Asynchronous Schemes for Stochastic and Misspecified Potential Games and Nonconvex Optimization

Jinlong Lei Uday V. Shanbhag *

Abstract

The distributed computation of equilibria and optima has seen growing interest in a broad collection of networked problems. We consider the computation of equilibria of convex stochastic Nash games characterized by a possibly nonconvex potential function. Since any stationary point of the potential function is a Nash equilibrium, there is an equivalence between asynchronous best-response (BR) schemes applied on Nash game and block-coordinate descent (BCD) schemes implemented on the potential function. We focus on two classes of stochastic Nash games: (**P1**): A potential stochastic Nash game, in which each player solves a parameterized stochastic convex program; and (**P2**): A misspecified generalization, where the player-specific stochastic program is complicated by a parametric misspecification with the unknown parameter being the solution to a stochastic convex optimization. In both settings, exact proximal BR solutions are generally unavailable in finite time since they necessitate solving stochastic programs. Consequently, we design two asynchronous inexact proximal BR schemes to solve problems (**P1**) and (**P2**), where in each iteration a single player is randomly chosen to compute an inexact proximal BR solution (via stochastic approximation) with delayed rival information while the other players keep their strategies invariant. In the misspecified regime (**P2**), each player possesses an extra estimate of the misspecified parameter by using a projected stochastic gradient (SG) algorithm with an increasing batch of sampled gradients. By imposing suitable conditions on the inexactness sequences, we prove that the iterates produced by both schemes converge almost surely to a connected subset of the set of Nash equilibria. When the player problems are strongly convex, an inexact pure BR scheme is shown to be convergent. In effect, we provide what we believe is amongst the first randomized BCD schemes for stochastic nonconvex (but block-wise convex) optimization with almost sure convergence properties. We further show that the associated gap function converges to zero in mean. These statements can be extended to allow for accommodating weighted potential games and generalized potential games. Finally, we present preliminary numerics based on applying the proposed schemes to congestion control and Nash-Cournot games.

1 Introduction

Nash games, rooted in the seminal work by [2], have seen wide applicability in a broad range of engineered systems, such as power grids, communication networks, transportation networks and sensor networks. In the $N-$player Nash game, each player maximizes a prescribed payoff over a player-specific strategy set, given the rival strategies. Nash's eponymous solution concept, Nash equilibrium (NE), requires that at an equilibrium, no player can improve its payoff by unilaterally deviating from its equilibrium strategy. Potential games represent

*The authors are with the Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA 16802 `jxl800,udaybag@psu.edu`. The work has been partly supported by NSF grant 1538605 and 1246887 (CAREER). Additionally, a preliminary version of this work appeared in the Proceedings of the IEEE Conference on Decision and Control [1].

an important subclass of Nash games, formally introduced in [3], that arise naturally in the modeling of many applications, ranging from congestion control [4], routing in communication networks [5], networked Cournot competition [6, 7, 8, 9], and a host of other control-theoretic problems [10, 11]. We refer the interested reader to the survey by [12] on potential games for additional references. An interesting extension is the class of *near potential games*, a class of games that are close to potential games and for which some learning dynamics (for instance, best-response, fictitious play, logit response) are discussed in [13]. Additional decomposition algorithms are proposed in [14] to solve *generalized* potential games in which the player-specific strategy set depends on the strategies selected by the other agents.

Motivation. While prior algorithmic efforts have considered deterministic Nash games (cf. [5, 6, 7]), there have also been recent attempts to contend with stochastic generalizations via stochastic gradient-response schemes (cf. [15, 16, 17, 9]). Yet, gradient-based schemes require access to rival strategies after every update, often undesirable in certain applications such as cellular networks. To this end, we consider *inexact* **proximal best-response** schemes applied to the stochastic Nash game (or equivalently *inexact* **block coordinate-descent** schemes applied to the stochastic potential function) that require rival strategies after taking an inexact proximal BR step, generally leading to lower communication complexity (measured by the amount of rival information). Further, in many regimes, the payoff functions are defined by parameters that are unavailable, e.g., parameters of inverse-demand functions. Accordingly, we allow for resolving **parametric misspecification** in player payoffs by equipping each player with a simultaneous learning step.

Problems of interest. We consider two classes of N-player potential stochastic Nash games with players indexed by i where $i \in \mathcal{N} \triangleq \{1, 2, \cdots, N\}$.

(P1): *Potential Stochastic Nash Games.* Suppose the ith player's strategy is denoted by x_i with a strategy set $X_i \subset \mathbb{R}^{n_i}$, implying that a feasible strategy x_i satisfies $x_i \in X_i$, and let $n \triangleq \sum_{i \in \mathcal{N}} n_i$. Additionally, suppose player i's objective (or negative of the utility function) is denoted by $f_i(x_i, x_{-i})$, which depends on its own strategy x_i and on the tuple of rival strategies $x_{-i} \triangleq \{x_j\}_{j \neq i}$. Suppose X and X_{-i} are defined as $X \triangleq \prod_{i=1}^N X_i$ and $X_{-i} \triangleq \prod_{j \neq i=1}^N X_j$, respectively. Given rival strategies x_{-i}, the ith player is faced by the following parameterized stochastic optimization problem:

$$\min_{x_i \in X_i} \quad f_i(x_i, x_{-i}) \triangleq \mathbb{E}_\xi \left[\psi_i(x_i, x_{-i}; \xi(\omega))\right], \tag{1}$$

where $\psi_i : X \times \mathbb{R}^d \to \mathbb{R}$ is a scalar-valued function and the random vector is $\xi : \Omega \to \mathbb{R}^d$ defined on the probability space $(\Omega, \mathcal{F}_x, \mathbb{P}_x)$. Our interest lies in a subclass of Nash games, qualified as *potential*, characterized by a potential function $P : X \to \mathbb{R}$ such that for any $i \in \mathcal{N}$ and for any $x_{-i} \in X_{-i}$:

$$P(x_i, x_{-i}) - P(x_i', x_{-i}) = f_i(x_i, x_{-i}) - f_i(x_i', x_{-i}), \quad \forall x_i, x_i' \in X_i. \tag{2}$$

Then the Nash game, in which the ith player solves (1) given x_{-i}, is called a *potential stochastic Nash game*. We aim to compute an NE, denoted by $x^* = \{x_i^*\}_{i=1}^N$ such that for any $i \in \mathcal{N}$, the following holds:

$$f_i(x_i^*, x_{-i}^*) \leq f_i(x_i, x_{-i}^*), \qquad \forall x_i \in X_i. \tag{3}$$

In other words, a feasible strategy tuple $x^* \in X$ is an Nash equilibrium if no player can improve its payoff by unilaterally deviating from its equilibrium strategy x_i^*.

(P2): *Misspecified Potential Stochastic Nash Games.* The frequently used assumption in game-theoretic models is that each player has perfect knowledge of the payoff function and is able to correctly forecast the choices of the other players. However, as pointed out by [18] in the context of Cournot oligopolies that firms are, in general, imperfectly aware of their environment. Therefore they may have an imperfect knowledge of the payoff. For example, players may employ the misspecified estimates of the demand function or the production capacity of their rivals. [18] introduced a learning process where firms update their conjectured demand functions according to the observed data when the game is played repeatedly. Subsequently, this notion was formalized in a series of papers for resolving parametric misspecification by [19, 20], [21], [22], amongst others. Recent work has examined the development of coupled stochastic approximation (SA) schemes for resolving misspecified stochastic optimization [23, 24], and stochastic Nash games [9]. In this work, we consider static stochastic Nash games complicated by a parameteric misspecification θ^*, in which the ith player's problem is represented as follows:

$$\min_{x_i \in X_i} \quad f_i(x_i, x_{-i}, \theta^*) \triangleq \mathbb{E}_\xi \left[\psi_i(x_i, x_{-i}, \theta^*; \xi(\omega))\right], \tag{4}$$

where $\theta^* \in \mathbb{R}^m$, $\xi : \Omega \to \mathbb{R}^d$ is defined on the probability space $(\Omega, \mathcal{F}_x, \mathbb{P}_x)$, and $\psi_i : X \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$ is a scalar-valued function. For instance, in the context of Nash-Cournot games [20], θ^* may represent the slope and intercept of a linear inverse demand (or price) function; see the example of Nash-Cournot games with misspecified parameters in Section 4.2. We consider the estimation of θ^* through solving a suitably defined convex stochastic program:

$$\min_{\theta \in \Theta} g(\theta) := \mathbb{E}_\eta \left[g(\theta, \eta)\right], \tag{5}$$

where $\Theta \in \mathbb{R}^m$ is a closed and convex set, $\eta : \Lambda \to \mathbb{R}^p$ is defined on the probability space $(\Lambda, \mathcal{F}_\theta, \mathbb{P}_\theta)$, and $g : \Theta \times \mathbb{R}^p \to \mathbb{R}$ is a scalar-valued function. We still consider the class of potential games and assume there exists a function $P(\cdot;\cdot) : X \times \Theta \to \mathbb{R}$ such that for any $i \in \mathcal{N}$ and every $x_{-i} \in X_{-i}$:

$$P(x_i, x_{-i}; \theta^*) - P(x_i', x_{-i}; \theta^*) = f_i(x_i, x_{-i}; \theta^*) - f_i(x_i', x_{-i}; \theta^*), \quad \forall x_i, x_i' \in X_i. \tag{6}$$

Then we refer to the stochastic Nash game (4) with the misspecified parameter θ^* being a solution to (5), as a misspecified potential stochastic Nash game. In such an instance, the problem of interest is to compute the correctly specified Nash equilibrium, defined as follows, holds for $i \in \mathcal{N}$:

$$f_i(x_i^*, x_{-i}^*, \theta^*) \leq f_i(x_i, x_{-i}^*, \theta^*), \qquad \forall x_i \in X_i. \tag{7}$$

Prior Research. We now discuss some relevant prior research on stochastic Nash games, best-response schemes for Nash games, coordinate-descent schemes for optimization problems, and distributed schemes for computing Nash equilibria.

(i) Stochastic Nash games. Both **(P1)** and **(P2)** belong to the class of static stochastic Nash games, as opposed to their dynamic variants, representing a set of Nash games in which the player objective functions are expectation-valued. Tractable sufficiency conditions for existence of equilibria to such games were provided in [25] in regimes where player objectives were convex and could be nonsmooth. SA schemes for a subclass of convex stochastic Nash games were presented under Lipschitzian assumptions in [15] via an iterative regularization technique while in [16, 17], Lipschitzian requirements were relaxed by utilizing a randomized smoothing approach. In both instances, almost sure convergence to an NE is guaranteed under suitable monotonicity requirements on the variational map. Though gradient-based schemes exist for solving stochastic Nash games are characterized by ease

of implementation and lower complexity in terms of each player step, such schemes are characterized by the following properties: (i) Players require rival strategies after every gradient step, necessitating a significant amount of communication; (ii) Convergence theory is reliant on a relatively strong monotonicity assumption on the gradient map; (iii) The schemes are synchronous. This motivates considering asynchronous best-response schemes, generally characterized by lower communication requirements and do not require a strong monotonicity property. Finally, [26] and [27] provide a characterization of rational behavior in a non-cooperative game, where a rational player acts optimally given the decisions of her competitors; i.e. rational players will play a best-response strategy. This represents a cornerstone of much of the discussion in the context of such techniques and further motivates the consideration of such schemes. However, we emphasize that schemes based on other strategies (such as gradient-response, etc.) are also of importance based on the setting being examined and the underlying assumptions imposed. Succinctly, we aim to develop implementable asynchronous proximal best-response schemes, a class of techniques that can cope with delays and requires far less communication.

(ii) Best-response (BR) and coordinate-descent schemes. In BR schemes, each player selects a BR strategy, given current rival strategies [28, 29]. [30] shows that for a class of games, it is best for a player, given that the others are repeatedly employing best-response, to also repeatedly employ a BR scheme. There have been efforts to extend BR schemes to engineering applications [31], where the BR can be expressed in a closed form. Recently, [32] have proposed several variants of the BR schemes to solve the two-stage noncooperative games with risk-averse players. Proximal BR schemes appear to have been first discussed in [33], where it is shown that the set of fixed points of the proximal BR map is equivalent to the set of Nash equilibria when the player-specific problem is convex. Additionally, [14] propose several regularized Gauss-Seidel BR schemes for generalized potential games and show that a limit point of the generated sequence is an NE when each player's subproblem is convex. Recall that stationary points of the potential function are NE of the original game when the player-specific problems are convex. Thus, BCD methods may be employed for obtaining either stationary points of the potential function (if nonconvex) or global minimizers (if convex), where the coordinates are partitioned into several blocks (each corresponding to a player in the associated Nash game) and at each iteration, a single block is chosen to update while the other blocks remain unchanged. Its original format dates back to [34], where blocks were updated *cyclically*. Convergence has been extensively studied for both convex and nonconvex regimes with either differentiable or nondifferentiable objectives [35]. Notice that the *asynchronous* BCD schemes, where at each epoch a single block is chosen, are, in essence, identical to asynchronous BR schemes. [36] considers a *randomized* BCD method that performs a gradient update on a randomly selected block. Extensions to convex nonsmooth regimes have been studied extensively [37] while in nonconvex regimes, [38, 39] provide convergence theory. [40] proposed an accelerated gradient method to solve nonconvex and possibly stochastic optimization, while [41] designed a randomized stochastic projected gradient algorithm to solve the constrained stochastic composite optimization. We summarize much of the prior work in Table 1, where we observe that there is no available a.s. convergence theory for (misspecified) potential stochastic Nash games (or nonconvex stochastic programs) via BR (or BCD) schemes.

(iii) Distributed computation of Nash equilibria. Recently, there has been an interest in considering settings where players compute their strategies in a distributed sense. [43] and [44] considered the networked aggregative games (where player payoffs are coupled via the aggregate strategy) with quadratic payoffs and proposed decentralized schemes for NE computation. In this setting, the communication graph is, in essence, the payoff dependence network. However, as has been pointed out in [45], a player is inherently limited in that it can communicate with (and be observed by) a few players in large-scale networked systems. Thus, in recent work,

Problem	Literature	Applicability	stochastic	method	rate	a.s.	convergence in mean	misspecification
Optimization	[38]	convex	✓	BSG	$\mathcal{O}(1/k)$	–	–	×
		nonconvex			–	–	✓	
	[40]	nonconvex	✓	gradient-based	$\mathcal{O}(1/\sqrt{k})$	–	✓	×
	[]ghadimi2016mini	nonconvex	✓	gradient-based	✓	–	✓	×
	[39]	nonconvex	×	gradient-based (asynchronous)	–	cluster point is an NE	–	×
	This work	nonconvex block-convex	✓	RBCD-based (asynchronous)	–	✓	✓	Section 2: × Section 3: ✓
Non-potential Game	[]jiang2017distributed	monotone	✓	gradient-based	$\mathcal{O}(1/k)$	✓	–	✓
	[]koshal2013regularized	monotone, Lipschitz	✓	gradient-based	–	✓	–	×
	[]yousefian2013regularized	monotone, non-Lipschitz	✓	gradient-based	–	✓	–	×
	[]yousefian2016self	strongly monotone, non-Lipschitz	✓	gradient-based	$\mathcal{O}(1/k)$	✓	✓	×
	[]pang2017two	contractive BR maps	✓	BR-based	linear	✓	–	×
	[]lei2017synchronous	contractive BR maps	✓	BR-based	linear	✓	✓	×
Potential Game	[]facchinei2011decomposition	generalized	×	BR-based (Gauss-Seidel)	–	cluster point is an NE	–	×
	This work	player-specific convex	✓	BR-based (asynchronous)	–	✓	✓	Section 2: × Section 3: ✓

Table 1: A list of some recent research papers on Nash games and (non)convex optimization.
The 3rd column "Applicability" specifies the major conditions required by the problem studied in the literature. In column 4 "stochastic" (or column 9 "misspecification"), ✓ means that the studied problem in the literature is stochastic (or with parametric misspecification), while × implies that the studied problem is deterministic (or without parametric misspecification); In columns 6-8, ✓implies that the reference has studied the corresponding convergence property, while dash – implies that the literature has not studied this property.

it is assumed that players form a (connected) communication network (which is not necessarily the same as the players' payoff dependence network) and exchange messages with the neighboring players to obtain estimates of unobserved information. In such a setting, [46] developed consensus-based distributed algorithms for aggregative games while [47] presented an asynchronous gossip-based algorithm. In a similar vein, [48] designed similar distributed algorithms for a class of generalized convex games while [49] considered similar settings under a strong monotonicity requirement on the concatenated gradient map. [50] designed a linearized ADMM-like scheme while continuous-time schemes were presented by [51] in which unobservable decisions are learnt via a consensus-based approach. Finally, [52] have developed distributed counterparts of (inexact) best-response and gradient-response for a range of stochastic Nash games.

Contributions: In [53] and [42], rate statements and iteration complexity bounds are provided for inexact proximal BR schemes for stochastic Nash games under a contractive requirement on the proximal BR map. However, even asymptotic guarantees are unavailable without such an assumption. Motivated by this gap, we aim to design convergent implementable asynchronous BR schemes such that at each epoch, a single player updates its strategy while the other players keep their strategies invariant. In our settings, each player-specific subproblem involves solving a stochastic program whose exact solution is generally unavailable in finite time, necessitating inexact solutions. Accordingly, we propose two classes of asynchronous inexact proximal BR schemes to compute NE of problems **(P1)** and **(P2)**, and make the following contributions: **(i).** In Section II, we propose an asynchronous inexact proximal BR scheme to solve **(P1)**. In each iteration, a single agent is randomly chosen to inexactly solve a stochastic optimization problem, given the delay-afflicted rival strategies via an SA scheme. By imposing suitable conditions on **(P1)** and on the inexactness sequence, in a regime that allows for uniformly bounded delays, we prove that the iterates converge a.s. to a connected subset of the set of Nash equilibria and that the gap function converges in mean to zero. Extensions are provided to the generalized stochastic potential games (with coupled strategy sets) and the weighted potential games. We further prove that asynchronous inexact pure BR schemes are convergent if player-specific problems are strongly convex. **(ii).** In Section III, we extend the regime to contend with the misspecified stochastic Nash game **(P2)** where every player updates its equilibrium strategy and its belief regarding the misspecified parameter (via variable sample-size SA schemes), given rival strategies afflicted by delays. Asymptotic guarantees analogous to Section II are provided and we additionally show that the belief regarding the misspecified parameter converges a.s. to its true counterpart. (iii) We provide some preliminary numerics on congestion games and Nash-Cournot games in Section IV, and conclude the paper in Section V.

Notations: When referring to a vector x, it is assumed to be a column vector while x^T denotes its transpose. Generally, $\|x\|$ denotes the Euclidean vector norm, i.e., $\|x\| = \sqrt{x^T x}$. For a nonempty closed convex set $X \subset \mathbb{R}^m$, we use $\Pi_X[x]$ to denote the Euclidean projection of a vector $x \in \mathbb{R}^m$ on X, i.e., $\Pi_X[x] = \min_{y \in X} \|x - y\|$. We write *a.s.* as the abbreviation for "almost surely". We use $\mathbb{E}[z]$ to denote the unconditional expectation of a random variable z. For a real number x, the floor function $\lfloor x \rfloor$ denotes the largest integer smaller than x. We use $[A]_{i,j}$ to denote the (i, j)-th entry of the matrix A.

2 Asynchronous Inexact Proximal Best-Response Schemes

In Section 2.1, we propose an asynchronous inexact proximal BR scheme to compute an equilibrium of the stochastic potential game **(P1)**. Then in Section 2.2 we introduce some basic assumptions, based on which, we proceed to prove the almost sure convergence and convergence in mean of the generated sequence to a Nash equilibrium in Section 2.3. In Section 2.4, we discuss some possible extensions including the generalized potential games allowing for coupled strategy sets, and weighted potential games. Finally, in Section 2.5, we show that in a delay-free regime, the asynchronous inexact pure BR scheme (i.e. without the proximal term) is convergent when the player-specific problem is strongly convex.

2.1 Algorithm Design

In standard potential games, a natural approach is an asynchronous BR method where in each iteration, one player updates its strategy by solving problem (1), given its rival strategies, referred to as the *best-response* problem. However, best-response schemes do not always lead to convergence to Nash equilibria. In fact, even in potential games where the potential function is player-wise convex, such convergence does not follow; see [14] for a simple counterexample. Accordingly, [14] propose a regularized BR scheme in which each player's objective is modified by adding a quadratic proximal term, and prove its convergence. Our research has been motivated by considering stochastic generalizations which do not follow immediately. Yet another reason for using proximal BR schemes may be drawn from the "momentum behavior" in economics, e.g. in the investment problems [54], the players may want to optimize their objective while staying close to their previous values. We then define the player i's proximal best-response problem as follows for some $\mu_i > 0$:

$$T_i(x) \triangleq \operatorname*{argmin}_{y_i \in X_i} \left[\mathbb{E}\left[\psi_i(y_i, x_{-i}; \xi(\omega))\right] + \frac{\mu_i}{2} \|y_i - x_i\|^2 \right]. \tag{8}$$

Since $T_i(x)$, the minimizer of a stochastic problem (8), is generally unavailable in finite time, we utilize Monte-Carlo sampling schemes in obtaining inexact solutions [55].

We assume that each player i always knows its current strategy, while is not immediately aware of rival strategies. Instead, its knowledge of each rival strategies may be afflicted by a rival-specific random delay, (see [56]).

We now propose an asynchronous inexact proximal best-response scheme (Algrithm 1) to compute an NE of this game. At time k, player i's strategy $x_i(k) \in \mathbb{R}^{n_i}$ is an estimate for its equilibrium strategy x_i^* and player i has access to delayed rival strategies $y^i(k) \triangleq (x_1(k - d_{i1}(k)), \cdots, x_N(k - d_{iN}(k)))$, where $d_{ij}(k)$ denotes the delay associated with player j's information, and $d_{ii}(k) = 0$. The scheme is defined as follows. At iteration $k \geq 0$, randomly pick a single i from $\mathcal{N}$ with probability $\mathbb{P}(i_k = i) = p_i > 0$. If $i_k = i$, then player i is chosen to initiate an update by computing an inexact proximal BR solution to problem (8) characterized by (9). We impose conditions on the inexactness sequence $\{\varepsilon_i(k)\}_{k \geq 1}$ when we proceed to investigate the convergence properties.

Algorithm 1 Asynchronous inexact proximal best-response scheme

---

Let $k := 0$, $x_{i,0} \in X_i$ for $i \in \mathcal{N}$. Additionally $0 < p_i < 1$ for $i \in \mathcal{N}$ such that $\sum_{i=1}^N p_i = 1$.

(S.1) Pick $i_k = i \in \mathcal{N}$ with probability p_i.

(S.2) If $i_k = i$, then player i updates $x_i(k+1) \in X_i$ as follows:

$$x_i(k+1) := T_i(y^i(k)) + \varepsilon_i(k+1), \tag{9}$$

where $\varepsilon_i(k+1)$ denotes the inexactness employed by player i at time $k+1$. Otherwise, $x_j(k+1) := x_j(k)$ if $j \notin i_k$.

(S.3) If $k > K$, stop; Else, $k := k+1$ and return to (S.1).

Remark 1 *In fact, in practical game-theoretic problems, players take actions in an asynchronous manner since there might not exist a global coordinator to ensure that players update simultaneously. The condition that $\mathbb{P}(i_k = i) = p_i > 0$ with $\sum_{i=1}^N p_i = 1$ accommodates the Poisson model employed by [57] and [58] as a special case. For $i \in \mathcal{N}$, player i is activated according to a local Poisson clock, which ticks according to a Poisson process with rate $\varrho_i > 0$. Suppose that there is a virtual global clock which ticks whenever any of the local Poisson clocks tick. Assume that the local Poisson clocks are independent, then the global clock ticks according to a Poisson process with rate $\sum_{i=1}^N \varrho_i$. Let Z_k denote the time of the k-th tick of the global clock. Since the local Poisson clocks are independent, with probability one, there is a single player whose Poisson clock ticks at time Z_k with probability $\mathbb{P}(i_k = i) = \frac{\varrho_i}{\sum_{i=1}^N \varrho_i} \triangleq p_i$. Further, the memoryless property of the Poisson process indicates that $\{i_k\}_{k \geq 0}$ is an independent and identically distributed (i.i.d.) sequence.*

2.2 Assumptions and Preliminary Results

For notational simplicity, let ξ denote $\xi(\omega)$ throughout the paper. We begin by imposing assumptions on X_i, f_i, ψ_i, and on the second moments of ψ_i.

Assumption 1 *Let the following hold.*
(a) For every $i \in \mathcal{N}$, the feasible set X_i is closed, compact, and convex;
(b) For every $i \in \mathcal{N}$, $f_i(x_i, x_{-i})$ is convex and continuously differentiable in $x_i \in X_i$ for every $x_{-i} \in X_{-i}$. In addition, there exists a Lipschitz constant $L_i > 0$ such that the following holds:

$$\|\nabla_{x_i} f_i(x) - \nabla_{x_i} f_i(x')\| \leq L_i \|x - x'\| \quad \forall x, x' \in X;$$

(c) For every $i \in \mathcal{N}$, all $x_{-i} \in X_{-i}$, and any $\omega \in \Omega$, $\psi_i(x_i, x_{-i}; \xi(\omega))$ is differentiable in x_i over an open set containing X_i such that $\nabla_{x_i} f_i(x_i, x_{-i}) = \mathbb{E}_\xi[\nabla_{x_i}\psi_i(x_i, x_{-i}; \xi)]$;
(d) For every $i \in \mathcal{N}$ and any $x \in X$, there exists a constant $M > 0$ such that $\mathbb{E}_\xi[\|\nabla_{x_i}\psi_i(x_i, x_{-i}; \xi)\|^2] \le M^2$.

It is seen that (c) and (d) pertain to the existence of a conditionally unbiased stochastic oracle and the boundedness of the conditional second moment of the sampled gradient generated by this oracle. Next, we assume the existence of a continuously differentiable potential function for the Nash game of interest.

Assumption 2 (Potential function) *There exists a potential function $P : X \to \mathbb{R}$ that is continuously differentiable over an open set containing X such that for any $i \in \mathcal{N}$ and any $x_{-i} \in X_{-i}$, equation* (2) *holds.*

We next make some assumptions on the delays as well as on the inexactness sequences $\{\varepsilon_i(k)\}$ utilized in Algorithm 1. We denote the σ-field of the entire information used by Algorithm 1 up to (and including) the update of $x(k)$ by $\mathcal{F}'_k$, and the σ-field generated from $\mathcal{F}'_k$ and the delays at time k by $\mathcal{F}_k \triangleq \sigma\{\mathcal{F}'_k, d_{ij}(k), i, j \in \mathcal{N}\}$. We will formally define $\mathcal{F}'_k$ after introducing the SA scheme (10).

Assumption 3 (Delay and inexactness sequences) *The following hold:*
(a) i_k is independent of $\mathcal{F}_k$ for all $k \ge 1$;
(b) for any $i \in \mathcal{N}$, the noise term $\{\varepsilon_i(k)\}$ satisfies the following condition:

$$\sum_{k=1}^{\infty} \mathbb{E}\left[\|\varepsilon_i(k+1)\|^2 \big| \mathcal{F}_k\right] < \infty, \ a.s., \ and \ \sum_{k=1}^{\infty} \mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] < \infty, \ a.s.;$$

(c) there exists a positive integer τ such that for any $i, j \in \mathcal{N}$ and any $k \ge 0$, $d_{ij}(k) \in \{0, \cdots, \tau\}$.

By Assumption 1, it is clear that $T_i(x)$ defined by (8) requires solving a strongly convex stochastic program. Thus, an approximation of the solution to the problem (8) with $x = y^i(k)$, characterized by (9), can be computed via the standard SA algorithm defined as follows for $t = 1, \ldots, j_i(k)$:

$$z_{i,t+1}(k) := \Pi_{X_i}\Big[z_{i,t}(k) - \gamma_{i,t}\left[\nabla_{x_i}\psi_i(z_{i,t}(k), y^i_{-i}(k); \xi_{i,t}(k)) + \mu_i(z_{i,t}(k) - x_i(k))\right]\Big], \tag{10}$$

where $\gamma_{i,t} = \frac{1}{\mu_i(t+1)}$ and $z_{i,t}(k)$ denotes the estimate of the proximal BR solution $T_i(y^i(k))$ at t-th inner step of the SA scheme (10) with the intinal value $z_{i,1}(k) = x_i(k)$. Set $x_i(k+1) = z_{i,j_i(k)}(k)$. Define $\xi_i(k) \triangleq (\xi_{i,1}(k), \cdots, \xi_{i,j_i(k)}(k))$, and $\mathcal{F}'_k \triangleq \sigma\{x(0), i_l, \xi_{i_l}(l), d_{i_l,j}(l), 0 \le l \le k-1, j \in \mathcal{N}\}$. Then by Algorithm 1, $x(k)$ is adapted to $\mathcal{F}'_k$ and $y^i(k)$ is adapted to $\mathcal{F}_k$. As a result, $T_i(y^i(k))$ is adapted to $\mathcal{F}_k$ by (8). Analogous to Lemma 3 in [42], the following result holds for the SA scheme (10).

Lemma 1 *Let Assumption 1 hold. Consider the asynchronous inexact proximal best-response scheme given by Algorithm 1. Assume that the random variables $\{\xi_{i,t}(k)\}_{1\le t\le j_i(k)}$ are i.i.d., and that for any $i \in \mathcal{N}$ the random vector $\xi_i(k)$ is independent of $\mathcal{F}_k$. Then we have the following for any $t : 1 \le t \le j_i(k)$.*

$$\mathbb{E}[\|z_{i,t}(k) - T_i(y^i_k)\|^2 | \mathcal{F}_k] \le Q_i/(t+1) \quad a.s.,$$

where $Q_i \triangleq \frac{2M^2}{\mu_i^2} + 2D^2_{X_i}$ and $D_{X_i} \triangleq \sup\{d(x_i, x'_i) : x_i, x'_i \in X_i\}$.

Remark 2 *(i) Let $\mathbf{1}_{[i_k=i]}$ denote the indicator function of the event $i_k = i$, defined as $\mathbf{1}_{[i_k=i]} = 1$ if $i_k = i$, and $= 0$, otherwise. Define $\Gamma_{i,0} \triangleq 1$ and $\Gamma_i(k) \triangleq 1 + \sum_{t=0}^{k-1} \mathbf{1}_{\{i_t=i\}}$ for all $k \geq 1$. Then the computation of $\Gamma_i(k)$ merely uses player i's local information and for every $\omega \in \Omega$, there exists a sufficiently large $\tilde{k}(\omega)$ that is possibly contingent on the sample path ω such that for any $i \in \mathcal{N}$:*

$$\Gamma_i(k) \geq \frac{kp_i}{2} + 1 \quad \forall k \geq \tilde{k}(\omega). \tag{11}$$

The proof can be found in Lemma 7 of [46].
(ii) Set $j_i(k) \triangleq \lfloor \Gamma_i(k)^{2(1+\delta)} \rfloor$ for some positive $\delta > 0$ and $x_i(k+1) = z_{i,j_i(k)}(k)$. Then by Lemma 1, we have that $\mathbb{E}[\|x_i(k+1) - T_i(x(k))\|^2 | \mathcal{F}_k] \leq \frac{Q_i}{j_i(k)+1} \leq \frac{Q_i}{\Gamma_i(k)^{2(1+\delta)}} \triangleq \alpha_i(k)^2$. Thus, $\lfloor \Gamma_i(k)^{2(1+\delta)} \rfloor$ steps of (10) suffice for obtaining a solution to (9) with $\varepsilon_i(k)$ satisfying $\mathbb{E}\left[\|\varepsilon_i(k+1)\|^2 \big| \mathcal{F}_k\right] \leq \alpha_i(k)^2$ a.s.. Then by the conditional Jensen's inequality, $\mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] \leq \alpha_i(k)$ a.s.. By invoking (11), we obtain that for all $i \in \mathcal{N}$, $\sum_{k=1}^{\infty} \alpha_{i,k}^2 < \infty$ a.s., and $\sum_{k=1}^{\infty} \alpha_{i,k} < \infty$ a.s.. Then Assumption 3(b) holds. □

The following result from Eqn. (18) in [59] establishes an equivalence between Nash equilibria of the stochastic Nash game (1) and solutions to the variational inequality problem (12).

Lemma 2 *Let Assumptions 1(a), 1(b), and 2 hold. Then x^* is an NE of the potential game (1) if and only if x^* is a solution to the following problem:*

$$\nabla_x P(x^*)^T (y - x^*) \geq 0 \quad \forall y \in X. \tag{12}$$

Further, the set of Nash equilibria is nonempty and compact.

2.3 Convergence Analysis

We now establish the almost sure convergence and convergence in mean of the iterates produced by Alg. 1. Parts of the proof are inspired by Theorem 4.1 in [39] and Theorem 4.3 in [14].

Theorem 1 (almost sure convergence to Nash equilibrium) *Let $\{x(k)\}$ be generated by Algorithm 1. Suppose Assumptions 1, 2 and 3 hold. We further assume that for every $i \in \mathcal{N}$, the parameter μ_i utilized in (8) satisfies $\mu_i > \frac{L_i}{2} + \frac{\sqrt{2}\tau L_i}{2}(\frac{L_i}{L_{\rm ave}} + \frac{L_{\rm ave}}{L_i})$, where $L_{\rm ave} \triangleq \sum_{i\in\mathcal{N}} L_i/N$. Then the following hold:*
(a) **(square summability):** *For any $i \in \mathcal{N}$, $\sum_{k=0}^{\infty} \|T_i(y^i(k)) - x(k)\|^2 < \infty \quad a.s.$.*
(b) **(cluster point is an NE):** *For almost all $\omega \in \Omega$, every limit point of $x(k,\omega)$ is a Nash equilibrium.*
(c) **(almost sure convergence to a connected subset of the set X^* of Nash equilibria):** *There exists a connected subset $X_c^* \subset X^*$ such that $d(x(k), X_c^*) \xrightarrow[k\to\infty]{} 0$ a.s..*

Proof. By Assumption 1(b), we have the following bound:

$$f_i\left(T_i(y^i(k)), x_{-i}(k)\right) \leq f_i(x(k)) + \nabla_{x_i} f_i(x(k))^T \left(T_i(y^i(k)) - x_i(k)\right) + \frac{L_i}{2}\left\|T_i(y^i(k)) - x_i(k)\right\|^2. \tag{13}$$

Since $T_i(y^i(k))$ is a global minimum of (8) and $x_i(k) \in X_i$, by the optimality condition we have that

$$
\begin{aligned}
0 &\le \big(\nabla_{x_i} f_i(T_i(y^i(k)), y^i_{-i}(k)) + \mu_i(T_i(y^i(k)) - x_i(k))\big)^T \big(x_i(k) - T_i(y^i(k))\big) \\
&= -\big(T_i(y^i(k)) - x_i(k)\big)^T \nabla_{x_i} f_i\left(T_i(y^i(k)), y^i_{-i}(k)\right) - \mu_i \|T_i(y^i(k)) - x_i(k)\|^2 \\
&= -\big(T_i(y^i(k)) - x_i(k)\big)^T \nabla_{x_i} f_i(x_i(k), y^i_{-i}(k)) - \mu_i \|T_i(y^i(k)) - x_i(k)\|^2 \\
&\quad - \big(\nabla_{x_i} f_i(T_i(y^i(k)), y^i_{-i}(k)) - \nabla_{x_i} f_i(x_i(k), y^i_{-i}(k))\big)^T \big(T_i(y^i(k)) - x_i(k)\big) \\
&\le -\nabla^T_{x_i} f_i(y^i(k)) \big(T_i(y^i(k)) - x_i(k)\big) - \mu_i \|T_i(y^i(k)) - x_i(k)\|^2,
\end{aligned} \tag{14}
$$

where the last inequality follows by $\big(\nabla_{x_i} f_i(x_i, x_{-i}) - \nabla_{x_i} f_i(x_i', x_{-i})\big)^T (x_i - x_i') \ge 0 \; \forall x_i, x_i' \in X_i, \; \forall x_{-i} \in X_{-i}$ from Assumption 1(b). Adding terms (13) and (14), we have the following inequality:

$$
\begin{aligned}
f_i\left(T_i(y^i(k)), x_{-i}(k)\right) \le{}& f_i(x(k)) + \big(\nabla_{x_i} f_i(x(k)) - \nabla_{x_i} f_i(y^i(k))\big)^T \big(T_i(y^i(k)) - x_i(k)\big) \\
&- (\mu_i - L_i/2) \|T_i(y^i(k)) - x_i(k)\|^2.
\end{aligned} \tag{15}
$$

By Assumption 1(b) and $ab \le \frac{a^2+b^2}{2}$, we obtain the following sequences of inequalities for any $C_i > 0$:

$$
\begin{aligned}
&\big(\nabla_{x_i} f_i(x(k)) - \nabla_{x_i} f_i(y^i(k))\big)^T \big(T_i(y^i(k)) - x_i(k)\big) \le \big\|\nabla_{x_i} f_i(x(k)) - \nabla_{x_i} f_i(y^i(k))\big\| \, \big\|T_i(y^i(k)) - x_i(k)\big\| \\
&\le L_i \|x(k) - y^i(k)\| \|T_i(y^i(k)) - x_i(k)\| \le \frac{L_i^2}{2C_i} \|x(k) - y^i(k)\|^2 + \frac{C_i}{2} \|T_i(y^i(k)) - x_i(k)\|^2.
\end{aligned} \tag{16}
$$

Since $y^i(k) \triangleq (x_1(k - d_{i1}(k)), \cdots, x_N(k - d_{iN}(k)))$ and $d_{ij}(k) \in \{0, 1, \cdots, \tau\}$, the following holds

$$
\begin{aligned}
\|x(k) - y^i(k)\|^2 &= \sum_{j=1}^N \|x_j(k) - x_j(k - d_{ij}(k))\|^2 = \sum_{j=1}^N \Big\| \sum_{h=k-d_{ij}(k)+1}^{k} (x_j(h) - x_j(h-1)) \Big\|^2 \\
&= \sum_{j=1}^N d_{ij}(k)^2 \Big\| \frac{1}{d_{ij}(k)} \sum_{h=k-d_{ij}(k)+1}^{k} (x_j(h) - x_j(h-1)) \Big\|^2 \\
&\le \sum_{j=1}^N d_{ij}(k)^2 \Big(\frac{1}{d_{ij}(k)} \sum_{h=k-d_{ij}(k)+1}^{k} \|x_j(h) - x_j(h-1)\|^2 \Big) \quad \text{(by Jensen's inequality)} \\
&\le \tau \sum_{j=1}^N \sum_{h=k-\tau+1}^{k} \|x_j(h) - x_j(h-1)\|^2 = \tau \sum_{h=k-\tau+1}^{k} \|x(h) - x(h-1)\|^2 \\
&= \underbrace{\tau \sum_{h=k-\tau+1}^{k} (h - k + \tau) \|x(h) - x(h-1)\|^2}_{\triangleq V_k} + \tau^2 \|x(k+1) - x(k)\|^2 \\
&\quad - \tau \sum_{h=k-\tau+2}^{k+1} \big(h - (k+1) + \tau\big) \|x(h) - x(h-1)\|^2.
\end{aligned} \tag{17}
$$

Suppose $C_i = \sqrt{2} L_i^2 \tau / L_{\text{ave}}$. Then by substituting (17) into (16), and by invoking (15), we obtain that

$$
\begin{aligned}
f_i\left(T_i(y^i(k)), x_{-i}(k)\right) \le{}& f_i(x(k)) + \frac{L_{\text{ave}}}{2\sqrt{2}} \big(V_k - V_{k+1} + \tau \|x(k+1) - x(k)\|^2\big) \\
&- \left(\mu_i - \frac{L_i + \sqrt{2} L_i^2 \tau / L_{\text{ave}}}{2}\right) \|T_i(y^i(k)) - x_i(k)\|^2.
\end{aligned} \tag{18}
$$

By Algorithm 1, we have that at time instance k,

$$\|x(k+1)-x(k)\|^2 = \|x_{i_k}(k+1)-x_{i_k}(k)\|^2 \le 2\|T_{i_k}(y^{i_k}(k)) - x_{i_k}(k)\|^2 + 2\|\varepsilon_{i_k}(k+1)\|^2. \tag{19}$$

By employing Assumptions 1(c), 1(d), and the Jensen's inequality, the following holds for any $x \in X$.

$$\begin{aligned}\|\nabla_{x_i} f_i(x_i, x_{-i})\| &= \|\mathbb{E}[\nabla_{x_i}\psi_i(x_i, x_{-i};\xi)]\| \le \mathbb{E}[\|\nabla_{x_i}\psi_i(x_i, x_{-i};\xi)\|] \\ &\le \sqrt{\mathbb{E}[\|\nabla_{x_i}\psi_i(x_i, x_{-i};\xi)\|^2]} \le M.\end{aligned} \tag{20}$$

Then by Algorithm 1 and by invoking Assumption 2, we may obtain the following bound:

$$\begin{aligned}&P(x(k+1)) - P(x(k)) = P(x_{i_k}(k+1), x_{-i_k}(k)) - P(x_{i_k}(k), x_{-i_k}(k)) \\ &= f_{i_k}(x_{i_k}(k+1), x_{-i_k}(k)) - f_{i_k}(x_{i_k}(k), x_{-i_k}(k)) \\ &= f_{i_k}\big(T_{i_k}(y^{i_k}(k)), x_{-i_k}(k)\big) - f_{i_k}(x(k)) + f_{i_k}(x_{i_k}(k+1), x_{-i_k}(k)) - f_{i_k}\big(T_{i_k}(y^{i_k}(k)), x_{-i_k}(k)\big) \\ &= f_{i_k}\big(T_{i_k}(y^{i_k}(k)), x_{-i_k}(k)\big) - f_{i_k}(x(k)) + \varepsilon_{i_k}(k+1)^T \nabla_{x_{i_k}} f_{i_k}(z_{i_k}(k+1), x_{-i_k}(k)) \quad \text{(by mean−value theorem)} \\ &\le f_{i_k}\big(T_{i_k}(y^{i_k}(k)), x_{-i_k}(k)\big) - f_{i_k}(x(k)) + M\|\varepsilon_{i_k}(k+1)\|, \quad \text{(by Cauchy−Schwarz inequality and (20))},\end{aligned} \tag{21}$$

where $z_{i_k}(k+1) = \vartheta_{i_k,k} x_{i_k}(k+1) + (1-\vartheta_{i_k,k}) T_{i_k}(y^{i_k}(k))$ for some $\vartheta_{i_k,k} \in (0,1)$. Therefore, by combining (18), (21), and (19), we have the following inequality:

$$\begin{aligned}P(x(k+1)) \le P(x(k)) &+ \frac{L_{\text{ave}}}{2\sqrt{2}}(V_k - V_{k+1}) + M\|\varepsilon_{i_k}(k+1)\| + \frac{L_{\text{ave}}\tau}{\sqrt{2}}\|\varepsilon_{i_k}(k+1)\|^2 \\ &- \Big(\mu_{i_k} - \frac{L_{i_k} + \sqrt{2}(L_{i_k}^2/L_{\text{ave}} + L_{\text{ave}})\tau}{2}\Big)\|T_{i_k}(y^{i_k}(k)) - x_{i_k}(k)\|^2\end{aligned} \tag{22}$$

Therefore, by rearranging the terms of (22) and taking expectations conditioned on $\mathcal{F}_k$, we obtain that

$$\begin{aligned}&\mathbb{E}\left[P(x(k+1)) + \frac{L_{\text{ave}}}{2\sqrt{2}}V_{k+1}\big|\mathcal{F}_k\right] \le \mathbb{E}\left[P(x(k)) + \frac{L_{\text{ave}}}{2\sqrt{2}}V_k\big|\mathcal{F}_k\right] + \frac{L_{\text{ave}}\tau}{\sqrt{2}}\sum_{i=1}^N \mathbb{E}\left[\|\varepsilon_i(k+1)\|^2\big|\mathcal{F}_k\right] \\ &+ M\sum_{i=1}^N \mathbb{E}\left[\|\varepsilon_i(k+1)\|\big|\mathcal{F}_k\right] - \mathbb{E}\Big[\Big(\mu_{i_k} - \frac{L_{i_k} + \sqrt{2}(L_{i_k}^2/L_{\text{ave}} + L_{\text{ave}})\tau}{2}\Big)\|T_{i_k}(y^{i_k}(k)) - x_{i_k}(k)\|^2\big|\mathcal{F}_k\Big].\end{aligned} \tag{23}$$

Since $T_i(y^i(k))\ \forall i \in \mathcal{N}$ is adapted to $\mathcal{F}_k$, and i_k is independent of $\mathcal{F}_k$, by Corollary 7.1.2 in [60][1] and $\mathbb{P}(i_k = i) = p_i$, the last term on the right-hand side of (23) is equivalent to

$$\begin{aligned}&\mathbb{E}_{i_k}\Big[\Big(\mu_{i_k} - \frac{L_{i_k} + \sqrt{2}(L_{i_k}^2/L_{\text{ave}} + L_{\text{ave}})\tau}{2}\Big)\|T_{i_k}(y^{i_k}(k)) - x_{i_k}(k)\|^2\Big] \\ &= \sum_{i=1}^N p_i\Big(\mu_i - \frac{L_i + \sqrt{2}(L_i^2/L_{\text{ave}} + L_{\text{ave}})\tau}{2}\Big)\|T_i(y^i(k)) - x_i(k)\|^2.\end{aligned} \tag{24}$$

[1]Let the random vectors $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ on $(\Omega, \mathcal{F}, \mathbb{P})$ be independent of one another and let f be a Borel function on $\mathbb{R}^{m\times n}$ with $|\mathbb{E}[f(X,Y)]| \le \infty$. If for any $x \in \mathbb{R}^m$, $g(x) = \begin{cases}\mathbb{E}[f(x,Y)] & \text{if } |\mathbb{E}[f(x,Y)]| \le \infty \\ 0 & \text{otherwise}\end{cases}$, then g is a Borel function with $g(X) = \mathbb{E}[f(X,Y)|\sigma(X)]$.

Since $x(k)$ and V_k are adapted to $\mathcal{F}_k$, by (23) and (24), we have the following:

$$\begin{aligned}&\mathbb{E}\left[P(x(k+1))+\frac{L_{\text{ave}}}{2\sqrt{2}}V_{k+1}\big|\mathcal{F}_k\right]\le\mathbb{E}\left[P(x(k))+\frac{L_{\text{ave}}}{2\sqrt{2}}V_k\big|\mathcal{F}_k\right]+M\sum_{i=1}^N\mathbb{E}\left[\|\varepsilon_i(k+1)\|\big|\mathcal{F}_k\right]\\&+\frac{L_{\text{ave}}\tau}{\sqrt{2}}\sum_{i=1}^N\left[\|\varepsilon_i(k+1)\|^2\big|\mathcal{F}_k\right]-\sum_{i=1}^N p_i\left(\mu_i-\frac{L_i+\sqrt{2}(L_i^2/L_{\text{ave}}+L_{\text{ave}})\tau}{2}\right)\|T_i(y^i(k))-x_i(k)\|^2.\end{aligned}\tag{25}$$

(a) By using $\mu_i>\frac{L_i}{2}+\frac{\sqrt{2}(L_i^2/L_{\text{ave}}+L_{\text{ave}})\tau}{2}$ and Assumption 3(b), we may then invoke Theorem 1 in [61], and conclude that for every $i\in\mathcal{N}$, $\sum\limits_{k=0}^{\infty}\|T_i(y^i(k))-x_i(k)\|^2<\infty,\ a.s.$.

(b) By result (a), we have the following for any $i\in\mathcal{N}$:

$$\lim_{k\to\infty}\|T_i\left(y^i(k)\right)-x_i(k)\|=0,\quad a.s.\quad .\tag{26}$$

Let $\bar{x}(\omega)$ be a cluster point of sequence $\{x(k,\omega)\}$. Then there exists a subsequence $\mathcal{K}(\omega)$ such that

$$\lim_{k\to\infty,k\in\mathcal{K}(\omega)}x(k,\omega)=\bar{x}(\omega).\tag{27}$$

Then by (26) and (27), we have that

$$\lim_{k\to\infty,k\in\mathcal{K}(\omega)}T_i(y^i(k,\omega))=\bar{x}_i(\omega),\quad\forall i\in\mathcal{N}.\tag{28}$$

We intend to show that $\bar{x}(\omega)$ is a Nash equilibrium. We proceed by contradication. Then there exists an $i\in\mathcal{N}$ and a vector $\bar{y}_i\in X_i$ such that $f_i(\bar{y}_i,\bar{x}_{-i}(\omega))<f_i\left(\bar{x}_i(\omega),\bar{x}_{-i}(\omega)\right)$. By definition, the directional derivative of f_i at point $(\bar{x}_i(\omega),\bar{x}_{-i}(\omega))$ with respect to x_i along the vector $q_i=\bar{y}_i-\bar{x}_i(\omega)$, denoted by $f_i'(\bar{x}_i(\omega),\bar{x}_{-i}(\omega);q_i)$, satisfies the following:

$$\begin{aligned}f_i'(\bar{x}_i(\omega),\bar{x}_{-i}(\omega);q_i)&\triangleq\inf_{\lambda>0}\frac{f_i(\bar{x}_i(\omega)+\lambda q_i,\bar{x}_{-i}(\omega))-f_i(\bar{x}_i(\omega),\bar{x}_{-i}(\omega))}{\lambda}\\&\le f_i(\bar{x}_i(\omega)+q_i,\bar{x}_{-i}(\omega))-f_i(\bar{x}_i(\omega),\bar{x}_{-i}(\omega))\quad\text{(by setting }\lambda\text{=1)}\\&=f_i(\bar{y}_i,\bar{x}_{-i}(\omega))-f_i(\bar{x}_i(\omega),\bar{x}_{-i}(\omega))<0.\end{aligned}$$

Since f_i is differentiable at x with respect to x_i by Assumption 1(b), the following holds

$$0>f_i'(\bar{x}_i(\omega),\bar{x}_{-i}(\omega);q_i)=(\bar{y}_i-\bar{x}_i(\omega))^T\nabla_{x_i}f_i(\bar{x}_i(\omega),\bar{x}_{-i}(\omega)).\tag{29}$$

Recall that $T_i(y_k^i)$ is defined as a global minimum of a convex optimization problem (8). Since $\bar{y}_i\in X_i$, by the optimality condition for a constrained convex programming, we obtain that

$$\mu_i\left(\bar{y}_i-T_i(y^i(k))\right)^T\left(T_i(y^i(k))-x_i(k)\right)+\left(\bar{y}_i-T_i(y^i(k))\right)^T\nabla_{x_i}f_i(T_i(y^i(k)),x_{-i}(k))\ge0.\tag{30}$$

Since $\nabla_{x_i}f_i(x)$ is continuous in x by Assumption 1(b), by taking limits to $k\to\infty,k\in\mathcal{K}(\omega)$, and using (26)-(28), we obtain that $(\bar{y}_i-\bar{x}_i(\omega))^T\nabla_{x_i}f_i(\bar{x}_i(\omega),\bar{x}_{-i}(\omega))\ge0$, which contradicts (29). Thus, $\bar{x}(\omega)$ is an NE, proving (b).

(c) We first validate that $\lim\limits_{k\to\infty}d(x(k),X^*)=0\ a.s.$. Assume false; then $\limsup\limits_{k\to\infty}d(x(k,\omega),X^*)>0$ with some positive probability, i.e. for all $\omega\in\widehat{\Omega}\subset\Omega$ where $\mathbb{P}(\widehat{\Omega})>0$. Then for $\omega\in\widehat{\Omega}$, by the boundedness of

$\{x(k,\omega)\}_{k\geq 0}$ we can extract a convergent subsequence $\{x(k,\omega)\}_{k\in\mathcal{K}(\omega)}$ such that $\lim_{k\in\mathcal{K}(\omega),k\to\infty} d(x(k,\omega),X^*) > 0$. This contradicts result (b). Hence $\lim_{k\to\infty} d(x(k),X^*) = 0\ a.s.$.

We now proceed to show a.s. convergence to a connected subset of X^*, denoted by X_c^*, through a contradiction argument. We assume the contrary that X_c^* is disconnected with some positive probability. Then there exist at least two closed connected sets $X_{c_1}^*$ and $X_{c_2}^*$ such that $X_c^* = X_{c_1}^* \cup X_{c_2}^*$ with $d(X_{c_1}^*, X_{c_2}^*) > 0$. By hypothesis, the sequence $\{x(k)\}$ cannot converge to either $X_{c_1}^*$ or $X_{c_2}^*$ a.s., and $x(k)$ visits $X_{c_1}^*, X_{c_2}^*$ infinitely often. Define $\rho \triangleq \frac{1}{3}d(X_{c_1}^*, X_{c_2}^*)$. By $d(x(k),X_c^*) \xrightarrow[k\to\infty]{} 0$, we know there exists k_0 such that

$$x(k) \in B(X_{c_1}^*,\rho) \cup B(X_{c_2}^*,\rho)\ \ \forall k \geq k_0, \tag{31}$$

where $B(A,\rho)$ denotes the ρ-neighborhood of A. Define n_0 and n_p, m_p for $p\geq 1$ as follows:

$$\begin{aligned} &n_0 \triangleq \inf\{k > k_0, d(x(k),X_{c_1}^*) < \rho\}, m_p \triangleq \inf\{k > n_{p-1}, d(x(k),X_{c_2}^*) < \rho\},\\ \text{and } &n_p \triangleq \inf\{k > m_p, d(x(k),X_{c_1}^*) < \rho\},\ \ p \geq 1. \end{aligned}$$

Then $\{n_p\}$ and $\{m_p\}$ are infinite sequences by the converse of result (c). By (31), we have $x(n_p) \in B(X_{c_1}^*,\rho)$ and $x(n_p - 1) \in B(X_{c_2}^*,\rho)$ for any $p \geq 1$. Then by $d(X_{c_1}^*, X_{c_2}^*) = 3\rho$, it follows that $\|x(n_p) - x(n_p - 1)\| > \rho\ \forall p \geq 1$ with some positive probability. Then we have the following:

$$\mathbb{E}[\|x(n_p) - x(n_p-1)\|] > 0, \tag{32}$$

where the unconditional expectation is taken w.r.t. the information of ξ and random delays up to time n_p. Hereafter, the unconditional expectation of a variable is taken w.r.t. to all historical random effects.

Also, by taking unconditional expectations on both sides of (25), we have the following:

$$\begin{aligned} &\mathbb{E}\left[P(x(k+1)) + \frac{L_{\text{ave}}}{2\sqrt{2}}V_{k+1}\right] \leq \mathbb{E}\left[P(x(k)) + \frac{L_{\text{ave}}}{2\sqrt{2}}V_k\right] + M\sum_{i=1}^N \mathbb{E}\left[\|\varepsilon_i(k+1)\|\right] \\ &+ \frac{L_{\text{ave}}\tau}{\sqrt{2}}\sum_{i=1}^N \mathbb{E}\left[\|\varepsilon_i(k+1)\|^2\right] - \sum_{i=1}^N p_i\Big(\mu_i - \frac{L_i + \sqrt{2}(L_i^2/L_{\text{ave}} + L_{\text{ave}})\tau}{2}\Big)\mathbb{E}[\|T_i(y^i(k)) - x_i(k)\|^2]. \end{aligned} \tag{33}$$

By using Assumption 3(b), we have that

$$\begin{aligned} &\sum_{k=1}^\infty \mathbb{E}[\|\varepsilon_i(k+1)\|] = \mathbb{E}\Big[\sum_{k=1}^\infty \mathbb{E}\left[\|\varepsilon_i(k+1)\|\big|\mathcal{F}_k\right]\Big] < \infty, \text{ and} \\ &\sum_{k=1}^\infty \mathbb{E}[\|\varepsilon_i(k+1)\|^2] = \mathbb{E}\Big[\sum_{k=1}^\infty \mathbb{E}\big[\|\varepsilon_i(k+1)\|^2\big|\mathcal{F}_k\big] < \infty. \end{aligned} \tag{34}$$

Then from (33) it follows that

$$\begin{aligned} &\sum_{k=0}^\infty\sum_{i=1}^N p_i\Big(\mu_i - \frac{L_i + \sqrt{2}(L_i^2/L_{\text{ave}} + L_{\text{ave}})\tau}{2}\Big)\mathbb{E}[\|T_i(y^i(k)) - x_i(k)\|^2] \leq \mathbb{E}\left[P(x_0) + \frac{L_{\text{ave}}}{2\sqrt{2}}V_0\right] \\ &- \liminf_{k\to\infty}\mathbb{E}\left[P(x(k)) + \frac{L_{\text{ave}}}{2\sqrt{2}}V_k\right] + M\sum_{i=1}^N\sum_{k=0}^\infty \mathbb{E}\left[\|\varepsilon_i(k+1)\|\right] + \frac{L_{\text{ave}}\tau}{\sqrt{2}}\sum_{i=1}^N\sum_{k=0}^\infty \mathbb{E}\left[\|\varepsilon_i(k+1)\|^2\right] < \infty, \end{aligned}$$

where the second inequality holds by the boundedness of $P(x(k))+V_k$ since $P(\cdot)$ is continuous and X is compact. Hence by using $\mu_i > \frac{L_i}{2} + \frac{\sqrt{2}(L_i^2/L_{\rm ave}+L_{\rm ave})\tau}{2}$ and Jensen's inequality, we have that

$$\lim_{k\to\infty} \mathbb{E}\left[\|T_i(y^i(k)) - x_i(k)\|\right] = 0 \;\; \forall i \in \mathcal{N}. \tag{35}$$

Note that $\|x(k+1)-x(k)\| \leq \sum_{i=1}^N (\|\varepsilon_i(k+1)\| + \|T_i(y^i(k)) - x_i(k)\|)$. Then by (34),we have that

$$\lim_{k\to\infty} \mathbb{E}[\|x(k+1)-x(k)\|] = 0. \tag{36}$$

This contradicts (32), and hence the converse assumption does not hold. Then result (c) is proved. □

Theorem 1 shows that the estimates generated by Algorithm 1 converge almost surely to the set of Nash equilibria. If the set of Nash equilibria contains isolated points, then for almost all $\omega \in \Omega$, $x(k,\omega)$ converges to an NE. Further, if the potential game (1) admits a unique NE, then the iterates converge almost surely to the NE. It is also worth emphasizing that while we use the term "best-response", since μ_i has to be sufficiently large for all i, this can be seen to be more akin to "better-response." In what follows, we discuss the convergence in mean of the iterates. Since the potential function is employed as a vehicle to analyze convergence of the iterates, a natural approach would have to employ the value of the potential function. However, the iterates may converge to stationary points which are not necessarily global minimizers and as a consequence, we need to select an appropriate metric to capture stationarity.

Note from Lemma 2 that a stationary point of $\min_{x\in X} P(x)$ is given by a solution to the variational inequality problem VI$(X, \nabla_x P)$ that requires an $x \in X$ such that $(y-x)^T \nabla_x P(x) \geq 0 \; \forall y \in X$. Suppose X^* denotes the set of solutions to VI(X,F). A merit function for ascertaining the departure from solvability of the VI is a gap function. It may be recalled from [62] that a function $G(x)$ is called a gap function if it satisfies two properties: (i) $G(\cdot)$ is sign restricted over the set X; (ii) $G(x) = 0$ if and only if x solves VI(X,F). We consider a primal gap function [62, Theorem 3.1] that has found a fair amount of applicability in the context of variational inequality problems.

Definition 1 *Let $X \subseteq \mathbb{R}^n$ be a nonempty, closed, and convex set. Let $F : X \to \mathbb{R}^n$ and let $G : X \to \mathbb{R}^+$ be defined by $G(x) \triangleq \sup_{y\in X} F(x)^T(x-y), \quad \forall x \in X$.*

The following result shows the mean convergence in the sense that the limit of $\mathbb{E}[G(x(k))]$ is zero. This is analogous to showing that expected sub-optimality tends to zero in the context of stochastic program.

Theorem 2 (Convergence in mean) *Let $\{x(k)\}$ be generated by Algorithm 1. Suppose Assumptions 1, 2 and 3 hold, and, in addition, that for every $i \in \mathcal{N}$, the parameter μ_i utilized in* (8) *satisfies $\mu_i > \frac{L_i}{2} + \frac{\sqrt{2}(L_i^2/L_{\rm ave}+L_{\rm ave})\tau}{2}$. Then we have that $\lim\limits_{k\to\infty} \mathbb{E}\big[G(x(k))\big] = 0$.*

Proof. By Assumption 2, we have the following for any $i \in \mathcal{N}$ and any $\bar{y}_i \in X_i$:

$$\begin{aligned}
&(x_i(k) - \bar{y}_i)^T \nabla_{x_i} P(x(k)) = (x_i(k) - \bar{y}_i)^T \nabla_{x_i} f_i(x(k)) \\
&= \big(x_i(k) - T_i(y^i(k))\big)^T \nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k)) + \big(T_i(y^i(k)) - \bar{y}_i\big)^T \nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k)) \\
&- (x_i(k) - \bar{y}_i)^T \left(\nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k)) - \nabla_{x_i} f_i(x(k))\right).
\end{aligned}$$

Then we have the following sequence of inequalities for any $i \in \mathcal{N}$ and any $\bar{y}_i \in X_i$:

$$
\begin{aligned}
&(x_i(k)-\bar{y}_i)^T \nabla_{x_i} P(x(k)) \leq \mu_i \left(\bar{y}_i - T_i(y^i(k))\right)^T \left(T_i(y^i(k)) - x_i(k)\right) \\
&+ \left(x_i(k) - T_i(y^i(k))\right)^T \nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k)) \\
&+ (\bar{y}_i - x_i(k))^T \left(\nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k)) - \nabla_{x_i} f_i(x(k))\right) \quad \text{(by (30))} \\
&\leq \mu_i \left\|\bar{y}_i - T_i(y^i(k))\right\| \left\|T_i(y^i(k)) - x_i(k)\right\| + \left\|x_i(k) - T_i(y^i(k))\right\| \|\nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k))\| \\
&+ \|\bar{y}_i - x_i(k)\| \left\|\nabla_{x_i} f_i(T_i(y^i(k)), x_{-i}(k)) - \nabla_{x_i} f_i(x(k))\right\| \quad \text{(by Cauchy−Schwarz inequality)} \\
&\leq (\mu_i D_{X_i} + M + L_i D_{X_i}) \left\|T_i(y^i(k)) - x_i(k)\right\|. \quad \text{(by Assumptions 1(a), 1(b) and (20))}.
\end{aligned}
$$

By summing these inequalities over i, it follows that

$$
\begin{aligned}
G(x(k)) &= \sup_{\bar{y} \in X} (x(k) - \bar{y})^T \nabla P(x(k)) = \sum_{i=1}^N \sup_{\bar{y}_i \in X_i} (x_i(k) - \bar{y}_i)^T \nabla_{x_i} P(x(k)) \\
&\leq \sum_{i=1}^N (\mu_i D_{X_i} + M + L_i D_{X_i}) \left\|T_i(y^i(k)) - x_i(k)\right\|.
\end{aligned} \tag{37}
$$

Then, by taking expectations on both sides of (37), we obtain that

$$
\mathbb{E}\left[G(x(k))\right] \leq \sum_{i=1}^N (\mu_i D_{X_i} + M + L_i D_{X_i}) \, \mathbb{E}\left[\left\|T_i(y^i(k)) - x_i(k)\right\|\right] \implies \lim_{k\to\infty} \mathbb{E}\left[G(x(k))\right] \leq 0 \quad \text{(by (35))}.
$$

However, $G(x(k)) \geq 0$ since $x(k) \in X$, implying the required result that $\lim_{k\to\infty} \mathbb{E}\left[G(x(k))\right] = 0$. □

We now define an alternative proximal gradient-response map as follows for $\mu_i > 0$:

$$
T_i^{\mu_i}(x) \triangleq \operatorname*{argmin}_{y_i \in X_i} \left[(y_i - x_i)^T \nabla_{x_i} f_i(x) + \frac{\mu_i}{2}\|y_i - x_i\|^2\right]. \tag{38}
$$

Since each player's subproblem is convex, definition (38) is equivalent to

$$
T_i^{\mu_i}(x) = \Pi_{X_i}\left[x_i - \frac{1}{\mu_i}\nabla_{x_i} f_i(x)\right]. \tag{39}
$$

Corollary 1 (a.s., mean convergence under proximal gradient-response map) *Consider Algorithm 1 with* (9) *replaced by the following variable sample-size projected gradient-response scheme:*

$$
x_i(k+1) = \Pi_{X_i}\left[x_i(k) - \frac{1}{\mu_i}\Big(\frac{1}{N_i(k)}\sum_{t=1}^{N_i(k)} \nabla_{x_i}\psi_i(x_i(k), y_{-i}^i(k); \xi_{i,t}(k))\Big)\right], \tag{40}
$$

where $N_i(k) \triangleq \lfloor \Gamma_i(k)^{2(1+\delta)} \rfloor$ for some positive $\delta > 0$ with $\Gamma_i(k)$ defined by (2), *and $\xi_{i,1}(k), \cdots, \xi_{i,N_i(k)}(k)$ are $N_i(k)$ realizations of the random vector ξ. Let Assumptions 1, 2, 3(a), and 3(c) hold, and $\mu_i > \frac{L_i}{2} + \frac{\sqrt{2}(L_i^2/L_{\rm ave} + L_{\rm ave})\tau}{2}$ for all $i \in \mathcal{N}$. Additionally, suppose that for every $i \in \mathcal{N}$, the random variables $\{\xi_{i,t}(k)\}_{1 \leq t \leq N_i(k)}$ are independent of $\mathcal{F}_k$. Then the results of Theorem 1 and Theorem 2 still hold.*

Proof. Define $\varepsilon_i(k+1) \triangleq x_i(k+1) - T_i^{\mu_i}(y^i(k))$. Then by using (39), (40), and the nonexpansive property of the projection operator we obtain that

$$
\|\varepsilon_i(k+1)\| \leq \frac{1}{\mu_i N_i(k)} \sum_{t=1}^{N_i(k)} \left\|\nabla_{x_i}\psi_i(y^i(k); \xi_{i,t}(k)) - \nabla_{x_i} f_i(y^i(k))\right\|.
$$

Thus, by Assumption 1(d), we obtain that

$$\mathbb{E}\left[\|\varepsilon_i(k+1)\|^2 \big| \mathcal{F}_k\right] \leq \frac{1}{\mu_i^2 N_i(k)} \mathbb{E}_\xi[\|\nabla_{x_i}\psi_i(y^i(k);\xi) - \nabla_{x_i} f_i(y^i(k))\|^2] \leq \frac{M^2}{\mu_i^2 N_i(k)},$$

and hence by the conditional Jensen's inequality, $\mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] \leq \frac{M}{\mu_i \sqrt{N_i(k)}}$. By noting that $N_i(k) \triangleq \left\lfloor \Gamma_i(k)^{2(1+\delta)} \right\rfloor$ and Remark 2(i), Assumption 3(b) holds. Since $T_i^{\mu_i}(y^i(k))$ is a global minimum of (38) and $x_i(k) \in X_i$, by the optimality condition, it follows that

$$0 \leq -\nabla_{x_i}^T f_i(y^i(k)) \left(T_i^{\mu_i}(y^i(k)) - x_i(k)\right) - \mu_i \|T_i^{\mu_i}(y^i(k)) - x_i(k)\|^2 \tag{41}$$

which is indeed the last inequality in Equation (14). Then by inequality (41), similar to the proof of Theorem 1 and Theorem 2, we conclude the result. $\square$

2.4 Generalized Potential games and Weighted Potential Games

We now consider the generalized Nash setting where the strategy sets are coupled in Section 2.4.1, and in Section 2.4.2 we consider the weighted potential game, a generalization of standard potential games.

2.4.1 Generalized potential Nash games

We now extend the separable constraint to the shared constraint regime, a special case of coupled constraints. Suppose there exists a nonempty closed set $C \in \mathbb{R}^n$ such that player i's feasible set $X_i(x_{-i}) = \{x_i \in X_i : (x_i, x_{-i}) \in C\}$ depends on the rivals' strategies x_{-i}, where $X_i \in \mathbb{R}^{n_i}$ are nonempty closed sets such that $\prod_{i=1}^N X_i \cap C$ is nonempty. We say that a point $x \in \mathbb{R}^n$ is feasible if $x_i \in X_i(x_{-i})$ for any $i \in \mathcal{N}$. The aim of player i is to choose a strategy x_i that solves the following parameterized stochastic program:

$$\min_{x_i \in X_i(x_{-i})} \quad f_i(x_i, x_{-i}) \triangleq \mathbb{E}_\xi\left[\psi_i(x_i, x_{-i}; \xi(\omega))\right]. \tag{42}$$

Assume that there exists a continuous potential function $P : C \bigcap \prod_{i=1}^N X_i \to \mathbb{R}$ such that for any $i \in \mathcal{N}$ and any x_{-i} with $X_i(x_{-i})$ being nonempty, we have the following equality:

$$P(x_i, x_{-i}) - P(x_i', x_{-i}) = f_i(x_i, x_{-i}) - f_i(x_i', x_{-i}), \;\; \forall x_i, x_i' \in X_i(x_{-i}). \tag{43}$$

Then the problem (42) is called a generalized potential stochastic game where the (generalized) NE x^* is a feasible point such that the following holds for any i:

$$f_i(x_i^*, x_{-i}^*) \leq f_i(x_i, x_{-i}^*) \quad \forall x_i \in X_i(x_{-i}^*).$$

Suppose that for any $i \in \mathcal{N}$, (i) $f_i(x)$ is continuously differentiable on C, (ii) the feasible set $X_i(x_{-i})$ is inner-semicontinous relative to the set of points x_{-i} for which $X_i(x_{-i})$ is nonempty [14, A2], (iii) $X_i(x_{-i})$ is convex for all x_{-i} for which $X_i(x_{-i})$ is nonempty, and $f_i(\cdot, x_{-i})$ is convex in $x_i \in X_i(x_{-i})$ for all x_{-i} for which $X_i(x_{-i})$ is nonempty. Then computing the proximal best-response solution $T_i(x)$ of (42) involves solving a strongly convex stochastic program. Let Algorithm 1 with $d_{ij}(k) = 0 \; \forall i, j \in \mathcal{N}$ be applied to the generalized stochastic potential game with feasible initial point x_0, where it is required that $x_i(k+1) \in X_i(x_{-i}(k))$, which can be guaranteed by using the projected stochastic gradient scheme to obtain the approximate proximal BR solutions. Similar to Lemma 4.1 in [14], we may show that $x(k+1)$ is feasible. We then conclude the following result, for which the proof is similar to that of Theorems 1 and 2.

Corollary 2 (Generalized potential stochastic Nash games) *Let Algorithm 1 be applied to the stochastic generalized Nash game* (42) *satisfying* (43)*, where x_0 is feasible, $y^i(k) = x(k)$, and it is required that $x_i(k+1) \in X_i(x_{-i}(k))$. Then the results of Theorems 1 and 2 still hold under suitable conditions.*

Nevertheless, we may be unable to manage the delayed regime since we cannot guarantee that the delay-afflicted rival strategies, denoted by $y^i_{-i}(k)$, allow for retaining feasibility; namely, for some i, k, the set $X_i(y^i_{-i}(k))$ may be empty. Consequently, player i may be unable to find a feasible strategy, given rival strategies, and hence Algorithm 1 is not well-defined in that step (S.2) cannot be implemented.

2.4.2 Weighted potential games

We now consider the weighted potential game, in which there exist positive numbers $w_1, \cdots, w_N$ such that, for any $i \in \mathcal{N}$ and any x_{-i} the following equality holds:

$$P(x_i, x_{-i}) - P(x_i', x_{-i}) = w_i \left(f_i(x_i, x_{-i}) - f_i(x_i', x_{-i}) \right) \ \ \forall x_i, x_i' \in X_i(x_{-i}). \tag{44}$$

Then by applying Algorithm 1, we obtain the following results.

Corollary 3 (Weighted potential stochastic Nash games) *Let Algorithm 1 be applied to the stochastic Nash game* (1)*, in which the objective function satisfies* (44)*. Suppose Assumptions 1, 2 and 3 hold. Let the parameter μ_i used in* (8) *satisfy $\mu_i > \frac{L_i}{2} + \frac{\sqrt{2}\tau L_i}{2}(\frac{L_{\text{ave}}}{L_i}\frac{w_{\text{ave}}}{w_i} + \frac{L_i}{L_{\text{ave}}}\frac{w_i}{w_{\text{ave}}})$, where $w_{\text{ave}} = \sum_{i\in\mathcal{N}} w_i/N$. Then the results of Theorem 1 and Theorem 2 hold.*

Proof. By multiplying both sides of (15) with w_i and rearranging the terms we obtain that

$$\begin{aligned} & w_i\Big(f_i\left(T_i(y^i(k)), x_{-i}(k)\right) - f_i(x(k))\Big) \\ & \leq w_i \left(\nabla_{x_i} f_i(x(k)) - \nabla_{x_i} f_i(y^i(k))\right)^T \left(T_i(y^i(k)) - x_i(k)\right) - w_i(\mu_i - L_i/2)\|T_i(y^i(k)) - x_i(k)\|^2. \end{aligned} \tag{45}$$

By substituting (16) and (17) into (45), the following holds for $C_i = \frac{\sqrt{2}L_i^2 w_i \tau}{L_{\text{ave}} w_{\text{ave}}}$:

$$\begin{aligned} w_i\Big(f_i\left(T_i(y^i(k)), x_{-i}(k)\right) - f_i(x(k))\Big) &\leq \frac{w_i L_i^2 \tau}{2C_i}(V_k - V_{k+1} + \tau\|x(k+1) - x(k)\|^2) \\ - w_i(\mu_i - L_i/2)\|T_i(y^i(k)) - x_i(k)\|^2 &+ \frac{w_i C_i}{2}\|T_i(y^i(k)) - x_i(k)\|^2. \end{aligned} \tag{46}$$

Then by (44), similar to (21) we have that

$$P(x(k+1)) - P(x(k)) \leq w_{i_k}\Big(f_{i_k}\left(T_{i_k}(y^{i_k}(k)), x_{-i_k}(k)\right) - f_{i_k}(x(k)) + M\|\varepsilon_{i_k}(k+1)\|\Big),$$

which incorporating with (19) and (46) yields the following inequality:

$$\begin{aligned} P(x(k+1)) - P(x(k)) \leq \frac{L_{\text{ave}} w_{\text{ave}}}{2\sqrt{2}}(V_k - V_{k+1}) + \frac{L_{\text{ave}} w_{\text{ave}}\tau}{\sqrt{2}}\|\varepsilon_{i_k}(k+1)\|^2 + M w_{i_k}\|\varepsilon_{i_k}(k+1)\| \\ - w_i\left((\mu_i - L_i/2) - \frac{L_{\text{ave}} w_{\text{ave}}\tau}{\sqrt{2}w_i} - \frac{L_i^2 w_i \tau}{\sqrt{2}L_{\text{ave}} w_{\text{ave}}}\right)\|T_{i_k}(y^{i_k}(k)) - x_{i_k}(k)\|^2. \end{aligned} \tag{47}$$

Since $\mu_i > \frac{L_i}{2} + \frac{\tau L_i}{2}(\frac{L_{\text{ave}} w_{\text{ave}}}{L_i w_i} + \frac{L_i w_i}{L_{\text{ave}} w_{\text{ave}}})$, by using (47), similar to the proof of Theorems 1 and 2, we obtain the results. □

2.5 Asynchronous inexact best-response scheme without delays

We now show that the asynchronous inexact pure best-response scheme (i.e. without a proximal term) is also applicable when the player-specific problem is strongly convex and each player may obtain its rivals' latest strategies without delay. We define best-response map of player i as follows:

$$\widehat{T}_i(x) \triangleq \operatorname*{argmin}_{y_i \in X_i} \left[f_i(y_i, x_{-i}) \right], \tag{48}$$

based on which, analogous to Algorithm 1, we design an asynchronous inexact best-response scheme.

Algorithm 2 Asynchronous inexact best-response scheme

Let $k := 0$, $x_{i,0} \in X_i$ for $i \in \mathcal{N}$. Additionally $0 < p_i < 1$ for $i \in \mathcal{N}$ such that $\sum_{i=1}^N p_i = 1$.

(S.1) Pick $i_k = i \in \mathcal{N}$ with probability p_i.

(S.2) If $i_k = i$, then player i updates $x_i(k+1) \in X_i$ as follows:

$$x_i(k+1) := \widehat{T}_i(x(k)) + \varepsilon_i(k+1),$$

where $\varepsilon_i(k+1)$ denotes the inexactness. Otherwise, $x_j(k+1) := x_j(k)$ if $j \notin i_k$.

(S.3) If $k > K$, stop; Else, $k := k+1$ and return to (S.1).

Theorem 3 *Let $\{x(k)\}$ be generated by Algorithm 2. Suppose Assumptions 1, 2, and 3(b) hold. Assume that for every $i \in \mathcal{N}$, $\sum_{k=1}^\infty \mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] < \infty$ a.s., and $f_i(x_i, x_{-i})$ is μ_i-strongly convex in $x_i \in X_i$ for all $x_{-i} \in X_{-i}$. Then for almost all $\omega \in \Omega$, every limit point of $\{x(k,\omega)\}$ is an NE.*

Proof. Since $\widehat{T}_i(x(k))$ is a global minimum of the problem (48), by $x_i(k) \in X_i$ and the optimality condition, we obtain that $\left(x_i(k) - \widehat{T}_i(x(k))\right)^T \nabla_{x_i} f_i(\widehat{T}_i(x(k)), x_{-i}(k)) \geq 0$. Then by the μ_i-strong convexity of $f_i(x_i, x_{-i})$ in $x_i \in X_i$, we have that

$$\begin{aligned} f_i(x(k)) &\geq f_i\left(\widehat{T}_i(x(k)), x_{-i}(k)\right) + \left(x_i(k) - \widehat{T}_i(x(k))\right)^T \nabla_{x_i} f_i\left(\widehat{T}_i(x(k)), x_{-i}(k)\right) + \frac{\mu_i}{2}\left\|\widehat{T}_i(x(k)) - x_i(k)\right\|^2 \\ &\geq f_i\left(\widehat{T}_i(x(k)), x_{-i}(k)\right) + \frac{\mu_i}{2}\left\|\widehat{T}_i(x(k)) - x_i(k)\right\|^2. \end{aligned}$$

By rearranging the terms and using Assumption 1(b), we have the following:

$$f_i\left(\widehat{T}_i(x(k)), x_{-i}(k)\right) \leq f_i(x(k)) - \frac{\mu_i}{2}\left\|\widehat{T}_i(x(k)) - x_i(k)\right\|^2.$$

Similar to (21), we can also show that

$$\begin{aligned} &P(x(k+1)) - P(x(k)) = f_{i_k}(x_{i_k}(k+1), x_{-i_k}(k)) - f_{i_k}(x_{i_k}(k), x_{-i_k}(k)) \\ &\leq f_{i_k}\left(\widehat{T}_{i_k}(x(k)), x_{-i_k}(k)\right) - f_{i_k}(x(k)) + M\|\varepsilon_{i_k}(k+1)\| \\ &\leq -\frac{\mu_{i_k}}{2}\left\|\widehat{T}_{i_k}(x(k)) - x_{i_k}(k)\right\|^2 + M\|\varepsilon_{i_k}(k+1)\|. \end{aligned}$$

By taking expectations conditioned on $\mathcal{F}_k$, similarly to (25), we conclude that

$$\mathbb{E}\left[P(x(k+1)) \big| \mathcal{F}_k\right] \leq P(x(k)) + M\sum_{i=1}^N \mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] - \sum_{i=1}^N \frac{p_i \mu_i}{2}\left\|\widehat{T}_i(x(k)) - x_i(k)\right\|^2.$$

Since $\sum_{k=1}^{\infty} \mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] < \infty \; a.s.$ for every $i \in \mathcal{N}$, we then use Theorem 1 in [61], and conclude that $\sum_{k=0}^{\infty} \left\|\widehat{T}_i(x(k)) - x_i(k)\right\|^2 < \infty \; a.s.$ for every $i \in \mathcal{N}$. Thus, $\lim_{k\to\infty} \|\widehat{T}_i(x(k)) - x_i(k)\| = 0 \; a.s.$ for every $i \in \mathcal{N}$. The rest of the proof is the same as that of Theorem 1(b). □

Remark 3 *Theorem 3 shows that in delay-free regimes, asynchronous inexact pure BR schemes retain almost sure convergence when player-specific problems are strongly convex. Consequently, in deterministic regimes, the exact BR scheme is convergent when each player's subproblem is strongly convex, complementing the findings from [14] in that the best-response schemes can lead to convergence to Nash equilibria when the player-specific problem is strongly convex rather than merely convex.*

3 Misspecified Potential Stochastic Nash Games

In this section, we consider the misspecified stochastic Nash game **(P2)**. A sequential approach for resolving such a problem relies on first estimating θ^* and subsequently estimating x^* based on the belief regarding θ^*. As pointed by [24, 9], this sequential approach is characterized by several shortcomings: (i) In any sequential approach, the computation of θ^* has to be completed in finite time; this is generally impossible since θ^* is defined as a solution to the stochastic program. (ii) If the learning of θ^* is terminated prematurely, this leads to an erroneous estimate $\hat{\theta}$. One then proceeds to compute a Nash equilibrium given $\hat{\theta}$, which results in an incorrect Nash equilibrium. As a result, the two-stage sequential method, in stochastic regimes, cannot provide asymptotically accurate solutions and at best provided approximate solutions. Motivated by these shortcomings, we propose a framework that combines the asynchronous inexact proximal best-response scheme with joint learning for the misspecified parameter θ^*. Under suitable conditions, we prove the almost sure convergence and the convergence in mean of the generated strategy vector to the set of Nash equilibria. Additionally, we show that for every player, its belief regarding the misspecified parameter converges almost surely to the true counterpart.

3.1 Algorithm Design and Assumptions

We impose the following conditions on the misspecified problem.

Assumption 4 *(a) For every $i \in \mathcal{N}$, X_i is a closed, compact, and convex set; $f_i(x_i, x_{-i}; \theta)$ is convex and continuously differentiable in x_i over an open set containing X_i for every $x_{-i} \in X_{-i}$ and every $\theta \in \Theta$.*
(b) For every $i \in \mathcal{N}$, $\nabla_{x_i} f_i(x; \theta^*)$ is Lipschitz continuous in x with Lipschitz constant L_x, i.e.,

$$\|\nabla_{x_i} f_i(x; \theta^*) - \nabla_{x_i} f_i(x'; \theta^*)\| \le L_x \|x - x'\| \quad \forall x, x' \in X.$$

Further, there exists a constant L_{θ^*} such that for any $x \in X$ and every $i \in \mathcal{N}$:

$$\|\nabla_{x_i} f_i(x_i, x_{-i}; \theta) - \nabla_{x_i} f_i(x_i, x_{-i}; \theta^*)\| \le L_{\theta^*} \|\theta - \theta^*\| \quad \forall \theta \in \Theta.$$

(c) $g(\theta)$ is strongly convex with convexity constant μ_g and is continuously differentiable in θ on an open set containing Θ with the gradient function being L_g-Lipschitz continuous, where $g(\theta)$ is defined in (5).
(d) There exists a function $P(\cdot;\cdot) : X \times \Theta \to \mathbb{R}$ such that for every $i \in \mathcal{N}$ and any $x_{-i} \in X_{-i}$, (6) *holds.*

We define $P(x) \triangleq P(x; \theta^*)$ as the potential function of the problem (4).

Assumption 5 *(a) For any $i \in \mathcal{N}$, all $x_{-i} \in X_{-i}$, any $\theta \in \Theta$ and any $\xi \in \mathbb{R}^d$, $\psi_i(x_i, x_{-i}; \theta; \xi)$ is differentiable in x_i over an open set containing X_i such that $\nabla_{x_i} f_i(x_i, x_{-i}; \theta) = \mathbb{E}[\nabla_{x_i} \psi_i(x_i, x_{-i}; \theta; \xi)]$. (b) For any $i \in \mathcal{N}$ and any $x \in X$, there exists a constant $M_1 > 0$ such that $\mathbb{E}[\|\nabla_{x_i} \psi_i(x_i, x_{-i}; \theta; \xi)\|^2] \le M_1^2$. (c) For any $\eta \in \mathbb{R}^p$, $g(\theta, \eta)$ is differentiable in θ over an open set containing Θ such that $\nabla g(\theta) = \mathbb{E}_\eta[\nabla g(\theta; \eta)]$.*

If $T_i(x, \theta)$ is defined as follows:

$$T_i(x,\theta) \triangleq \underset{y_i \in X_i}{\operatorname{argmin}} \left[f_i(y_i, x_{-i}; \theta) + \frac{\mu}{2}\|y_i - x_i\|^2 \right], \quad \mu > 0, \tag{49}$$

then $T_i(x, \theta)$ is uniquely defined by invoking Assumption 4(a). Additionally, we may claim the Lipschitz continuity of $T_i(x, \cdot)$ based on the following Lemma, akin to the result proved by [63].

Lemma 3 *Define $L_t \triangleq \frac{\mu L_{\theta^*}}{\mu^2 + L_x^2}\left(1 - L_x/\sqrt{\mu^2 + L_x^2}\right)^{-1}$. Then for any $i \in \mathcal{N}$ and any $x \in X$:*

$$\|T_i(x,\theta) - T_i(x,\theta^*)\| \le L_t \|\theta - \theta^*\| \quad \forall \theta \in \Theta. \tag{50}$$

Proof. By the first-order optimality condition of (49), $T_i(x, \theta)$ is a fixed point of the map $\Pi_{X_i}\left[y_i - \alpha\left(\nabla_{x_i} f_i(y_i, x_{-i}; \theta) + \mu(y\right.\right.$
Then by using the nonexpansivity property of the projection operator, the triangle inequality, and Assumption 4(b), we have that

$$\begin{aligned}
&\|T_i(x,\theta) - T_i(x,\theta^*)\| = \Big\| \Pi_{X_i}\left[T_i(x,\theta) - \alpha\left(\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta) + \mu(T_i(x,\theta) - x_i)\right)\right] \\
&\qquad - \Pi_{X_i}\left[T_i(x,\theta^*) - \alpha\left(\nabla_{x_i} f_i(T_i(x,\theta^*), x_{-i}; \theta^*) + \mu(T_i(x,\theta^*) - x_i)\right)\right] \Big\| \\
&\le \Big\| (1 - \alpha\mu)\left(T_i(x,\theta) - T_i(x,\theta^*)\right) - \alpha\left(\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta^*) - \nabla_{x_i} f_i(T_i(x,\theta^*), x_{-i}; \theta^*)\right) \\
&\qquad - \alpha\left(\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta) - \nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta^*)\right) \Big\| \le \alpha L_{\theta^*}\|\theta - \theta^*\| \\
&+ \Big\| (1 - \alpha\mu)\left(T_i(x,\theta) - T_i(x,\theta^*)\right) - \alpha\left(\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta^*) - \nabla_{x_i} f_i(T_i(x,\theta^*), x_{-i}; \theta^*)\right) \Big\|.
\end{aligned} \tag{51}$$

By recalling that $f_i(x_i, x_{-i}; \theta)$ is convex $x_i \in X_i$ for all $x_{-i} \in X_{-i}$, $\theta \in \Theta$, and $\nabla_{x_i} f_i(x; \theta^*)$ is Lipschitz continuous in x with Lipschitz constant L_x, we conclude that for $\alpha = \frac{\mu}{\mu^2 + L_x^2}$,

$$\begin{aligned}
&\Big\| (1 - \alpha\mu)\left(T_i(x,\theta) - T_i(x,\theta^*)\right) - \alpha\left(\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta^*) - \nabla_{x_i} f_i(T_i(x,\theta^*), x_{-i}; \theta^*)\right) \Big\|^2 \\
&\le \|(1 - \alpha\mu)(T_i(x,\theta) - T_i(x,\theta^*)\|^2 + \alpha^2 \left\|\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta^*) - \nabla_{x_i} f_i(T_i(x,\theta^*), x_{-i}; \theta^*)\right\|^2 \\
&- 2\alpha(1 - \alpha\mu)\left(T_i(x,\theta) - T_i(x,\theta^*)\right)^T \left(\nabla_{x_i} f_i(T_i(x,\theta), x_{-i}; \theta^*) - \nabla_{x_i} f_i(T_i(x,\theta^*), x_{-i}; \theta^*)\right) \\
&\le \left((1 - \alpha\mu)^2 + \alpha^2 L_x^2\right) \|T_i(x,\theta) - T_i(x,\theta^*)\|^2 = \frac{L_x^2}{\mu^2 + L_x^2} \|T_i(x,\theta) - T_i(x,\theta^*)\|^2 .
\end{aligned}$$

This incorporated with (51) implies that the following holds for $\alpha = \frac{\mu}{\mu^2 + L_x^2}$:

$$\left(1 - \frac{L_x}{\sqrt{\mu^2 + L_x^2}}\right)\|T_i(x,\theta) - T_i(x,\theta^*)\| \le \frac{\mu L_{\theta^*}}{\mu^2 + L_x^2}\|\theta - \theta^*\|.$$

The result follows by the definition of L_t. $\square$

We propose an asynchronous inexact proximal BR scheme that is coupled with learning (Algrithm 3) to compute an NE of the misspecified potential stochastic game. Player i at time k utilizes an estimate $x_i(k)$ of its equilibrium strategy x_i^*, an estimate $\theta_i(k)$ of the unknown parameter θ^*, and has access to possibly delay-afflicted rival strategies $y^i(k) \triangleq (x_1(k-d_{i1}(k)), \cdots, x_N(k-d_{iN}(k)))$ with delays $d_{ij}(k), j \in \mathcal{N}$. The scheme is defined as follows: At major iteration $k \geq 0$, he index i is selected randomly from $\mathcal{N}$ with probability $\mathbb{P}(i_k = i) = p_i > 0$. If $i_k = i$, then player i is chosen to initiate an update by computing an inexact proximal BR solution to the problem (49) characterized by (52), and updating $\theta_i(k+1)$ via the variable sample-size SA scheme (53) with $N_i(k)$ sampled gradients. We impose conditions on the inexactness sequence $\{\varepsilon_i(k)\}_{k\geq 1}$ and further specify the selection of $N_i(k)$ in the convergence analysis.

Algorithm 3 Asynchronous inexact proximal best-response scheme with stochastic learning

Let $k := 0$, $x_i(0) \in X_i$ and $\theta_i(0) \in \Theta$ for $i \in \mathcal{N}$. Additionally $0 < p_i < 1$ for $i \in \mathcal{N}$ such that $\sum_{i=1}^N p_i = 1$.

(S.1) Pick $i_k = i \in \mathcal{N}$ with probability p_i.

(S.2) If $i_k = i$, then player i updates $x_i(k+1) \in X_i$ and $\theta_i(k+1) \in \Theta$ as follows:

$$x_i(k+1) := T_i(y^i(k), \theta_i(k)) + \varepsilon_i(k+1), \tag{52}$$

$$\theta_i(k+1) := \Pi_\Theta\Big[\theta_i(k) - \frac{\beta_i}{N_i(k)} \sum_{p=1}^{N_i(k)} \nabla g\left(\theta_i(k), \eta_{i,p}(k)\right)\Big], \tag{53}$$

where $\varepsilon_i(k+1)$ denotes the inexactness, $\nabla g\left(\theta_i(k), \eta_{i,p}(k)\right), p = 1, \cdots, N_i(k)$ denotes the sampled gradient, and $\beta_i = \frac{1}{L_g}$; Otherwise, $x_j(k+1) := x_j(k), \theta_{j,k+1} = \theta_j(k)$ if $j \neq i_k$.

(S.3) If $k > K$, stop; Else, $k := k+1$ and return to (S.1).

We then list the following conditions concerning the delays, observation noise of the gradient function $\nabla g(\theta)$ as well as the inexactness sequence $\{\varepsilon_i(k)\}$ utilized in Algorithm 3. We denote the σ-field of the entire information used by Algorithm 3 up to (and including) the updates of $x_i(k), \theta_i(k)$ for all $i \in \mathcal{N}$ by $\mathcal{F}_k'$, and the σ-field generated from $\mathcal{F}_k'$ and the delays at step k by $\mathcal{F}_k \triangleq \sigma\big\{\mathcal{F}_k', d_{ij}(k), i, j \in \mathcal{N}\big\}$. We will further define $\mathcal{F}_k'$ after introducing the SA scheme (54).

Assumption 6 *(a) $\{i_k\}$ is an i.i.d. sequence, where i_k is independent of $\mathcal{F}_k$ for all $k \geq 1$.*
(b) For any $i \in \mathcal{N}$, the noise term $\{\varepsilon_i(k)\}$ satisfies the following condition:

$$\sum_{k=1}^\infty \mathbb{E}\left[\|\varepsilon_i(k+1)\|^2 \big| \mathcal{F}_k\right] < \infty, \text{ and } \sum_{k=1}^\infty \mathbb{E}\left[\|\varepsilon_i(k+1)\| \big| \mathcal{F}_k\right] < \infty \ \ a.s..$$

(c) There exists a positive integer τ such that for any $i, j \in \mathcal{N}$ and any $k \geq 0$, $d_{ij}(k) \in \{0, \cdots, \tau\}$.
(d) Define $e_{i,p}(k) \triangleq \nabla g\left(\theta_i(k), \eta_{i,p}(k)\right) - \nabla g\left(\theta_i(k)\right)$. There exists a constant $M_2 > 0$ such that for any $k \geq 0$ and $p = 1, \cdots, N_i(k)$, $\mathbb{E}\left[\|e_{i,p}(k)\|^2 | \mathcal{F}_k\right] \leq M_2^2$.

Analogous to the computation of the inexact best-response (9) in Algorithm 1, we still utilize SA to compute (52). By (49) it is seen that the computation of $T_i(x, \theta)$ requires solving a strongly convex stochastic program. Thus, an inexact solution to the problem (49), characterized by (52), can also be computed via the SA algorithm defined as follows for $t = 1, \ldots, j_i(k)$:

$$x_{i,t+1}(k) := \Pi_{X_i}\Big[z_{i,t}(k) - \gamma_{i,t}\left[\nabla_{x_i}\psi_i(z_{i,t}(k), y^i_{-i,k}; \theta_i(k); \xi_{i,t}(k)) + \mu(z_{i,t}(k) - x_i(k))\right]\Big], \tag{54}$$

where $\gamma_{i,t} = \frac{1}{\mu(t+1)}$, and $z_{i,t}(k)$ denotes the estimate of the proximal BR solution $T_i(y^i(k), \theta_i(k))$ at t-th inner step of the SA scheme (54) with the initial value $z_{i,1}(k) = x_i(k)$. Set $x_i(k+1) = z_{i,j_i(k)}(k)$. Define $\xi_i(k) \triangleq (\xi_{i,1}(k), \cdots, \xi_{i,j_i(k)}(k))$, and $\eta_i(k) \triangleq (\eta_{i,1}(k), \cdots, \eta_{i,N_i(k)}(k))$, and $\mathcal{F}'_k \triangleq \sigma\{x(0), i_l, \xi_{i_l}(l), \eta_{i_l}(l), d_{i_l,j}(l), 0 \le l \le k-1, j \in \mathcal{N}\}$. Then by Algorithm 3, $x(k)$ and $\theta(k)$ are adapted to $\mathcal{F}'_k$, and hence $y^i(k)$ is adapted to $\mathcal{F}_k$. Thus, $T_i(y^i(k), \theta_i(k))$ is adapted to $\mathcal{F}_k$ by its definition (49). Then by Assumptions 4 and 5, we obtain the same bound as that of Lemma 1. Consequently, Assumption 6(b) is satisfied by setting $j_i(k) = \lfloor \Gamma_i(k)^{2(1+\delta)} \rfloor$, where $\delta > 0$ and $\Gamma_i(k)$ is defined in Remark 2.

3.2 Convergence Analysis

We begin by proving a supporting Lemma, an extension of the analogous determinstic (error-free) result from [64], which will be used in the convergence analysis of Algorithm 3.

Lemma 4 *Suppose Assumption 4 (c) holds. Let $\theta, y \in \Theta$ and suppose θ^+ and $c_\Theta(\theta)$ are defined by .*

$$\theta^+ := \Pi_\Theta \left(\theta - \frac{1}{L_g} (\nabla_\theta g(\theta) + u) \right) \text{ and } c_\Theta(\theta) \triangleq L_g(\theta - \theta^+), \tag{55}$$

respectively. Then the following holds.

$$g(\theta^+) - g(y) \le c_\Theta(\theta)^T(\theta - y) - \frac{1}{2L_g}\|c_\Theta(\theta)\|^2 - u^T(\theta^+ - y) - \frac{\mu_g}{2}\|\theta - y\|^2.$$

Proof. We begin by recalling the projection inequality

$$\left(\theta^+ - \left(\theta - \frac{1}{L_g} (\nabla_\theta g(\theta) + u) \right) \right)^T (\theta^+ - y) \le 0.$$

Consequently, we have that

$$\nabla_\theta g(\theta)^T(\theta^+ - y) \le c_\Theta(\theta)^T(\theta^+ - y) - u^T(\theta^+ - y). \tag{56}$$

Then by the μ_g-strong convexity and L_g-smoothness of $g(\cdot)$, we may now derive the following bound.

$$\begin{aligned} g(\theta^+) - g(y) &= g(\theta^+) - g(\theta) + g(\theta) - g(y) \\ &\le \nabla_\theta g(\theta)^T(\theta^+ - \theta) + \frac{L_g}{2}\|\theta^+ - \theta\|^2 + \nabla_\theta g(\theta)^T(\theta - y) - \frac{\mu_g}{2}\|\theta - y\|^2 \\ &= \nabla_\theta g(\theta)^T(\theta^+ - y) + \frac{1}{2L_g}\|c_\Theta(\theta)\|^2 - \frac{\mu_g}{2}\|\theta - y\|^2 \\ &\overset{(56)}{\le} c_\Theta(\theta)^T(\theta^+ - y) - u^T(\theta^+ - y) + \frac{1}{2L_g}\|c_\Theta(\theta)\|^2 - \frac{\mu_g}{2}\|\theta - y\|^2 \\ &= c_\Theta(\theta)^T(\theta - y) - \frac{1}{2L_g}\|c_\Theta(\theta)\|^2 - u^T(\theta^+ - y) - \frac{\mu_g}{2}\|\theta - y\|^2. \qquad \square \end{aligned}$$

Theorem 4 (almost sure convergence for inexact BR with learning) *Let $\{x(k)\}$ and $\{\theta(k)\}$ be generated by Algorithm 3. Suppose Assumptions 4, 5 and 6 hold. We further assume that μ used in* (49) *satisfies $\mu > \frac{L_x}{2} + \sqrt{3}L_x\tau$, and for every $i \in \mathcal{N}$, $N_i(k) = \lfloor \Gamma_i(k)^{2(1+\delta)} \rfloor$ for some $\delta > 0$ with $\Gamma_i(k)$ defined by* (2)*. Then the sequences $\{\theta_i(k)\}$ and $\{x_i(k)\}$ satisfy the following.*
(a) For any $i \in \mathcal{N}$, $\sum_{k=1}^\infty \|\theta_i(k) - \theta^*\|^2 < \infty$ a.s., and $\sum_{k=1}^\infty \|\theta_i(k) - \theta^*\| < \infty$ a.s..

(b) For any $i \in \mathcal{N}$, $\sum_{k=0}^{\infty} \|T_i(y^i(k);\theta^*) - x_i(k)\|^2 < \infty \quad a.s.$.
(c) For almost all $\omega \in \Omega$, every limit point of $x(k,\omega)$ is a Nash equilibrium.
(d) There exists a connected subset $X_c^* \subset X^*$ such that $d(x(k), X_c^*) \xrightarrow[k\to\infty]{} 0\ a.s.$.

Proof. (a) For $i_k = i$, by $e_{i,p}(k)$ defined in Assumption 6(d) we can rewrite (53) as follows:

$$\theta_i(k+1) = \Pi_\Theta\left[\theta_i(k) - \frac{1}{L_g}\left(\nabla g(\theta_i(k)) + u_k\right)\right], \tag{57}$$

where $u_k \triangleq \frac{1}{N_i(k)}\sum_{p=1}^{N_i(k)} e_{i,p}(k)$. From Lemma 4, by setting $\theta = \theta_i(k)$, $u = u_k$, and $y = \theta^*$, we have that $\theta^+ = \theta_i(k+1)$, $c_\Theta(\theta_i(k)) = L_g(\theta_i(k) - \theta_i(k+1))$, implying the following inequality.

$$\begin{aligned}-c_\Theta(\theta_k)^T(\theta_i(k)-\theta^*) &\le -\underbrace{(g(\theta_i(k+1)) - g(\theta^*))}_{\ge 0} - \frac{1}{2L_g}\|c_\Theta(\theta_i(k))\|^2 - \frac{\mu_g}{2}\|\theta_i(k)-\theta^*\|^2 - u_k^T(\theta_i(k+1)-\theta^*)\\ &\le -\frac{1}{2L_g}\|c_\Theta(\theta_i(k))\|^2 - \frac{\mu_g}{2}\|\theta_i(k)-\theta^*\|^2 - u_k^T(\theta_i(k+1)-\theta^*).\end{aligned} \tag{58}$$

Since $c_\Theta(\theta_i(k)) = L_g(\theta_i(k) - \theta_i(k+1))$, we may bound $\|\theta_i(k+1)-\theta^*\|^2$ as follows.

$$\begin{aligned}&\|\theta_i(k+1)-\theta^*\|^2 = \|\theta_k - \frac{1}{L_g}c_\Theta(\theta_i(k)) - \theta^*\|^2 = \|\theta_i(k)-\theta^*\|^2 + \frac{1}{L_g^2}\|c_\Theta(\theta_i(k))\|^2 - \frac{2}{L_g}c_\Theta(\theta_i(k))^T(\theta_i(k)-\theta^*)\\ &\overset{(58)}{\le} \|\theta_i(k)-\theta^*\|^2 + \frac{1}{L_g^2}\|c_\Theta(\theta_i(k))\|^2 - \frac{2}{L_g}\left(\frac{1}{2L_g}\|c_\Theta(\theta_i(k))\|^2 + \frac{\mu_g}{2}\|\theta_i(k)-\theta^*\|^2 + u_k^T(\theta_i(k+1)-\theta^*)\right)\\ &= \left(1-\frac{\mu_g}{L_g}\right)\|\theta_i(k)-\theta^*\|^2 - \frac{2}{L_g}u_k^T(\theta_i(k+1)-\theta^*)\\ &= \left(1-\frac{\mu_g}{L_g}\right)\|\theta_i(k)-\theta^*\|^2 - \frac{2}{L_g}u_k^T(\theta_i(k+1)-\bar\theta_i(k+1)) - \frac{2}{L_g}u_k^T(\bar\theta_i(k+1)-\theta^*),\end{aligned}$$

where $\bar\theta_i(k+1) \triangleq \Pi_\Theta\left[\theta_i(k) - \frac{1}{L_g}\nabla_\theta g(\theta_i(k))\right]$. By (57) and the non-expansivity of the Euclidean projector, one may obtain that $-u_k^T(\theta_i(k+1)-\bar\theta_i(k+1)) \le \|u_k^T\|\,\|\theta_i(k+1)-\bar\theta_i(k+1)\| \le \|u_k\|^2/L_g$. Therefore,

$$\|\theta_i(k+1)-\theta^*\|^2 \le \left(1-\frac{\mu_g}{L_g}\right)\|\theta_i(k)-\theta^*\|^2 + \frac{2}{L_g^2}\|u_k\|^2 - \frac{2}{L_g}u_k^T(\bar\theta_i(k+1)-\theta^*).$$

Taking expectations conditioned on $\mathcal{F}_k$ on both sides of the above equation, we obtain the next inequality since $\theta_i(k)$ and $\bar\theta_i(k+1)$ are adapted to $\mathcal{F}_k$.

$$\begin{aligned}\mathbb{E}[\|\theta_i(k+1)-\theta^*\|^2 \mid \mathcal{F}_k] &\le \left(1-\frac{\mu_g}{L_g}\right)\|\theta_i(k)-\theta^*\|^2 + \frac{2}{L_g^2}\mathbb{E}[\|u_k\|^2 \mid \mathcal{F}_k] \quad (\text{by } \mathbb{E}[u_k|\mathcal{F}_k]=0)\\ &\le \left(1-\frac{\mu_g}{L_g}\right)\mathbb{E}[\|\theta_i(k)-\theta^*\|^2] + \frac{2M_2^2}{L_g^2N_i(k)}, \quad (\text{by Assumption } 6(d)).\end{aligned}$$

While for $i_k \ne i$, $\mathbb{E}\left[\|\theta_i(k+1)-\theta^*\|\middle|\mathcal{F}_k\right] = \|\theta_i(k)-\theta^*\|^2$. Then by $\mathbb{P}(i_k = i) = p_i$, we obtain that

$$\mathbb{E}\left[\|\theta_i(k+1)-\theta^*\|^2\middle|\mathcal{F}_k\right] \le (1-p_i\mu_g/L_g)\|\theta_i(k)-\theta^*\|^2 + \frac{2p_iM_2^2}{L_g^2N_i(k)}. \tag{59}$$

By employing the conditional variant of Jensen's inequality to (59), we may conclude the following.

$$
\begin{aligned}
\mathbb{E}\left[\|\theta_i(k+1)-\theta^*\|\big|\mathcal{F}_k\right] &\le \sqrt{1-p_i\mu_g/L_g}\|\theta_i(k)-\theta^*\| + \frac{\sqrt{2p_i}M_2}{L_g\sqrt{N_i(k)}} \quad {\scriptstyle (\text{by } \sqrt{a^2+b^2}\le a+b \text{ for } a,b\ge 0)} \\
&= \|\theta_i(k)-\theta^*\| - \left(1-\sqrt{1-p_i\mu_g/L_g}\right)\|\theta_i(k)-\theta^*\| + \frac{\sqrt{2p_i}M_2}{L_g\sqrt{N_i(k)}}.
\end{aligned}
$$

Thus, by using Theorem 1 in [61], and $\sum_{k=1}^{\infty} 1/\sqrt{N_i(k)} < \infty, a.s.$ by Remark 2, we obtain that $\sum_{k=1}^{\infty}\|\theta_i(k)-\theta^*\| < \infty\ a.s.$, and hence $\sum_{k=1}^{\infty}\|\theta_i(k)-\theta^*\|^2 < \infty\ a.s.$.

(b) Note that $\nabla_{x_i} f_i(x;\theta^*)$ is Lipschitz continuous in $x \in X$ with Lipschitz constant L_x by Assumption 4(b). Then similar to (18) we obtain the following inequality for any $C > 0$:

$$
\begin{aligned}
f_i\left(T_i(y^i(k),\theta^*), x_{-i}(k);\theta^*\right) + V_{k+1} &\le f_i(x(k);\theta^*) + V_k + \frac{L_x^2\tau^2}{2C}\|x(k+1)-x(k)\|^2 \\
&\quad - \left(\mu - \frac{L_x+C}{2}\right)\|T_i(y^i(k),\theta^*)-x_i(k)\|^2,
\end{aligned} \tag{60}
$$

where $V_k \triangleq \frac{L_x^2\tau}{2C}\sum_{h=k-\tau+1}^{k}(h-k+\tau)\|x(h)-x(h-1)\|^2$. By Assumptions 5(a), 5(b), and Jensen's inequality, the following holds for any $x \in X$:

$$
\|\nabla_{x_i} f_i(x_i,x_{-i};\theta)\| = \|\mathbb{E}[\nabla_{x_i}\psi_i(x_i,x_{-i};\theta;\xi)]\| \le \sqrt{\mathbb{E}[\|\nabla_{x_i}\psi_i(x_i,x_{-i};\theta;\xi)\|^2]} \le M_1.
$$

Then by the mean-value theorem and Cauchy-Schwarz inequality, we have that

$$
\begin{aligned}
&f_i\left(x_i(k+1),x_{-i}(k);\theta^*\right) - f_i\left(T_i(y^i(k),\theta^*),x_{-i}(k);\theta^*\right) \\
&= \left(x_i(k+1)-T_i(y^i(k),\theta^*)\right)^T\nabla_{x_i} f_i\left(z_i(k+1),x_{-i}(k);\theta^*\right) \le M_1\|x_i(k+1)-T_i(y^i(k),\theta^*)\|.
\end{aligned} \tag{61}
$$

where $z_i(k+1) = \vartheta_{i,k}x_i(k+1) + (1-\vartheta_{i,k})T_i(y^i(k),\theta^*)$ for some $\vartheta_{i,k} \in (0,1)$. By the triangle inequality, Lemma 3, and (52), we may conclude that

$$
\begin{aligned}
\|x_i(k+1)-T_i(y^i(k),\theta^*)\| &\le \|T_i(y^i(k),\theta_i(k))-T_i(y^i(k),\theta^*)\| + \|\varepsilon_i(k+1)\| \\
&\le L_t\|\theta_i(k)-\theta^*\| + \|\varepsilon_i(k+1)\|.
\end{aligned} \tag{62}
$$

We then substitute (62) in (61) to obtain the following bound:

$$
f_i\left(x_i(k+1),x_{-i}(k);\theta^*\right) - f_i\left(T_i(y^i(k),\theta^*),x_{-i}(k);\theta^*\right) \le M_1L_t\|\theta_i(k)-\theta^*\| + M_1\|\varepsilon_i(k+1)\|. \tag{63}
$$

Then by Algorithm 3, Assumption 4(d), (60) and (63), we may obtain the following bound:

$$
\begin{aligned}
&P(x(k+1)) - P(x(k)) = f_{i_k}(x_{i_k}(k+1),x_{-i_k}(k);\theta^*) - f_{i_k}(x_{i_k}(k),x_{-i_k}(k);\theta^*) \\
&= f_{i_k}\left(x_{i_k}(k+1),x_{-i_k}(k);\theta^*\right) - f_{i_k}\left(T_{i_k}(y^{i_k}(k),\theta^*),x_{-i_k}(k);\theta^*\right) \\
&+ f_{i_k}\left(T_i(y^{i_k}(k),\theta^*),x_{-i_k}(k);\theta^*\right) - f_{i_k}(x(k);\theta^*) \\
&\le M_1L_t\|\theta_{i_k}(k)-\theta^*\| + M_1\|\varepsilon_{i_k}(k+1)\| + V_k - V_{k+1} \\
&- \left(\mu - \frac{L_x+C}{2}\right)\|T_i(y^i(k),\theta^*)-x_i(k)\|^2 + \frac{L_x^2\tau^2}{2C}\|x(k+1)-x(k)\|^2.
\end{aligned} \tag{64}
$$

By the update rule in Algorithm 3 and (62), we have that

$$\begin{aligned}\|x(k+1)-x(k)\| &= \|x_{i_k}(k+1)-x_{i_k}(k)\| = \|x_{i_k}(k+1)-T_{i_k}(y^{i_k}(k),\theta^*)+T_{i_k}(y^{i_k}(k),\theta^*)-x_{i_k}(k)\| \\ &\le L_t\|\theta_{i_k}(k)-\theta^*\|+\|\varepsilon_{i_k}(k+1)\|+\|T_{i_k}(y^{i_k}(k),\theta^*)-x_{i_k}(k)\|,\end{aligned}$$

which when combined with (64) and $(a+b+c)^2 \le 3(a^2+b^2+c^2)$ yields the following inequality:

$$\begin{aligned}P(x(k+1))+V_{k+1} \le{}& P(x(k))+V_k+M_1L_t\|\theta_{i_k}(k)-\theta^*\|+M_1\|\varepsilon_{i_k}(k+1)\|+\frac{3L_x^2\tau^2}{2C}\|\varepsilon_{i_k}(k+1)\|^2 \\ &-\underbrace{\left(\mu-\frac{L_x+C}{2}-\frac{3L_x^2\tau^2}{2C}\right)}_{\triangleq \widetilde{C}}\|T_{i_k}(y^{i_k}(k),\theta^*)-x_{i_k}(k)\|^2+\frac{3L_x^2L_t^2\tau^2}{2C}\|\theta_{i_k}(k)-\theta^*\|^2 \\ \le{}& P(x(k))+V_k-\widetilde{C}\|T_{i_k}(y^{i_k}(k),\theta^*)-x_{i_k}(k)\|^2+M_1L_t\|\theta_{i_k}(k)-\theta^*\| \\ &+\frac{3L_x^2L_t^2\tau^2}{2C}\|\theta_{i_k}(k)-\theta^*\|^2+M_1\sum_{i=1}^N\|\varepsilon_i(k+1)\|+\frac{3L_x^2\tau^2}{2C}\sum_{i=1}^N\|\varepsilon_i(k+1)\|^2.\end{aligned} \tag{65}$$

Note that $x(k), V_k, \theta_i(k), T_i(y^i(k);\theta^*)\ \forall i\in\mathcal{N}$ are adapted to $\mathcal{F}_k$, and i_k is independent of $\mathcal{F}_k$. Then by taking expectations conditioned on $\mathcal{F}_k$ of (65), and using Corollary 7.1.2 in [60] and $\mathbb{P}(i_k=i)=p_i$, it follows that

$$\begin{aligned}\mathbb{E}\left[P(x(k+1))+V_{k+1}\middle|\mathcal{F}_k\right] \le{}& P(x(k))+V_k-\widetilde{C}\sum_{i=1}^N p_i\|T_i(y^i(k),\theta^*)-x_i(k)\|^2 \\ &+M_1L_t\sum_{i=1}^N p_i\|\theta_i(k)-\theta^*\|+\frac{3L_x^2L_t^2\tau^2}{2C}\sum_{i=1}^N p_i\|\theta_i(k)-\theta^*\|^2 \\ &+M_1\sum_{i=1}^N\mathbb{E}\left[\|\varepsilon_i(k+1)\|\middle|\mathcal{F}_k\right]+\frac{3L_x^2\tau^2}{2C}\sum_{i=1}^N\mathbb{E}\left[\|\varepsilon_i(k+1)\|^2\middle|\mathcal{F}_k\right].\end{aligned} \tag{66}$$

By setting $C=\frac{\sqrt{3}L_x\tau}{2}$ we derive $\frac{C}{2}+\frac{3L_x^2\tau^2}{2C}=\sqrt{3}L_x\tau$. Thus, by taking $\mu>\frac{L_x}{2}+\sqrt{3}L\tau$ it follows that $\widetilde{C}>0$. Therefore, by using Theorem 1 of [61], Assumption 6(b) and result (a), we have that $\sum_{k=1}^\infty\sum_{i=1}^N p_i\|T_i(y^i(k),\theta^*)-x_i(k)\|^2<\infty\ a.s.$. Then by $p_i\in(0,1)$ we obtain (b).

(c) The proof is similar to that of Theorem 1(b).

(d) Since $\mathbb{E}\left[\|\varepsilon_i(k+1)\|\middle|\mathcal{F}_k\right], \|\theta_i(k)-\theta^*\|$, and $\|T_i(y^i(k);\theta^*)-x_i(k)\|^2$ are nonnegative for $k\ge 1$, by Assumption 6(b), results (a) and (b) we have the following for any $i\in\mathcal{N}$:

$$\begin{aligned}&\sum_{k=1}^\infty\mathbb{E}[\|\varepsilon_i(k+1)\|]=\mathbb{E}\left[\sum_{k=1}^\infty\mathbb{E}\left[\|\varepsilon_i(k+1)\|\middle|\mathcal{F}_k\right]\right]<\infty, \\ &\sum_{k=1}^\infty\mathbb{E}[\|\theta_i(k)-\theta^*\|]<\infty, \text{ and } \sum_{k=1}^\infty\mathbb{E}\left[\|T_i(y^i(k);\theta^*)-x_i(k)\|^2\right]<\infty.\end{aligned}$$

Thus, by the Jensen's inequality we have the following for any $i\in\mathcal{N}$:

$$\lim_{k\to\infty}\mathbb{E}[\|\varepsilon_i(k+1)\|]=0,\ \lim_{k\to\infty}\mathbb{E}[\|\theta_i(k)-\theta^*\|]=0, \text{ and } \lim_{k\to\infty}\mathbb{E}\left[\|T_i(y^i(k);\theta^*)-x_i(k)\|\right]=0. \tag{67}$$

Then by the triangle inequality and (62), we obtain that

$$\begin{aligned}\mathbb{E}[\|x_i(k+1)-x_i(k)\|] &\leq \mathbb{E}[\|x_i(k+1)-T_i(y^i(k),\theta^*)\|]+\mathbb{E}[\|x_i(k)-T_i(y^i(k),\theta^*)\|] \\ &\leq L_t\mathbb{E}[\|\theta_i(k)-\theta^*\|]+\mathbb{E}[\|\varepsilon_i(k+1)\|]+\mathbb{E}[\|x_i(k)-T_i(y^i(k),\theta^*)\|] \to 0, \text{ as } k\to\infty.\end{aligned}$$

This implies that $\lim_{k\to\infty}\mathbb{E}[\|x(k+1)-x(k)\|]=0$. The result follows by proceeding as in Theorem 1(c). □

Remark 4 *Note that in Algorithm 3, we have utilized a variable sample-size stochastic approximation framework for updating each player's estimate of $\{\theta_i(k)\}$, where each player initiates its update from an independently selected starting point. The proposed scheme is similar to that developed in [65] but we provide a distinct proof, inspired by [64], through which the required summability requirements are proven. Related schemes under differing assumptions with similar linear convergence rates have been studied over the last decade (cf. [66, 67]). Accelerated variants of such schemes have also been studied in smooth [68] and nonsmooth [69] regimes. One might go a step further and note that players can utilize any update rule as long as it produces a sequence of iterates satisfying the requirement that for every $i\in\mathcal{N}$, $\sum_{k=1}^{\infty}\|\theta_i(k)-\theta^*\|<\infty$ a.s..*

Theorem 4 shows that the estimates of the equilibrium strategy and the misspecified parameter generated by Algorithm 3 converge almost surely to the set of Nash equilibria and to θ^*, respectively. Define the gap function $G(x;\theta^*)\triangleq\sup_{y\in X}\nabla P(x;\theta^*)^T(x-y)$. The following result shows the convergence in mean of $x(k)$, characterized by the convergence of $G(x(k))$ to zero in the mean sense.

Theorem 5 (Convergence in mean of gap function.) *Let $\{x(k)\}$ and $\{\theta(k)\}$ be generated by Algorithm 3. Suppose Assumptions 4, 5 and 6 hold, and, in addition, that $\mu>\frac{L_x}{2}+\sqrt{3}L_x\tau$, $\beta_i\in(0,2\mu_g/L_g^2)$ and $N_i(k)=\lceil\Gamma_i(k)^{2(1+\delta)}\rceil$ for some $\delta>0$. Then $\lim_{k\to\infty}\mathbb{E}\big[G(x(k);\theta^*)\big]=0$.*

Proof. Similar to the proof of Theorem 2, we have the following bound:

$$\mathbb{E}\left[G(x(k);\theta^*)\right]\leq\sum_{i=1}^{N}(\mu D_{X_i}+M+L_xD_{X_i})\,\mathbb{E}\left[\left\|T_i(y^i(k);\theta^*)-x_i(k)\right\|\right] \implies \lim_{k\to\infty}\mathbb{E}\left[G(x(k);\theta^*)\right]\leq 0 \quad \text{(by (67))}.$$

However, $G(x(k);\theta^*)\geq 0$ since $x(k)\in X$, implying that $\lim_{k\to\infty}\mathbb{E}\left[G(x(k);\theta^*)\right]=0$. □

Remark 5 *(i) Algorithm 3 may also be extended to the generalized and weighted potential games with misspecified parameters since Algorithm 1 is applicable to the generalized and weighted potential games as shown in Section 2.4.*
(ii) The recent work by [9] also considered the misspecified convex stochastic Nash games. This work was distinct in both its motivation and contributions.

- *[9] consider monotone Nash games while we consider stochastic potential games. Note that instances of one do not necesarily lie within the other.*
- *The update of equilibrium strategies in [9] utilizes projected gradient response while here we take inexact proximal best-response steps.*
- *The update for the misspecified parameter in [9] is based on a projected SG algorithm with a single sampled gradient per step and with a decreasing step-size; here, we utilize an* **increasing sample-size** *projected SG scheme with a* **constant** *step-size.*

4 Preliminary numerics

In this section, we empirically validate the performance of Algorithm 1 and Algorithm 3 on the problem of congestion control and misspecified stochastic Nash-Cournot games, respectively.

4.1 Congestion Control

We consider a congestion control problem on a connected network characterized by a set of nodes $\mathcal{V} = \{1, \cdots, V\}$ and a set of links $\mathcal{L} = \{1, \cdots, L\}$ connecting the nodes. There are N users in the network, where each player i aims at sending a flow rate $x_i \in C_i = \{x_i \in \mathbb{R} : 0 \leq x_i \leq x_{i,\max}\}$ from the source node s_i to the destination node d_i through a path $\mathcal{L}_i$ in the network. The upper bound $x_{i,\max}$ on user i's flow rate might represent a player-specific physical limitation. The payoff function of player i takes as the difference of a player-specific pricing function and a utility function U_i associated to the flow x_i parameterized by uncertainty ξ_i, ζ_i:

$$\psi_i(x_i, x_{-i}; \xi_i, \zeta_i) = \sum_{l \in \mathcal{L}_i} P_l \left(\sum_{j : l \in \mathcal{L}_j} x_j \right) - U_i \left(x_i, \xi_i, \zeta_i \right).$$

The first term can be interpreted as the price that player i pays for the network resources with P_l depending on the aggregated flows on the link l. Suppose that $P_l, l \in \mathcal{L}$ is convex and $U_i, i \in \mathcal{N}$ is concave on $[0, x_{i,max}]$. Typical examples for the pricing and utility functions are given by the following:

$$P_l = \frac{a_l}{b_l - \sum_{j : l \in \mathcal{L}_j} x_j} \text{ and } U_i = \xi_i \log(1 + x_i + \zeta_i),$$

where ξ_i, ζ_i are random variables. Suppose each link $l \in \mathcal{L}$ in the network has a positive capacity c_l. Let us introduce a routing matrix $A \in \mathbb{R}^{L \times N}$, where $[A]_{l,i} = 1$ if $l \in \mathcal{L}_i$, and $[A]_{l,i} = 0$, otherwise. The capacity constraints of all links can be expressed in the vector form as $Ax \leq c$ with $c = col\{c_l\}_{l=1}^{L}$. For a fixed feasible x_{-i}, we derive the bound of the user i's flow rate x_i denoted by

$$0 \leq X_i(x_{-i}) \leq \min_{l \in \mathcal{L}_i} \{c_l - \sum_{j \neq i} A_{l,j} x_j\}.$$

The ith user aims at solving the following problem:

$$\min_{x_i \in C_i \cap X_i(x_{-i})} \quad f_i(x_i, x_{-i}) \triangleq \mathbb{E}\left[\psi_i(x_i, x_{-i}; \xi_i, \zeta_i)\right].$$

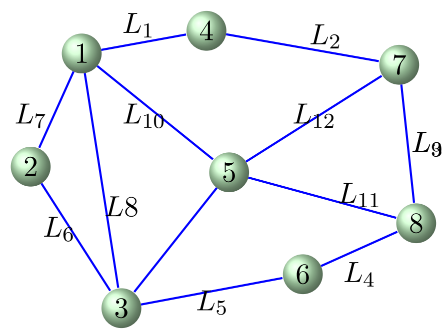
Thus, the resulting problem is a generalized potential game with the coupled constraint $X = \{x \in \mathbb{R}^n : Ax \leq c\}$ and the potential function defined as follows:

$$P(x) = \sum_{l \in \mathcal{L}} \frac{a_l}{b_l - \sum_{j : l \in \mathcal{L}_j} x_j} - \sum_{i \in \mathcal{N}} \mathbb{E}\left[U_i \left(x_i, \xi_i, \zeta_i\right)\right].$$

Further, it is shown in Theorem 3.1 of [70] that the congestion control problem has a unique inner NE under appropriately chosen parameters.

We conducted numerical simulations for a network of $V = 8$ nodes and $L = 12$ links shown in Figure 1. The parameters a_l, b_l of the utility function P_l and the capacity constraint c_l associated to link $l \in \mathcal{L}$ are given in Figure 1 as well. There are $N = 8$ users sending flows through the network depicted in Figure 1. The link paths of user $i \in \mathcal{N}$ as well as local parameters $\zeta_i, \xi_i, x_{i,\max}, p_i$ are given in Table 2, where $U[\tau_1, \tau_2]$

denotes the uniform distribution over the interval $[\tau_1, \tau_2]$. For any $k \geq 0, i, j \in \mathcal{N}$, the communication delays $d_{ij}(k)$ are independently generated from a uniform distribution on the set $\{0, 1, \cdots, \tau\}$ with $\tau = 4$. We carry out simulations for Algorithm 1, where the inexact solution (9) satisfying Assumption 3 are computed via the SA scheme (SA$_{i,k}$) with $j_{i,k} = \lfloor \Gamma_{i,k}^3 \rfloor$ and $\mu = 1$. The estimates of each users equilibrium flow rates are shown in Figure 2, which demonstrates the almost sure convergence of the iterates generated by Algorithm 1. Figure 3 displays the trajectory of the mean gap function $\mathbb{E}[G(x_k)]$ calculated by averaging across 50 sample paths, which demonstrates convergence in mean of the estimates generated by Algorithm 1.



Links l / Parameters	1	2	3	4	5	6	7	8	9	10	11	12
a_l	5	4	3	5	4	3	5	4	3	5	4	3
b_l	6	10	8	6	9	5	6	5	6	6	8	9
c_l	5	8	6	5	8	4	5	4	4	5	7	8

Figure 1: A network with 8 nodes and 12 links.

User i	Link path	ξ_i	ζ_i	$x_{i,\max}$	p_i
1	L_1, L_2, L_{12}	$U[10, 12]$	$U[0, 1]$	3	1/8
2	L_3, L_4, L_5	$U[10, 12]$	$U[0, 1]$	4	1/8
3	L_{10}, L_{11}, L_{12}	$U[10, 12]$	$U[0, 1]$	4	1/8
4	L_6, L_9, L_{12}	$U[10, 12]$	$U[0, 1]$	3	1/8
5	L_5, L_8	$U[10, 12]$	$U[0, 1]$	5	1/8
6	L_1, L_2, L_7	$U[10, 12]$	$U[0, 1]$	3	1/8
7	L_3, L_{10}, L_{11}	$U[10, 12]$	$U(0, 1)$	4	1/8
8	L_6	$U[10, 12]$	$U[0, 1]$	3	1/8

Table 2: Link paths and local parameters of all users

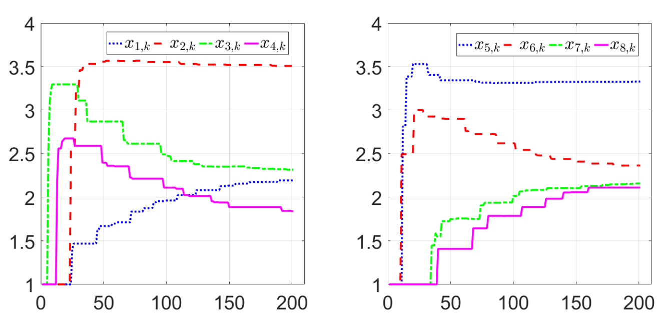


Figure 2: Flow rates of players (a single sample path)

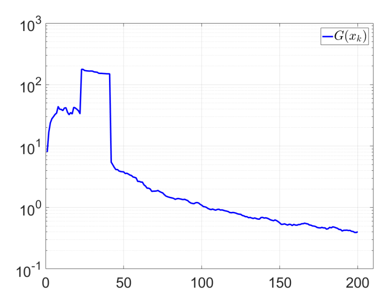


Figure 3: Trajectory of $\mathbb{E}[G(x_k)]$

4.2 Nash-Cournot Games with Misspecified Parameters

We apply Algorithm 3 to the networked Nash-Cournot game [8, 7]. Suppose there are N firms, regarded as the set of players $\mathcal{N} = \{1, \ldots, N\}$, competing over L markets denoted by $\mathcal{L} = \{1, \cdots, L\}$. Firm $i \in \mathcal{N}$ sells its products $x_i = (x_{i,1}, \cdots, x_{i,n_i}) \in \mathbb{R}^{n_i}$ to each connected market with n_i denoting the number of markets connected to firm i. We use matrix $A_i \in \mathbb{R}^{L \times n_i}$ to specify the participation of firm i in the markets, where $[A_i]_{j,p} = 1$ if firm i delivers its production $x_{i,p}$ to market j, and $[A_i]_{j,p} = 0$, otherwise. The production cost function of firm i is given by $c_i(x_i; \xi_i) = (c_i + \xi_i)^T x_i$ for some positive parameter $c_i \in \mathbb{R}^{n_i}$ and random disturbance ξ_i with mean zero. Denote by $A = [A_1, \cdots, A_N]$, by $Ax = \sum_{i=1}^N A_i x_i \in \mathbb{R}^L$ and by $S_j = [Ax]_j$ the aggregated products of all connected firms delivered to market j, where $[Ax]_j$ denotes the j-th entry of the vector Ax. Furthermore, the price of products sold in market $j \in \mathcal{L}$ is assumed to follow a linear function corrupted by noise:

$$p_j(S_j; \zeta_j) = a_j^* + \zeta_j - b_j^* S_j,$$

where $a_j^* > 0, b_j^* > 0$ are the pricing parameters, and the random disturbance ζ_j is zero-mean. Then firm $i \in \mathcal{N}$ has a stochastic payoff function defined as follows:

$$\psi_i(x; \theta^*; \xi_i, \zeta_i) = c_i(x_i; \xi_i) - \sum_{j \in \mathcal{L}} p_j(S_j; \zeta_j)[A_i x_i]_j = (c_i + \xi_i)^T x_i - \Big(a^* + \zeta - B^* AX\Big)^T A_i x_i,$$

where $a^* = col\{a_1^*, \cdots, a_L^*\}$, $\zeta = col\{\zeta_1, \cdots, \zeta_L\}$, $B^* = diag\{b_1^*, \cdots, b_L^*\}$, and $\theta^* = (a^*, B^*)$ is unknown to all companies. Suppose firm $i \in \mathcal{N}$ has finite production capacity $X_i = \{x_i \in \mathbb{R}^{n_i} : 0 \leq x_i \leq \text{cap}_i\}$. In this networked Cournot competition, firm i minimizes $c_i^T x_i - (A_i x_i)^T a^* + (A_i x_i)^T B^* \sum_{i=1}^N A_i x_i$ over X_i. If $P(x)$ is defined as

$$P(x) \triangleq \sum_{i=1}^N c_i^T x_i - \left(\sum_{i=1}^N A_i x_i\right)^T a^* + \Big(col\,\{A_i x_i\}_{i=1}^N\Big)^T \chi \Big(col\,\{A_i x_i\}_{i=1}^N\Big),$$

where $\chi = \frac{1}{2}(\mathbf{I}_N + \mathbf{J}_N) \otimes B^*$. Then for any $i \in \mathcal{N}$ and for any $x_{-i} \in X_{-i}$, equation (6) holds for all $x_i, x_i' \in X_i$. Thus, the Nash-Cournot game admits a potential function $P(x)$. By definition of $f_i(x_i, x_{-i}; \theta^*)$, $\nabla_{x_i} f_i(x_i, x_{-i}; \theta^*)$ depends on $a_j^*, b_j^*, j \in \mathcal{L}$ if firm i sells its products to market j. Each firm i can observe the historic data about the aggregated sales S_j in market j and the price of products $p_j = a_j^* + \zeta_j - b_j^* S_j$ if the market j is connected to firm i. As such, firm i is able to learn the pricing parameters a_j^*, b_j^* of the connected market j through solving the following problem:

$$\min_{a_j \geq 0, b_j \geq 0} \mathbb{E}[(a_j - b_j S_j - p_j)^2] \tag{68}$$

In the numerical investigation, there are $V = 13$ firms to sell their products to $L = 7$ markets with the network shown in Figure 4. Suppose each component of cap_i and the cost pricing parameter c_i of the firm $i \in \mathcal{N}$ satisfy uniform distributions specified by $U[5, 8]$ and $U[2, 4]$. The pricing parameters a_j^*, b_j^* of market $j \in \mathcal{L}$ are drawn from uniform distributions $U[4, 6]$ and $U[0.2, 0.4]$, respectively. Suppose the random variables $\xi_i, i \in \mathcal{N}$ and $\zeta_j, j \in \mathcal{L}$ are drawn from uniform distributions $U[-c_i^*/8, c_i^*/8]$ and $U[-a_j^*/8, a_j^*/8]$, respectively. Suppose the historic aggregated sales S_j in market $j \in \mathcal{L}$ satisfies the uniform distribution $U[0, 5]$. For any $k \geq 0, i, j \in \mathcal{N}$, the communication delays $d_{ij}(k)$ are independently generated from a uniform distribution on the set $\{0, 1, \cdots, \tau\}$ with $\tau = 4$. We carry out simulations for Algorithm 3, where the inexact solution (52) satisfying Assumption 6 is

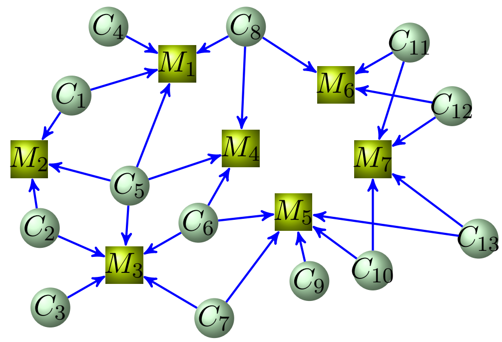
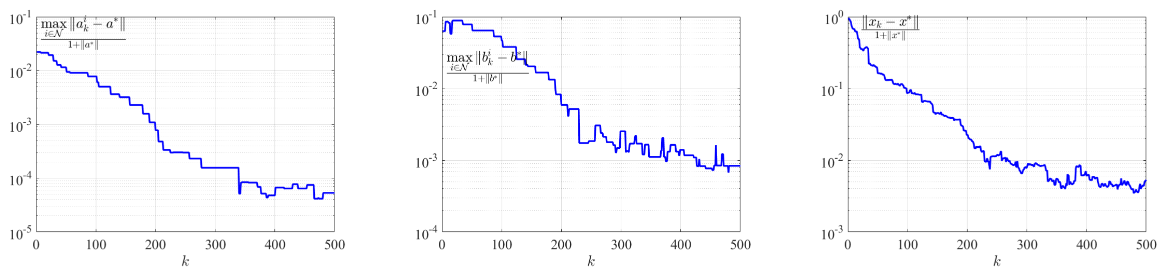


Figure 4: Networked Nash-Cournot: An edge from C_i to M_j implies firm C_i sells its products to market M_j.



(a) Learning a^*

(b) Learning b^*

(c) Learning x^*

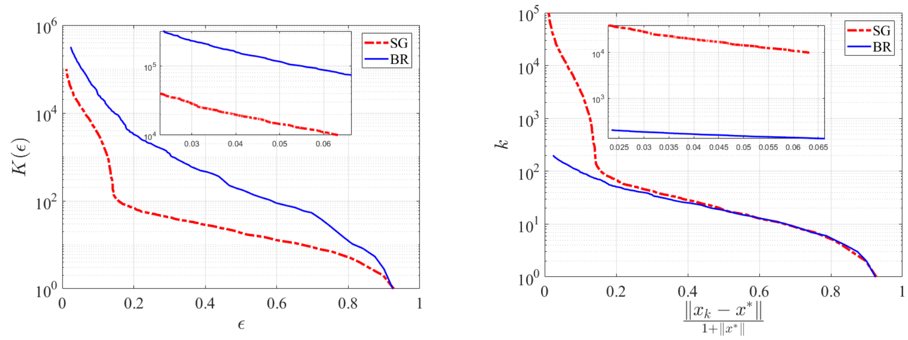
Figure 5: The estimates of a^*, b^* and x^* (a single sample path)

computed via the SA scheme (54) with $N_{i,k} = j_{i,k} = \lfloor \Gamma_{i,k}^3 \rfloor$, $p_i = 1/N \; \forall i \in \mathcal{N}$, $\beta = 0.1$ and $\mu = 5$. The scaled errors of learning schemes for the unknown parameters a^*, b^* and the Nash equilibrium x^* are provided in Figs. 5, where a_k^i and b_k^i denotes the estimates of a^* and b^* given by firm i at time k. The figure demonstrates the almost sure convergence of Algorithm 3.

Comparison with the asynchronous SG method: Suppose there are no communication delays among the players, i.e., $\tau = 0$. Set $p_i = 1/N \; \forall i \in \mathcal{N}$, $\beta = 0.1$ and $\mu = 5$. Let $N_{i,k} = \lfloor \Gamma_{i,k}^3 \rfloor$ steps of the SA scheme (54) be taken at major iteration k to obtain an inexact solution to (8), where $\Gamma_{i,k}$ is defined in Lemma 1. Set $j_{i,k} = \lfloor \Gamma_{i,k}^3 \rfloor$ in equation (52). We then carry out simulations for both Algorithm 3 and the asynchronous SG algorithm, which indeed is Algorithm 3 with equations (52) and (53) replaced by

$$
\begin{aligned}
x_{i,k+1} &= \Pi_{X_i}\left[x_{i,k} - \gamma_{i,k} \nabla_{x_i} \psi_i(x_k, \theta_{i,k}; \xi_{i,k})\right], \\
\theta_{i,k+1} &= \Pi_{\Theta}\left[\theta_{i,k} - \beta \gamma_{i,k} \nabla g\left(\theta_{i,k}, \eta_{i,k}\right)\right],
\end{aligned}
$$

where $\gamma_{i,k} = \frac{1}{\Gamma_{i,k}^{0.6}}$. We compare the two methods for the estimates of the equilibrium strategy x^* in terms of (i) the total number of the gradients steps (iteration complexity), and (ii) the communication overhead for achieving the same accuracy. Let $K(\epsilon)$ denotes the smallest total number of SG steps the players have carried out to make $\mathbb{E}\left[\frac{\|x_k - x^*\|}{1+\|x^*\|}\right] < \epsilon$. The empirically observed relationship between ϵ and $K(\epsilon)$ for both methods are shown in Figure 6(a), where the empirical errors are calculated by averaging across 50 trajectories. From the figure, it is seen that the iteration complexity are of the same orders while the constant of SG is superior to that of Algorithm 3. Since SG requires the rivals' newest information to carry out a single gradient step, the resulting communication overhead



(a) Iteration Complexity (b) Communication Overhead

Figure 6: Comparison of Stochastic Gradient Algorithm and Inexact BR Algorithm 3

is proportional to the total number of gradient steps. In contrast, Algorithm 3 carries out multiple gradient steps without requiring the most recent rivals' information. Further, the communication overhead of the two methods are shown in Figure 6(b). From the results in Figure 6, upon termination, Algorithm 3 requires approximately about 10 times more gradient steps than the standard SG method while characterized by approximately 500 times less communication overhead.

5 Concluding Remarks

This paper develops an asynchronous inexact proximal best-response scheme (combined with joint learning) to compute the Nash equilibrium of a stochastic potential Nash game (possibly corrupted by misspecification). When player-specific problems are convex, we show that the estimates generated by the proposed schemes converge almost surely to a connected subset of the NE set with uniformly bounded delays. Since the game is characterized by a possibly nonconvex potential function, the schemes can be viewed as randomized block coordinate descent schemes for a stochastic nonconvex optimization problem which is block-wise convex. We show that the gap function converges to zero in mean for both schemes as well. Furthermore, we prove almost sure convergence of the asynchronous inexact BR scheme in the delay-free regime when the player-specific problems are strongly convex. Finally, we apply the developed methods to the congestion control problem and the Nash-Cournot game, and demonstrate the simulation results.

References

[1] J. Lei and U. V. Shanbhag, "A randomized inexact proximal best-response scheme for potential stochastic Nash games," in *56th IEEE Conference on Decision and Control, CDC 2017*, 2017.

[2] J. F. Nash, Jr., "Equilibrium points in n-person games," *Proc. Nat. Acad. Sci. U. S. A.*, vol. 36, pp. 48–49, 1950.

[3] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[4] F. P. Kelly, A. K. Maulloo, and D. K. Tan, “Rate control for communication networks: shadow prices, proportional fairness and stability,” *Journal of the Operational Research society*, vol. 49, no. 3, pp. 237–252, 1998.

[5] E. Altman, Y. Hayel, and H. Kameda, “Evolutionary dynamics and potential games in non-cooperative routing,” in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. WiOpt 2007. 5th International Symposium on*. IEEE, 2007, pp. 1–5.

[6] A. Kannan and U. V. Shanbhag, “Distributed computation of equilibria in monotone Nash games via iterative regularization techniques,” *SIAM Journal on Optimization*, vol. 22, no. 4, pp. 1177–1205, 2012.

[7] P. Yi and L. Pavel, “A distributed primal-dual algorithm for computation of generalized Nash equilibria with shared affine coupling constraints via operator splitting methods,” *arXiv preprint arXiv:1703.05388*, 2017.

[8] C.-K. Yu, M. van der Schaar, and A. Sayed, “Distributed learning for stochastic generalized Nash equilibrium problems,” *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3893 – 3908, 2017.

[9] H. Jiang, U. V. Shanbhag, and S. P. Meyn, “Distributed computation of equilibria in misspecified convex stochastic Nash games,” *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 360–371, 2018.

[10] J. R. Marden, G. Arslan, and J. S. Shamma, “Cooperative control and potential games,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 6, pp. 1393–1407, 2009.

[11] G. Scutari, S. Barbarossa, and D. P. Palomar, “Potential games: A framework for vector power control problems with coupled constraints,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, vol. 4, 2006, pp. 241–244.

[12] D. González-Sánchez and O. Hernández-Lerma, “A survey of static and dynamic potential games,” *Science China Mathematics*, vol. 59, no. 11, pp. 2075–2102, 2016.

[13] O. Candogan, A. Ozdaglar, and P. A. Parrilo, “Dynamics in near-potential games,” *Games and Economic Behavior*, vol. 82, pp. 66–90, 2013.

[14] F. Facchinei, V. Piccialli, and M. Sciandrone, “Decomposition algorithms for generalized potential games,” *Computational Optimization and Applications*, vol. 50, no. 2, pp. 237–262, 2011.

[15] 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, 2013.

[16] F. Yousefian, A. Nedic, and U. V. Shanbhag, “A regularized smoothing stochastic approximation (RSSA) algorithm for stochastic variational inequality problems,” in *Simulation Conference (WSC), 2013 Winter*. IEEE, 2013, pp. 933–944.

[17] F. Yousefian, A. Nedić, and U. V. Shanbhag, “Self-tuned stochastic approximation schemes for non-Lipschitzian stochastic multi-user optimization and Nash games,” *IEEE Transactions on Automatic Control*, vol. 61, no. 7, pp. 1753–1766, 2016.

[18] A. P. Kirman, "Learning by firms about demand conditions," in *Adaptive economic models*. Elsevier, 1975, pp. 137–156.

[19] G. I. Bischi, A. K. Naimzada, and L. Sbragia, "Oligopoly games with local monopolistic approximation," *Journal of economic behavior & organization*, vol. 62, no. 3, pp. 371–388, 2007.

[20] G.-I. Bischi, L. Sbragia, and F. Szidarovszky, "Learning the demand function in a repeated cournot oligopoly game," *International Journal of Systems Science*, vol. 39, no. 4, pp. 403–419, 2008.

[21] F. Szidarovszky, "Global stability analysis of a special learning process in dynamic oligopolies," *Journal of Economic Research*, vol. 9, no. 2, pp. 175–190, 2004.

[22] D. Leonard and K. Nishimura, "Nonlinear dynamics in the Cournot model without full information," *Annals of Operations Research*, vol. 89, pp. 165–173, 1999.

[23] H. Jiang and U. V. Shanbhag, "On the solution of stochastic optimization and variational problems in imperfect information regimes," *SIAM Journal on Optimization*, vol. 26, no. 4, pp. 2394–2429, 2016.

[24] H. Ahmadi and U. V. Shanbhag, "On the resolution of misspecified convex optimization and monotone variational inequality problems," *arXiv preprint arXiv:1408.5532*, 2014.

[25] U. Ravat and U. V. Shanbhag, "On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games," *SIAM J. Optim.*, vol. 21, no. 3, pp. 1168–1199, 2011.

[26] B. D. Bernheim, "Rationalizable strategic behavior," *Econometrica*, vol. 52, no. 4, pp. 1007–28, 1984.

[27] D. G. Pearce, "Rationalizable strategic behavior and the problem of perfection," *Econometrica*, vol. 52, no. 4, pp. 1029–50, 1984.

[28] D. Fudenberg and D. K. Levine, *The theory of learning in games*, ser. MIT Press Series on Economic Learning and Social Evolution. Cambridge, MA: MIT Press, 1998, vol. 2.

[29] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1999, vol. 23.

[30] N. Nisan, M. Schapira, G. Valiant, and A. Zohar, "Best-response mechanisms." in *International Conference on Supercomputing (ICS)*, 2011, pp. 155–165.

[31] G. Scutari and D. P. Palomar, "Mimo cognitive radio: A game theoretical approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 761–780, 2010.

[32] J.-S. Pang, S. Sen, and U. V. Shanbhag, "Two-stage non-cooperative games with risk-averse players," *Mathematical Programming*, vol. 165, no. 1, pp. 235–290, 2017.

[33] F. Facchinei and J. S. Pang, *Nash equilibria: The variational approach*, ser. Convex Optimization in Signal Processing and Communications. Cambridge University Press (Cambridge, England), 2009.

[34] D. d'Esopo, "A convex programming procedure," *Naval Research Logistics Quarterly*, vol. 6, no. 1, pp. 33–42, 1959.

[35] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of optimization theory and applications*, vol. 109, no. 3, pp. 475–494, 2001.

[36] Y. Nesterov, "Efficiency of coordinate descent methods on huge-scale optimization problems," *SIAM Journal on Optimization*, vol. 22, no. 2, pp. 341–362, 2012.

[37] P. Richtárik and M. Takáč, "Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function," *Mathematical Programming*, vol. 144, no. 1-2, pp. 1–38, 2014.

[38] Y. Xu and W. Yin, "Block stochastic gradient iteration for convex and nonconvex optimization," *SIAM Journal on Optimization*, vol. 25, no. 3, pp. 1686–1716, 2015.

[39] D. Davis, "The asynchronous PALM algorithm for nonsmooth nonconvex problems," *ArXiv:1604.00526*, 2016.

[40] S. Ghadimi and G. Lan, "Accelerated gradient methods for nonconvex nonlinear and stochastic programming," *Mathematical Programming*, vol. 156, no. 1-2, pp. 59–99, 2016.

[41] S. Ghadimi, G. Lan, and H. Zhang, "Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization," *Mathematical Programming*, vol. 155, no. 1-2, pp. 267–305, 2016.

[42] J. Lei, U. V. Shanbhag, J.-S. Pang, and S. Sen, "On synchronous, asynchronous, and randomized best-response schemes for computing equilibria in stochastic Nash games," *arXiv preprint arXiv:1704.04578*, 2017.

[43] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash equilibria in constrained deterministic mean field control," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3315–3329, 2015.

[44] F. Parise, S. Grammatico, B. Gentile, and J. Lygeros, "Network aggregative games and distributed mean field control via consensus theory," *arXiv preprint arXiv:1506.07719*, 2015.

[45] J. R. Marden, "Learning in large–scale games and cooperative control," Ph.D. dissertation, UNIVERSITY OF CALIFORNIA Los Angeles, 2007.

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

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

[48] M. Zhu and E. Frazzoli, "Distributed robust adaptive equilibrium computation for generalized convex games," *Automatica*, vol. 63, pp. 82–91, 2016.

[49] T. Tatarenko, W. Shi, and A. Nedich, "Geometric convergence of gradient play algorithms for distributed Nash equilibrium seeking," *arXiv:1809.07383*, 2018.

[50] W. Shi and L. Pavel, "Lana: an ADMM-like Nash equilibrium seeking algorithm in decentralized environment," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 285–290.

[51] M. Ye and G. Hu, "Distributed Nash equilibrium seeking by a consensus based approach," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4811–4818, 2017.

[52] J. Lei and U. V. Shanbhag, "Distributed variable sample-size gradient-response and best-response schemes for stochastic nash games over graphs," *arXiv:1811.11246*, 2018.

[53] U. V. Shanbhag, J.-S. Pang, and S. Sen, "Inexact best-response schemes for stochastic Nash games: Linear convergence and iteration complexity analysis," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*, 2016, pp. 3591–3596.

[54] M. Grinblatt, S. Titman, and R. Wermers, "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior," *The American economic review*, pp. 1088–1105, 1995.

[55] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming*, ser. MPS/SIAM Series on Optimization. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2009, vol. 9.

[56] D. Fudenberg, Y. Ishii, and S. D. Kominers, "Delayed-response strategies in repeated games with observation lags," *Journal of Economic Theory*, vol. 150, pp. 487–514, 2014.

[57] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," *IEEE Transactions on Signal processing*, vol. 57, no. 7, pp. 2748–2761, 2009.

[58] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. SI, pp. 2508–2530, 2006.

[59] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 35–49, 2010.

[60] Y. S. Chow and H. Teicher, *Probability theory: independence, interchangeability, martingales*. Springer Science & Business Media, 2012.

[61] H. Robbins and D. Siegmund, "A convergence theorem for non negative almost supermartingales and some applications," in *Herbert Robbins Selected Papers*. Springer, 1985, pp. 111–135.

[62] T. Larsson and M. Patriksson, "A class of gap functions for variational inequalities," *Mathematical Programming*, vol. 64, no. 1-3, pp. 53–79, 1994.

[63] S. Dafermos, "Sensitivity analysis in variational inequalities," *Mathematics of Operations Research*, vol. 13, no. 3, pp. 421–434, 1988.

[64] S. Bubeck, "Convex optimization: Algorithms and complexity," *Foundations and Trends® in Machine Learning*, vol. 8, no. 3-4, pp. 231–357, 2015.

[65] R. H. Byrd, G. M. Chin, J. Nocedal, and Y. Wu, "Sample size selection in optimization methods for machine learning," *Mathematical programming*, vol. 134, no. 1, pp. 127–155, 2012.

[66] U. V. Shanbhag and J. H. Blanchet, "Budget-constrained stochastic approximation," in *Proceedings of the 2015 Winter Simulation Conference, Huntington Beach, CA, USA, December 6-9, 2015*, 2015, pp. 368–379.

[67] A. Jofré and P. Thompson, “On variance reduction for stochastic smooth convex optimization with multiplicative noise,” *Math. Program.*, vol. 174, no. 1-2, pp. 253–292, 2019.

[68] M. Schmidt, N. L. Roux, and F. R. Bach, “Convergence rates of inexact proximal-gradient methods for convex optimization,” in *Advances in neural information processing systems*, 2011, pp. 1458–1466.

[69] A. Jalilzadeh, U. V. Shanbhag, J. H. Blanchet, and P. W. Glynn, “Optimal smoothed variable sample-size accelerated proximal methods for structured nonsmooth stochastic convex programs,” *arXiv preprint arXiv:1803.00718*, 2018.

[70] T. Alpcan and T. Basar, “A game-theoretic framework for congestion control in general topology networks,” in *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, vol. 2. IEEE, 2002, pp. 1218–1224.