Probabilistic Analysis of Simulation-Based Games

YEVGENIY VOROBEYCHIK
Computer and Information Science
University of Pennsylvania
yev@seas.upenn.edu

The field of Game Theory has proved to be of great importance in modeling interactions between self-interested parties in a variety of settings. Traditionally, game theoretic analysis relied on highly stylized models to provide interesting insights about problems at hand. The shortcoming of such models is that they often do not capture vital detail. On the other hand, many real strategic settings, such as sponsored search auctions and supply-chains, can be modeled in high resolution using simulations. Recently, a number of approaches have been introduced to perform analysis of game-theoretic scenarios via simulation-based models. The first contribution of this work is the asymptotic analysis of Nash equilibria obtained from simulation-based models. The second contribution is to derive expressions for probabilistic bounds on the quality of Nash equilibrium solutions obtained using simulation data. In this vein, we derive very general distribution-free bounds, as well as bounds which rely on the standard normality assumptions, and extend the bounds to infinite games via Lipschitz continuity. Finally, we introduce a new maximum-a-posteriori estimator of Nash equilibria based on game-theoretic simulation data and show that it is consistent and almost surely unique.

Categories and Subject Descriptors: I.6.6 [**Simulation and Modeling**]: Simulation Output Analysis; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Game theory, Simulation and Modeling

Additional Key Words and Phrases: Game theory, simulation, Nash equilibrium

1. INTRODUCTION

In analyzing economic systems, especially those composed primarily of autonomous computational agents, the complexities must typically be distilled into a stylized model, with the hope that the resulting model captures the key components of the system and is nevertheless amenable to analytics. Increasingly, the boundaries of analytic tractability are being pushed, particularly as the systems, as well as the participating agents, are being more precisely engineered, and as the computational barriers that had once rendered complex mechanisms impractical are now coming down en masse. (Consider, for example, combinatorial auctions, which had in the past been shunned because of the complexity of the winner determination problem, but have now become ubiquitous in academic literature, as well as practice [Cramton et al. 2006].) The advantage of the increasingly precise models of Economic micro-systems, such as complex (e.g., combinatorial) auctions, is the ability to model these quite closely in a stylized way. The disadvantage, of course, is the added complexity in analyzing high-resolution models of strategic interactions.

Recently, a number of approaches have been introduced to analyzing game-theoretic models specified using simulations [Vorobeychik 2008; Wellman et al. 2005; Reeves 2005; Vorobeychik et al. 2006; Vorobeychik and Porche 2009; Wellman et al. 2008]. Many of these approaches focus on estimating Nash equilibria based on a finite number of simulation samples for a collection of joint strategic choices of the players. Often, an estimate of the payoff matrix is constructed from simulation data and a Nash equilibrium of the estimated game is used as an estimate of the actual Nash equilibrium. In this work, we focus on precisely such a setting, assuming (with one exception) that estimates of the payoff matrices of all players are available based on i.i.d. payoff samples for each profile of player strategies in the game. Intuitively, since the law of large numbers guarantees strong consistency of the estimated payoff matrices, we should see similar results in Nash equilibrium estimates based on these. Indeed, we obtain analogous convergence results in several senses in the context of Nash equilibrium estimation. Furthermore, we present the first (to our knowledge) analytical expressions for probabilistic bounds on Nash equilibrium estimates obtained from simulations, first by making no distributional assumptions on noise, and thereafter using the standard normality assumption. Finally, we present an alternative Nash equilibrium estimator which uses information about the distribution of noise. This estimator preserves the consistency property of data-based Nash equilibria, but is, in addition, almost surely unique when defined over the set of pure strategy profiles. Significantly, our proposed estimator provides, in a particular precise sense, the best Nash equilibrium estimates.

2. RELATED WORK

2.1 Simulation-Based Game Theoretic Analysis

Stylized analytic models of strategic interactions have been studied by game theorists for many years. For example, auction theory [Krishna 2002] has a deep understanding of a variety of auction mechanisms. Nevertheless, complex auction variants have recently emerged that are rather difficult to analyze in closed-form. For example, game-theoretic solutions to many interesting variants of combinatorial auctions that allow bidders to bid on any subset of a finite set of items have proved elusive. Furthermore, simulation models of various other strategic settings, such as supply chains [Collins et al. 2004] and combat [Lauren and Stephen 2002], are available and some effort has already been devoted to analyzing these (e.g., [Wellman et al. 2005; Vorobeychik and Porche 2009]). As yet, however, there is little statistical understanding of how game-theoretic solutions behave when the games are represented by simulations and constructed using Monte Carlo techniques (i.e., when payoffs are sampled from simulations for various joint strategic choices available to the players in the game). In this paper, we provide initial results, with a focus on finite games (games with finite sets of actions available to players), while entertaining several simple extensions to well-behaved infinite games.

2.2 Sample Average Approximation

It goes without saying that the results of this paper have the flavor of classical statistical analysis such as confidence bounds and estimator consistency. Similarly, there is a clear relationship, insofar as we are evaluating the probability that a particular strategic choice of a player is the best one available, to the literature on simulation optimization and probabilistic analysis of optimal outcomes of a finite set of choices [Kim and Nelson 2007] (indeed, a part of this work builds on Chang and Huang [2000] who provide confidence intervals for a best of N random choices under strong distributional assumptions).

This paper is also closely related to the literature on *sample average approximation (SAA)* [Shapiro 2001]. A superficial relationship to SAA is the same as to other simulation optimization techniques: a part of the problem of analyzing game-theoretic (Nash) equilibria is to assess each player's *incentives to deviate*, that is, the likelihood that the particular choice of each player is close to the optimal one. It turns out, however, that a deeper and more interesting relationship exists. One way to characterize equilibria is as global minima of a *game-theoretic regret* function, which, for any joint strategic choices of all players, measures the amount that each player can gain by selecting his truly optimal action (given that others stick to their strategy). In our setting, regret values are not available exactly, but can be estimated from simulations. As such, we can (and do) ask analogous questions about convergence of sample-based optima to those addressed in the SAA literature. A significant difference between the two settings is that in classic SAA analysis, the interest is on optimizing an expected value of a response function, with simulations yielding sample mean approximations of this expectation for every instance of a domain variable. In our case, we do not actually obtain samples of regret. Rather, the simulations yield *payoff samples*, and regret can be estimated by evaluating possible deviations from the prescribed strategic choices of all players. (A formal distinction between our setup and SAA can be found in Appendix A.) We use the structure of finite games to show that law of large numbers extends to the game-theoretic regret function (defined on an infinite mixed-strategy space) and use this result to recover results analogous to standard convergence [Kleywegt et al. 2001] and bounds [Shapiro and Homem-de-Mello 2001] of the sample average approximation technique.

3. SIMULATION-BASED AND EMPIRICAL GAMES

In this section we formalize the notion of strategic interactions (or games) between a set of rational agents. We denote a normal-form game by $\Gamma_N = [I, R, u(\cdot)]$, (the subscript N denotes a *normal-form* game, as compared to simulation-based games described below) where I is the (finite) set of players with $m = |I|$ the number of players (here $|I|$ denotes the cardinality of a finite set I), R the joint strategy set (pure or mixed, depending on context) with R_i the set of strategies of player i, and $u(\cdot)$ the function that maps joint strategy profiles $r \in R$ to a vector of payoff entries for all players, that is, $u(r) = \{u_1(r), \ldots, u_m(r)\}$, where $u_i(r)$ denotes the (deterministic) payoff function of player i. We assume that all payoffs $u_i(r)$ are finite. We use notation r_{-i} for a profile of strategies other than that of player i and $R_{-i} = \prod_{j \neq i} R_j$.

We let A_i denote the set of player i's pure strategies, with $A = A_1 \times \cdots \times A_m$ the joint pure strategy set. One may think of pure strategies as atomic actions of players in the game, such as a choice of a specific bid in an auction or a decision to confess in prisoners' dilemma. We denote by S_i the set of i's mixed strategies and use S to denote the joint mixed strategy set. Mixed strategies are probability distributions over pure (atomic) strategies. A key assumption is that each player selects an action to play according to s *independently* of selections made by all the other players. (For example, players cannot communicate at the time when they randomly draw their actions. An alternative concept of *correlated* strategies has also been studied [Osborne and Rubinstein 1994], although we do not deal with it here.) If $s \in S$ is a mixed strategy profile, then $s(a)$ denotes the probability that a pure strategy profile $a \in A$ is played under s. Analogously, $s_i(a_i)$ is the probability that player i's mixed strategy selects action a_i and $s_{-i}(a_{-i})$ is the probability that $a_{-i} \in A_{-i}$ is selected under the joint mixed strategy of players other than i. For $s \in S$, we define

$$u_i(s) = \sum_{a \in A} u_i(a)s(a).$$

In this paper, unless stated otherwise, we focus on finite games in the sense that the joint set of pure strategies A is finite. We also assume throughout that $A \subseteq R \subseteq S$.

For some of the results below, it will be necessary to have a specific topology of the mixed strategy space in mind. Having in mind that a mixed strategy profile is just a probability distribution over a finite set of action for each player, we let the set of mixed strategy profiles be $S \subset [0,1]^k$, where $k = \sum_{i \in I} |A_i|$. Thus, a mixed strategy profile $s \in S$ is just a vector of probabilities, with s_{ij} denoting the weight (probability) a player i places on the corresponding action j. Thus, if action sets are identical, we can view a mixed strategy profile as a matrix with players i as rows and actions j as columns. Naturally, since s must be a valid probability distribution, the constraint on $s \in S$ is that

$$\sum_{j \in A_i} s_{ij} = 1 \ \ \forall\, i \in I.$$

An important strategic element of a normal-form game is *game-theoretic regret* (or simply *regret*) of a profile $r \in R$, denoted by $\epsilon(r)$, which is the most any player can gain by deviating from r_i to any strategy in R_i. Formally,

$$\epsilon(r) = \max_{i \in I} \max_{r'_i \in R_i} u_i(r'_i, r_{-i}) - u_i(r) = \max_{i \in I} \max_{a_i \in A_i} u_i(a_i, r_{-i}) - u_i(r). \qquad (1)$$

REMARK 3.1. *To see that it is sufficient to take the maximum just over the set of pure strategies, let a_i^* be the strategy in A_i that yields the highest payoff to i and suppose that it is unique. Then for any s that doesn't put all its weight on a_i^* we can increase player i's payoff by shifting the weight from some suboptimal a'_i to a_i^*. In general, any optimal s^* can only have positive weights on optimal a_i^* which must necessarily yield identical payoffs. More generally, this would be true whenever R includes all pure strategy profiles, i.e., when $A \subset R$.*

Note that since the maximum above is taken over the set of pure strategies, if that set is finite, the computational complexity of it is insignificant (just a linear search, in principle, would do). While we do not deal with intractably large or infinite pure strategy sets directly in this work, the problem of approximating best responses and Nash equilibria in such settings is addressed elsewhere [Jordan et al. 2008; Vorobeychik and Wellman 2008].

In specifying a normal-form model of a strategic setting, the analyst has in mind predicting what players who are faced with decisions in such a strategic context will do. The outcomes of strategic interactions—that is, the ultimate decisions made by the players—are commonly assumed for the purposes of prediction to be *rational* (alternatively, *strategically stable*) in the sense that every player is playing optimally *given* the choices of other players. This notion of strategic stability is formalized as a *Nash equilibrium* solution concept for games.

DEFINITION 3.2. *A* Nash equilibrium *of the normal-form game* Γ_N *is a profile* $r \in R$ *such that for every player* $i \in I$*,*

$$u_i(r) \geq u_i(r'_i, r_{-i}) \quad \forall r'_i \in R_i.$$

We can alternatively define the Nash equilibrium concept using the game-theoretic regret: r is a Nash equilibrium if and only if $\epsilon(r) = 0$.

In an approximation context—for example, when a game is too large to compute a Nash equilibrium exactly, or when only payoff estimates are available—a common concept is that of an ϵ-Nash equilibrium.

DEFINITION 3.3. *An* ϵ-Nash equilibrium *of the normal-form game* Γ_N *is a profile* $r \in R$ *such that for every player* $i \in I$*,*

$$u_i(r) + \epsilon \geq u_i(r'_i, r_{-i}) \quad \forall r'_i \in R_i.$$

REMARK 3.4. *From the perspective of each player, the incentives to deviate may be smaller then* ϵ*. However, we utilize the worst-case bound in evaluating stability of a strategy profile to deviations by* any *player.*

We may again relate ϵ-Nash (or approximate Nash) equilibria to the regret function by noting that any profile $r \in R$ is an $\epsilon(r)$-Nash equilibrium. (Recent computational complexity results suggest that the problem of computing exact Nash equilibria is hard [Daskalakis et al. 2006], while Lipton et al. [2003] demonstrate that there exist approximate Nash equilibria with small support which can be computed in quasi-polynomial time.)

A wide variety of games analyzed in connection with practical domains (such as auctions) are modeled as *symmetric*, a notion we now define formally.

DEFINITION 3.5. *A* normal-form *game* $[I, R, u(\cdot)]$ *is* symmetric *if for all players* $i, j \in I$*,* $R_i = R_j$*, and* $u_i(r_i, r_{-i}) = u_j(r_j, r_{-j})$ *whenever* $r_i = r_j$ *and* $r_{-i} = r_{-j}$*. In simulation-based games, the latter requirement applies to expected payoffs, with expectation taken with respect to simulation noise.*

Informally, in a symmetric game all players have the same sets of strategies and their payoff functions are identical.

REMARK 3.6. *A symmetric game is more general than a coordination game, in which payoffs are the same for all players no matter what they play. A symmetric game merely requires that the payoffs of any two players are the same when they play an identical strategy and face identical strategies by the opponents.*

Since all payoff functions are identical, the vector of payoffs $u(\cdot)$ can be replaced by a single symmetric payoff function, which we denote by $v(\cdot)$. Additionally we use a_1 (analogously, r_1 and s_1) to denote a single player's strategy in a symmetric strategy profile a (r, s) and A_1 (R_1, S_1) to denote the symmetric strategy sets of players. Similarly, a_{m-1} (r_{m-1}, s_{m-1}) and A_{m-1} (R_{m-1}, S_{m-1}) denote a strategy profile and set of strategy profiles for $m - 1$ players. A well-known example of a symmetric game is a prisoner's dilemma. Significantly, every finite symmetric game is guaranteed to have a *symmetric* Nash equilibrium, which is a Nash equilibrium strategy profile r with the property that $r_i = r_j$ for all $i, j \in I$ (that is, all players play the same strategy) [Cheng et al. 2004]. Now, consider a symmetric strategy profile r in a symmetric game. In order to compute its game-theoretic regret (e.g., to verify if it is a Nash equilibrium), we need only check deviations for a single player, since utility functions of all players are identical. Note that if a game is symmetric, a regret of a *symmetric* strategy profile (that is, a strategy profile r in which all players play an identical strategy r_1) can be simplified slightly:

$$\epsilon(r) = \max_{a_1 \in A_1} v(a_1, r_{m-1}) - v(r).$$

Since symmetric profiles in symmetric games are especially clean and compelling (for example, they do not require an assumption that players coordinate who plays which strategy in an asymmetric profile), they frequently are the subject of game theoretic analysis in general [Krishna 2002], and simulation-based game theoretic analysis in particular [Vorobeychik and Wellman 2008].

In this paper we are interested in analyzing game-theoretic models in which player payoff functions are specified using simulations; we refer to these models as *simulation-based games*. A simulation-based game retains all the basic elements of the normal-form game, but makes the notion of payoff functions $u(\cdot)$ somewhat more precise in a way pertinent to analysis. Specifically, a payoff function in a simulation-based game is represented by an *oracle*, $\mathcal{O}$, which can be queried with any pure strategy profile $a \in A$ to produce a possibly noisy sample payoff vector U. In simulation-based games, we presume to have a simulation model of the player payoffs, and we must run this simulation to obtain a noisy sample of player payoffs for any fixed joint strategy profile. For example, we may imagine an agent-based combat simulation in which the analyst can set strategic parameters of the adversaries and run the simulation to obtain a sample outcome of a battle or a campaign. Other examples include simulation-based game theoretic analyses of supply-chains [Vorobeychik et al. 2006; Wellman et al. 2005] and simultaneous ascending auctions [Wellman et al. 2008].

Formally, a simulation-based game is denoted by $\Gamma_S = [I, R, \mathcal{O}]$ (the subscript S denotes a simulation-based game) where the oracle (simulator) $\mathcal{O}$ produces a sample vector of payoffs $U = (U_1, \ldots, U_m)$ to all players for a (typically pure) strategy profile r. We assume throughout that $E[U] = u(a)$, that is, the expected value of samples from the oracle is the actual (expected) payoff vector for the corresponding pure strategy profile.

REMARK 3.7. *Observe that there are two sources of randomness in games derived from simulations. One is the randomness in the Monte Carlo simulation experiments themselves, for example, cost and revenue from operating a supply chain given realized demand. The other is randomness that is a part of players'* mixed strategies. *The actual expected payoff to player i given a mixed strategy profile s involves the expectation taken with respect to* both *of these sources of noise. Typically, when we use the term "actual payoff", we mean actual expected payoff.*

We denote an estimate of a payoff to player i for profile a based on $n(a)$ i.i.d.

samples from $\mathcal{O}$ by

$$\hat{u}_{i,n}(a) = \frac{1}{n(a)} \sum_{j=1}^{n(a)} U_i(a)^j,$$

where each $U(a)^j$ is generated by invoking the oracle with profile a as input. The vector of payoff estimates for all players we denote by $\hat{u}_n(a)$ (or just $\hat{u}(a)$ where we would like to talk about any estimate of the payoff at a).

Using the estimates $\hat{u}(a)$, we can construct an estimated payoff matrix for a finite game from simulation data. We often call this estimated game an *empirical game* to allude to the empirical nature of its payoffs. Since an *empirical* or estimated game can be viewed as a normal-form game in its own right, all the concepts defined on normal-form games apply. Thus, we obtain *empirical regret* as

$$\hat{\epsilon}(r) = \max_i \max_{a_i \in A_i} \hat{u}_i(a_i, r_{-i}) - \hat{u}_i(r),$$

and, similarly, an *empirical* (ϵ-)Nash equilibrium is just a (ϵ-)Nash equilibrium of the empirical game.

One key distinction between $\mathcal{O}$ and $u(\cdot)$ is, thus, that $u(\cdot)$ is presumed to provide easy access to exact payoff evaluations, whereas $\mathcal{O}$ evaluates payoffs with noise. Another distinction, no less vital but somewhat more subtle, is that by specifying the payoff function as an oracle, we in practice resign ourselves to the fact that payoffs are not available in any easily tractable form and the game must of necessity be analyzed using simulation experiments. Thus, for example, even though the payoff function may have a closed-form specification, the Nash equilibria of the game (or some useful qualitative properties thereof) cannot be obtained using analytic means.

Implicit to the discussion of both the simulation-based games and empirical games is that they are defined with respect to an *underlying* game $\Gamma_{N \leftarrow S}$ characterized by the set of players I, a set of strategy profiles R, and the payoff function $u(\cdot)$ from which the oracle, in effect, is taking noisy samples. Given $\Gamma_{N \leftarrow S}$, the *true* regret of a profile $r \in R$, $\epsilon(r)$, in both the simulation-based and the empirical game is evaluated with respect to $u(\cdot)$ of this underlying game.

4. NASH EQUILIBRIUM ESTIMATION IN SIMULATION-BASED GAMES

Suppose that we are given a simulation-based game and our mission is to estimate its Nash equilibrium (or the entire set of Nash equilibria). Suppose further that every pure strategy profile $a \in A$ has been sampled at least once. Then the most direct—although not necessarily optimal—method for estimating Nash equilibria is the following:

(1) Estimate the payoff matrix of the game based on simulation data; let $\hat{u}(r)$ be such an estimate

(2) Numerically compute Nash equilibria of the estimated game (e.g., using the GAMBIT toolbox [McKelvey et al. 2005])

A level of sophistication can be added if we use variance reduction techniques rather than sample means to estimate payoffs of the underlying game [Reeves 2005;

Wellman et al. 2005]. For example, control variates, conditioning, and quasi-random sampling can achieve a considerable increase in sampling efficiency [Ross 2001; L'Écuyer 1994]. In this paper, however, we assume that we use sample means as payoff estimates—as described above.

Observe that the method for estimating Nash equilibria we just described uses only a part of the available information. Specifically, it does not use any information about the sampling noise. Intuitively, if such information can be used, we can perhaps reduce estimation variance. We follow this intuition to define a *maximum a posteriori (MAP) equilibrium estimator* in Section 7, which produces estimates based on a payoff prior, a noise distribution, and the simulation data.

5. CONSISTENCY RESULTS ABOUT NASH EQUILIBRIA IN EMPIRICAL GAMES

In much of what was described above, the set of Nash equilibria of the empirical game is used as the estimator of Nash equilibria of the underlying game. This seems intuitively to be a sensible approach, and we now confirm this by demonstrating convergence (in several senses) of sets of Nash equilibria computed in a finite empirical game to the set of equilibria of the underlying game. In this section we assume that all samples taken of the entire payoff matrix are independent of each other (that is, we obtain n i.i.d. samples for any fixed player i and profile a). However, we do not require that samples are independent for different entries in the payoff matrix. We use the notation Γ_n to refer to the game $[I, R, \{\hat{u}_{i,n}(\cdot)\}]$, whereas Γ denotes the underlying game, $[I, R, \{u_i(\cdot)\}]$, with $R = A$ (the set of pure strategy profiles) or $R = S$ (the set of mixed strategy profiles), depending on context below. Thus, $\hat{\epsilon}_n(r)$ is the (empirical) regret function computed with respect to the game Γ_n.

Our first result is that the empirical regrets converge uniformly on the joint mixed strategy space.

Theorem 5.1. *Suppose that $|I| < \infty, |A| < \infty$. $\hat{\epsilon}_n(s) \to \epsilon(s)$ a.s. uniformly on S.*

The proof of this theorem is in the appendix.

The implication of our first result is the next corollary, which suggests a slight modification of the Nash equilibrium estimator we have considered. Let $\mathcal{N}$ denote the set of all Nash equilibria of Γ. If we define $\mathcal{N}_{n,\gamma} = \{s \in S : \hat{\epsilon}_n(s) \leq \gamma\}$, we have the following corollary to Theorem 5.1:

Corollary 5.2. *Suppose that $|I| < \infty, |A| < \infty$. For every $\gamma > 0$, there is M such that $\forall n \geq M$, $\mathcal{N} \subset \mathcal{N}_{n,\gamma}$ a.s.*

Proof. Since $\epsilon(s) = 0$ for every $s \in \mathcal{N}$, we can find M large enough such that $\Pr\{\sup_{n \geq M} \sup_{s \in \mathcal{N}} \hat{\epsilon}_n(s) < \gamma\} = 1$. □

By the corollary, for any game with a finite set of pure strategies and for any $\epsilon > 0$, all Nash equilibria lie in the set of empirical ϵ-Nash equilibria if enough samples have been taken. This result suggests that if our goal is to estimate the set of Nash equilibria, we may in some settings effectively use the set of approximate empirical Nash equilibria. This estimation may be especially useful if the analyst is interested, for example, in the worst-case outcome of the entire set of equilibria, as is the case with *strong implementation* of social choice rules [Osborne and Rubinstein 1994].

Next, we show that when the number of samples is large enough, every Nash equilibrium of Γ_n is close to *some* Nash equilibrium of the underlying game. For the exposition that follows, we need a bit of additional notation. Let (Z, d) be a metric space, $X, Y \subset Z$, and define *directed Hausdorff distance* from X to Y to be

$$h_D(X, Y) = \sup_{x \in X} \inf_{y \in Y} d(x, y).$$

Observe that $U \subset X \Rightarrow h_D(U, Y) \leq h_D(X, Y)$. For the purpose of the next theorem, Z is the space of mixed strategy profiles $S \subset [0, 1]^k$, with k as defined in Section 3, and $d(s, s') = \|s - s'\|_\infty$ for $s, s' \in S$. (The specific choice of a norm in our case is not significant, since all norms are equivalent on a finite dimensional normed vector space.) Let $\mathcal{N}_n$ denote all Nash equilibria of the game Γ_n. We can then obtain the following general result, the proof of which is in the appendix.

THEOREM 5.3. *Suppose $|I| < \infty$ and $|A| < \infty$. Then $h_D(\mathcal{N}_n, \mathcal{N}) \to 0$ a.s.*

In words, every empirical Nash equilibrium is eventually arbitrarily close to *some* Nash equilibrium of the underlying game. Note that we have not proved the converse, and, indeed, it is not difficult to conceive of a counterexample. Consider, simply, a game with a constant payoff for every strategy profile. Every mixed strategy profile in this game is a Nash equilibrium. However, sampling noise would induce *generic* games, and it is well-known that the number of Nash equilibria in generic finite games is finite and, there are some mixed strategy profiles that are not close to any one of these. (More precisely, certain distributions of sampling noise—for example, Gaussian noise—would induce generic games. The term *generic* is used in Game Theory to denote a game in which a small change to any one of the payoffs does not introduce new Nash equilibria or remove existing ones [Webb 2006].)

The significance of the last result is that it does confirm a very important intuition that empirical Nash equilibria are reasonable Nash equilibrium estimates: the analyst who uses simulation-based game models is unlikely to be lead astray to spurious conclusions if enough samples are taken.

6. PROBABILISTIC BOUNDS ON EQUILIBRIUM APPROXIMATION QUALITY

The fact that we can demonstrate certain kinds of asymptotic convergence of empirical Nash equilibria affirms that these estimators are indeed reasonable. For practical purposes, however, that is not entirely sufficient: we would like also to establish statistical confidence in our results based on simulation data. Much of the past work on game-theoretic analysis using simulations had resorted to worst case sensitivity analysis procedures [Walsh et al. 2002] and indirect evidence [Wellman et al. 2005]. While these can be used to support certain claims, they are not entirely satisfactory, and it certainly seems possible to engage in statistical analysis of empirical equilibria given appropriate distributional assumptions. Reeves [2005] discusses sensitivity analysis procedure in the same spirit as our derivations which follow. His analysis involved sampling payoffs from a normal distribution centered at the sample mean and deriving an empirical probability distribution that particular strategies are played in an actual equilibrium.

In this section, we introduce a statistical framework for sensitivity analysis of solutions based on empirical games. Specifically, we derive probabilistic confidence bounds on the quality of empirical equilibria given various assumptions about the nature of noise and the structure of the underlying game.

6.1 Distribution-Free Bounds

We begin by deriving distribution-free bounds on the probability that the specified profiles are ϵ-Nash equilibria for some fixed ϵ. While most general, these bounds are likely the least useful of the bounds presented in this paper, as they will in practice rarely be very tight.

THEOREM 6.1. *Suppose that for a subset $\mathcal{A} \subset A$*

$$\Pr\left\{\max_{a\in\mathcal{A}} |u_i(a) - \hat{u}_i(a)| \geq \gamma\right\} \geq \delta$$

for all $i \in I$. Then

$$\Pr\{|\epsilon(s) - \hat{\epsilon}(s)| \geq 2\gamma\} \leq m(K+1)\delta,$$

for all $s \in S$ such that $\{a : s(a) > 0\} \subset \mathcal{A}$, where $K = \max_{i\in I} |A_i|$.

The proof of this result can be found in the appendix.

The bound above is rather abstract, and it is helpful to consider two concrete examples. First, suppose that the random variable $U_i(a)$ for the profile $a \in A$ has finite variance $\sigma^2 < \infty$. Then a simple application of Chebyshev's inequality combined with the union bound gives us $\delta = \frac{|\mathcal{A}|\sigma^2}{n\gamma^2}$. Similarly, if the random variable $U_i(a) \in [a, b]$, we obtain the above result with $\delta = |\mathcal{A}| \exp\left\{-\frac{\gamma^2 n}{2(b-a)^2}\right\}$. Perhaps the most interesting cases are (a) when $\mathcal{A} = \{a\}$, yielding a bound for a specific pure strategy profile, and (b) when $\mathcal{A} = A$, yielding the bound on any mixed strategy profile.

We can now use the result above to obtain a probabilistic bound on $\epsilon(r)$ by noting that

$$\Pr\{\epsilon(r) \geq \hat{\epsilon}(r) + 2\gamma\} = \Pr\{\epsilon(r) - \hat{\epsilon}(r) \geq 2\gamma\} \leq \Pr\{|\epsilon(r) - \hat{\epsilon}(r)| \geq 2\gamma\}. \quad (2)$$

6.2 Confidence Bounds for Finite Games with Normal Noise

Suppose that the game is finite and it is feasible to sample the entire payoff matrix of the game, with payoffs for every pure strategy profile sampled independently. Extension to using common random numbers is described in Section 6.4. Further, suppose that the game is symmetric and we are interested in probabilistic bounds on *symmetric* strategy profiles (that is, strategy profiles s in which all players play an identical strategies). Recall that we will use notation $v(a)$ to denote a payoff function in symmetric games, and denote each player's strategy in a symmetric strategy profile by r_1, while referring to identical sets of strategies by R_1. A subscript $m-1$ will denote a strategy profile and products of strategy sets for an arbitrary subset of $m-1$ players.

To derive a generic probabilistic bound for a pure strategy profile $a \in A$, suppose that we have an improper prior on $v(a)$ for all $a \in A$, and the sampling noise is Gaussian with known variance $\sigma^2(a)$ (i.e., variance may be different for different strategy profiles).

The results below build on the derivation of the distribution of the maximum of k variables based on samples of these with zero-mean additive Gaussian noise by Chang and Huang [2000], who demonstrate that if we start with an improper prior over the actual payoffs and observe samples distorted by Gaussian noise, the posterior distributions of $v(\cdot)|\hat{v}(\cdot)$ are Gaussian random variables with mean $\hat{v}(\cdot)$ and variance $\sigma^2(a)$. Furthermore, if payoffs are sampled independently, the actual (posterior) payoffs $v(\cdot)$ are also independently distributed, conditional on $\hat{v}(\cdot)$. Given these assumptions, the following general bound can derived (in the sequel we omit conditioning on $\hat{v}(\cdot)$ for brevity):

THEOREM 6.2. *Suppose that payoffs for every $a \in A$ are sampled independently with zero-mean additive Gaussian noise, and suppose that we have an improper prior on $v(a)$. Then, for any symmetric strategy profile $r \in R$ and $\epsilon \geq 0$,*

$$\Pr\left(\epsilon(r) \leq \epsilon\right) = \int_{\mathbb{R}} \prod_{b \in R_1 \setminus r_1} \Pr(v(b, r_{m-1}) \leq u + \epsilon) f_{v(r)}(u) du,$$

where $f_{v(r)}(u)$ is the pdf of $N(\hat{v}(r), \sigma(r))$.

PROOF.

$$\begin{aligned}
&\Pr\left(\max_{b \in R_1 \setminus r_1} v(b, r_{m-1}) - v(r) \leq \epsilon\right) \\
&= E_{v(r)} \Pr(\max_{b \in R_1 \setminus r_1} v(b, r_{m-1}) - v(r) \leq \epsilon | v(r)) \\
&= E_{v(r)} \left[\prod_{b \in R_1 \setminus r_1} \Pr(v(b, r_{m-1}) - v(r) \leq \epsilon | v(r))\right] \\
&= \int_{\mathbb{R}} \prod_{b \in R_1 \setminus r_1} \Pr(v(b, r_{m-1}) \leq u + \epsilon) f_{v(r)}(u) du.
\end{aligned}$$

□

The posterior distribution of the actual mean under the assumption of Gaussian noise was derived in Chang and Huang [2000]:

$$\Pr\left(v(a) \leq c\right) = 1 - \Phi\left[\frac{\sqrt{n(a)}(\hat{v}(a) - c)}{\sigma(a)}\right], \tag{3}$$

where $a \in A$ and $\Phi(\cdot)$ is $N(0,1)$ distribution function. Plugging this into the expression in Theorem 6.2 we get

$$\Pr\left(\epsilon(a) \leq \epsilon\right) = \int_{\mathbb{R}} \prod_{b \in A_1 \setminus a_1} \left(1 - \Phi\left[\frac{\sqrt{n(a)}(\hat{v}(a) - (u + \epsilon))}{\sigma(a)}\right]\right) f_{v(a)}(u) du \tag{4}$$

for any $\epsilon \geq 0$.

Having derived bounds on pure strategy profiles, it is not difficult to extend the results to bounds on mixed strategy profiles, as we do in the following theorem.

THEOREM 6.3. *Suppose that payoffs for every $a \in A$ are sampled independently with zero-mean additive Gaussian noise, and suppose that we have an improper* *prior on $v(a)$. Let $s \in S$ be a mixed strategy profile. Then,*

$$\Pr\left(\epsilon(s) \le \epsilon\right) = \int_{\mathbb{R}} \prod_{b \in A_1} \left[\Pr(W(b) \le u + \epsilon)\right] f_{W^*}(u) du,$$

where

$$\Pr(W(b) \le u + \epsilon) = 1 - \Phi\left[\frac{\sum_{c \in A_{m-1}} \hat{u}(b,c) s_{m-1}(c) - u - \epsilon}{\sqrt{\sum_{c \in A_{m-1}} \frac{\sigma^2(b,c)(s_1(c))^2}{n(b,c)}}}\right]$$

and

$$W^* \sim N\left(\sum_{a \in A} \hat{u}(a)s(a), \sum_{a \in A} \frac{\sigma^2(a)(s_1(a))^2}{n(a)}\right).$$

Proof. Since the posterior distribution of $v(a)$ for each a is Normal and since a mixed strategy profile $s \in S$ induces a linear combination of $v(a)$, the resulting payoffs are also distributed normally, where mean is just the linear combination of means, and variance is $\sum_{a \in A} \frac{\sigma^2(a)s^2(a)}{n(a)}$. This, in combination with Theorem 6.2 yields the result. □

6.3 Confidence Bounds for Infinite Games Based on Finite Game Restrictions

Suppose that we are trying to estimate a Nash equilibrium for a symmetric game $\Gamma = [I, R, v(\cdot)]$ with $R \subset \mathbb{R}^n$ infinite. Let $R^l \subset R$ be finite and define $\Gamma_l = [I, R^l, v(\cdot)]$ to be a finite restriction of Γ. To draw any conclusions about Γ based on its finite restriction we must make some assumptions about the structure of the actual payoff functions on the infinite domain. We assume that the payoff function $v(\cdot)$ satisfies the Lipschitz condition with Lipschitz coefficient B.

Define $d(R_1^l)$ to be the maximum distance from a point in R_1^l to its closest neighbor in R_1 for any player:

$$d(R_1^l) = \sup_{r_1 \in R_1} \inf_{r_1' \in R_1^l} \{\|r_1 - r_1'\|\} < \infty,$$

where $\|\cdot\|$ denotes Euclidean norm in $\mathbb{R}^n$. Then if r is an ϵ-Nash equilibrium of Γ_l with probability $1-\alpha$, then it is an $(\epsilon + Bd(R_1^l))$-Nash equilibrium of Γ with probability at least $1-\alpha$. Consequently, we have the following bound:

Theorem 6.4. *Suppose that payoffs for every $r \in R^l$ are sampled independently with zero-mean Gaussian noise, and suppose that we have an improper prior on $v(r)$. Furthermore, suppose that $v(r)$ are Lipschitz continuous with a coefficient at most B. Then,*

$$\Pr\left(\sup_{t \in R_1} v(t, r_{m-1}) - v(r) \le \epsilon\right) \ge \int_{\mathbb{R}} \prod_{t \in R_1^l} \Pr(v(t, r_{m-1}) \le u + \epsilon - Bd) f_{v(r)}(u) du.$$

PROOF.

$$\begin{aligned}
&\Pr\left(\sup_{t \in R_1} v(t, r_{m-1}) - v(r) \leq \epsilon\right) \\
&\geq \Pr\left(\max_{t \in R_1^l} v(t, r_{m-1}) - v(r) \leq \epsilon - Bd(R_1^l)\right) \\
&= \int_{\mathbb{R}} \prod_{t \in R_1^l} \Pr(v(t, r_{m-1}) \leq u + \epsilon - Bd(R_1^l)) f_{v(r)}(u) du.
\end{aligned}$$

□

6.4 Extension to Using Common Random Numbers

In the derivation of tight confidence bounds in Section 6.2 we had been working with a strong simplifying assumption that payoffs for all pure strategy profiles $a \in R$ are sampled independently. In some situations these assumptions are reasonable. For example, the game payoff function simulator maybe a binary executable provided by a third party (e.g., the Trading Agent Competition simulation-based analysis [Wellman et al. 2005]), and some variance reduction techniques that would introduce dependence (e.g., common random numbers) are difficult to implement. In general, however, this assumption entails a rather expensive sampling process, and it is quite desirable to use variance reduction techniques to reduce the simulation cost. One widely used and effective variance reduction method is to enforce common random numbers (CRN) in each replication. In our context, it would mean either (a) running a single simulation to obtain a *vector* of payoffs $u(a)$, but retain the assumption that each profile a is sampled independently, or (b) use a common random variables to obtain a sequence of samples of the entire payoff matrix.

While the convergence results and distribution-free bounds (Sections 5 and 6.1 respectively) do not require our independence assumptions (or symmetry assumptions) and hold even after introducing CRN, independence is required for our tight Gaussian bounds in Sections 6.2 and 6.3. We now generalize these results. For the foregoing discussion, we assume that we are interested in analyzing a symmetric strategy profile of a symmetric game with true (expected) payoff function $v(\cdot)$. The results can be generalized to non-symmetric games or strategy profiles in a straightforward way using, for example, Bonferroni inequality.

Let a be a pure strategy profile for which we want to obtain a regret bound. Select an arbitrary player i and fix a_{-i}. Since the number of pure strategies of player i is finite, we number them arbitrarily as $\{a_i^0, a_i^1, \ldots, a_i^L\}$, where $L = |A_i \setminus a_i|$ with $a_i^0 = a_i$ (the pure strategy played by i under a), and overload $\mathbf{a}_i$ to denote a vector of all pure strategies of player i. We let $v(\mathbf{a}_i, a_{-i})$ now be a vector of payoffs with each entry corresponding to $v(a_i^l, a_{-i})$ for $l = 0, \ldots, L$. Similarly, let $\hat{v}(\mathbf{a}_i, a_{-i})$ denote a vector of payoff estimates for $\mathbf{a}_i$.

As a first step, we note that given an improper prior on the actual payoff function $v(a)$, if the payoff estimate vector $\hat{v}(\mathbf{a}_i, a_{-i})$ has a multivariate normal distribution with mean vector $v(\mathbf{a}_i, a_{-i})$ and (known) covariance matrix $\Sigma(\mathbf{a}_i, a_{-i})$ with $\Sigma(a_i^{kl}, a_{-i})$ the covariance between payoffs of actions a_i^k and a_i^l, the posterior distribution of actual (mean) payoffs is also multivariate normal with mean $\hat{v}(\mathbf{a}_i, a_{-i})$ and covariance matrix $\Sigma(\mathbf{a}_i, a_{-i})/n$, where n is the number of samples taken to obtain the estimated vector $\hat{v}(\cdot)$ [DeGroot 2004]. Next, we use the distribution of the maximum of a multivariate normal distribution derived by Arellano-Valle and Genton [2008] to obtain the desired bound.

Since we are interested in obtaining a bound on $\epsilon(a) = \max_{b \in A_i \setminus a_i} v(b, a_{-i}) - v(a)$, it is easiest to work with the difference $\Delta v(\mathbf{a}_i, a_{-i}) = v(\mathbf{a}_i, a_{-i}) - v(a)$ instead of individual payoffs. Let $\Delta \hat{v}(\mathbf{a}_i, a_{-i}) = \hat{v}(\mathbf{a}_i, a_{-i}) - \hat{v}(a)$. Since $\Delta \hat{v}(a_i^0, a_{-i})$ is then just a constant 0, we let $\mathbf{a}_i$ in this notation be a truncated vector with strategy a_i omitted. Now, define

$$\mathbf{B} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \\ -1 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

where $\mathbf{B}$ has L rows, and let $\Delta\Sigma(\mathbf{a}_i, a_{-i}) = B\Sigma(\mathbf{a}_i, a_{-i})B^T$. This is the convariance matrix of the posterior of the difference $v(\mathbf{a}_i, a_{-i}) - v(a)$.

Let $\Delta \hat{v}(\mathbf{a}_i^{-l}, a_{-i})$ be the vector of sample mean payoff differences for actions other than a_i^l, let $\Delta\Sigma(\mathbf{a}_i^{ll}, a_{-i}) = \Delta\sigma^2(a_i^l, a_{-i})$ (i.e., the llth entry of $\Delta\Sigma(\mathbf{a}_i, a_{-i})$), define $\Delta\Sigma(\mathbf{a}_i^{-ll}, a_{-i}) = \Delta\Sigma(\mathbf{a}_i^{l-l}, a_{-i})$ to be covariance submatrices between payoff differences for action a_i^l and action vector $\mathbf{a}_i^{-l}$, and define $\Delta\Sigma(\mathbf{a}_i^{-l-l}, a_{-i})$ to be the covariance submatrix of the payoff difference for the action vector $\mathbf{a}_i^{-l}$. Further, define

$$\Delta \hat{v}(\mathbf{a}_i, a_{-i}) = \begin{pmatrix} \Delta \hat{v}(\mathbf{a}_i^{-l}, a_{-i}) \\ \Delta \hat{v}(a_i^l, a_{-i}) \end{pmatrix}$$

and

$$\Delta\Sigma(\mathbf{a}_i, a_{-i}) = \begin{pmatrix} \Delta\Sigma(\mathbf{a}_i^{-l-l}, a_{-i}) & \Delta\Sigma(\mathbf{a}_i^{-ll}, a_{-i}) \\ \Delta\Sigma(\mathbf{a}_i^{l-l}, a_{-i}) & \Delta\Sigma(\mathbf{a}_i^{ll}, a_{-i}) \end{pmatrix}.$$

Finally, define

$$\Delta \hat{v}(\mathbf{a}_i^{-l.l}, a_{-i}, x) = \Delta \hat{v}(\mathbf{a}_i^{-l}, a_{-i}) + (x - \Delta \hat{v}(a_i^l, a_{-i})) \frac{\Delta\Sigma(\mathbf{a}_i^{-ll}, a_{-i})}{\Delta\sigma_i^2(a_i^l, a_{-i})}$$

and

$$\Delta\Sigma(\mathbf{a}_i^{-l-l.l}, a_{-i}) = \Delta\Sigma(\mathbf{a}_i^{-l-l}, a_{-i}) - \frac{\Delta\Sigma(\mathbf{a}_i^{-ll}, a_{-i})\Delta\Sigma(\mathbf{a}_i^{-ll}, a_{-i})^T}{\Delta\sigma_i^2(a_i^l, a_{-i})}.$$

THEOREM 6.5. *Suppose that payoffs for every $a \in A$ are sampled with zero-mean additive Gaussian noise (not necessarily independently), and suppose that we have an improper prior on $v(a)$. Then, for any symmetric strategy profile $a \in A$,*

$$\Pr(\epsilon(a) \le \epsilon) = \int_{-\infty}^{\epsilon+u} \sum_{l=1}^{L} \phi_1^i(x)\Phi_{L-1}^{-i}(x)dx,$$

where $\phi_1^i(\cdot)$ and $\Phi_{L-1}^{-i}(\cdot)$ refer to a pdf of a single-variable normal distribution and a cdf of a multivariate normal distribution with $L-1$ variables respectively, where the former has mean $\Delta \hat{v}(a_i^l, a_{-i})$ and variance $\Delta\sigma_i^2(a_i^l, a_{-i})/n$, while the latter has mean (vector) $\Delta \hat{v}(\mathbf{a}_i^{-l.l}, a_{-i}, x)$ and covariance matrix $\Delta\Sigma(\mathbf{a}_i^{-l-l.l}, a_{-i})/n$.

Proof. Applying Corollary 4 from Arellano-Valle and Genton [2008] we have

$$\Pr\left(\max_{b\in A_1\setminus a_1}[v(b,a_{m-1})-v(a)]\leq\epsilon\right) = \int_{-\infty}^{\epsilon}\sum_{l=1}^{L}\phi_1^i(x)\Phi_{L-1}^{-i}(x)dx,$$

where $\phi_1^i(\cdot)$ and $\Phi_{L-1}^{-i}(\cdot)$ refer to a pdf of a single-variable normal distribution and a cdf of a multivariate normal distribution with $L-1$ variables respectively, where the former has mean $\Delta\hat{v}(a_i^l, a_{-i})$ and variance $\Delta\sigma_i^2(a_i^l, s_{-i})/n$, while the latter has mean (vector) $\Delta\hat{v}(\mathbf{a}_i^{-l.l}, s_{-i}, x)$ and covariance matrix $\Delta\Sigma(\mathbf{a}_i^{-l-l.l}, a_{-i})/n$. □

To complete the discussion, we simply note that if the actual pure strategy payoff matrix has a multivariate normal (posterior) distribution, extension to a distribution of payoffs for any mixed strategy profile (either s or (a_i, s_{-i}), where in the latter case player i plays a pure and others play a mixed strategy profile s_{-i}) is direct and yields a multivariate normal distribution, since a mixed strategy applies an affine transformation on the random payoff matrix.

6.5 Discussion

From a practical perspective, the most useful bounds are those derived with the assumption that payoffs are sampled with Gaussian noise (Sections 6.2 and 6.4). While the distribution-free bounds are quite general in principle, they are typically quite loose. More significantly, they can typically be applied to finite games only (or infinite games with strong assumptions about the actual payoff functions), and, additionally, they become less useful as the game size increases, since the bounds on mixed strategy profiles are linear in the size of the set of all pure strategy profiles. The bounds in Sections 6.2 and 6.4 are, in contrast, tight. Additionally, offline experiments with non-Gaussian noise and estimated (rather than known) variance suggest that neither assumption is very critical in practice. Indeed, a variation on the infinite-game bounds described in Section 6.3 has been used for probabilistic assessment of estimated equilibrium outcomes in Vorobeychik et al. [2006].

A central assumption of all of the above discussion has been that we observe samples of every pure strategy profile in the game. By the very nature of typical simulation-based games, we would be required to run separate simulations for every strategy profile, so running time is bounded from below by the size of the game. If we suppose that a game consists of m players, each with L strategies, the game size (the number of strategy profiles) is m^L. In the previous section we introduced symmetric games, which allow substantial reduction in the the size of the game: a game with m players with L strategies each now includes "only" $\binom{m+L+1}{m}$ distinct strategy profiles. Thus, for example, a symmetric game with 5 players, 10 strategies, would have 4368 strategy profiles, whereas an asymmetric with the same parameters explodes to $9,765,625$.

Besides the sheer size of the game, a significant factor in practice is the actual simulation running time, which may range from a fraction of a second [Vorobeychik and Wellman 2008] to nearly an hour [Wellman et al. 2005] for a single simulation experiment. Typically, even with the use of variance reduction techniques, it is not feasible to simulate the entire game. However, it is also not strictly necessary to do so in order to apply the bounds above. For example, suppose that a pure symmetric strategy profile a has been identified in a symmetric game as a likely candidate Nash equilibrium estimator. In order to probabilistically bound its regret, we need only to simulate single-player deviations from a. In this case, the number of players is irrelevant, and the number of strategies can be quite large. The simulation effort required to bound regret of mixed strategy profiles may be substantially greater, although here too we need not require samples for the entire game: if mixed strategies have *low support* (few pure strategies with strictly positive probability), only a relatively small submatrix of the game needs to be explored.

7. MAP SYMMETRIC EQUILIBRIUM ESTIMATION

To this point, the techniques for estimating Nash equilibria based on empirical games have involved using either Nash equilibria of the empirical game or profiles with low empirical regret. Neither of these techniques, however, takes into account available distributional information about the game. Intuitively, such information could be of great value, as we can use it to establish precise bounds on the probability that each profile is a Nash (or an ϵ-Nash) equilibrium of the underlying game, as shown in Section 6. We now attempt to utilize such distributional information in defining another equilibrium estimator, one which uses a profile most likely to be a Nash equilibrium of the underlying game, although restrict our analysis to symmetric games and symmetric strategy profiles. While certainly not the only method for using distributional information in the empirical game to estimate Nash equilibria, this method has some appealing properties which we demonstrate below. (An alternative, closely related method, was suggested by Jordan et al. [2008]. Their method uses the profile which is most likely to have the smallest $\epsilon(r)$ as the estimate of the best approximate Nash equilibrium.) We note that the definition of the estimator and the derivation of its properties require the strong independence assumptions of Section 6.2.

Assume now, as in much of this paper, that the sets of pure strategies of all players are finite, that is, $|A| < \infty$. Furthermore, assume that we have a symmetric game and are interested in estimating a symmetric Nash equilibrium. Under the assumption of an improper prior on actual payoffs and Gaussian noise with known finite variance, we derived above exact probabilistic bounds on ϵ-Nash equilibria of symmetric pure and mixed strategy profiles. By setting the ϵ in these expressions to 0, the bounds describe the probability that a profile r is a Nash equilibrium.

For a symmetric profile r define

$$P_n(r) = \Pr\{\epsilon(r) = 0\} = \int_{\mathbb{R}} \prod_{b \in R_1 \setminus r_1} \left(1 - \Phi\left[\frac{\sqrt{n(r)}(\hat{v}(r) - u))}{\sigma(r)}\right]\right) f_{v(r)}(u)du, \quad (5)$$

where $v(r)$ is the symmetric payoff function and r_1 and R_1 are the symmetric strategy and strategy set of each player. Since every player plays an identical strategy under a symmetric profile r, we focus our attention on the strategy $r_1 \in R_1$ from which the symmetric profiles are composed. We now define the maximum a posteriori (MAP) Nash equilibrium estimator $\hat{r}_{MAP}$ to be

$$\hat{r}_{MAP} = \arg\max_{r_1 \in R_1} P_n(r_1), \quad (6)$$

where we abuse notation slightly by overloading $P_n(r_1)$ to mean $P_n(r)$ with r a symmetric strategy profile in which each player plays a strategy r_1. For this expression to be well-defined, we need to ensure that the maximum actually exists. If R_1 is finite, that is entirely obvious. Let us take a more general case where R_1 is the set of mixed strategies S_1. The following lemma takes us most of the way.

LEMMA 7.1. *Suppose that $\sigma^2(a) > 0$ for all $a \in A$. Then $P_n(s_1)$ is continuous on $s_1 \in S_1$.*

Since this and other results in this section are rather involved, we relegate the proofs to the online companion. Based on this result, we can now readily confirm that the expression for the MAP estimator is well-defined.

THEOREM 7.2. *Suppose that $\sigma_i^2(a) > 0$ for all $i \in I$, $a \in A$. Then the maximum in Equation (6) exists when $R_1 = S_1$.*

PROOF. By Lemma 7.1, $P_n(s_1)$ is continuous on S_1. Since S_1 is a closed and bounded subset of a $\mathbb{R}^{|A_1|}$ (being a simplex), it is compact in $\mathbb{R}^{|A_1|}$. By Weierstrass theorem, the maximum exists on S_1. □

COROLLARY 7.3. *The maximum in Equation 6 exists when R_1 is a closed subset of S_1, with the case of finite $R_1 \subset S_1$ being a (trivial) special case.*

PROOF. This follows because a closed subset of a compact set is compact. □

Thus, while it may well be computationally challenging to obtain the exact estimate $\hat{r}_{MAP}$, we may be able to approximate it well by restricting the search, for example, to the set of pure symmetric strategy profiles.

Having verified that the MAP estimator is well-defined, we demonstrate that the estimate which it produces for a particular underlying game is *almost surely unique* when restricted to the space of pure strategy profiles.

Suppose we have a game in which every pure strategy profile has been sampled at least n times. The following theorem states that with probability 1 there are no two profiles $a, a' \in A$ with the same value of $P_n(a)$.

THEOREM 7.4. *Let $P_n(r)$ be as defined in Equation 5 and $r \in R = A$, and suppose that the noise distribution of the payoff samples is absolutely continuous with respect to Lebesgue measure. Then there are no two profiles $a, a' \in S$ such that $P_n(a) = P_n(a')$.*

COROLLARY 7.5. *Let $P(a)$ be as defined in Equation 5 and suppose that the noise distribution of the payoff samples is absolutely continuous with respect to Lebesgue measure. The estimate $\hat{a}_{MAP} = \max_{a_1 \in A_1} P_n(a)$ is almost surely unique.*

We now ascertain that the MAP estimator is consistent in a certain sense, when the maximum is taken over a (sub)set of pure strategies. Before formally stating the consistency result, we first prove the two lemmas about consistency of the actual probability of zero-regret for equilibrium and non-equilibrium profiles. Although these lemmas may be of independent interest, our primary goal is to use them in proving the consistency of the MAP estimator below.

LEMMA 7.6. *Suppose that a symmetric strategy profile s is not a Nash equilibrium and variance of noise for every strategy profile is strictly positive. Then $P_n(s) \to 0$ a.s.*

Lemma 7.7. *Suppose a symmetric strategy profile a is a pure strategy Nash equilibrium of a generic game (that is, pure strategy payoff matrix has no ties). Then* $P_n(a) \to 1$ *a.s.*

These lemmas imply the following theorem.

Theorem 7.8. *Suppose that the game is generic and* $\hat{a}_{MAP} = \max_{a_1 \in A_1} P_n(a)$ *is a MAP estimator (6), and suppose that the game has at least one pure strategy Nash equilibrium profile. Then there is a symmetric pure strategy Nash equilibrium* a^* *such that each player's strategy is* $\hat{a}_{MAP} = a_1^*$ *after a finite number of steps (and forever after).*

Proof. Fix $\epsilon > 0$. Since Lemma 7.6 guarantees that any profile a which is not a Nash equilibrium will have $P_n(a) \leq \epsilon$ after a finite number of steps and Lemma 7.7 ensures that any Nash equilibrium profile a' will have $P_n(a') \geq 1 - \epsilon$ after a finite number of steps, we need only set $\epsilon < 1/2$ to obtain a separation between equilibrium and non-equilibrium profiles after a finite number of steps. □

8. CONCLUSION

Recently, a number of approaches have been introduced to perform analysis of game-theoretic scenarios via simulation-based models. We contribute to this line of research by presenting a statistical analysis of game-theoretic solution estimates obtained using simulations. Specifically, we provide an asymptotic convergence analysis of Nash equilibria of empirically derived games and present expressions for probabilistic bounds on the quality of Nash equilibrium approximations given simulation data. In this vein, we derive very general distribution-free bounds, as well as bounds which rely on the standard normality assumptions, and extend our bounds to infinite games via Lipschitz continuity. Finally, we introduce a new maximum-a-posteriori estimator of Nash equilibria based on game-theoretic simulation data and show that it is consistent and almost surely unique when defined over the set of symmetric pure strategy profiles.

ACKNOWLEDGMENTS

I am truly indebted to the reviewers for their time and the numerous thoughtful comments on previous drafts. I also thank Michael Wellman, Christopher Kiekintveld, and Patrick Jordan for their input at very early stages of this research and for their ideas that have left an indelible imprint on it.

REFERENCES

Arellano-Valle, R. B. and Genton, M. G. 2008. On the exact distribution of the maximum of absolutely continuous dependent random variables. *Statistics and Probability Letters 78*, 27–35.

Chang, Y.-P. and Huang, W.-T. 2000. Generalized confidence intervals for the largest value of some functions of parameters under normality. *Statistica Sinica 10*, 1369–1383.

Cheng, S.-F., Reeves, D. M., Vorobeychik, Y., and Wellman, M. P. 2004. Notes on equilibria in symmetric games. In *AAMAS-04 Workshop on Game-Theoretic and Decision-Theoretic Agents.* New York.

Collins, J., Arunachalam, R., Sadeh, N., Eriksson, J., Finne, N., and Janson, S. 2004. The supply chain management game for the 2005 trading agent competition. Tech. rep., Carnegie Mellon University.

Cramton, P., Shoham, Y., and Steinberg, R., Eds. 2006. *Combinatorial Auctions*. MIT Press.

Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. 2006. The complexity of computing a Nash equilibrium. In *Thirty-Eighth ACM Symposium on Theory of Computing*. 71–78.

DeGroot, M. 2004. *Optimal Statistical Decisions*. Wiley-Interscience.

Jordan, P., Vorobeychik, Y., and Wellman, M. P. 2008. Searching for approximate equilibria in empirical games. In *Seventh International Joint Conference on Autonomous Agents and Multiagent Systems*. 1063–1070.

Keener, R. 2004. *Statistical Theory: A Medley of Core Topics*. University of Michigan Department of Statistics.

Kim, S.-H. and Nelson, B. L. 2007. Recent advances in ranking and selection. In *Proceedings of the 2007 Winter Simulation Conference*, S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, Eds. Piscataway, NJ: IEEE, Inc, 162–172.

Kleywegt, A. J., Shapiro, A., and Homem-de-Mello, T. 2001. The sample average approximation method for stochastic discrete optimization. *SIAM Journal of Optimization 12,* 2, 479–502.

Krishna, V. 2002. *Auction Theory*. Academic Press.

Lauren, M. K. and Stephen, R. T. 2002. Map-aware non-uniform automata (MANA): A New Zealand approach to scenario modeling. *Journal of Battlefield Technology 5,* 1, 27–31.

L'Écuyer, P. 1994. Efficiency improvement and variance reduction. In *Proceedings of the 1994 Winter Simulation Conference*, J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila, Eds. Piscataway, NJ: IEEE, Inc, 122–132.

Lipton, R. J., Markakis, E., and Mehta, A. 2003. Playing large games using simple strategies. In *Fourth ACM Conference on Electronic Commerce*. 36–41.

McKelvey, R. D., McLennan, A. M., and Turocy, T. L. 2005. Gambit: Software tools for game theory, version 0.2005.06.13.

Osborne, M. J. and Rubinstein, A. 1994. *A Course in Game Theory*. MIT Press.

Reeves, D. M. 2005. Generating trading agent strategies: Analytic and empirical methods for infinite and large games. Ph.D. thesis, University of Michigan.

Ross, S. M. 2001. *Simulation*, 3rd ed. Academic Press.

Shapiro, A. 2001. Monte Carlo simulation approach to stochastic programming. In *Proceedings of the 2001 Winter Simulation Conference*, P. Farrington, H. Nembhard, D. Surrock, and G. Evans, Eds. Piscataway, NJ: IEEE, Inc, 428–431.

Shapiro, A. and Homem-de-Mello, T. 2001. On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal of Optimization 11,* 1, 70–86.

Vorobeychik, Y. 2008. Mechanism design and analysis using simulation-based game models. Ph.D. thesis, University of Michigan.

Vorobeychik, Y., Kiekintveld, C., and Wellman, M. P. 2006. Empirical mechanism design: methods, with an application to a supply chain scenario. In *Seventh ACM Conference on Electronic Commerce*. 306–315.

Vorobeychik, Y. and Porche, I. 2009. Game theoretic methods for analysis of combat simulations. Working paper.

Vorobeychik, Y. and Wellman, M. P. 2008. Stochastic search methods for Nash equilibrium approximation in simulation-based games. In *Seventh International Conference on Autonomous Agents and Multiagent Systems*. 1055–1062.

Walsh, W. E., Das, R., Tesauro, G., and Kephart, J. O. 2002. Analyzing complex strategic interactions in multi-agent systems. In *AAAI-02 Workshop on Game Theoretic and Decision Theoretic Agents*.

Webb, J. N. 2006. *Game Theory: Decisions, Interaction, and Evolution*. Springer.

Wellman, M. P., Estelle, J. J., Singh, S., Vorobeychik, Y., Kiekintveld, C., and Soni, V. 2005. Strategic interactions in a supply chain game. *Computational Intelligence 21,* 1, 1–26.

Wellman, M. P., Osepayshvili, A., MacKie-Mason, J. K., and Reeves, D. M. 2008. Bidding strategies for simultaneous auctions. *B.E. Journal of Theoretical Economics (Topics) 8,* 1.

Appendix

A. FORMAL COMPARISON WITH SAA

In the related work section we informally explained the distinction between our work and sample average approximation. We now make the comparison formally. The problem in the SAA setup is to approximate

$$\min_{r \in R} E[U(r, X)],$$

where U is some target function of r and X a random variable, with

$$\min_{r \in R} \frac{1}{n} \sum_{j=1}^{n} U(r, X_j).$$

In our case, the problem is (essentially) to approximate

$$\min_{r \in R} \max_{i \in I} \max_{a_i \in A_i} E[U_i(a_i, r_{-i}, X) - U_i(r, X)]$$

with

$$\min_{r \in R} \max_{i \in I} \max_{a_i \in A_i} \frac{1}{n} \sum_{j=1}^{n} [U_i(a_i, r_{-i}, X_j) - U_i(r, X_j)].$$

B. PROOF OF THEOREM 5.1

First, we need the following straightforward fact, the proof of which we omit:

Claim B.1. *Let X be compact and $f_i(x)$ continuous on X. Then*

$$|\max_{x \in X} f_1(x) - \max_{x \in X} f_2(x)| \leq \max_{x \in X} |f_1(x) - f_2(x)|.$$

Proof of Theorem 5.1. By the Strong Law of Large Numbers, $\hat{u}_{i,n}(a) \to u_i(a)$ a.s. for all $i \in I, a \in A$. That is, $\Pr\{\lim_{n \to \infty} \hat{u}_{i,n}(a) = u_i(a)\} = 1$, or, equivalently [Keener 2004], for any $\alpha > 0$ and $\delta > 0$, there is $M(i,a) > 0$ such that

$$\Pr\left\{ \sup_{n \geq M(i,a)} |\hat{u}_{i,n}(a) - u_i(a)| < \frac{\delta}{2|A|} \right\} \geq 1 - \alpha.$$

By taking $M = \max_{i \in I} \max_{a \in A} M(i,a)$,[1] we have

$$\Pr\left\{ \max_{i \in I} \max_{a \in A} \sup_{n \geq M} |\hat{u}_{i,n}(a) - u_i(a)| < \frac{\delta}{2|A|} \right\} \geq 1 - \alpha.$$

[1] Note that since A is finite, M is finite.

Thus, by the claim, for any $n \geq M$,

$$\sup_{n\geq M} |\hat{\epsilon}_n(s) - \epsilon(s)| \leq$$

$$\max_{i\in I, a_i\in A_i} \sup_{n\geq M} |\hat{u}_{i,n}(a_i, s_{-i}) - u_i(a_i, s_{-i})| + \max_{i\in I} \sup_{n\geq M} |\hat{u}_{i,n}(s) - u_i(s)| \leq$$

$$\max_{i,a_i} \sum_{b\in A_{-i}} \sup_{n\geq M} |\hat{u}_{i,n}(a_i, b) - u_i(a_i, b)| s_{-i}(b) + \max_i \sum_{b\in A} \sup_{n\geq M} |\hat{u}_{i,n}(b) - u_i(b)| s(b) \leq$$

$$\max_{i,a_i} \sum_{b\in A_{-i}} \sup_{n\geq M} |\hat{u}_{i,n}(a_i, b) - u_i(a_i, b)| + \max_i \sum_{b\in A} \sup_{n\geq M} |\hat{u}_{i,n}(b) - u_i(b)| <$$

$$\max_{i\in I} \max_{a_i\in A_i} \sum_{b\in A_{-i}} \left(\frac{\delta}{2|A|}\right) + \max_{i\in I} \sum_{b\in A} \left(\frac{\delta}{2|A|}\right) \leq \delta$$

with probability at least $1 - \alpha$. Note that since $s_{-i}(a)$ and $s(a)$ are bounded between 0 and 1, we were able to drop them from the expressions above to obtain a bound that will be valid independent of the particular choice of s. Furthermore, since the above result can be obtained for an arbitrary $\alpha > 0$ and $\delta > 0$, we have $\Pr\{\lim_{n\to\infty} \hat{\epsilon}_n(s) = \epsilon(s)\} = 1$ uniformly on S. □

C. PROOF OF THEOREM 5.3

We first note that the function $\epsilon(s)$ is continuous in a finite game.

LEMMA C.1. *Let S be a mixed strategy set defined on a finite game. Then $\epsilon : S \to \mathbb{R}$ is continuous.*

We prove the result using uniform continuity of $u_i(s)$ and preservation of continuity under maximum. We omit the simple proofs of the claims below.

CLAIM C.2. *A function $f : \mathbb{R}^k \to \mathbb{R}$ defined by $f(t) = \sum_{i=1}^k z_i t_i$, where z_i are constants in $\mathbb{R}$, is uniformly continuous in t.*

CLAIM C.3. *Let $f(a, b)$ be uniformly continuous in $b \in B$ for every $a \in A$, with $|A| < \infty$. Then $V(b) = \max_{a\in A} f(a, b)$ is uniformly continuous in b.*

PROOF OF LEMMA C.1. Now, recall that $\epsilon(s) = \max_i[\max_{a_i\in A_i} u_i(a_i, s_{-i}) - u_i(s)]$. By the claims C.2 and C.3, $\max_{a_i\in A_i} u_i(a_i, s_{-i})$ is uniformly continuous in s_{-i} and $u_i(s)$ is uniformly continuous in s. Since the difference of two uniformly continuous functions is uniformly continuous, and since this continuity is preserved under maximum by our second claim, we have the desired result. □

Define $B_S(x, \delta)$ to be an open ball in $S \subset [0, 1]^k$ with center $x \in S$ and radius δ. Let

$$\mathcal{N}_\delta = \bigcup_{x\in\mathcal{N}} B_S(x, \delta),$$

that is, the union of open balls of radius δ with centers at Nash equilibria of Γ.

CLAIM C.4. *$\bar{\epsilon} = \min_{s\in S\setminus\mathcal{N}_\delta} \epsilon(s)$ exists and $\bar{\epsilon} > 0$.*

PROOF. Since $\mathcal{N}_\delta$ is an open subset of compact S, it follows that $S \setminus \mathcal{N}_\delta$ is compact. As we had also proved in Lemma C.1 that $\epsilon(s)$ is continuous, existence follows from the Weierstrass theorem. That $\bar{\epsilon} > 0$ is clear since $\epsilon(s) = 0$ if and only if s is a Nash equilibrium of Γ. □

PROOF OF THEOREM 5.3. Now, by Theorem 5.1 and Claim C.4, for any $\alpha > 0$ there is M such that

$$\Pr\{\sup_{n \geq M} \sup_{s \in S} |\hat{\epsilon}_n(s) - \epsilon(s)| < \bar{\epsilon}\} \geq 1 - \alpha.$$

Consequently, for any $\delta > 0$,

$$\Pr\{\sup_{n \geq M} h_D(\mathcal{N}_n, \mathcal{N}_\delta) = 0\} = \Pr\{\forall n \geq M \ \mathcal{N}_n \subset \mathcal{N}_\delta\} \geq$$
$$\Pr\{\sup_{n \geq M} \sup_{s \in \mathcal{N}_n} \epsilon(s) < \bar{\epsilon}\} \geq \Pr\{\sup_{n \geq M} \sup_{s \in S} |\hat{\epsilon}_n(s) - \epsilon(s)| < \bar{\epsilon}\} \geq 1 - \alpha.$$

Since this holds for an arbitrary $\alpha > 0$ and $\delta > 0$ and since $h_D(\mathcal{N}_\delta, \mathcal{N}) = \delta$, the desired result follows. □

D. PROOF OF THEOREM 6.1

The first lemma provides a (deterministic) bound on $\epsilon(r)$, where r is some strategy profile in the set of joint strategies, given that we have a bound on the quality of the payoff function approximation for every point in the domain.

LEMMA D.1. *Let $u_i(r)$ be the underlying payoff function of player i and $\hat{u}_i(r)$ be an approximation of $u_i(r)$ for each $i \in I$. Suppose that $|u_i(r) - \hat{u}_i(r)| \leq \delta$ for strategy profiles r and $\forall (a_i, r_{-i}) : i \in I, a_i \in A_i$. Then $|\epsilon(r) - \hat{\epsilon}(r)| \leq 2\delta$.*

PROOF. We use Claim B.1 to provide a bound on $|\epsilon(r) - \hat{\epsilon}(r)|$:

$$|\epsilon(r) - \hat{\epsilon}(r)| = |\max_{i \in I} \max_{a_i \in A_i} [\hat{u}(a_i, r_{-i}) - \hat{u}(r)] - \max_{i \in I} \max_{a_i \in A_i} [u_i(a_i, r_{-i}) - u_i(r)]| \leq$$
$$\max_{i \in I} \max_{a_i \in A_i} |[\hat{u}(a_i, r_{-i}) - \hat{u}(r)] - [u_i(a_i, r_{-i}) - u_i(r)]| =$$
$$\max_{i \in I} \max_{a_i \in A_i} |[\hat{u}(a_i, r_{-i}) - u_i(a_i, r_{-i})] + [u_i(r) - \hat{u}(r)]| \leq$$
$$\max_{i \in I} \max_{a_i \in A_i} |\hat{u}(a_i, r_{-i}) - u_i(a_i, r_{-i})| + |\hat{u}(r) - u_i(r)|.$$

The result now follows from the assumption that $|u_i(r) - \hat{u}_i(r)| \leq \delta$ for strategy profiles r and $\forall (a_i, r_{-i}) : a_i \in A_i$. □

The next step is to show that if a bound on approximate utilities holds pointwise on some strategy space $R \subset S$, we can establish a bound on regrets.

LEMMA D.2. *Suppose* $\Pr\{|u_i(r) - \hat{u}_i(r)| \geq \gamma\} \leq \delta \ \ \forall \ i \in I, r \in R$. *Then*

$$\Pr\{|\epsilon(r) - \hat{\epsilon}(r)| \geq 2\gamma\} \leq m(K+1)\delta,$$

where $K = \max_{i \in I} |A_i|$. Note that the bound is uniform since it does not depend on the specific profile $r \in R$.

PROOF. Let $\hat{\epsilon}_n(r)$ denote the empirical regret with respect to the empirical game with at least n payoff samples taken for every strategy profile, and let $\epsilon_i(r)$ denote the player i's contribution to regret (that is, the benefit player i can gain by deviating from r.

$$\begin{aligned}
\Pr\{|\hat{\epsilon}_n(r) - \epsilon(r)| \geq 2\gamma\} &= \Pr\{|\max_{i\in I} \hat{\epsilon}_{i,n}(r) - \max_{i\in I} \epsilon_i(r)| \geq 2\gamma\} \\
&\leq \Pr\{\max_{i\in I} |\hat{\epsilon}_{i,n}(r) - \epsilon_i(r)| \geq 2\gamma\} \\
&\leq \Pr\{\exists i \in I : |\hat{\epsilon}_{i,n}(r) - \epsilon_i(r)| \geq 2\gamma\} \\
&\leq \sum_{i\in I} \Pr\{|\hat{\epsilon}_{i,n}(r) - \epsilon_i(r)| \geq 2\gamma\} \\
&\leq \sum_{i\in I} \Pr\{\exists r_i' \in r_i \cup A_i : |\hat{u}_{i,n}(r_i', r_{-i}) - u_i(r_i', r_{-i})| \geq \gamma\} \\
&\leq \sum_{i\in I} \left(\sum_{a_i\in A_i} \Pr\{|\hat{u}_{i,n}(a_i, r_{-i}) - u_i(a_i, r_{-i})| \geq \gamma\} \right) \\
&\qquad + \sum_{i\in I} \Pr\{|\hat{u}_{i,n}(r) - u_i(r)| \geq \gamma\} \\
&\leq \sum_{i\in I} (K+1)\delta \leq m(K+1)\delta,
\end{aligned}$$

where line 5 above follows from Lemma D.1. □

Next, we use the bound on the greatest difference between approximate and actual payoffs on $\mathcal{A}$ to produce the desired bound for payoff approximation of any mixed strategy profile.

LEMMA D.3. *Suppose that for a subset $\mathcal{A} \subset A$*

$$\Pr\left\{\max_{a\in\mathcal{A}} |u_i(a) - \hat{u}_i(a)| \geq \gamma\right\} \geq \delta$$

for all $i \in I$. Then

$$\Pr\{|u_i(s) - \hat{u}_i(s)| \geq \gamma\} \leq \delta \;\forall i \in I, s \in S.$$

PROOF.

$$\begin{aligned}
&\Pr\{|u_i(s) - \hat{u}_i(s)| \geq \gamma\} = \\
&\quad = \Pr\{|u_i(s) - \hat{u}(s)| \geq \gamma | \sup_{a\in\mathcal{A}} |u_i(a) - \hat{u}_i(a)| < \gamma\} \\
&\qquad \times \Pr\{\sup_{a\in\mathcal{A}} |u_i(a) - \hat{u}_i(a)| < \gamma\} \\
&\quad + \Pr\{|u_i(s) - \hat{u}_i(s)| \geq \gamma | \exists a \in \mathcal{A} : |u_i(a) - \hat{u}_i(a)| \geq \gamma\} \\
&\qquad \times \Pr\{\exists a \in \mathcal{A} : |u_i(a) - \hat{u}_i(a)| \geq \gamma\}.
\end{aligned}$$

By the condition of the lemma, $\Pr\{\exists a \in \mathcal{A} : |u_i(a) - \hat{u}_i(a)| \geq \gamma\} \leq \delta$, and, thus, the second part of the sum is bounded by δ. Furthermore, observe that if

$\sup_{a\in\mathcal{A}} |u_i(a) - \hat{u}_i(a)| < \gamma$

$$\begin{aligned} |u_i(s) - \hat{u}_i(s)| &= |\sum_{a\in A}(u_i(a) - \hat{u}_i(a))s(a)| = |\sum_{a\in\mathcal{A}}(u_i(a) - \hat{u}_i(a))s(a)| \\ &\leq \sum_{a\in\mathcal{A}} s(a)|u_i(a) - \hat{u}_i(a)| < \sum_{a\in A} s(a)\gamma = \gamma, \end{aligned}$$

where the last step follows since s is a probability distribution over A. Consequently, $\Pr\{|u_i(s) - \hat{u}(s)| \geq \gamma | \sup_{a\in A} |u_i(a) - \hat{u}_i(a)| < \gamma\} = 0$ and the desired result follows. □

To complete the proof of the theorem in question, simply combine the results of Lemma D.2 and Lemma D.3.

ONLINE COMPANION: PROOFS FOR SECTION 7

Proof of Lemma 7.1

First, we prove the following claim.

CLAIM D.4. *Let s be a symmetric mixed strategy profile. Then* $\Pr\{W(b, s_{m-1}) \leq d\}$ *is continuous in s.*

PROOF. Fix $s \in S$ and note recall that

$$\Pr\{W(b, s_{m-1}) \leq d\} = 1 - \Phi \left[\frac{\sum_{c \in A_{m-1}} \hat{v}(b,c) s_{m-1}(c) - d}{\sqrt{\sum_{c \in A_{m-1}} \frac{\sigma^2(b,c) s_{m-1}^2(c)}{n(b,c)}}} \right].$$

Now, since $\Phi(\cdot)$ is a continuous function, and both the numerator and the denominator inside are continuous in s and s', and, furthermore, since the denominator is strictly positive for any valid mixed strategy profile $s \in S$, the expression inside the absolute value sign is continuous. □

PROOF OF LEMMA 7.1. We now use this result to prove the lemma. First, we note that the continuity of the density function in the expression for $P_n(r_1)$ can be established in essentially the same way as the continuity of $\Pr\{W(b, s_{m-1}) \leq d\}$, with the lone substantial exception that we in that case rely on the continuity of the normal density function rather than the normal distribution function. The remaining argument proceeds identically. Furthermore, the fact that a finite product of continuous functions is continuous produces an integrand (trivially integrable) which is continuous in s. □

Proof of Theorem 7.4

First we state a few basic general facts about continuous functions that are strictly increasing. Since these results are rather straightforward mathematical facts, we omit the proofs.

CLAIM D.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous strictly increasing function. Then set $B \subset \mathbb{R}$ is open if and only if $f(B)$ is open. Furthermore B is non-empty if and only if $f(B)$.*

CLAIM D.6. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous strictly increasing function and $B \subset \mathbb{R}$ be a set such that $\lambda(f(B)) > 0$, where λ is the Lebesgue measure on $\mathbb{R}$. Then $\lambda(B) > 0$. Defining $\nu = \lambda \circ f$, this implies that $\nu << \lambda$.*

Our next step is to show that the function which comprises a portion of $P(a)$ is strictly increasing (its continuity is clear). From this and the above claims we will show the desired result. Let

$$K(u) = \prod_{b \in A_1} \left[1 - \Phi \left[\frac{\sum_{c \in A_{m-1}} \hat{v}(b,c) s_{m-1}(c) - u}{\sqrt{\sum_{c \in A_{m-1}} \frac{\sigma^2(b,c) s_{m-1}^2(c)}{n(b,c)}}} \right] \right]$$

and note that $K(u)$ is strictly increasing in u since Φ is a standard normal distribution and is strictly decreasing in u. Let $f(u, \mu)$ be a normal density function with a fixed variance and mean μ.

Claim D.7. *The function* $h(\mu) = \int_{\mathbb{R}} K(u)f(u,\mu)du$ *is strictly increasing in* μ.

Proof. Let $\mu > \mu'$ and let $g(u) = f(u,\mu) - f(u,\mu')$. Then

$$\int_{\mathbb{R}} [K(u)f(u,\mu) - K(u)f(u,m\mu')]du = \int_{\mathbb{R}} K(u)[f(u,\mu) - f(u,\mu')]du =$$

$$= \int_{-\infty}^{\mu'} K(u)g(u)du + \int_{\mu'}^{\mu} K(u)g(u)du + \int_{\mu}^{\infty} K(u)g(u)du.$$

Since $K(u)$ is strictly increasing in u,

$$\int_{\mu'}^{\mu} K(u)g(u)du > \int_{\mu'}^{\mu} K(\mu')g(u)du = K(\mu')\int_{\mu'}^{\mu} g(u)du = 0,$$

where the last equality follows because the two normal distributions involved have identical variance:

$$\begin{aligned}\int_{\mu'}^{\mu} g(u)du &= (F(\mu,\mu) - F(\mu',\mu)) - (F(\mu,\mu') - F(\mu,\mu)) \\ &= \Pr_{\mu}\{\mu' \leq u \leq \mu\} - \Pr_{\mu'}\{\mu' \leq u \leq \mu\} = 0,\end{aligned}$$

where $F(\cdot,\mu)$ is the cdf of Normal with mean μ. Similarly,

$$\int_{-\infty}^{\mu'} K(u)g(u)du + \int_{\mu}^{\infty} K(u)g(u)du > \int_{\mu}^{\infty} K(u)g(u)du - \int_{\mu}^{\infty} K(u)g(u)du = 0.$$

To get the inequality above, note that since Normal is symmetric and $K(u)$ is strictly increasing,

$$\int_{-\infty}^{\mu'} K(u)g(u)du > -\int_{\mu}^{\infty} K(u)g(u)du.$$

It then follows that

$$\int_{\mathbb{R}} [K(u)f(u,m) - K(u)f(u,m')]du > 0$$

and, thus, $\int_{\mathbb{R}} K(u)f(u,\mu)du$ is strictly increasing in μ. □

Having all the pieces at our disposal, we now proceed to prove the result towards which we have been building.

Proof of Theorem 7.4. Let $\mu = \hat{v}(a)$ be distributed according to F for each player. Without loss of generality, keep the payoffs for all deviations constant. (Since we show that even when these are constant the measure of the resulting set will be zero, this will certainly continue to hold for the product measure since all payoffs are assumed to be independently sampled.) Defining $h(\mu) = \int_{\mathbb{R}} K(u)f(u,\mu)du$, we showed above that it is strictly increasing, and its continuity is immediate from definition. Since for any $c \in \mathbb{R}$, $\lambda(c) = 0$, defining $\nu = \lambda \circ h$, we have that $\nu(c) = 0$ by Claim D.6.

Now, the probability that some two fixed pure strategy profiles a, a' yield $P_n(a) = P_n(a')$ is (by conditioning on the value of a') $E_{c \sim G(P(a'))}[F(P_n(a) = P_n(a'))|P_n(a') = c] = E_{c \sim G(P_n(a'))}[F(P_n(a) = c)]$, where G is the probability distribution on $P_n(a')$

induced by the prior distribution. Since $F(P_n(a) = c) = 0$ as argued above for any $c \in \mathbb{R}$, the resulting expectation and, therefore, the prior probability, that $P_n(a) = P_n(a')$ is zero. And, finally, if the probability is zero for any two fixed a, a', then it is zero for any finite set of these. □

Proof of Lemma 7.6

For simplicity, we assume (without loss of generality) that every strategy profile has the same variance σ^2. Additionally (also without loss of generality), we assume that s does not correspond to some pure strategy profile—that is, at least two pure strategy profiles a, a' are played with positive probability under s. Let $f_n(u)$ be the density function of the normal distribution $N(\hat{v}(s), \sigma^2(s))$, where $\hat{v}(s) = \sum_{a \in A} \hat{v}(a)s(a)$ and $\sigma^2(s) = (\sigma^2/n)\sum_{a\in A} s(a)^2$.

The idea of the proof is that SLLN (strong law of large numbers) implies convergence of payoffs for every mixed strategy profile. In particular, this will be the case for s and a deviation (b, s_{m-1}) of an arbitrary player which strictly improves his actual payoff (which must exist since s is not a Nash equilibrium). Since n can also be large enough to make the variance of the posterior Normal distribution arbitrarily small, and, thus, the posterior probability that $\epsilon(s) = 0$ arbitrarily small.

Formally, the goal is to show that for any $\epsilon > 0$, $\alpha > 0$, there is $N \geq 1$ such that $\Pr\{\sup_{n \geq N} P_n(s) < \epsilon\} \geq 1-\alpha$. Let $b \in A_1$ be a profile which yield a strictly higher payoff to i than s_1 when others play s_{m-1}. Let $\Delta = v(b, s_{m-1}) - v(s) > 0$. Fix $\epsilon > 0$ and $\alpha > 0$. Define $\delta = \frac{\Delta}{4} > 0$. Now, by SLLN, there is N_1 large enough such that both $\sup_{n\geq N_1} |\hat{v}_n(s) - v(s)| < \delta$ and $\sup_{n \geq N_1} |\hat{v}_n(b, s_{m-1}) - v(b, s_{m-1})| < \delta$ with probability at least $1 - \alpha$.

Since the variance of the posterior Normal distribution of $v(s)$ is $\frac{\sigma^2}{n}$, we can find N_2 large enough so that for any function $G(u) \leq 1$, for δ defined as above, and for any fixed $\gamma > 0$,

$$\sup_{n\geq N_2} \int_{\mathbb{R}} G(u) f_n(u) du \leq \sup_{n \geq N_2} \int_{\hat{v}_n(s)-\delta}^{\hat{v}_n(s)+\delta} G(u) f_n(u) du + \gamma.$$

Therefore, we have that

$$\begin{aligned}
\sup_{n\geq N_2} \Pr_n\{\epsilon(s) = 0\} &= \sup_{n\geq N_2} \int_{\hat{v}_n(s)-\delta}^{\hat{v}_n(s)+\delta} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}(\hat{v}_n(b, s_{m-1}) - u))\right] f_n(u) du + \gamma \\
&\leq \sup_{n\geq N_2} \int_{\hat{v}_n(s)-\delta}^{\hat{v}_n(s)+\delta} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}(\hat{v}_n(b, s_{m-1}) - \hat{v}_n(s) - \delta))\right] f_n(u) du + \gamma \\
&\leq \sup_{n\geq N_2} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}(\hat{v}_n(b, s_{m-1}) - \hat{v}_n(s) - \delta))\right](1-\gamma) + \gamma.
\end{aligned}$$

Then, for $N_3 = \max\{N_1, N_2\}$, $\hat{v}_n(b, s_{m-1}) - \hat{v}_n(s) \geq 2\delta$ and, consequently,

$$\sup_{n\geq N_3} \Pr_n\{\epsilon(s) = 0\} \leq \sup_{n\geq N_3} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}\delta)\right](1-\gamma) + \gamma.$$

Since δ and σ are fixed and strictly positive, we can find N_4 large enough so that

$\Phi(\cdot)$ is arbitrarily close to 1. In particular, we can ensure that

$$1 - \Phi(\frac{\sqrt{n}}{\sigma}\delta) \leq \gamma.$$

Letting $N = \max\{N_3, N_4\}$ and combining gives us

$$\sup_{n \geq N} P_n(s) \leq \gamma(1-\gamma) + \gamma \leq 2\gamma$$

with probability at least $1 - \alpha$. Setting $\gamma = \epsilon/2$ yields the desired result.

Proof of Lemma 7.7

For simplicity (and as above), we assume (without loss of generality) that every player and profile has the same variance σ^2. Let $f_n(u)$ be the density function of the normal distribution $N(\hat{v}(a), \sigma^2/n)$.

Fix $\alpha > 0$ and $\epsilon > 0$. Our goal is to show that there is a finite N such that $\sup_{n \geq N} P_n(a) \geq 1-\epsilon$ with probability at least $1-\alpha$. Define $\Delta = \min_{b \in A_1 \setminus a_1} v(a) - v(b, a_{m-1})$. Set $\delta = \Delta/4 < 0$ for the analysis below. By SLLN, there is $N_1 < \infty$ such that $\sup_{n \geq N_1} |\hat{v}_n(a) - v(a)| \leq \delta$, as well as $\sup_{n \geq N_1} |\hat{v}_n(b, a_{m-1}) - v(b, a_{m-1})| \leq \delta$ for every $b \in A_1 \setminus a_1$ simultaneously with probability at least $1 - \alpha$. Now, since for any function $G(u)$,

$$\int_{\mathbb{R}} G(u) f_n(u) du \geq \int_{\hat{v}_n(a)-\delta}^{\hat{v}_n(a)+\delta} G(u) f_n(u) du,$$

we have that

$$\begin{aligned}
\sup_{n \geq N_1} \Pr_n(\epsilon(a) = 0) &\geq \sup_{n \geq N_1} \int_{\mathbb{R}} \prod_{b \in A_1 \setminus a_1} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}(\hat{v}_n(b, a_{m-1}) - u))\right] f_n(u) du \\
&\geq \sup_{n \geq N_1} \int_{\mathbb{R}} \prod_{b \in A_1 \setminus a_1} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}(\hat{v}_n(b, a_{m-1}) - \hat{v}_n(a) + \delta))\right] f_n(u) du \\
&= \sup_{n \geq N_1} (1-\gamma) \prod_{b \in A_1 \setminus a_1} \left[1 - \Phi(\frac{\sqrt{n}}{\sigma}(\hat{v}_n(b, a_{m-1}) - \hat{v}_n(a) + \delta))\right].
\end{aligned}$$

Since $\hat{v}_n(b, a_{m-1}) - \hat{v}_n(a) \leq -2\delta$ for all $b \in A_1 \setminus a_1$ simultaneously, we obtain that

$$\begin{aligned}
\sup_{n \geq N_1} \Pr_n(\epsilon(a) = 0) &\geq \sup_{n \geq N_1} (1-\gamma) \prod_{b \in A_1 \setminus a_1} \left[1 - \Phi(-\frac{\sqrt{n}}{\sigma}\delta))\right] \\
&\geq \sup_{n \geq N_1} (1-\gamma) \left[1 - \Phi(-\frac{\sqrt{n}}{\sigma}\delta))\right]^{|A_1|-1}.
\end{aligned}$$

Since $\sigma > 0$ and $\delta > 0$ are both fixed, we can find N_2 large enough so that

$$\Phi(-\frac{\sqrt{n}}{\sigma}\delta)) \leq \gamma,$$

and by allowing $N = \max\{N_1, N_2\}$ we have $\sup_{n \geq N} P_n(a) \geq (1-\gamma)^{|A_1|}$ with probability at least $1 - \alpha$. By choosing $\gamma = 1 - (1-\epsilon)^{\frac{1}{|A_1|}}$, we obtain the desired result.