{y}_i^{(k)}, \ \hat{\lambda}_i^{(k+1)} = 2\lambda_i^{(k+1)} - \tilde{\lambda}_i^{(k)}$;

end
for *edge* $(j,i) \in \mathcal{E}_g$ **do**
- $\mu_{ji}^{(k+1)} = \tilde{\mu}_{ji}^{(k)} + \frac{\tau_{3i}}{2}\hat{y}_{ji}^{(k+1)}, \ \hat{\mu}_{ji}^{(k+1)} = 2\mu_{ji}^{(k+1)} - \tilde{\mu}_{ji}^{(k)}$;
- $z_{ji}^{(k+1)} = \tilde{z}_{ji}^{(k)} + \frac{\tau_{4i}}{2}\hat{\lambda}_{ji}^{(k+1)}, \ \hat{z}_{ji}^{(k+1)} = 2z_{ji}^{(k+1)} - \tilde{z}_{ji}^{(k)}$;

end
for *player* $i \in \mathcal{N}$ **do**
- $\bar{y}_i^{(k+1)} = \mathrm{Pj}_{\hat{\mathcal{X}}_i}[\hat{y}_i^{(k+1)} - \frac{\tau_{1i}}{2}(\mathcal{R}_i^T A_i^T \hat{\lambda}_i^{(k+1)} + \rho_\mu \hat{y}_{iL}^{(k+1)} + \hat{\mu}_{iB}^{(k+1)})]$;
- $\bar{\lambda}_i^{(k+1)} = \mathrm{Pj}_{\mathbb{R}_+^m}[\hat{\lambda}_i^{(k+1)} + \tau_{2i}(A_i(\bar{y}_i^{i(k+1)} - \frac{1}{2}\hat{y}_i^{i(k)}) - \frac{\rho_z}{2}\hat{\lambda}_{iL}^{(k+1)} - \frac{1}{2}\hat{z}_{iB}^{(k+1)})]$;

end
for *edge* $(j,i) \in \mathcal{E}_g$ **do**
- $\bar{\mu}_{ji}^{(k+1)} = \hat{\mu}_{ji}^{(k+1)} + \tau_{3i}(\bar{y}_{ji}^{(k+1)} - \frac{1}{2}\hat{y}_{ji}^{(k+1)})$;
- $\bar{z}_{ji}^{(k+1)} = \hat{z}_{ji}^{(k+1)} + \tau_{4i}(\bar{\lambda}_{ji}^{(k+1)} - \frac{1}{2}\hat{\lambda}_{ji}^{(k+1)})$;

end
K-M updates: $\tilde{\psi}^{(k+1)} = \tilde{\psi}^{(k)} + 2\gamma^{(k)}(\bar{\psi}^{(k+1)} - \psi^{(k+1)})$;
Return: $\{y_i^{(k)}\}$.

$T_i^{(k)}$ denotes the number of noise realizations that player i observes at the kth iteration when implementing the inexact solver.

Theorem 2. *Consider an SGNEP given in (3), and suppose Assumptions 1 to 6 hold. Moreover, $(\gamma^{(k)})_{k\in\mathbb{N}}$ is a sequence such that $\gamma^{(k)} \in [0,1]$ and $\sum_{k\in\mathbb{N}} \gamma^{(k)}(1-\gamma^{(k)}) = +\infty$. If the sequence $(\tilde{\psi}^{(k)})$ generated by the inexact solver satisfies*

(i) $(\|\tilde{\psi}^{(k)}\|_{\mathcal{K}})_{k\in\mathbb{N}}$ is bounded a.s.;
(ii) $\sum_{k\in\mathbb{N}} \gamma^{(k)}\mathbb{E}[\varepsilon^{(k)} \mid \mathcal{F}^{(k)}] < \infty$, a.s. ,

then $(\boldsymbol{y}^{(k)})_{k\in\mathbb{N}}$ and $(\boldsymbol{\lambda}^{(k)})_{k\in\mathbb{N}}$ generated by Algorithm 1 satisfy a.s. $\lim_{k\to\infty} \boldsymbol{y}^{(k)} = (\mathbf{1}_N \otimes y^*)$ and $\lim_{k\to\infty} \boldsymbol{\lambda}^{(k)} = (\mathbf{1} \otimes \lambda^*)$, where y^* is a v-SGNE to the original SGNEP (3) and (y^*, λ^*) together is a solution to the KKT conditions (5) of the SGNEP.

Before proceeding, it is worth highlighting why we need to have both the condition (i) and (ii) to hold in Theorem 2. With the condition (ii), i.e., $\sum_{k\in\mathbb{N}} \gamma^{(k)}\mathbb{E}[\varepsilon^{(k)} \mid \mathcal{F}^{(k)}] < \infty$ a.s., one can show that $(\|\tilde{\psi}^{(k)}\|_{\mathcal{K}})_{k\in\mathbb{N}}$ is bounded a.s. A similar proof for deterministic cases can be found in [29, Prop. 5.34]. Nevertheless, under the partial-information setting, ensuring the condition (ii) requires the fulfillment of condition (i), i.e., the almost-sure boundedness of $(\|\tilde{\psi}^{(k)}\|_{\mathcal{K}})_{k\in\mathbb{N}}$. Thus, we should prove the condition (i) using a more primitive condition than the condition (ii), and keep both (i) and (ii) in the statement of Theorem 2.

B. *Construction of a Desirable Inexact Solver*

As we discussed at the end of Section III, it is challenging to solve the augmented best-response subproblems which involve the expected-value objectives precisely (the argmin problems in the first player for-loop of Algorithm 1). Moreover, Theorem 2 suggests that we can still obtain a v-SGNE by solving these augmented best-response subproblems not precisely but up to some prescribed accuracy. In this subsection, we consider a specific scenario-based solver using the projected stochastic subgradient method [32][33], and study the explicit conditions that the projected stochastic subgradient solver should satisfy to be a feasible inexact solver in the context of distributed SGNEP with only partial-decision information.

We first make an assumption regarding the unbiasedness and finite-variance properties of a general projected stochastic subgradient method. Throughout the paper, we use k to index the major iterations (the iteration of the v-SGNE seeking Algorithm 1) and t to index the minor iterations (the iteration of the inexact solver in the first player for-loop of Algorithm 1). Furthermore, at each major iteration k, for each player i, let the augmented scenario-based objective function be denoted by $\hat{J}_i^{(k)}(v_i; \xi_{i,t}^{(k)}) := J_i(v_i; y_i^{-i(k+1)}, \xi_{i,t}^{(k)}) + (\tilde{\varphi}_i^{(k)})^T y_i^i + \frac{1}{2\tau_{1i}}\|v_i - \tilde{y}_i^{i(k)}\|_2^2$, and the augmented expected-value objective function be denoted by $\hat{\mathbb{J}}_i^{(k)}(v_i) := \mathbb{J}_i(v_i; y_i^{-i(k+1)}) + (\tilde{\varphi}_i^{(k)})^T v_i + \frac{1}{2\tau_{1i}}\|v_i - \tilde{y}_i^{i(k)}\|_2^2$, where $\tilde{\varphi}_i^{(k)} := \frac{1}{2}(A_i^T \tilde{\lambda}_i^{(k)} + \tilde{\mu}_{iB}^{i(k)} + \rho_\mu \tilde{y}_{iL}^{i(k)})$. Note that $\hat{\mathbb{J}}_i^{(k)}(\cdot)$ is the objective in the first player-loop of Algorithm 1 that needs to be inexactly solved. Here, the vector $\tilde{\varphi}_i^{(k)}$ represents some augmented terms that enforce the consensus constraints and the global resource constraints. For brevity, the local estimates of the other players' decisions $y_i^{-i(k+1)}$ are omitted from the arguments of the augmented functions defined above. Let $T_i^{(k)}$ denote the total number of the projected stochastic subgradient steps taken in the kth major iteration. The subgradient of the scenario-based objective function at the kth major iteration and the tth minor iteration is denoted by $g_{i,t}^{(k)} \in \partial_{y_i^i}\hat{J}_i^{(k)}(y_{i,t}^{i(k+1)}; \xi_{i,t}^{(k)})$, where $t = 0, 1, \ldots, T_i^{(k)} - 1$.

Assumption 7. *For each player $i \in \mathcal{N}$, at each major iteration k and minor iteration t of Algorithm 1, there exists a $g_{i,t}^{(k)} \in \partial_{y_i^i}\hat{J}_i^{(k)}(y_{i,t}^{i(k+1)}; \xi_{i,t}^{(k)})$ such that the following two statements hold:*

(i) (Unbiasedness) $\mathbb{E}[g_{i,t}^{(k)} \mid \sigma\{\mathcal{F}_k, \xi_{i,[t]}^{(k)}\}]$ is almost surely a subgradient of the expected-value augmented objective $\hat{\mathbb{J}}_i^{(k)}(\cdot)$ at $y_{i,t}^{i(k+1)}$, where $\xi_{i,[t]}^{(k)} := \{\xi_{i,0}^{(k)}, \ldots, \xi_{i,t-1}^{(k)}\}$ with $\xi_{i,[0]}^{(k)} := \varnothing$;
(ii) (Finite variance) $\mathbb{E}[\|g_{i,t}^{(k)}\|_2^2 \mid \mathcal{F}_k] \le \alpha_{g,i}^2\|\tilde{\psi}^{(k)}\|_2^2 + \beta_{g,i}^2$ a.s. for some positive constants $\alpha_{g,i}$ and $\beta_{g,i}$.

We refer the reader to the first paragraph of Section IV-A for the definitions of the stack vector $\tilde{\psi}^{(k)}$ and the σ-field $\mathcal{F}_k$ as a reminder. The proposed projected stochastic subgradient solver for the first player for-loop of Algorithm 1 is given in Algorithm 2.

The following lemma discusses the convergence rate of

Algorithm 2: Projected Stochastic Subgradient Inexact Solver

For each player $i \in \mathcal{N}$**, at the** k**th major iteration of Algorithm 1:**

Initialize: $y_{i,0}^{i(k+1)} := \tilde{y}_i^{i(k)}$;

for $t = 0$ *to* $T_i^{(k)} - 1$ **do**

Set $\kappa_{i,t} := \frac{2\tau_{1i}}{t+2}$;

$y_{i,t+1}^{i(k+1)} := \mathrm{Pj}_{\mathcal{X}_i^B}[y_{i,t}^{i(k+1)} - \kappa_{i,t} \cdot g_{i,t}^{(k)}]$;

end

Return: $y_i^{i(k+1)} := y_{i,T_i^{(k)}}^{i(k+1)}$.

Algorithm 2 as a minor updating routine inside Algorithm 1. We use $y_{i,*}^{i(k+1)}$ to denote the accurate minimizer of the expected-value augmented function $\hat{\mathbb{J}}_i^{(k)}(\cdot)$.

Lemma 1. *Suppose Assumptions 1 to 7 hold. For each player* $i \in \mathcal{N}$*, the step size at the* t*th minor iteration is set to be* $\kappa_{i,t} := \frac{2\tau_{1i}}{t+2}$*. Then, for any* $T = 1, \ldots, T_i^{(k)}$*, the distance between the approximate solution and the accurate solution satisfies* $\mathbb{E}[\|y_{i,T}^{i(k+1)} - y_{i,*}^{i(k+1)}\|_2^2 \mid \mathcal{F}_k] \leq (2\tau_{1i})^2 T^{-1}(\alpha_{g,i}^2\|\tilde{\psi}^{(k)}\|_2^2 + \beta_{g,i}^2)$ *a.s.*

For each player $i \in \mathcal{N}$, after the kth major iteration of Algorithm 1 where player i implements $T_i^{(k)}$ projected stochastic subgradient steps in Algorithm 2, $\mathbb{E}\big[\|y_i^{i(k+1)} - y_{i,*}^{i(k+1)}\|_2^2 \mid \mathcal{F}_k\big] \leq (2\tau_{1i})^2 (T_i^{(k)})^{-1}(\alpha_{g,i}^2\|\tilde{\psi}^{(k)}\|_2^2 + \beta_{g,i}^2)$. Based on this result, it is straightforward to derive an upper bound for the approximate error $\varepsilon^{(k)} := \|\mathscr{R}(\tilde{\psi}^{(k)}) - \mathscr{R}_*(\tilde{\psi}^{(k)})\|_{\mathcal{K}}$. As will be shown later, this upper bound can be treated as a function of $\underline{T}^{(k)} := \min\{T_i^{(k)} : i \in \mathcal{N}\}$ which we can tune to provide a desirable sequence of approximation accuracies.

Lemma 2. *Consider the error sequence* $(\varepsilon^{(k)})_{k\in\mathbb{N}}$ *generated by Algorithm 1 using Algorithm 2 as the inexact solver. Suppose Assumptions 1 to 7 hold, and the updating step size of the projected stochastic subgradient method of player* $i \in \mathcal{N}$ *at the* t*th minor iteration is* $\kappa_{i,t} := \frac{2\tau_{1i}}{t+2}$*. Then there exist some positive constants* α_ψ *and* β_ψ *such that the following relation holds a.s.:*

$$\mathbb{E}\big[\varepsilon^{(k)} \mid \mathcal{F}_k\big] \leq (\underline{T}^{(k)})^{-1/2}(\alpha_\psi\|\tilde{\psi}^{(k)}\|_{\mathcal{K}} + \beta_\psi). \tag{15}$$

We define $\gamma_T^{(k)} := \gamma^{(k)}(\underline{T}^{(k)})^{-1/2}$. From Theorem 2, it suffices to have the sequence $(\gamma_T^{(k)})_{k\in\mathbb{N}}$ summable and $(\|\tilde{\psi}^{(k)}\|)_{k\in\mathbb{N}}$ bounded. To this end, we next focus on proving the conditions needed to guarantee the boundedness of $(\tilde{\psi}^{(k)})_{k\in\mathbb{N}}$ and finally derive the sufficient conditions to ensure the convergence.

Theorem 3. *Consider the sequence* $(\tilde{\psi}^{(k)})_{k\in\mathbb{N}}$ *generated by Algorithm 1 using Algorithm 2 as the inexact solver. Suppose Assumptions 1 to 7 hold, and the updating step size of the projected stochastic subgradient method of each player* $i \in \mathcal{N}$ *at the* t*th minor iteration is* $\kappa_{i,t} := \frac{2\tau_{1i}}{t+1}$*. In addition, the sequence* $(\gamma^{(k)})_{k\in\mathbb{N}}$ *satisfies* $0 \leq \gamma^{(k)} \leq 1$*,* $\sum_{k\in\mathbb{N}}\gamma^{(k)}(1-\gamma^{(k)}) = +\infty$*, and the sequence* $(\gamma_T^{(k)})_{k\in\mathbb{N}}$ *is absolutely summable. Then* $(\|\tilde{\psi}^{(k)}\|_{\mathcal{K}})$ *is bounded a.s., and* $\sum_{k\in\mathbb{N}}\gamma^{(k)}\mathbb{E}[\varepsilon^{(k)} \mid \mathcal{F}_k] < \infty$ *a.s. As a result, the sequence* $(\tilde{\psi}^{(k)})_{k\in\mathbb{N}}$ *will converge to a fixed point of* $\mathscr{R}_*$ *and the associated sequence* $(y^{(k)})_{k\in\mathbb{N}}$ *will converge to a v-SGNE of the problem* (3)*.*

V. CASE STUDY AND NUMERICAL SIMULATIONS

In this section, we apply the proposed algorithm to solve a stochastic Nash-Cournot production game in networked regimes. Suppose there are N manufacturers/players competing over m different markets. For each player i in this network, it supplies n_i markets with $x_i \in \mathbb{R}^{n_i}$ units of commodities respectively. The basic setup is almost the same as the one considered in [28, Sect. V] except for the definition of market prices. In this example, the market prices is defined as $P(Ax;\xi_i) = w - \Sigma Ax + \xi_i$, which maps the total quantities of supply $Ax \in \mathbb{R}^m$ to the market unit prices, where $A := [A_1, \ldots, A_N]$, $x := [x_i]_{i\in\mathcal{N}}$, and the uncertainty is captured by the zero-mean random vector ξ_i. We omit the other details for brevity and refer the interested readers to [28, Sect. V]. Overall, each player $i \in \mathcal{N}$, given the supply strategies of others (x_{-i}), aims to solve the following optimization problem (16):

$$\begin{aligned}&\underset{x_i\in X_i}{\text{minimize}}\ \mathbb{E}[x_i^T Q_i x_i + q_i^T x_i - (w - \Sigma Ax + \xi_i)^T A_i x_i]\\ &\text{subject to } A_i x_i \leq c - \textstyle\sum_{j\in\mathcal{N}_{-i}} A_j x_j.\end{aligned} \tag{16}$$

We consider a game network with $N = 10$ players and $m = 5$ markets. Each player i has the dimension n_i of its decision vector chosen uniformly and at random from $\{2,3,4,5\}$. Each entry of vector c is fixed to be 2; the local feasible set of each player $i \in \mathcal{N}$ is the direct product of n_i connected compact intervals $[0, b_{ij}]$ with $b_{ij} \sim U[10, 10.1]$; each random noise ξ_i has the distribution $U[-0.2, 0.2]$. For the remaining parameters, we draw each entry of the vector w from $U[3,5]$ and each diagonal entry of Σ from $U[0.5, 1]$. Assuming Q_i is diagonal, each diagonal entry of Q_i has $[Q_i]_{jj} \sim U[1,2]$, and each entry of q_i has $[q_i]_j \sim U[0.1, 0.6]$. The communication graph consists of an undirected circle and 5 randomly selected edges. We choose $\rho_\mu = 10$, $\boldsymbol{\tau}_1 = 0.02 \otimes I_{Nn}$, $\boldsymbol{\tau}_2 = 0.1 \otimes I_{Nm}$, $\boldsymbol{\tau}_3 = 0.5 \otimes I_{En}$, and $\boldsymbol{\tau}_4 = 0.5 \otimes I_{Em}$, such that the operator $\mathcal{R}^T\tilde{\mathbb{F}} + \frac{\rho_\mu}{2}L_n$ is a maximally monotone operator and the matrix Φ is positive definite. By combining these specifications and the arguments in [28, Sect. V], we can verify the fulfillment of the technical assumptions.

We fix the parameter $(\gamma^{(k)})_{k\in\mathbb{N}}$ to be $\frac{1}{2}$ and choose the count of minor steps taken per major iteration to be $T^{(k)} = 0.001 \times k^{2.1} + 20$, $k + 20$, and 20, respectively. The performances of the proposed algorithm are shown in Fig. 1. We use the thick and semi-transparent lines to illustrate the real fluctuation of the metrics throughout the iterations, while using the thin lines to exhibit the simple moving averages of the metrics with a window size of 50. In Fig. 1, we let $y_j^{(k)}$ denote the stack of player j's local decision and local estimates at the kth iteration, and y^* the generalized Nash equilibrium of the game. The average of the normalized

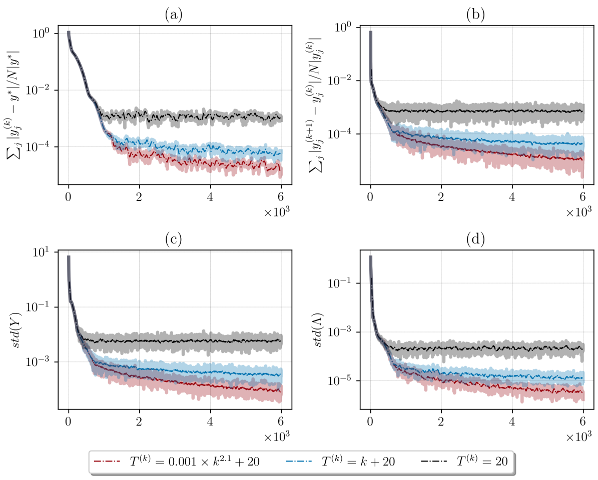

Fig. 1: Performances of Alg. 1 in a Nash-Cournot Game

distances to the v-SGNE is presented in Fig. 1(a), where the unique v-SGNE is computed using the centralized method in [34]. Fig. 1(b) shows the relative length of the updating step at each iteration. Fig. 1(c) and (d) exhibit how the sums of the standard deviations of the local estimates $\{y_j\}$ and $\{\lambda_j\}$ evolve over the iterations, respectively. The curves of $T^{(k)} \propto k^{2.1}$ illustrate a steady convergence towards the v-GNE as suggested in Theorem 3, while the trajectories of $T^{(k)} = 20$ stop decreasing after some iterations. The curves of $T^{(k)} \propto k$ also keep descending yet with a gentler trend compared with those of $T^{(k)} \propto k^{2.1}$, which suggests the possibility of some relaxations to the current conditions in Theorems 2 and 3.

VI. Conclusion and Future Directions

In this paper, we study the stochastic generalized Nash equilibrium problem and propose a distributed stochastic algorithm under the partial-decision information setting based on solving augmented best-response subproblems induced by the Douglas-Rachford scheme. The proposed algorithm is proved to converge to a true variational stochastic generalized Nash equilibrium if the sequence of inertial step sizes and the inverse of the number of realizations per major iteration decrease altogether at a proper rate. This raises the question if there exists a less conservative bound for this decreasing rate such that the proposed algorithm can still converge yet with a faster convergence rate and fewer observations per major iteration. Another interesting work remains concerning the convergence rate analysis of the proposed algorithm when the operators of interest are nonexpansive or merely quasinonexpansive.

References

[1] J. F. Nash *et al.*, "Equilibrium points in n-person games," *Proceedings of the national academy of sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[2] F. Facchinei, A. Fischer, and V. Piccialli, "On generalized Nash games and variational inequalities," *Operations Research Letters*, vol. 35, no. 2, pp. 159–164, 2007.

[3] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *Annals of Operations Research*, vol. 175, no. 1, pp. 177–211, 2010.

[4] A. Kannan, U. V. Shanbhag, and H. M. Kim, "Strategic behavior in power markets under uncertainty," *Energy Systems*, vol. 2, no. 2, pp. 115–141, 2011.

[5] ——, "Addressing supply-side risk in uncertain power markets: stochastic Nash models, scalable algorithms and error analysis," *Optimization Methods and Software*, vol. 28, no. 5, pp. 1095–1138, 2013.

[6] E. Nikolova and N. E. Stier-Moses, "A mean-risk model for the traffic assignment problem with stochastic travel times," *Operations Research*, vol. 62, no. 2, pp. 366–382, 2014.

[7] F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[8] J. Koshal, A. Nedic, and U. V. Shanbhag, "Regularized iterative stochastic approximation methods for stochastic variational inequality problems," *IEEE Transactions on Automatic Control*, vol. 58, no. 3, pp. 594–609, 2012.

[9] R. Bot, P. Mertikopoulos, M. Staudigl, and P. Vuong, "Mini-batch forward-backward-forward methods for solving stochastic variational inequalities," *Stochastic Systems*, 2020.

[10] F. Yousefian, A. Nedić, and U. V. Shanbhag, "On smoothing, regularization, and averaging in stochastic approximation methods for stochastic variational inequality problems," *Mathematical Programming*, vol. 165, no. 1, pp. 391–431, 2017.

[11] S. Cui and U. V. Shanbhag, "On the analysis of reflected gradient and splitting methods for monotone stochastic variational inequality problems," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 4510–4515.

[12] A. Kannan and U. V. Shanbhag, "Optimal stochastic extragradient schemes for pseudomonotone stochastic variational inequality problems and their variants," *Computational Optimization and Applications*, vol. 74, no. 3, pp. 779–820, 2019.

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

[14] F. Parise, S. Grammatico, B. Gentile, and J. Lygeros, "Distributed convergence to Nash equilibria in network and average aggregative games," *Automatica*, vol. 117, p. 108959, 2020.

[15] P. Yi and L. Pavel, "An operator splitting approach for distributed generalized Nash equilibria computation," *Automatica*, vol. 102, pp. 111–121, 2019.

[16] ——, "Distributed generalized Nash equilibria computation of monotone games via double-layer preconditioned proximal-point algorithms," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 299–311, 2018.

[17] B. Franci and S. Grammatico, "A distributed forward-backward algorithm for stochastic generalized Nash equilibrium seeking," *IEEE Transactions on Automatic Control*, 2020.

[18] S. Cui, B. Franci, S. Grammatico, U. V. Shanbhag, and M. Staudigl, "A relaxed-inertial forward-backward-forward algorithm for stochastic generalized Nash equilibrium seeking," *arXiv preprint arXiv:2103.13115*, 2021.

[19] J. Lei, U. V. Shanbhag, J.-S. Pang, and S. Sen, "On synchronous, asynchronous, and randomized best-response schemes for stochastic Nash games," *Mathematics of Operations Research*, vol. 45, no. 1, pp. 157–190, 2020.

[20] L. Pavel, "Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. 1584–1597, 2019.

[21] M. Bianchi, G. Belgioioso, and S. Grammatico, "A fully-distributed proximal-point algorithm for Nash equilibrium seeking with linear convergence rate," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 2303–2308.

[22] G. Belgioioso, A. Nedich, and S. Grammatico, "Distributed generalized nash equilibrium seeking in aggregative games on time-varying networks," *IEEE Transactions on Automatic Control*, 2020.

[23] B. Franci and S. Grammatico, "Stochastic generalized Nash equilibrium seeking under partial-decision information," *arXiv preprint arXiv:2011.05357*, 2020.

[24] Y. Huang and J. Hu, "Distributed computation of stochastic GNE with partial information: An augmented best-response scheme," *arXiv preprint arXiv:2109.12290*, 2021.

[25] D. P. Palomar and Y. C. Eldar, *Convex optimization in signal processing and communications*. Cambridge university press, 2010.

[26] U. Ravat and U. V. Shanbhag, “On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games,” *SIAM Journal on Optimization*, vol. 21, no. 3, pp. 1168–1199, 2011.

[27] A. A. Kulkarni and U. V. Shanbhag, “On the variational equilibrium as a refinement of the generalized Nash equilibrium,” *Automatica*, vol. 48, no. 1, pp. 45–55, 2012.

[28] Y. Huang and J. Hu, “Distributed solution of GNEP over networks via the Douglas-Rachford splitting method,” in *2021 IEEE 60th Conference on Decision and Control (CDC)*. IEEE, 2021, to appear, a full version is available at https://arxiv.org/abs/2103.09393.

[29] H. H. Bauschke, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed., ser. CMS Books in Mathematics, Ouvrages de mathématiques de la SMC, 2017.

[30] H. E. Bell, “Gershgorin’s theorem and the zeros of polynomials,” *The American Mathematical Monthly*, vol. 72, no. 3, pp. 292–295, 1965.

[31] H. Robbins and D. Siegmund, “A convergence theorem for non-negative almost supermartingales and some applications,” in *Optimizing methods in statistics*. Elsevier, 1971, pp. 233–257.

[32] S. Boyd and A. Mutapcic, “Stochastic subgradient methods,” *Lecture Notes for EE364b, Stanford University*, 2008.

[33] S. Lacoste-Julien, M. Schmidt, and F. Bach, “A simpler approach to obtaining an O(1/t) convergence rate for the projected stochastic subgradient method,” *arXiv preprint arXiv:1212.2002*, 2012.

[34] G. Belgioioso and S. Grammatico, “Projected-gradient algorithms for generalized equilibrium seeking in aggregative games are pre-conditioned forward-backward methods,” in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 2188–2193.