Distributed Stochastic Nash Equilibrium Learning in Locally Coupled Network Games with Unknown Parameters

Yuanhanqing Huang   HUAN1282@PURDUE.EDU
Elmore Family School of Electrical and Computer Engineering, Purdue University, USA

Jianghai Hu   JIANGHAI@PURDUE.EDU
Elmore Family School of Electrical and Computer Engineering, Purdue University, USA

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

In stochastic Nash equilibrium problems (SNEPs), it is natural for players to be uncertain about their complex environments and have multi-dimensional unknown parameters in their models. Among various SNEPs, this paper focuses on locally coupled network games where the objective of each rational player is subject to the aggregate influence of its neighbors. We propose a distributed learning algorithm based on the proximal-point iteration and ordinary least-square estimator, where each player repeatedly updates the local estimates of neighboring decisions, makes its augmented best-response decisions given the current estimated parameters, receives the realized objective values, and learns the unknown parameters. Leveraging the Robbins-Siegmund theorem and the law of large deviations for M-estimators, we establish the almost sure convergence of the proposed algorithm to solutions of SNEPs when the updating step sizes decay at a proper rate.

1. Introduction

Nash equilibrium problems, rooted in the seminal work by Nash et al. (1950), model and describe the interactions among multiple decision-makers or players where they aim at optimizing their own payoffs given the strategies of others. In a stochastic Nash equilibrium problem (SNEP), players take the uncertainty in their payoffs into account when deciding on their actions (Facchinei and Kanzow (2010); Shanbhag (2006)). This type of problem can be applied to model a considerable number of applications such as power markets (Kannan et al. (2011, 2013)), engagement of multiple humanitarian organizations in disaster relief (Nagurney et al. (2020)), and the traffic assignment of strategic risk-averse users (Nikolova and Stier-Moses (2014)), to name a few.

The past decade has witnessed significant progress in the distributed solution of Nash equilibrium problems (NEPs) under both deterministic and stochastic setups (Yi and Pavel (2019); Pavel (2019); Bianchi et al. (2020); Shi and Pavel (2017)). It is often assumed that either each player is allowed to communicate with the players affecting its objective or each player maintains a local estimate of the decisions of all the other players. To model network games with better scalability, considerable effort has been spent on studying network games with special structures, such as average aggregative games (AAGs) and network aggregative games (NAGs) (Parise and Ozdaglar (2021)). Here, we focus on locally coupled network games, where the objective of each player depends on its own decision and some linear transformation of the decisions of its neighbors, determined by an underlying communication network. In particular, if the influence of its neighbors can

be expressed as a convex combination of their decisions, it corresponds to NAGs, the properties and distributed solutions of which have been extensively investigated in Parise et al. (2015, 2020).

Most of the aforementioned work designs distributed solutions under the fundamental assumption that each player has perfect knowledge of the payoff function. Nevertheless, in general, players may not be perfectly aware of their environments, and the outcomes of their actions may not always coincide with the predictions (Kirman (1975); Frydman (1982)). Consequently, each player needs to modify its model or the parameters of its models in light of the observations it makes (Esponda et al. (2021)). On the other end of the spectrum, there is an emerging research interest in finding Nash equilibria through bandit/zeroth-order online learning schemes, assuming the players are completely oblivious to the game mechanism, perhaps even ignoring its existence (Bravo et al. (2018); Héliou et al. (2021); Tatarenko and Kamgarpour (2020), etc.). In this work, we consider the setting where the players are aware of their own objectives' functional forms and feasibility constraints while uncertain about some parameters in their objectives, and they participate in sequential repetitions of the same game while learning the parameters over this process.

To learn equilibria while confronted with unknown parameters, the authors of Jiang et al. (2017) present two distributed schemes to solve SNEP without observations of others' strategies. The first scheme is based on the stochastic gradient method which constructs the learning problems independent of the computation of Nash equilibria (NEs); the second scheme solves stochastic Nash-Cournot games via iterative fixed-point methods under a common knowledge assumption concerning the cost functions and strategy sets of their competitors. The authors of Lei and Shanbhag (2020) further extend the first scheme in Jiang et al. (2017) and design an asynchronous inexact proximal best-response solution to the unknown problems. Nevertheless, in most practical applications and online learning settings, the parameter learning and NE seeking processes tangle with each other and players are unwilling to share their local information over the whole network. On that account, the authors of Meigs et al. (2017, 2019) instead consider the case where the parameter estimation process is intrinsically coupled with the strategy update process. Moreover, they postulate that each player can observe the necessary information for parameter learning (e.g. aggregates of neighbors' strategies in NAGs) without the common-knowledge assumption. The learning algorithm considered in this paper is similar to the one in Meigs et al. (2019), while we extend the results by proposing a solution that can handle a more general class of games with multi-dimensional unknown parameters in objectives. Furthermore, we establish the convergence without requiring the contractiveness of the NE-seeking algorithm, and the theoretical analysis can be further extended to the solutions of generalized Nash equilibrium problems (GNEPs) Facchinei and Kanzow (2010).

In this paper, we develop a distributed learning algorithm that guarantees almost-sure convergence to stochastic Nash equilibria (SNEs) in locally coupled network games with unknown parameters. We assume that after all players determine their decisions, each player can observe its own realized objective value and the decisions made by its neighbors. At each iteration, every player selects its decision indicated by the solution of its augmented best-response function parameterized by the current parameter estimates along with some random exploration vector. Then each player receives feedback about the objective values and neighbors' decisions and updates its parameters via an ordinary least squares estimator (OLSE). Furthermore, unlike most of the existing work that enjoys contractive iterations in the NE seeking dynamics, the fixed-point iteration operator considered in this work only satisfies (quasi)nonexpansiveness, due to the partial-information setting (and the global resource constraints in GNEPs). By leveraging the Robbins-Siegmund theorem, we establish the main convergence theorem for the proposed algorithm and discuss the conditions needed to en-

sure the convergence to solutions. We derive an upper bound for the asymptotic convergence rate of the OLSE and discuss the proper choice of step sizes to guarantee the convergence. The technical proofs and complementary examples and discussions are included in Huang and Hu (2022)

Basic Notations: For a set of matrices $\{V_i\}_{i \in S}$, we let $\text{blkd}(V_1, \ldots, V_{|S|})$ or $\text{blkd}(V_i)_{i \in S}$ denote the diagonal concatenation of these matrices, $[V_1, \ldots, V_{|S|}]$ their horizontal 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 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. $\iota_S(x)$ is defined to be the indicator function of a set $\mathcal{S}$, i.e., if $x \in \mathcal{S}$, then $\iota_S(x) = 0$; otherwise, $\iota_S(x) = +\infty$. $N_S(x)$ denotes the normal cone to the set $S \subseteq \mathbb{R}^n$ at the point x: if $x \in S$, then $N_S(x) := \{u \in \mathbb{R}^n \mid \sup_{z \in S} \langle u, z - x \rangle \leq 0\}$; otherwise, $N_S(x) := \varnothing$. If $S \in \mathbb{R}^n$ is a closed and convex set, the map $\text{Pj}_S : \mathbb{R}^n \to S$ denotes the projection onto S, i.e., $\text{Pj}_S(x) := \text{argmin}_{v \in S} \|v - x\|_2$. We use $\rightrightarrows$ to indicate a point-to-set map. For an operator $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, $\text{Zer}(T) := \{x \in \mathbb{R}^n \mid Tx \ni \mathbf{0}\}$ and $\text{Fix}(T) := \{x \in \mathbb{R}^n \mid Tx \ni x\}$ denote its zero set and fixed point set, respectively. We denote $\text{dom}(T)$ the domain of the operator T and $\text{gra}(T)$ the graph of it. The resolvent and reflected resolvent of T are defined as $J_T := (I + T)^{-1}$ and $R_T := 2J_T - I$, respectively.

2. Preliminaries of Stochastic Locally Coupled Network Games

2.1. Formulation of Locally Coupled Network Games

We consider a game played among a group of self-interested players indexed by $\mathcal{N} := \{1, \ldots, N\}$, whose interactions are specified by an underlying communication network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. We use (i, j) to denote a directed edge having player i as its tail and j as its head. Although each edge $e \in \mathcal{E}$ admits certain direction, the communication through the edge e are undirected, i.e., each player i can send messages to both its in-neighbors $\mathcal{N}_i^+ := \{j \in \mathcal{N} \mid (j, i) \in \mathcal{E}\}$ and out-neighbors $\mathcal{N}_i^- := \{j \in \mathcal{N} \mid (i, j) \in \mathcal{E}\}$, the cardinalities of which are denoted by N_i^+ and N_i^-, respectively.

Assumption 1 *(Communicability) The underlying communication graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is undirected and connected. Furthermore, it has no self-loops.*

The goal of each player i is to minimize an expected-value objective defined by $\mathbb{J}_i(x_i; x_i^+, w_i^*) := \mathbb{E}[J_i(x_i; s_i(x_i^+; \xi_i, w_i^*))]$ which depends on its own decision $x_i \in \mathbb{R}^{n_i}$ and the decisions of its in-neighbors $x_i^+ := [x_j]_{j \in \mathcal{N}_i^+}$. It is worth mentioning that x_i^+ are treated as parametric inputs of $\mathbb{J}_i$. We use s_i to denote the neighboring aggregate function $s_i(x_i^+; \xi_i, w_i^*) := w_{ii}^* + \sum_{j \in \mathcal{N}_i^+} w_{ji}^{*T} x_j + \xi_i$, where $w_{ii}^* \in \mathbb{R}$ and $w_{ji}^* \in \mathbb{R}^{n_j}$ are some constant parameters, and we let the random variable $\xi_i : \Omega \to \mathbb{R}$ capture uncertainty in s_i. The local decision x_i made by player i is subject to a set of local feasibility constraints $\mathcal{X}_i \subseteq \mathbb{R}^{n_i}$. We further define $w_i^* := [w_{ji}^*]_{j \in \{i\} \cup \mathcal{N}_i^+}$, $w^* := [w_i^*]_{i \in \mathcal{N}}$, $n := \sum_{i \in \mathcal{N}} n_i$, $n_i^+ = \sum_{j \in \mathcal{N}_i^+} n_j$, and $\mathcal{X} := \prod_{i \in \mathcal{N}} \mathcal{X}_i$. The feasible parameter set of player i is denoted by $\mathcal{W}_i$, and $w_i^* \in \mathcal{W}_i$. Altogether, the local stochastic optimization problem of player i can be formally written as:

$$\text{minimize}_{x_i \in \mathcal{X}_i} \mathbb{J}_i(x_i; x_i^+, w_i^*) \tag{1}$$

The solution concept of the problem described in (1) we focus on in this paper is stochastic Nash equilibria (SNEs) (Ravat and Shanbhag (2011)), whose definition is given as follows:

Definition 1 *The collective decision $x^* \in \mathcal{X}$ is an SNE if no player can benefit by unilaterally deviating from x^*, i.e., $\forall i \in \mathcal{N}$, $\mathbb{J}_i(x_i^*; x_i^{*+}, w_i^*) \leq \mathbb{J}_i(x_i; x_i^{*+}, w_i^*)$ for any $x_i \in \mathcal{X}_i$.*

We then make the following regularity assumptions concerning the objective functions, feasible sets, and solution sets. In particular, Assumption 2 (iii) is imposed to facilitate our later analysis regarding the convergence of the parameter learning and SNE seeking algorithm.

Assumption 2 *(Local Objectives) For each $i \in \mathcal{N}$, given any fixed sample $\omega_i \in \Omega_i$ and the precise parameters w_i^*, the scenario-based and expected-value objectives J_i and $\mathbb{J}_i$ satisfy:*

(i) $J_i(x_i; s_i(x_i^+, \xi_i; w_i^*))$ is convex in x_i given any fixed x_i^+;

(ii) $J_i(x_i; s_i(x_i^+, \xi_i; w_i^*))$ is proper and lower-semicontinuous in x_i and x_i^+;

(iii) $\mathbb{J}_i$ can be written as $\mathbb{J}_i(x_i; x_i^+, \hat{w}_i) = \mathbb{f}_i(x_i; x_i^+) + \mathbb{g}_i(x_i; x_i^+, \hat{w}_i)$, where for any fixed x_i^+, $\mathbb{f}_i$ is convex in x_i and $\mathbb{g}_i$ is differentiable in x_i. Moreover, $\mathbb{g}_i$ is continuous in $\hat{w}_i$ and $\nabla_{x_i}\mathbb{g}_i$ is Lipschitz in $\hat{w}_i$ with the constant $\alpha_{g,i}\|x_i^+\|_2 + \beta_{g,i}$ $(\alpha_{g,i}, \beta_{g,i} \geq 0)$ on $\mathcal{W}_i$ for any fixed $x_i \in \mathcal{X}_i$ and x_i^+.

Assumption 3 *(Feasible Sets) For each $i \in \mathcal{N}$, $\mathcal{X}_i$ is nonempty, compact, convex, and satisfies Slater's constraint qualification (CQ).*

Assumption 4 *(Existence of SNE) The SNEP considered admits a nonempty set of SNEs.*

By stacking the partial gradients $\partial_{x_i}\mathbb{J}_i(x_i; x_i^+, w_i^*)$, we can construct the so-called pseudo-gradient operator $\mathbb{F}_{w^*} : \mathcal{X} \rightrightarrows \mathbb{R}^n$ as $\mathbb{F}_{w^*} : x \mapsto [\partial_{x_i}\mathbb{J}_i(x_i; x_i^+, w_i^*)]_{i \in \mathcal{N}}$. This operator plays a significant role in regulating different types of games, analyzing the properties of solution sets, etc. Games with maximally monotone pseudo-gradient $\mathbb{F}_{w^*}$ are called monotone games. As has been shown in (Palomar and Eldar, 2010, Prop. 12.4, Sect. 12.2.3), to compute SNEs of (1), we can instead solve the corresponding generalized variational inequality (GVI): find a pair of vectors (x^*, g^*) such that $x^* \in \mathcal{X}$ and $g^* \in \mathbb{F}_{w^*}(x^*)$ and $(x - x^*)^T g^* \geq 0, \forall x \in \mathcal{X}$. The Karush-Kuhn-Tucker (KKT) problem of the corresponding GVI can be written as follows:

$$\mathbf{0} \in \partial_{x_i}\mathbb{J}_i(x_i; x_i^+, w_i^*) + N_{\mathcal{X}_i}(x_i), \forall i \in \mathcal{N}, \tag{2}$$

under the proper CQ. To motivate our analysis, we briefly discuss one typical example of locally coupled network games below and another one in (Huang and Hu, 2021, Appendix G).

Example 1 *(Scalar linear quadratic games (Parise and Ozdaglar (2019))) There is a finite set of players indexed by $i = 1, \ldots, N$, each making a scalar non-negative bounded strategy x_i to optimizing its quadratic objective $J_i(x_i; x_i^+) := \frac{1}{2}(x_i)^2 + (K_i \sum_{j \in \mathcal{N}_i^+} w_{ij}x_j - a_i)x_i$, where $K_i, a_i \in \mathbb{R}$, and w_{ij} indicates the influence of player j's decision on the objective function of player i. This model has been applied to investigate various economic settings including the private provision of public goods and games with local payoff complementarities but global substitutability. For more examples of locally coupled network games satisfying the assumptions discussed, see Parise and Ozdaglar (2019, 2021) and the references therein.*

2.2. Distributed Solution via Proximal-Point Algorithm with Precise Parameters

We start by proposing a distributed solution for the SNE problem with precise knowledge of the involved parameters. Some iterative algorithms for computing NEs with an emphasis on algorithms amenable to decomposition have been proposed, such as the best-response (BR) iteration Meigs

et al. (2019) and the proximal BR iteration (Palomar and Eldar, 2010, Sec. 12.6). In these algorithms, the implementation of the fixed-point scheme can be carried out in a distributed and independent manner, and each player makes decisions based on others' decisions from the last iteration. Nevertheless, the convergence properties rely on the contractiveness of their fixed-point iterations, which requires additional regularities on the objectives and network structures.

To tackle a more general class of games, we design the fixed-point iteration in light of the proximal-point algorithm (PPA) and the Krasnosel'skii-Mann algorithm (KM) (Bauschke, 2017, Thm. 23.41, Thm. 5.15). Given a finite-dimensional maximally monotone operator A, the resolvent $J_A := (I + A)^{-1}$ is firmly nonexpansive. Combining PPA and KM yields the following fixed-point iteration prototype $x^{(k+1)} := x^{(k)} + \gamma^{(k)}(J_A x^{(k)} - x^{(k)})$, which will generate a sequence converging to a point in $\text{Zer}(A)$, the zero set of A. Here, $(\gamma^{(k)})_{k \in \mathbb{N}}$ satisfies $\gamma^{(k)} \in [0, 1]$ and $\sum_{n \in \mathbb{N}} \gamma^{(k)}(1 - \gamma^{(k)}) = +\infty$.

For each player $i \in \mathcal{N}$, we endow it with a local estimate y_i^j for the decision of each of its in-neighbors $j \in \mathcal{N}_i^+$. In what follows, we use y_i^i to denote the local decision of player i, $y_i^+ := [y_i^j]_{j \in \mathcal{N}_i^+}$ the stack of the local estimates of its in-neighbors' decisions, and $y_i^- := [y_j^i]_{j \in \mathcal{N}_i^-}$ the stack of local estimates of y_i^i maintained by player i's out-neighbors. Let $y_i := [y_i^i; [y_i^j]_{j \in \mathcal{N}_i^+}]$, and $y := [y_i]_{i \in \mathcal{N}}$. The feasible region of the stack vector y_i is given by $\tilde{\mathcal{X}}_i := \mathcal{X}_i \times \mathbb{R}^{n_i^+}$. Let $\tilde{\mathcal{X}} := \prod_{i \in \mathcal{N}} \tilde{\mathcal{X}}_i \subseteq \mathbb{R}^{\tilde{n}}$, where $\tilde{n} := n + \sum_{i \in \mathcal{N}} n_i^+$. With the introduction of local estimates, we can construct the extended pseudo-gradient $\tilde{\mathbb{F}}_{w^*} : \tilde{\mathcal{X}} \rightrightarrows \mathbb{R}^n$ as $\tilde{\mathbb{F}}_{w^*} : y \mapsto [\partial_{y_i^i} \mathbb{J}_i(y_i^i; y_i^+, w_i^*)]_{i \in \mathcal{N}}$, and the selection matrix $\mathcal{R}$ as $\mathcal{R} := \text{blkd}(\{\mathcal{R}_i\}_{i \in \mathcal{N}})$, where each $\mathcal{R}_i := [I_{n_i}, \mathbf{0}_{n_i \times n_i^+}]$.

The use of local estimate y_i^j can be interpreted as introducing a "pseudo-player" j_i into the network. Each pseudo-player j_i is connected to player j, where player j and pseudo-player j_i for all $i \in \mathcal{N}_j^-$ constitute a connected component. We then conceptually disconnect the edges in $\mathcal{E}$, which gives rise to a new dependency network $\tilde{\mathcal{G}}$ with N such connected components. Let L denote the Laplace matrix of the network $\tilde{\mathcal{G}}$, which has eigenvalue zero with multiplicity N and other eigenvalues greater than zero. Each zero eigenvalue is associated with an eigenvector corresponding to the consensus within a connected component. By further extending each entry of L to a square matrix with proper dimension, we obtain a square matrix $\tilde{L} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ that we can leverage to obtain a compact form of the fixed-point iteration. We define the following operator $\mathbb{T}$ whose zeros correspond to exactly the SNEs of (1):

$$\mathbb{T} : y \mapsto \partial(\sum_{i \in \mathcal{N}} \mathbb{J}_i(y_i^i; y_i^+, w_i^*) + \iota_{\mathcal{X}_i}(y_i^i)) + \rho \tilde{L} y = \mathcal{R}^T \tilde{\mathbb{F}}_{w^*}(y) + N_{\tilde{\mathcal{X}}}(y) + \rho \tilde{L} y, \quad (3)$$

where $\rho \in \mathbb{R}_{++}$ is a constant controlling the contribution of local estimation errors. The formal statement of the equivalence is given below and the proof is reported in (Huang and Hu, 2022, Appendix A).

Theorem 2 *Suppose Assumptions 1 to 3 hold, and there exists $y^* \in \text{Zer}(\mathbb{T})$. Then $\tilde{L}y^* = \mathbf{0}$ and the tuple $\{y_i^{i*}\}_{i \in \mathcal{N}}$ satisfies the KKT conditions (2) for an SNE. Conversely, if the problem (1) has a solution $\{y_i^{i\dagger}\}_{i \in \mathcal{N}}$, then there exist local estimates $y_i^{j\dagger}$ such that their stack $y^\dagger \in \text{Zer}(\mathbb{T})$.*

Since the matrix $\tilde{L}$ couples the updates of all y_i and thus the resolvent of $\mathbb{T}$ can not be computed distributedly, we introduce a design matrix $\Phi := \tau^{-1} - \rho \tilde{L}$ and compute the resolvent of $\Phi^{-1} \mathbb{T}$ instead, where $\tau := \text{blkd}(\tau_1, \ldots, \tau_N)$ and each $\tau_i := \text{blkd}(\tau_{i0} \otimes I_{n_i}, \{\tau_{ij} \otimes I_{n_j}\}_{j \in \mathcal{N}_i^+}) \in \mathbb{S}_{++}^{n_i + n_i^+}$

5

is a diagonal matrix with all diagonal entries (step sizes) positive. Moreover, to ensure that the zero set of $\mathbb{T}$ and $\Phi^{-1}\mathbb{T}$ are equivalent and the associated Hilbert space is well-defined, the selected step sizes in τ should be sufficiently small such that Φ is positive definite. According to the Gershgorin circle theorem (Bell (1965)), for each player i, it suffices to set the step size of the local decision $\tau_{i0}^{-1} > 2\rho N_i^-$ and the step sizes of the local estimate $\tau_{ij}^{-1} > 2$. Let $\mathcal{K}$ be the Hilbert space obtained by endowing the vector space $\mathbb{R}^{\tilde{n}}$ with the inner product $\langle y, y' \rangle_{\mathcal{K}} = \langle \Phi y, y' \rangle$. With all the introduced elements, we can compute a zero of $\Phi^{-1}\mathbb{T}$ by utilizing the following iteration:

$$\tilde{y}^{(k+1)} := J_{\Phi^{-1}\mathbb{T}}(y^{(k)}), \; y^{(k+1)} := y^{(k)} + \gamma^{(k)}(\tilde{y}^{(k+1)} - y^{(k)}). \tag{4}$$

The detailed implementation of (4) consists of the optimization of augmented best-response objectives $\hat{\mathbb{J}}_i^{(k)}$'s and some linear updates, where $\hat{\mathbb{J}}_i^{(k)}(\tilde{y}_i^i; w_i^*) := \mathbb{J}_i(\tilde{y}_i^i; \tilde{y}_i^{+(k+1)}, w_i^*) + \rho(\sum_{j \in \mathcal{N}_i^-} y_i^{i(k)} - y_j^{i(k)})^T \tilde{y}_i^i + \frac{1}{2\tau_{i0}}\|\tilde{y}_i^i - y_i^{i(k)}\|_2^2$, which is omitted here for brevity. In the following, we describe a modified version that will be used throughout the learning dynamics in Section 2.3. With the introduction of the local estimates y_i^j and the extended pseudogradient $\tilde{\mathbb{F}}_{w^*}$, the operator $\mathbb{T}$ is no longer maximally monotone and the resolvent $J_{\Phi^{-1}\mathbb{T}}$ does not possess the firmly nonexpansive property in general. We denote the greatest (resp. smallest) out-neighbor count in $\mathcal{G}$ by $\bar{N}^-$ (resp. $\underline{N}^-$), i.e., $\bar{N}^- := \max\{N_i^- : i \in \mathcal{N}\}$ (resp. $\underline{N}^- := \min\{N_i^- : i \in \mathcal{N}\}$). To prove the convergence for the iteration with precise parameters, we will need to impose the following assumption on the pseudo-gradient $\mathbb{F}_{w^*}$ and the extended pseudo-gradient $\tilde{\mathbb{F}}_{w^*}$.

Assumption 5 *(Regularity of Pseudo-Gradient) At least one of the following statements holds:*

(i) The operator $\mathcal{R}^T \tilde{\mathbb{F}}_{w^*} + \rho \tilde{L}$ is maximally monotone;

(ii) The pseudogradient $\mathbb{F}_{w^*}$ is strongly monotone and Lipschitz continuous, i.e., there exist $\eta > 0$ and $\theta_1 > 0$, such that $\forall x, x' \in \mathbb{R}^n$, $\langle x - x', \mathbb{F}_{w^*}(x) - \mathbb{F}_{w^*}(x') \rangle \geq \eta\|x - x'\|^2$ and $\|\mathbb{F}_{w^*}(x) - \mathbb{F}_{w^*}(x')\| \leq \theta_1\|x - x'\|$. The operator $\mathcal{R}^T \tilde{\mathbb{F}}_{w^*}$ is Lipschitz continuous, i.e., there exists $\theta_2 > 0$, such that $\forall y, y' \in \mathbb{R}^{nN}$, $\|\tilde{\mathbb{F}}_{w^*}(y) - \tilde{\mathbb{F}}_{w^*}(y')\| \leq \theta_2\|y - y'\|$. Moreover, the weight of $\rho \tilde{L}$ satisfies $\rho \geq \frac{1}{\sigma_1}\left(\frac{\bar{N}^-+1}{\underline{N}^-+1}\frac{(\theta_1+\theta_2)^2}{4\eta} + \theta_2\right)$.

From one perspective, the convergence result under Assumption 5 (i) directly follows from the monotone operator theory and firmly nonexpansive fixed-point iterations. Nevertheless, verifying the fulfillment of this assumption is cumbersome and often can only be done numerically (See (Johansson and Rantzer, 2012, Sec. 4.2.3) for examples). From another perspective, for a monotone game, i.e., games with monotone $\mathbb{F}_{w^*}$, when restricted to the consensus subspace $y_i^j = y_j^j$ for all $i \in \mathcal{N}$ and $j \in \mathcal{N}_i^+$, $\mathbb{T}$ enjoys maximally monotonicity and its resolvent $J_{\Phi^{-1}\mathbb{T}}$ possesses firmly nonexpansiveness. Controlling the growth rate of $\tilde{\mathbb{F}}_{w^*}$ w.r.t. local estimates with Lipschitz continuity, Assumption 5 (ii) extends the above statement and allows for a violation of the consensus constraints if the missing monotonicity of $\tilde{\mathbb{F}}$ can be compensated by the measure of violation $\frac{\rho}{2}y^T\tilde{L}y$. A weaker concept that emerges under (ii) is quasinonexpansiveness: a general operator A is called quasinonexpansive if $\forall x \in \text{dom } A$ and $\forall y \in \text{Fix}(A)$, $\|Ax - y\| \leq \|x - y\|$. Now we are ready to formulate the theorem providing sufficient conditions for the proposed algorithm to work with perfect information of its model parameters. The proof is reported in (Huang and Hu, 2022, Appendix A).

Theorem 3 *Suppose Assumptions 1 to 5 hold, and τ is properly chosen such that Φ is positive definite. Then $J_{\Phi^{-1}\mathbb{T}}$ is a continuous quasinonexpansive operator, and the sequence $(y^{(k)})_{k \in \mathbb{N}}$*

generated by the fixed-point iteration (4) *will converge to a zero of* $\mathbb{T}$, *where* $\gamma^{(k)} \in [0,1]$ *and* $\sum_{n \in \mathbb{N}} \gamma^{(k)}(1 - \gamma^{(k)}) = +\infty$.

2.3. SNE Problems with Unknown Parameters

Subroutine 1: Distributed Nash Equilibrium Seeking

Available variables: $\{y_i^{(k)}\}$, $\{\hat{w}_i^{(k)}\}$;

At the k-th iteration, each player $i \in \mathcal{N}$:

　　Receive $\{y_j^{j(k)}\}_{j \in \mathcal{N}_i^+}$ from its in-neighbors and $\{y_j^{i(k)}\}_{j \in \mathcal{N}_i^-}$ from its out-neighbors;

　　$\tilde{y}_i^{j(k+1)} = y_i^{j(k)} - \tau_{ij}\rho(y_i^{j(k)} - y_j^{j(k)}), \quad \forall j \in \mathcal{N}_i^+;$

　　$\tilde{y}_i^{i(k+1)} = \underset{\tilde{y}_i^i \in \mathcal{X}_i}{\operatorname{argmin}}\Big\{\mathbb{J}_i(\tilde{y}_i^i; \tilde{y}_i^{+(k+1)}, \hat{w}_i^{(k)}) + \rho\big(\sum_{j \in \mathcal{N}_i^-} y_i^{i(k)} - y_j^{i(k)}\big)^T \tilde{y}_i^i + \frac{1}{2\tau_{i0}}\|\tilde{y}_i^i - y_i^{i(k)}\|_2^2\Big\};$

　　$y_i^{(k+1)} = y_i^{(k)} + \gamma^{(k)}(\tilde{y}_i^{(k+1)} - y_i^{(k)}).$

We now shift to the setting where at each iteration k, each player i has no access to the precise parameter w_i^* while it maintains a parameter estimate $\hat{w}_i^{(k)} := [\hat{w}_{ji}^{(k)}]_{j \in \{i\} \cup \mathcal{N}_i^+}$. We investigate the following simple learning dynamics where at each iteration, each player i first makes a decision which is determined by the optimizer of the given augmented objective using the estimated parameters and certain random exploration factor, then observes the realized objective value and the decisions made by its neighbors, and finally updates the estimates the unknown parameters. Note that instead of minimizing the augmented objective $\hat{\mathbb{J}}_i^{(k)}(\tilde{y}_i^i; w_i^*)$, player i now makes a best-response decision w.r.t. $\hat{\mathbb{J}}_i^{(k)}(\tilde{y}_i^i; \hat{w}_i^{(k)})$. This substitution in parameters gives rise to an estimated operator $\mathbb{T}^{(k)}$ for $\mathbb{T}$, with $\tilde{\mathbb{F}}_{w^*}$ replaced by $\tilde{\mathbb{F}}_{\hat{w}^{(k)}}$. For notational brevity, we let the exact iterations be denoted by $\mathscr{R}_* := J_{\Phi^{-1}\mathbb{T}}$ and $\mathscr{P}_* := I + \gamma^{(k)}(\mathscr{R}_* - I)$, and the estimated iterations (based on parameter estimates) be denoted by $\mathscr{R}^{(k)} := J_{\Phi^{-1}\mathbb{T}^{(k)}}$ and $\mathscr{P}^{(k)} := I + \gamma^{(k)}(\mathscr{R}^{(k)} - I)$. The SNE seeking dynamics is then described by $y^{(k+1)} := \mathscr{P}^{(k)}y^{(k)}$, and the intermediate result is given by $\tilde{y}^{(k+1)} := \mathscr{R}^{(k)}y^{(k)}$. The detailed implementation is included in Subroutine 1.

3. Parameter Learning Model

In this section, we consider the proper way to generate the estimate sequence $(\hat{w}_i^{(k)})_{k \in \mathbb{N}}$ when each player i has no clue or is uncertain about the parameters w_i^* inside its objective $\mathbb{J}_i$. Assume that each player has access to a bandit feedback system, which returns the realized objective function value based on the decision profile of the whole player network. To enable each player to perform ordinary least squares estimation to learn $\hat{w}_i^{(k)}$ at each iteration k, we make the following assumption:

Assumption 6 *(Parameter Learning) For each player $i \in \mathcal{N}$ and at each iteration $k \in \mathbb{N}$, the following conditions hold:*

　　(i) $(\xi_i^{(k)})_{k \in \mathbb{N}}$ is a sequence of real independent random variables with expectation zero and range bounded;

　　(ii) The scenario-based function J_i is invertible in s_i;

　　(iii) The feasible parameter set $\mathcal{W}_i \subseteq \mathbb{R}^{n_i^+ + 1}$ is convex and compact. In addition, for any $\hat{w}_i \in \mathcal{W}_i$, the augmented objective $\hat{\mathbb{J}}_i^{(k)}(\cdot; \hat{w}_i)$ is a strictly convex function on $\mathcal{X}_i$.

Note that the uniformly "thin" tail of the random variable ξ_i is essential to apply the large deviations result for later convergence analysis, and Assumption 6 (i) prescribes a sufficient condition where this requirement is satisfied. Condition (ii) enables each player to recover the values s_i based on the observed objective values. Condition (iii) ensures that $\hat{\mathbb{J}}_i^{(k)}(\cdot; \hat{w}_i)$ with $\hat{w}_i \in \mathcal{W}_i$ admits a unique argmin solution and the resolvent $J_{\Phi^{-1}\mathbb{T}}$ is well-defined and single-valued in $\mathcal{K}$.

To estimate the unknown parameters, each player $i \in \mathcal{N}$ picks the pivot point $y_i^{i(k+1)} \in \mathcal{X}_i$ where this player seeks to observe its payoff and estimate the parameters. Moreover, player i draws a random exploration vector $\delta_i^{(k)} : \Omega \to \mathbb{R}^{n_i}$ and plays $\check{y}_i^{i(k+1)} := y_i^{i(k+1)} + \delta_i^{(k)}$. The random vector $\delta_i^{(k)}$ should satisfy the following assumption.

Assumption 7 *(Random Exploration) For each player $i \in \mathcal{N}$, $(\delta_i^{(k)})_{k \in \mathbb{N}}$ is a sequence of independent identically distributed (i.i.d.) random variables with zero mean, bounded range, and positive definite covariance matrix.*

Nevertheless, the feasibility issue will arise with the introduction of the random exploration $\delta_i^{(k)}$. In the spirit of Agarwal et al. (2010); Bravo et al. (2018), we assume in the following that each local feasible set $\mathcal{X}_i$ is a convex body in $\mathbb{R}^{n_i}$, i.e., it has a nonempty topological interior. Moreover, we will introduce a "safe net", and adjust the chosen pivot point $y_i^{i(k+1)}$ to reside within a suitably shrunk zone of $\mathcal{X}_i$. In details, let $\mathbb{B}_{r_i}(p_i)$ be an r_i-ball centered at some $p_i \in \mathcal{X}_i$ so that $\mathbb{B}_{r_i}(p_i) \subseteq \mathcal{X}_i$. Then, instead of directly perturbing $y_i^{i(k+1)}$ by $\delta_i^{(k)}$, we consider the feasibility adjustment $\tilde{\delta}_i^{(k)} := \delta_i^{(k)} - \bar{\delta}_i r_i^{-1}(y_i^{i(k+1)} - p_i)$, where $\bar{\delta}_i := \max\{\|\delta_i(\omega)\|_2 : \omega \in \Omega\}$ satisfies $\bar{\delta}_i r_i^{-1} < 1$. Each player i plays $y_i^{i(k+1)} + \tilde{\delta}_i^{(k)}$ instead of $y_i^{i(k+1)} + \delta_i^{(k)}$. This adjustment moves each pivot $\mathcal{O}(\bar{\delta}_i)$-closer to the interior base point p_i with $y_{i\delta}^{i(k+1)} = y_i^{i(k+1)} - \bar{\delta}_i r_i^{-1}(y_i^{i(k+1)} - p_i)$, and then perturbs $y_{i\delta}^{i(k+1)}$ by $\delta_i^{(k)}$. Given the fact that $\check{y}_i^{i(k+1)} := y_{i\delta}^{i(k+1)} + \delta_i^{(k)} = (1 - \bar{\delta}_i r_i^{-1})y_i^{i(k+1)} + \bar{\delta}_i r_i^{-1}(p_i + r_i \bar{\delta}_i^{-1} \delta_i^{(k)}) \in \mathcal{X}_i$ and $p_i + r_i \bar{\delta}_i^{-1} \delta_i^{(k)} \in \mathbb{B}_{r_i}(p_i)$, feasibility of the query point is then ensured. After the above feasibility adjustment, let $s_i^{(k)} = w_{ii}^* + \sum_{j \in \mathcal{N}_i^+} w_{ji}^{*T} \check{y}_j^{j(k+1)} + \xi_i^{(k)}$ and $\ell_i^{(k)} = [1; [\check{y}_j^{j(k+1)}]_{j \in \mathcal{N}_i^+}]$. Based on the observed values available at the k-th iteration, the OLSE $\hat{w}_i^{(k+1)}$ is given by

$$\hat{w}_i^{(k+1)} := \text{argmin}_{w_i \in \mathcal{W}_i} \frac{1}{k+1} \sum_{t=0}^{k} (s_i^{(t)} - \langle \ell_i^{(t)}, w_i \rangle)^2. \tag{5}$$

The complete parameter learning dynamics is given in Subroutine 2.

Subroutine 2: Unknown Parameter Estimation

At the k-th iteration:

Each player $i \in \mathcal{N}$:

    Randomly picks an exploration factor $\delta_i^{(k)}$;

    Makes its decision to play $\check{y}_i^{i(k+1)} := y_i^{i(k+1)} + \delta_i^{(k)} - \bar{\delta}_i r_i^{-1}(y_i^{i(k+1)} - p_i)$;

    Observes $J_i(\check{y}_i^{i(k+1)}; s_i^{(k)})$ and receives $\{\check{y}_j^{j(k+1)}\}_{j \in \mathcal{N}_i^+}$ from its in-neighbors;

    Estimates the unknown parameters $\hat{w}_i^{(k+1)}$ by solving (5).

---

4. Learning Dynamics and Convergence Analysis

Assembling the updating steps of NE seeking and those of parameter estimation together, the learning dynamics of NE seeking with unknown parameters in objectives is described in Algorithm 3.

Algorithm 3: Distributed Learning of v-SGNE with Unknown Parameters

---

Initialize: $\{y_i^{(0)}\}$, $\{\hat{w}_i^{(0)}\}$ with $y_i^{(0)} \in \mathcal{X}_i$ and $\hat{w}_i^{(0)} \in \mathcal{W}_i$;
Iterate until convergence:
1) NE seeking updating step: run Subroutine 1;
2) parameter estimation updating step: run Subroutine 2;
Return: $\{y_i^{(k)}\}$, $\{\hat{w}_i^{(k)}\}$.

We start by establishing the following convergence result for a fixed-point iteration with the K-M scheme and a general continuous quasinonexpansive fixed-point iteration operator $\mathscr{R}_* : \mathcal{H} \to \mathcal{H}$ and its approximates $(\mathscr{R}^{(k)})_{k \in \mathbb{N}}$:

$$x^{(k+1)} := x^{(k)} + \gamma^{(k)}(\mathscr{R}^{(k)}x^{(k)} - x^{(k)}), \tag{6}$$

where $\mathcal{H}$ is a finite-dimensional Hilbert space, with its inner product and norm denoted by $\langle,\rangle_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}}$, respectively. Before proceeding, we introduce the following notations to facilitate the later discussion and analysis. The estimated iteration error and its norm in $\mathcal{H}$ at each iteration k are defined as: $\epsilon^{(k)} := \mathscr{R}^{(k)}(x^{(k)}) - \mathscr{R}_*(x^{(k)})$ and $\varepsilon^{(k)} := \|\epsilon^{(k)}\|_{\mathcal{H}}$. A residual function $\mathrm{res}(x) := \|x - \mathscr{R}_*(x)\|_{\mathcal{H}}$ is introduced such that $\mathrm{res}(x^*) = 0$ is a necessary condition for $x^* \in \mathrm{Fix}(\mathscr{R}_*)$. The proof of the following convergence theorem is reported in (Huang and Hu, 2022, Appendix B).

Theorem 4 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ be a sequence of sub-σ-fields of $\mathcal{F}$. Suppose for $k = 0, 1, \ldots$, $x^{(k)}$ is an $\mathcal{F}_k$-measurable random vector generated by the inexact fixed-point iteration (6), where $\mathscr{R}_* : \mathcal{H} \to \mathcal{H}$ is a continuous quasinonexpansive operator and $(\mathscr{R}^{(k)})_{k \in \mathbb{N}}$ denotes a sequence of its approximates subject to stochasticity. Moreover, suppose the sequence $(\gamma^{(k)})_{k \in \mathbb{N}}$ satisfies $0 \leq \gamma^{(k)} \leq 1$ and $\sum_{k \in \mathbb{N}} \gamma^{(k)}(1 - \gamma^{(k)}) = +\infty$. If the approximates $(\mathscr{R}^{(k)})_{k \in \mathbb{N}}$ and the generated $(x^{(k)})_{k \in \mathbb{N}}$ satisfy the following two conditions:*

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

then $(x^{(k)})_{k \in \mathbb{N}}$ will almost surely converge to a fixed point of $\mathscr{R}_*$.

We now consider the specific iteration for the locally coupled network games discussed in Section 2.3. Let $(y^{(k)})_{k \in \mathbb{N}}$ denote the sequence generated by applying $(\mathscr{P}^{(k)})_{k \in \mathbb{N}}$, i.e., $y^{(k+1)} := \mathscr{P}^{(k)} \circ \cdots \circ \mathscr{P}^{(0)}(y^{(0)})$. We further define $y_*^{(k+1)} := \mathscr{P}_*(y^{(k)})$ for each $k \in \mathbb{N}$. For brevity, we shall write $\{\delta_i^{(k)}\}$ in replacement of $\{\delta_i^{(k)}\}_{i \in \mathcal{N}}$ and similarly for other sets indexed by $\mathcal{N}$, unless otherwise specified. For each $k \geq 2$, define the sub-σ-field $\mathcal{F}_k$ as follows:

$$\mathcal{F}_k := \sigma\{y^{(0)}, \hat{w}^{(0)}, \{\delta_i^{(0)}\}, \{\xi_i^{(0)}\}, \ldots, \{\delta_i^{(k-2)}\}, \{\xi_i^{(k-2)}\}\}; \tag{7}$$

and define $\mathcal{F}_0 := \sigma\{y^{(0)}\}$ and $\mathcal{F}_1 := \sigma\{y^{(0)}, \hat{w}^{(0)}\}$. In the next theorem, we will prove that the sequence produced by Algorithm 3 satisfies the conditions required in Theorem 4. Define $\|\Delta\hat{w}_i^{(k)}\|_2 := \|\hat{w}_i^{(k)} - w_i^*\|_2$, $\|\Delta\hat{w}^{(k)}\|_2 := \|\hat{w}^{(k)} - w^*\|_2$, and the estimated iteration error $\epsilon^{(k)} :=$

$\mathscr{R}^{(k)}(y^{(k)}) - \mathscr{R}_*(y^{(k)})$ and its norm $\varepsilon^{(k)} := \|\epsilon^{(k)}\|_{\mathcal{K}}$, with $\mathscr{R}^{(k)}$ and $\mathscr{R}_*$ given in Sec. 2.3. The detailed proof is reported in (Huang and Hu, 2022, Appendix B).

Theorem 5 *Consider the sequence $(y^{(k)})_{k\in\mathbb{N}}$ generated by Algorithm 3. Suppose Assumptions 1 to 6 hold, and given a sequence $(\gamma^{(k)})$ where $0 \leq \gamma^{(k)} \leq 1$ and $\sum_{k\in\mathbb{N}} \gamma^{(k)}(1 - \gamma^{(k)}) = +\infty$, the sequence $(\gamma^{(k)}\mathbb{E}[\|\Delta\hat{w}^{(k)}\|_2 \mid \mathcal{F}_k])_{k\in\mathbb{N}}$ is a.s. absolutely summable. Then $(\|y^{(k)}\|_{\mathcal{K}})_{k\in\mathbb{N}}$ is bounded a.s., and the estimated iteration error satisfies $\sum_{k\in\mathbb{N}} \gamma^{(k)}\mathbb{E}[\varepsilon^{(k)} \mid \mathcal{F}_k] < \infty$ a.s.*

To establish the convergence of Algorithm 3, Theorems 4 and 5 suggest that the sequence $(\gamma^{(k)}\mathbb{E}[\|\Delta\hat{w}^{(k)}\|_2 \mid \mathcal{F}_k])_{k\in\mathbb{N}}$ should fulfill the summability assumption a.s. Our aim in what follows will be proving the following asymptotic convergence rate result of OLSE and investigating the relation between the estimation error $\|\Delta\hat{w}^{(k)}\|_2$ and the total number of observations made until the k-th iteration by utilizing the law of large deviation for OLSEs. We refer the interested readers to (Huang and Hu, 2022, Appendix C) for the detailed proof.

Theorem 6 *Suppose Assumptions 6 and 7 hold, and each player $i \in \mathcal{N}$ at the iteration $k \in \mathbb{N}$ uses Subroutine 2 to obtain an estimate $\hat{w}_i^{(k)}$. Let $\alpha_2 \in (0, \frac{1}{2})$ be an arbitrary constant. Then on a sample set $\hat{\Omega}$ with probability one, for any $\hat{\omega} \in \hat{\Omega}$, there exists a sufficiently large index $K_{lse}(\hat{\omega}) \in \mathbb{N}$ such that for all $k > K_{lse}(\hat{\omega})$, $\mathbb{E}[\|\Delta\hat{w}_i^{(k)}\|_2 \mid \mathcal{F}_k](\hat{\omega}) \leq C_{\Delta,i} k^{\alpha_2 - 1/2}$ holds for each $i \in \mathcal{N}$, where $C_{\Delta,i}$ is a constant independent of $\hat{\omega}$ and k.*

To conclude, we note that as long as the sequence of step sizes is chosen as $\gamma^{(k)} := 1/k^{\alpha_1}$ with $1/2 < \alpha_1 \leq 1$, there always exists a feasible $\alpha_2 = \frac{1}{2}(\alpha_1 - \frac{1}{2}) \in (0, \frac{1}{4}]$, such that by Theorem 6 for any $\hat{\omega} \in \hat{\Omega}$ and $k > K_{lse}(\hat{\omega})$, $\gamma^{(k)}\mathbb{E}[\|\Delta\hat{w}_i^{(k)}\|_2 \mid \mathcal{F}_k](\hat{\omega}) \leq C_{\Delta,i} k^{-\frac{1}{2}\alpha_1 - \frac{3}{4}}$ for all i. This together with Theorems 4 and 5 implies the almost-sure convergence of Algorithm 3, i.e., $y_i^{i(k)} \overset{\text{a.s.}}{\to} x_i^*$ and $\hat{w}^{(k)} \overset{\text{a.s.}}{\to} w_i^*$ for all i, where $x^* := [x_i^*]_{i\in\mathcal{N}}$ denotes an SNE of (1).

5. Conclusion and Future Directions

This paper develops a distributed solution to find Nash equilibria in stochastic locally coupled network games with unknown parameters by combining the proximal-point algorithm for Nash equilibrium seeking and the ordinary least square estimator for parameter learning. Almost-sure convergence of the solution algorithm is established, which can be further extended to handle generalized Nash equilibrium problems and iterations using inexact solvers. There remain several open problems. In the learning dynamics, to fulfill the identifiability condition for the estimator, each player is required to add random exploration factors to its decisions, and the actual decisions (perturbed by random exploration factors) it plays throughout the iteration will eventually bounce within some ϵ-neighborhood of a true Nash equilibrium, instead of converging to it. Hence, one of our future directions is to design learning dynamics such that the actual sequences of play can converge to the true Nash equilibria. Another potential future direction resides in considering an estimator which can better deal with the nonlinear parameter estimation and can work more efficiently in an online-learning fashion with suitable guarantees on convergence rate. In addition, even though we can extend the current analysis and similarly prove the convergence to a generalized Nash equilibrium when taking locally coupled constraints and global resource constraints, the actual action sequence may violate these coupled constraints during the iterations, which prevents the application of the proposed solution in some practical situations. We intend to address these questions in future work.

Acknowledgments

References

Alekh Agarwal, Ofer Dekel, and Lin Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *COLT*, pages 28–40. Citeseer, 2010.

Heinz H Bauschke. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics, Ouvrages de mathématiques de la SMC. 2nd ed. 2017. edition, 2017. ISBN 3-319-48310-2.

Howard E Bell. Gershgorin's theorem and the zeros of polynomials. *The American Mathematical Monthly*, 72(3):292–295, 1965.

Mattia Bianchi, Giuseppe Belgioioso, and Sergio Grammatico. Fast generalized Nash equilibrium seeking under partial-decision information. *arXiv preprint arXiv:2003.09335*, 2020.

Mario Bravo, David S Leslie, and Panayotis Mertikopoulos. Bandit learning in concave n-person games. *arXiv preprint arXiv:1810.01925*, 2018.

Ignacio Esponda, Demian Pouzo, and Yuichi Yamamoto. Asymptotic behavior of bayesian learners with misspecified models. *Journal of Economic Theory*, 195:105260, 2021.

Francisco Facchinei and Christian Kanzow. Generalized Nash equilibrium problems. *Annals of Operations Research*, 175(1):177–211, 2010.

Roman Frydman. Towards an understanding of market processes: individual expectations, learning, and convergence to rational expectations equilibrium. *The American Economic Review*, 72(4): 652–668, 1982.

Amélie Héliou, Matthieu Martin, Panayotis Mertikopoulos, and Thibaud Rahier. Zeroth-order nonconvex learning via hierarchical dual averaging. In *International Conference on Machine Learning*, pages 4192–4202. PMLR, 2021.

Yuanhanqing Huang and Jianghai 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.

Yuanhanqing Huang and Jianghai Hu. Distributed stochastic nash equilibrium learning in locally coupled network games with unknown parameters, 2022. URL https://arxiv.org/abs/2204.00100.

Hao Jiang, Uday V Shanbhag, and Sean P Meyn. Distributed computation of equilibria in misspecified convex stochastic Nash games. *IEEE Transactions on Automatic Control*, 63(2):360–371, 2017.

Rolf Johansson and Anders Rantzer. *Distributed decision making and control*. Springer, 2012.

Aswin Kannan, Uday V Shanbhag, and Harrison M Kim. Strategic behavior in power markets under uncertainty. *Energy Systems*, 2(2):115–141, 2011.

Aswin Kannan, Uday V Shanbhag, and Harrison M Kim. Addressing supply-side risk in uncertain power markets: stochastic Nash models, scalable algorithms and error analysis. *Optimization Methods and Software*, 28(5):1095–1138, 2013.

Alan P Kirman. Learning by firms about demand conditions. In *Adaptive economic models*, pages 137–156. Elsevier, 1975.

Jinlong Lei and Uday V Shanbhag. Asynchronous schemes for stochastic and misspecified potential games and nonconvex optimization. *Operations Research*, 68(6):1742–1766, 2020.

Emily Meigs, Francesca Parise, and Asuman Ozdaglar. Learning dynamics in stochastic routing games. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 259–266. IEEE, 2017.

Emily Meigs, Francesca Parise, and Asuman Ozdaglar. Learning in repeated stochastic network aggregative games. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 6918–6923. IEEE, 2019.

Anna Nagurney, Mojtaba Salarpour, June Dong, and Ladimer S Nagurney. A stochastic disaster relief game theory network model. In *SN Operations Research Forum*, volume 1, pages 1–33. Springer, 2020.

John F Nash et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

Evdokia Nikolova and Nicolás E Stier-Moses. A mean-risk model for the traffic assignment problem with stochastic travel times. *Operations Research*, 62(2):366–382, 2014.

Daniel P Palomar and Yonina C Eldar. *Convex optimization in signal processing and communications*. Cambridge university press, 2010.

Francesca Parise and Asuman Ozdaglar. A variational inequality framework for network games: Existence, uniqueness, convergence and sensitivity analysis. *Games and Economic Behavior*, 114:47–82, 2019.

Francesca Parise and Asuman Ozdaglar. Analysis and interventions in large network games. *Annual Review of Control, Robotics, and Autonomous Systems*, 4:455–486, 2021.

Francesca Parise, Basilio Gentile, Sergio Grammatico, and John Lygeros. Network aggregative games: Distributed convergence to Nash equilibria. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 2295–2300. IEEE, 2015.

Francesca Parise, Sergio Grammatico, Basilio Gentile, and John Lygeros. Distributed convergence to Nash equilibria in network and average aggregative games. *Automatica*, 117:108959, 2020.

Lacra Pavel. Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach. *IEEE Transactions on Automatic Control*, 65 (4):1584–1597, 2019.

Uma Ravat and Uday V Shanbhag. On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games. *SIAM Journal on Optimization*, 21(3):1168–1199, 2011.

Vinayak Uday V Shanbhag. *Decomposition and sampling methods for stochastic equilibrium problems*. PhD thesis, Stanford University, 2006.

Wei Shi and Lacra Pavel. LANA: an ADMM-like Nash equilibrium seeking algorithm in decentralized environment. In *2017 American Control Conference (ACC)*, pages 285–290. IEEE, 2017.

Tatiana Tatarenko and Maryam Kamgarpour. Bandit online learning of Nash equilibria in monotone games. *arXiv preprint arXiv:2009.04258*, 2020.

Peng Yi and Lacra Pavel. An operator splitting approach for distributed generalized Nash equilibria computation. *Automatica*, 102:111–121, 2019.