Nash Equilibria in Random Games

Imre Bárány[*]
barany@renyi.hu

Santosh Vempala[†]
vempala@math.mit.edu

Adrian Vetta[‡]
vetta@math.mcgill.ca

Abstract

We consider Nash equilibria in 2-player random games and analyze a simple Las Vegas algorithm for finding an equilibrium. The algorithm is combinatorial and always finds a Nash equilibrium; on $m \times n$ payoff matrices, it runs in time $O(m^2 n \log\log n + n^2 m \log\log m)$ with high probability. Our main tool is a polytope formulation of equilibria.

1 Introduction

The complexity of finding a Nash equilibrium in a 2-player game is perhaps the outstanding open problem in algorithmic game theory [28]. In a 2-player game, the first player, namely Alice, has m pure strategies $\mathcal{S} = \{\sigma_1, \sigma_2 \ldots, \sigma_m\}$ while the second player, namely Bob, has n pure strategies $\mathcal{T} = \{\tau_1, \tau_2 \ldots, \tau_n\}$. We are given two payoff matrices A and B for Alice and Bob, respectively. Here, the ijth entry of A is the payoff to Alice when she plays σ_i and Bob plays τ_j. A Nash equilibrium is a pair of mixed strategies (probability distributions) x, y such that given x, the distribution y on $\mathcal{T}$ maximizes $x^T By$, the payoff to Bob and simultaneously, given y, the distribution x on $\mathcal{S}$ maximizes $x^T Ay$, the payoff to Alice.

On the one hand, it is easy to check that a given pair of mixed strategies forms an equilibrium. On the other hand, the best algorithms for finding a Nash equilibrium of an arbitrary 2-player game have exponential complexity. Moreover, there is some evidence that finding an equilibrium is unlikely to be NP-hard [26, 27]. Determining the complexity of finding equilibria has lead to much research in a variety of directions, e.g., a quasi-polynomial-time algorithm to find an *approximate* Nash equilibrium due to Lipton et al. [22]; an investigation into the complexity of finding pure strategy Nash equilibria in succinctly specified games by Fabrikant et al. [14]; a polynomial-time algorithm of Papadimitriou and Roughgarden for finding Nash equilibria in multi-player symmetric game in which each player has a small number of strategies [30]; a proof that the Lemke-Howson algorithm takes exponential time with all possible initial pivots [32].

In this paper, we consider 2-player games where the two payoff matrices are chosen randomly. Our motivation is the question of whether finding Nash equilibria is any easier in random games compared to general games, that is, easier "on average". In a random game, every entry in each of the matrices is drawn independently according to some probability distribution. We consider the uniform distribution on an interval and the standard Normal distribution $N(0, 1)$. In the first case, the distribution of any set of k entries of a payoff matrix is uniform in a k-dimensional cube, while in the second case it is a k-dimensional Normal. In fact, there has been much work in this direction for a special case of 2-player games, the zero-sum case. This case is equivalent to linear programming. Motivated by the question of explaining the success of the simplex algorithm, Borgwardt [7], Smale [34] and Megiddo [25] studied linear programs where the constraints are chosen randomly from spherically symmetric distributions and showed that variants of the simplex algorithm run in polynomial time. Besides simplex, other simple methods (e.g., the perceptron algorithm) also work for random linear programs, demonstrating that they have considerably more structure than arbitrary linear programs.

Here we show random games are indeed much simpler than general games. Specifically, we show that with high probability, there is a Nash equilibria in which the supports of the mixed strategies of both players have small cardinality. We remark that in a random game the supports of each player will have the same cardinality in a Nash equilibrium. As a result, the following naive heuristic is a Las Vegas algorithm for finding Nash equilibria: exhaustively check for Nash equilibria with supports of cardinality $i = 1, 2, \ldots$ until an equilibrium is found. In fact, with high probability only two phases will be required!

The key to our result is a reformulation of the problem

[*] Rényi Institute of Mathematics, Hungarian Academy of Sciences, and Department of Mathematics, University College London. Supported in part by Hungarian NSF Grants No T037846 and T046246.

[†] Massachusetts Institute of Technology. Supported in part by NSF award CCR-0307536.

[‡] McGill University. Supported in part by NSERC grant 28833-04 and FQRNT grant NC-98649.

in terms of random polytopes. For convenience, we will assume that $m = n$. We will see in Section 2 that, given a mixed strategy for Alice (Bob), the supports of a best response strategy for Bob (Alice) are precisely those supports that induce facets with non-negative normal vectors in an associated random polytope. Consequently, the algorithmic problem of finding Nash equilibria can be tackled by considering problems relating to the number of points on the convex hull of a set of n random points in d dimensions. In Section 3 we extend analysis of such random polytopes for our purpose. This allows us to examine the quality of our algorithm in Section 4. In particular, consider the convex hull of n random points in d dimensions. Let N_1 denote the expected number of points that lie on the boundary of the convex hull (this will differ for the Normal and uniform distributions). Then, our main theorem can be stated as follows.

Theorem 1. *The probability that a random game contains no Nash equilibria with supports of size at most d is less than*

$$f(d)\left(\frac{1}{n}+\frac{1}{(N_1)^2}\right)$$

where $f(d)$ is a function of d alone.

It is known that N_1 grows as a function of n for both the distributions we consider and so we get our algorithmic result as a corollary.

Corollary 1. *There is a combinatorial algorithm that finds a Nash equilibrium in a random game and, with high probability, runs in time $O(n^3 \log\log n)$.* □

In related work, McLennan and Berg [24] have studied the expected number of Nash equilibria for certain random games. Finally, we remark that extending our work to allow matrix entries to have arbitrary means would give a polynomial-time randomized algorithm for finding approximate Nash equilibria in arbitrary games.

2 A Geometric Interpretation of Nash Equilibria

We begin by showing how Nash equilibria are closely related to random polytopes. For any subset of Alice's strategies, $S \subseteq \mathcal{S}$, with $|S| = d$, let $B^S = \{b_1^S, \ldots, b_n^S\}$ be the set of subcolumns of B induced by S, i.e., whose entries correspond to the rows of S. We will view the elements of B^S as points in $\mathbb{R}^d$, and denote by $\mathsf{conv}(B^S)$ the polytope corresponding to the convex hull of these points. Similarly A^T is the set of subrows of A induced by $T \subseteq \mathcal{T}$.

Lemma 1. *A point $x \in B^S$ is a best response to some mixed strategy ν with support S if and only if x maximizes $\nu \cdot y$ among all points $y \in \mathsf{conv}(B^S)$.*

Proof. Any mixed strategy λ for Bob induces a point $y \in \mathsf{conv}(B^S)$ where $y(\lambda) = \sum_{i=1}^n \lambda_i b_i^S$. The expected payoff to Bob of the mixed strategy λ is then just $\nu \cdot y$. □

We will say that a facet of the polytope is *useful* if its normal vector is nonnegative. Similarly a face of any dimension is called useful if it is contained in some useful facet. Our interest in useful facets comes from the fact that only nonnegative vectors induce feasible mixed strategies. In particular, we obtain from Lemma 1 the following corollary.

Corollary 2. *Let $S = \{\sigma_1, \ldots, \sigma_d\}$ and $\nu \geq 0$. Then $\nu/||\nu||_1$ is a mixed strategy on S for Alice, and the vertices of the face in $\mathsf{conv}(B^S)$ with normal $\nu \geq 0$ are best response strategies for Bob.* □

Next, consider subsets $S \subseteq \mathcal{S}, T \subseteq \mathcal{T}$ that are supports of mixed strategies that form a Nash equilibrium. Let $B^S(T)$ be the set of vectors in B^S that correspond to T, and let $A^T(S)$ be the set of vectors in A^T that correspond to the rows of S. To induce a Nash equilibrium, we need that each $x \in B^S(T)$ is a best response to some mixed strategy ν with support S, and each $y \in A^T(S)$ is a best response to some mixed strategy λ with support T. In other words, the points in $A^T(S)$ lie on a useful facet of $\mathsf{conv}(A^T)$ and the points in $B^S(T)$ lie on a useful facet of $\mathsf{conv}(B^S)$. Since the entries of A and B are chosen at random, the corresponding point sets will be in general position and so for a Nash equilibrium, we must have $|S| = |T|$.

So Nash equilibria in random games are closely related to the polytopes produced by random points. In particular, we need to study the faces of random polytopes. Towards this end, let N_i be the expected number of faces of dimension $i-1$ induced by $\mathsf{conv}(\mathcal{P})$ where $\mathcal{P}$ is a set of n random points (N_1 is number vertices, N_d number of facets). Again, since the points are in general position any such face is induced by exactly i points. We have the following simple relationship between number of faces of a given dimension and the number of points on the convex hull.

Lemma 2. *The number of faces of $conv(\mathcal{P})$ of dimension s is at least $\frac{1}{s+1}\binom{d-1}{s}$ times the number of points on the convex hull.*

Proof. Every vertex of the convex hull is contained in at least d facets. Take a vertex x and consider any d facets containing x. These induce a set of $\binom{d-1}{s}$ faces of dimension s that contain x. The union of these sets may contain at most d copies of any face. Moreover, as all the points are in general position, every s-dimensional face contains exactly $s+1$ points. Thus, summing over all vertices, we obtain the result. □

It follows that $N_{s+1} \geq \frac{1}{s+1}\binom{d-1}{s} N_1$. Our interest is in the expected number of useful faces of dimension $i-1$, denoted N_i^+. We will use the following notation. We write $A \lesssim B$ if there is a fixed function $f(d)$, of the dimension d alone, such that $A \leq f(d)B$. We will need the simple observation that $N_s^+ \gtrsim N_1^+ \gtrsim N_1/2^d$.

3 Random Polytopes under Gaussian and Uniform Distributions

In this section we present the main technical results we will need for analysing our Las Vegas algorithm. First we need to develop some understanding of the behaviour of N_i. There has already been a large amount of work studying N_i for various distributions; for a rather comprehensive survey see [39]. We begin with two basic results regarding the uniform and Normal distributions, respectively.

Theorem 2. *[12] Given n points drawn independently and uniformly from a d-dimensional unit cube, the expected number of points on the convex hull is $\gtrsim (\log n)^{d-1}$.* □

Theorem 3. *[31] Given n points in $\mathbb{R}^d$ with coordinates drawn independently from $N(0,1)$, the expected number of points on the convex hull is $\gtrsim (\log n)^{\frac{1}{2}(d-1)}$.* □

We will be be interested in the following more general question. Let $\mathcal{P} = \{x_1, \ldots, x_n\}$ be a set of i.i.d. random points from a distribution with density function f. Let F be the set of subsets of points that induces facets of $\mathsf{conv}(\mathcal{P})$, and let Y_1 and Y_2 be two subsets of X. What is the probability that both Y_1 and Y_2 induce facets, i.e., $\mathsf{P}(Y_1, Y_2 \in F)$? The probability that one subset Y_1 induces a facet is well-understood for many important distributions, including the Gaussian and the cube. We prove the following refinements (that may be of independent interest).

Lemma 3. *Suppose f is the Normal density or the uniform density over a cube. If Y_1, Y_2 are disjoint subsets of $\mathcal{P}$, then*

$$\mathsf{P}(Y_1, Y_2 \in F) \lesssim \mathsf{P}(Y_1 \in F)^2$$

Lemma 4. *Suppose f is the Normal density or the uniform density over a cube. If Y_1, Y_2 are subsets of $\mathcal{P}$ with $|Y_1 \cap Y_2| > 0$, then*

$$\mathsf{P}(Y_1, Y_2 \in F) \lesssim \frac{\mathsf{P}(Y_1 \in F)^2}{\mathsf{P}(Y_1 \cap Y_2 \textit{ is a face of } \mathsf{conv}(\mathcal{P}))}$$

The proof of these lemmas for the Normal distribution can be carried out directly using the density function. We give such a proof of Lemma 3 at the end of this section (the proof of Lemma 4 is similar). For a cube, however, things are more complicated and we will use an *economic cap-covering* [6]. The next theorem is implied by the results of [4].

Theorem 4. *Assume K is a convex body in $\mathbb{R}^d$, $\mathsf{vol}(K) = 1$ and $0 < \varepsilon < \varepsilon_0(d)$ where $\varepsilon_0(d)$ is a constant depending only on d. Let $K(\varepsilon)$ denote the set of points from K that can be cut off by a cap of volume ε. Then there are convex sets $C_1, \ldots, C_m$ and pairwise disjoint convex sets $C'_1, \ldots, C'_m$ such that $C'_i \subset C_i$ for every i such that every cap of volume ε is contained in one of the C_i and, further,*

$$\bigcup_1^m C'_i \subset K(\varepsilon) \subset \bigcup_1^m C_i$$

$$\mathsf{vol}(C'_i) \gtrsim \varepsilon \textit{ and } \mathsf{vol}(C_i) \lesssim \varepsilon.$$

□

The meaning is that $K(\varepsilon)$ can be economically covered, that is, covered without much overlap, by the sets C_i where each C_i has volume a constant times ε. This implies that $\mathsf{vol}(K(\varepsilon)) \sim m\varepsilon$. The extra condition that each cap of volume ε is contained in one of the C_i will be very useful when estimating an integral in the proof of Lemma 5.

When K is the unit cube, the sets C_i will be axis-parallel dyadic boxes:

$$\prod_1^d [0, 2^{-f_i}]$$

with integers $f_i \geq 0$ summing to f and $2^{-f} \approx \varepsilon$. This is only the covering near a single vertex, the origin, of the unit cube, which is extended to a complete covering using the symmetries of the cube in the obvious way. One can define cap coverings for other distributions as well, replacing the volume of each set C_i by its measure. Such a covering exists for the Normal distribution if the complement of a suitably large ball is deleted.

We will use cap coverings to prove the following generalization of Lemma 4. We focus now on the case when K is the unit cube, but the proof applies whenever a cap covering exists.

Lemma 5. *Let S, S^*, T, T^* be sets of cardinality d with $S \cap S^* = X$, $T \cap T^* = Y$, $|X| = s > 0$, $|Y| = t > 0$. We denote by $K = Q^{|S \cup S^*|} \subset \mathbb{R}^{2d-s}$ the unit cube in $\mathbb{R}^{2d-s}$. Assume $x_1, \ldots, x_n$ are uniform, independent, random points from K. Assume $T = \{x_1, \ldots, x_d\}$ and $T^* = \{x_1, \ldots, x_t, x_{d+1}, \ldots, x_{2d-t}\}$. Let $\mathcal{F}(S)$ denote the facets of the convex hull of the x_is projected onto R^S, and similarly for $\mathcal{F}(S^*)$. Then,*

$$\mathsf{P}[T \in \mathcal{F}(S) \textit{ and } T^* \in \mathcal{F}(S^*)] \lesssim \frac{(\ln n)^{2d-s-1}}{n^{2d-t}}.$$

Proof. Let $\mathcal{P} = \{x_1, \ldots, x_n\}$ as before. It follows from the main theorem in [5] that the random polytope $\mathsf{conv}(\mathcal{P})$ contains $K(c(\ln n)/n)$ with probability at least $1 - n^{-3d}$ if the constant c is chosen large enough. Note that if $T \in$

$\mathcal{F}(S)$, then T is a face of the convex hull of $\mathsf{conv}(\mathcal{P})$ as well.

We fix a cap covering M_f given by Theorem 4 for every large enough f for the unit cube. We write $L|S$ and $L|S^*$ for the projection of a set $L \subset \mathbb{R}^{|S \cup S^*|}$ onto $R^{|S|}$ and $R^{|S^*|}$, resp. We need a special minimal cap $C(T)$ of K that contains T: namely, writing $C(T|S)$ for the minimal cap containing $T|S$ in $Q^{|S|}$ we let

$$C(T) = C(T|S) \times Q^{|S^* \setminus S|}.$$

Similarly,

$$C(T^*) = C(T^*|S^*) \times Q^{|S \setminus S^*|}.$$

Note that both $C(T)$ and $C(T^*)$ are caps of K and $C(T) \cap C(T^*)$ is non-empty. This implies that $C(T) \cap C(T^*)$ contains a vertex of K.

Let $\chi[E]$ be the indicator of the event E. Define D to be the event that $C(T) \cap C(T^*)$ contains the origin. Now, writing P for the probability in the Lemma, we have

$$\begin{aligned} P &= \int_K \dots \int_K \chi[T \in \mathcal{F}(S)]\, \chi[T^* \in \mathcal{F}(S^*)] dx_1 \dots dx_n \\ &\le 2^d \int_K \dots \int_K \chi[D] \times \\ &\quad \times \chi[T \in \mathcal{F}(S)]\, \chi[T^* \in \mathcal{F}(S^*)] dx_1 \dots dx_n. \end{aligned}$$

By the remark at the beginning of this proof we can restrict integration to the subset where $\mathsf{vol}(C(T)) \le c(\ln n)/n$ and the same for $C(T^*)$. With this we only lose n^{-3d} in probability. We replace this restricted integral by a double sum in the following way.

Given $x_1, \dots, x_d$ and $C(T)$ with $0 \in C(T)$, let f be the largest integer with $C(T) \subset C_f$ for some $C_f \in M_f$ where M_f is the cap covering for $\varepsilon = 2^{-f}$. Of course, the C_f that matters is of the form

$$\prod_{i \in S \cup S^*} [0, 2^{-f_i}]$$

with $f_i = 0$ when $i \in S^* \setminus S$. Let M_f^1 be the set of these elements of M_f. Obviously, $\mathsf{vol}(C(T)) \ge c(d)2^{-f}$ with a suitable small $c(d) > 0$, and similarly for $C(T^*)$. Analogously, we get a C_g^* from M_g for each $C(T^*)$, and the ones that matter are collected in $M_g^2 \subset M_g$.

We integrate then on each $C_f \in M_f^1$ and $C_g^* \in M_g^2$ for $f, g \ge f_0$. Assuming $g \ge f$, the integrand is at most

$$(1 - c(d)2^{-f})^{n-(2d-t)} (\mathsf{vol}(C_f)\mathsf{vol}(C_g^*))^{d-t} (\mathsf{vol}(C_f \cap C_g^*))^t$$

since $x_1, \dots, x_t$ come from $C_f \cap C_g^*$, $x_{t+1}, \dots, x_d$ from C_f, and $x_{d+1}, \dots, x_{2d-t}$ from C_g^*, while the rest of the x_i come from $K \setminus C(T)$. Summing this for all $g \ge f \ge f_0$, and all $C_f \in M_f^1$, all $C_g^* \in M_g^2$ we are lead to the sum

$$\sum_{f \ge f_0} \sum_{g \ge f} \exp\{-c(d)2^{-f}n/2\} 2^{-f(d-t)} 2^{-g(d-t)} S_{f,g}$$

where

$$S_{f,g} = \sum_{M_f^1} \sum_{M_g^2} (\mathsf{vol}(C_f \cap C_g^*))^t$$

Now $C_f \cap C_g^*$ is an axis-parallel dyadic box in K, so for some $h \ge g$, it is equal to a unique $C_h \in M_h$. Consequently,

$$S_{f,g} \le \sum_{h \ge g} 2^{-ht} \sum_{C_h \in M_h} N(f,g)$$

where $N(f,g)$ is the number of pairs C_f and C_g^* with $C_h = C_f \cap C_g^*$. For a fixed C_h, it is not hard to see that $N(f,g) \lesssim (2h - f - g + 1)^s$. Further, $|M_h| \approx h^{2d-s-1}$. Thus

$$S_{f,g} \lesssim \sum_{h \ge g} 2^{-ht} h^{2d-s-1} (2h - g - f + 1)^s.$$

The rest is a computation: one shows first that $S_{f,g}$ is dominated by the term when $h = g$. (This is done by comparing the $h+1$st and hth terms.) Then the next sum

$$\begin{aligned} & \textstyle\sum_{g \ge f} \exp\{-c(d)2^{-f}n/2\} 2^{-f(d-t)} 2^{-g(d-t)} S_{f,g} \\ \lesssim & \textstyle\sum_{g \ge f} [\exp\{-c(d)2^{-f}n/2\} 2^{-f(d-t)} \\ & \quad \times 2^{-g(d-t)} 2^{-gt} g^{2d-s-1} (g - f + 1)^s] \end{aligned}$$

is again dominated by the term when $g = f$ (again by comparing consecutive terms), implying that

$$P \lesssim \sum_{f \ge f_0} \exp\{-c(d)2^{-f}n/2\} 2^{-f(2d-t)} f^{2d-s-1}.$$

Finally, the last sum is dominated by the term when $2^{-f} = 1/n$. (For f with $2^{-f} < 1/n$ this is done by comparing consecutive terms, again. For f with $2^{-f} > 1/n$ the factor $\exp\{-c(d)2^{-f}n/2\}$ is very small except for finitely many terms.) So we have

$$P \lesssim \frac{(\ln n)^{2d-s-1}}{n^{2d-t}}.$$

$\square$

Remark 1. This method works when $T \cap T^* = \emptyset$. Then $(\mathsf{vol}(C_f \cap C_g^*))^t = 1$ since $t = 0$ and M_h does not appear at all. With a similar computation one could prove Claim 1 (see later) in the following form:

$$\begin{aligned} & \mathsf{P}[T \in \mathcal{F}(S) \text{ and } T^* \in \mathcal{F}(S^*)] \\ & \quad \lesssim \mathsf{P}[T \in \mathcal{F}(S)]\, \mathsf{P}[T^* \in \mathcal{F}(S^*)]. \end{aligned}$$

Remark 2. For general convex bodies the outcome depends on $S_{f,g}$ and then on $N(f,g)$. For smooth convex bodies $N(f,g)$ is a constant, and the computation is simpler.

We now give a more direct proof of Lemma 3 for the case of the Normal distribution.

Proof. We prove the lemma directly for the Normal density. We write

$$\begin{aligned}&\mathsf{P}(Y_1, Y_2 \in F)\\ &= \int\limits_{x_1,\ldots x_n \in \mathbb{R}^d} \chi[Y_1 \in F]\, \chi[Y_2 \in F]\, df(x_1)\cdots df(x_n)\end{aligned}$$

Assume that $Y_1 = \{x_1, \ldots, x_d\}, Y_2 = \{x_{d+1}, \ldots, x_{2d}\}$ For a subset Y, let $H(Y)$ be the hyperplane spanning Y and $V(Y)$ be measure of the distribution in the halfspace bounded by this hyperplane not containing the origin. Then the above probability can be bounded as

$$\begin{aligned}&\mathsf{P}(Y_1, Y_2 \in F) \qquad (1)\\ &\leq\; 2 \int\limits_{x_1,\ldots,x_{2d}, V(Y_2)\geq V(Y_1)} (1 - V(Y_1))^{n-2d}\, df(x_1)\cdots df(x_{2d})\end{aligned}$$

We now estimate this when f is the standard Normal density in $\mathbb{R}^d$. Clearly, $V(Y_1)$ depends only on the distance of $H(Y_1)$ from the origin. So, it will be convenient to parametrize in terms of hyperplanes, and positions on them. This is achieved by the Blaschke-Petkantschin formula. For a spherically symmetric density function f we have,

$$\begin{aligned}&\int\limits_{x_1,\ldots,x_d \in \mathbb{R}^d} g(x_1, \ldots, x_d)\, df(x_1)\cdots df(x_d)\\ &= \Delta(d) \int\limits_{H:\text{(d-1)-flat}} \int\limits_{x_1,\ldots,x_d \in H} g(x_1, \ldots, x_d) df_H(x_1)\cdots df_H(x_d) d\mu(H).\end{aligned}$$

Here $\Delta(d)$ is a function of d alone and $d\mu(H)$ is the measure induced on $(d-1)$-flats determined by picking d points from f. By spherical symmetry, $d\mu(H) = dhdu$ where u is a unit vector (normal to H) and h is the distance of H from the origin. For the Normal density, the measure of a random point on a flat is determined by its position on the flat and the distance of the flat to the origin, i.e., $df_H(x_1) = df_h(x_1)$. The integrand on the RHS of (1) depends only on the distance of $H(Y_1)$ from the origin. Thus, applying the formula twice, once for $x_{d+1}, \ldots, x_{2d}$ and then for $x_1, \ldots, x_d$, we have

$$\begin{aligned}&\int\limits_{x_1,\ldots,x_{2d}, V(Y_2)\geq V(Y_1)} (1 - V(Y_1))^{n-2d}\, df(x_1)\ldots df(x_{2d})\\ &= \Delta(d)^2 \int\limits_{t_1 \leq t_2} (1 - V(t_1))^{n-2d}\, d\nu(t_1) d\nu(t_2)\\ &\leq \frac{\Delta(d)^2 \mathsf{vol}(S_d)^2}{(2\pi)^d} \int\limits_{t_1} (1 - V(t_1))^{n-2d} e^{-dt_1^2} t_1^{2(d-1)}\, dt_1\\ &\leq \frac{\Delta(d)^2 \mathsf{vol}(S_d)^2}{(2\pi)^d} \int\limits_{t} \left(1 - \frac{e^{-\frac{1}{2}t^2}}{2\sqrt{2\pi}t}\right)^{n-2d} e^{-dt^2} t^{2d-2}\, dt\\ &\lesssim n^{-2d} (\ln n)^{d-1}\end{aligned}$$

Since

$$\mathsf{P}(Y_1 \in F) \gtrsim \frac{(\ln n)^{(d-1)/2}}{\binom{n}{d}}$$

the lemma follows for the Normal density. □

4 A Las Vegas Algorithm

We now have the tools needed to analyse our very simple algorithm. Recall that the algorithm exhaustively checks for Nash equilibria with supports of size 1, then for Nash equilibria with supports of size 2, etc., until we find a Nash equilibrium. There are $\binom{n}{d}^2$ pairs of supports of size d and determining whether a pair of supports induce a Nash equilibrium can be done in polynomial-time via a linear program. Thus, provided that our game has a Nash equilibrium induced by supports of constant size, we obtain a polynomial time algorithm. We will show that with high probability, a random game has a 2×2 equilibrium.

More generally, we consider the probability that there is no $d \times d$ Nash equilibrium in a random game. Let S and T be strategy subsets of Alice and Bob respectively, where $|S| = |T| = d$. We let $\mathcal{F}(T) = \{S_1, S_2, \ldots, S_p\}$ the set of facets of the polytope $\mathsf{conv}(A^T)$ with nonnegative normal vectors (here $\mathsf{E}(p) = N_d^+$); each S_i corresponds to the set of rows that induce the facet. We also define $\bar{\mathcal{F}}(T)$ to be the set of all faces contained in facets of $\mathcal{F}(T)$. To avoid confusion between row and column vectors, the sets $\mathcal{G}(S)$ and $\bar{\mathcal{G}}(S)$ are defined similarly w.r.t. the polytope $\mathsf{conv}(B^S)$. Then $S \in \mathcal{F}(T)$ iff Alice's strategies induced by S are all best responses to some mixed strategy by Bob on the strategy set induced by T. Note that S and T induce a Nash equilibrium, denoted by $S \leftrightarrow T$, if and only if $S \in \mathcal{F}(T)$ and $T \in \mathcal{G}(S)$. We denote by $\mathcal{E}_{ST}$ the event that $S \leftrightarrow T$, and by χ_{ST} the indicator variable for this event. By the independence of the payoff matrices for Alice and Bob, the probability that $S \leftrightarrow T$ is exactly the product of the probabilities that $S \in \mathcal{F}(T)$ and $T \in \mathcal{G}(S)$.

We begin with the expected number of $d \times d$ Nash equilibria.

Lemma 6. *The expected number of Nash equilibria in a random game is*

$$\mathsf{E}\left(\sum_{S,T:|S|=|T|=d} \chi_{ST}\right) = (N_d^+)^2.$$

Proof. Given S let us first evaluate the probability that $T \in \mathcal{G}(S)$ for some T. We have seen that this is the case only if $B^S(T)$ induces a useful facet in B^S. Thus, this probability is exactly

$$\frac{N_d^+}{\binom{n}{d}}.$$

Similarly, we also have that

$$\mathsf{P}(S \in \mathcal{F}(T)) = \frac{N_d^+}{\binom{n}{d}}.$$

By the independence of the payoff matrices, we obtain $\mathsf{P}(S \leftrightarrow T) = \left(\frac{N_d^+}{\binom{n}{d}}\right)^2$. Summing up over all pairs, S, T, the lemma follows. □

Let μ denote the above expectation. We are interested in the probability that the random variable $Z = \sum_{S,T} \chi_{ST} > 0$. We will use the notation

$$\triangle = \sum_{(S,T),(S',T'):(S,T)\bowtie(S',T')} \mathsf{P}(\mathcal{E}_{ST} \wedge \mathcal{E}_{S'T'}),$$

where $\bowtie$ signifies that the events $\mathcal{E}_{ST}$ and $\mathcal{E}_{S'T'}$ are dependent. We remark that such a dependency arises if and only if either $S \cap S' \neq \emptyset$ or $T \cap T' \neq \emptyset$. Our interest in $\triangle$ arises because, applying standard techniques [2], we obtain:

$$\mathsf{P}(Z = 0) \leq \frac{Var(Z)}{\mathsf{E}(Z)^2} \leq \frac{\mathsf{E}(Z) + \triangle}{\mathsf{E}(Z)^2} \tag{2}$$

The next lemma bounds Δ.

Lemma 7. *In a random game*

$$\triangle \lesssim \mu^2 \left(\frac{1}{n} + \frac{1}{(N_1^+)^2}\right).$$

Proof. Let $X = S^* \cap S$ and $Y = T^* \cap T$, $s = |X|$ and $t = |Y|$. We remark that only the cardinalities of these intersections will be of consequence. Recall that $X \in \bar{\mathcal{F}}(T^*)$ means that the vertices corresponding to X form a useful face (i.e., are all best responses) in the game induced by the strategies of T^*. So

$$\begin{aligned}
\triangle &= \sum_{s,t} \sum_{\substack{(S,T),(S^*,T^*):\\ |X|=s,|Y|=t}} \mathsf{P}(S \leftrightarrow T \text{ and } S^* \leftrightarrow T^*) \\
&= \sum_{s,t} \sum_{\substack{(S,T),(S^*,T^*):\\ |X|=s,|Y|=t}} [\mathsf{P}(S \in \mathcal{F}(T) \text{ and } S^* \in \mathcal{F}(T^*)) \\
&\qquad\qquad \times \mathsf{P}(T \in \mathcal{G}(S) \text{ and } T^* \in G(S^*))) \\
&= 2\sum_{s\geq 1} \sum_{\substack{(S,T),(S^*,T^*):\\ |X|=s,|Y|=0}} [\mathsf{P}(S \in \mathcal{F}(T) \text{ and } S^* \in \mathcal{F}(T^*)) \\
&\qquad\qquad \times \mathsf{P}(T \in \mathcal{G}(S) \text{ and } T^* \in \mathcal{G}(S^*))] \\
&\quad + \sum_{s,t\geq 1} \sum_{\substack{(S,T),(S^*,T^*):\\ |X|=s,|Y|=t}} [\mathsf{P}(S \in \mathcal{F}(T) \text{ and } S^* \in \mathcal{F}(T^*)) \\
&\qquad\qquad \times \mathsf{P}(T \in \mathcal{G}(S) \text{ and } T^* \in \mathcal{G}(S^*))] \\
&= 2\triangle_1 + \triangle_2.
\end{aligned}$$

The following claims will be useful in bounding these terms.

Claim 1. *For n sufficiently larger than d, if $T \cap T^* = 0$ then*

$$\mathsf{P}(T \in \mathcal{G}(S) \wedge T^* \in \mathcal{G}(S^*)) \lesssim \left(\frac{N_d^+}{\binom{n}{d}}\right)^2$$

Proof. Let $\mathcal{G}^{-T}(S^*)$ be the set of useful facets of the polytope induced by the rows of S^* if we ignore the d points corresponding to the columns of T; define $\mathcal{G}^{-T^*}(S)$ similarly. Then clearly

$$\begin{aligned}
&\mathsf{P}(T \in \mathcal{G}(S) \wedge T^* \in \mathcal{G}(S^*)) \\
&\quad\leq \quad \mathsf{P}(T \in \mathcal{G}^{-T^*}(S) \wedge T^* \in \mathcal{G}^{-T}(S^*)) \\
&\quad= \quad \mathsf{P}(T \in \mathcal{G}^{-T^*}(S) | T^* \in \mathcal{G}^{-T}(S^*)) \mathsf{P}(T^* \in \mathcal{G}^{-T}(S^*))
\end{aligned}$$

But $\mathsf{P}(T \in \mathcal{G}^{-T^*}(S) | T^* \in \mathcal{G}^{-T}(S^*))$ is maximised when $S = S^*$. The result then follows from Lemma 3. □

Similarly, applying Lemma 4, or Lemma 5, we obtain

Claim 2. *For n sufficiently larger than d,*

$$\mathsf{P}(T \in \mathcal{G}(S) \textit{ and } T^* \in \mathcal{G}(S^*)) \lesssim \frac{(N_d^+)^2 \binom{n}{t}}{\binom{n}{d}^2 N_t^+}$$

□

We are now ready to complete the proof of Lemma 7. First we bound $\triangle_1$ from above. Observe that since $t = 0$

the events $S \in \mathcal{F}(T)$ and $S^* \in \mathcal{G}(T^*)$ are independent. Therefore,

$$\begin{aligned}\mathsf{P}(S \in \mathcal{F}(T) \text{ and } S^* \in \mathcal{F}(T^*)) &= \mathsf{P}(S \in \mathcal{F}(T))^2 \\ &= \left(\frac{N_d^+}{\binom{n}{d}}\right)^2\end{aligned}$$

Applying Claim 1 we have

$$\begin{aligned}\triangle_1 &= \sum_{s\geq 1} \sum_{\substack{(S,T),(S^*,T^*): \\ |X|=s,|Y|=0}} [\mathsf{P}(S \in \mathcal{F}(T) \text{ and } S^* \in \mathcal{F}(T^*)) \\ &\qquad\qquad \times \mathsf{P}(T \in \mathcal{G}(S) \text{ and } T^* \in \mathcal{G}(S^*))] \\ &\lesssim \sum_{s\geq 1} \sum_{\substack{(S,T),(S^*,T^*): \\ |X|=s,|Y|=0}} \frac{(N_d^+)^4}{\binom{n}{d}^4}\end{aligned}$$

However,

$$\begin{aligned}&\frac{(N_d^+)^4}{\binom{n}{d}^4} \sum_{s:d\geq s\geq 1} \sum_{\substack{(S,T),(S^*,T^*): \\ |X|=s,|Y|=0}} 1 \\ &= \frac{\mu^2}{\binom{n}{d}^4} \sum_{s:d\geq s\geq 1} \binom{n}{d}\binom{d}{s}\binom{n-d}{d-s}\binom{n}{d}\binom{n-d}{d} \\ &\lesssim \frac{\mu^2}{n}\end{aligned}$$

Thus,

$$\triangle_1 \lesssim \frac{\mu^2}{n}$$

Next, we work towards bounding $\triangle_2$. Applying Claim 2, we obtain

$$\begin{aligned}\triangle_2 &= \sum_{s,t\geq 1} \sum_{\substack{(S,T),(S^*,T^*): \\ |X|=s,|Y|=t}} [\mathsf{P}(S \in \mathcal{F}(T) \text{ and } S^* \in \mathcal{F}(T^*)) \\ &\qquad\qquad \times \mathsf{P}(T \in \mathcal{G}(S) \text{ and } T^* \in \mathcal{G}(S^*))] \\ &\lesssim \sum_{s,t\geq 1} \sum_{\substack{(S,T),(S^*,T^*): \\ |X|=s,|Y|=t}} \frac{(N_d^+)^4\binom{n}{s}\binom{n}{t}}{\binom{n}{d}^4 N_s^+ N_t^+}\end{aligned}$$

Now we have

$$\mu^2 \sum_{s,t\geq 1} \frac{\binom{d}{s}\binom{d}{t}}{N_s^+ N_t^+} \lesssim \frac{\mu^2}{(N_1^+)^2}$$

Thus

$$\triangle_2 \lesssim \frac{\mu^2}{(N_1^+)^2}$$

and this completes the proof of Lemma 7. □

Theorem 1 now follows from (2) and Lemma 7. Then, by Theorems 2 and 3, which imply that N_1^+ becomes much larger than $f(d)$ as n increases, we obtain our result for random games in which the payoff entries are either uniformly or Normally distributed. In fact, it is enough to look for 2×2 equilibria.

It follows that the run time of the algorithm is $O(m^2 n \log\log n + n^2 m \log\log m)$, with high probability. To see this observe that we need to calculate the convex hull of n points for each pair of strategies of Alice, and calculate the convex hull of m points for each pair of strategies of Bob. Since we can find the convex hull of k points in 2 dimensions in time $O(k \log h)$, where h is the number of points on the convex hull [21], this takes time $O(m^2 n \log\log n + n^2 m \log\log m)$. If a pair of strategies for Alice and a pair of strategies for Bob mutually induce facets with non-negative normals in the convex hull associated with the other pair then we have a Nash equilibrium. The normals to these facets also give the probability distributions on the strategy supports (at the Nash equilibrium). Thus our algorithm is entirely combinatorial.

5 Concluding Remarks

We have show that finding equilibria on average is easy. This raises several questions: (i) Can we extend the analysis to more general distributions? Our result is unaffected by linear transformations of the payoff matrices. So, for example, each matrix can be chosen from an arbitrary Gaussian. (ii) Does our algorithm have polynomial expected running time? (iii) We crucially use the fact that the mean of each entry is zero. Is this necessary? An algorithm that works for Gaussian entries with arbitrary means (and time polynomial in the largest variance), akin to smoothed analysis [35], would give a polynomial-time randomized algorithm for finding approximate Nash equilibria in arbitrary games [19]: add random Gaussians to the entries of the given payoff matrices; an equilibrium of the perturbed game will be an approximate equilibrium of the original game with high probability, given that the variance of the Gaussians is small enough. The current best algorithm for finding approximate equilibria has quasi-polynomial complexity [22].

Finally, we observe that finding approximate equilibria in random games is quite easy: for both the distributions we consider, with high probability there will be many pure strategy approximate equilibria and hence one of them can be found in sublinear time.

References

[1] F. Affentranger and J. Wieacker, "On the convex hull of a uniform random points in a simple d-polytope",

Discrete and Computational Geometry, **6**, pp291-305, 1991.

[2] N. Alon and J. Spencer, *The Probabilistic Method*, 2nd edition, Wiley, 2000.

[3] D. Avis and K. Fukuda, “A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra”, *Discrete Computational Geometry*, **8**, pp295-313, 1992.

[4] I. Bárány, “Intrinsic volumes and f-vectors of random polytopes”, *Math. Ann.* **285**, pp671-699, 1989.

[5] I. Bárány and L. Dalla, “Few points to generate a random polytope”, *Mathematika* **44**, pp325-331, 1997.

[6] I. Bárány and D. Larman, “Convex bodies, economic cap coverings, random polytopes”, *Mathematika* **35**, pp274-291, 1988.

[7] K. Borgwardt, *The Simplex Method: a probabilistic analysis*, Springer-Verlag, 1987.

[8] C. Buchta, “On the average number of maxima in a set of vectors”, *Information Processing Letters*, **33**, pp63-65, 1989.

[9] C. Buchta, “A remark on random approximations of simple polytopes”, *Anz. sterreich. Akad. Wiss. Math.-Natur. Kl.*, **126**, pp17-20, 1989.

[10] V. Conitzer and T. Sandholm, “Complexity results about Nash equilibria”, *Proceedings of IJCAI*, pp765-771, 2003.

[11] L. Devroye, “A note on finding convex hulls via maximal vectors”, *Information Processing Letters*, **11(1)**, pp53-56, 1980.

[12] R. Dwyer, “On the convex hull of random points in a polytope”, *Journal of Applied Probability*, **25**, pp688-699, 1988.

[13] R. Dwyer, “Convex hulls of samples from spherically symmetric distributions”, *Discrete Applied Mathematics*, **31**, pp113-132, 1991.

[14] A. Fabrikant, C. Papadimitriou and K. Talwar, “The complexity of pure-strategy Nash Equilibria”, *Proceedings of 36th STOC*, pp604-612, 2004.

[15] W. Feller, *An Introduction to Probability Theory and it Applications (Vol. I)*, Wiley, 1950.

[16] H. Groemer, “Limit theorems for the convex hull of random points in higher dimensions”, *Pacific Journal of Mathematics*, **45(2)**, pp525-533, 1973.

[17] I. Heuter, “Limit theorems for the convex hull of random points in higher dimensions”, *Transactions of the American Mathematical Society*, **351**, pp4337-4363, 1999.

[18] S. Janson, “Poisson approximation for large deviations”, *Random Structures and Algorithms*, **1**, pp221-230, 1990.

[19] A. Kalai, personal communication, 2005.

[20] J. Kingman, “Random secants of a convex body”, *Journal of Applied Probability*, **6**, pp66-672, 1969.

[21] D. Kirkpatrick and R. Seidel, “The ultimate planar convex hull algorithm”, *SIAM J. Comp.*, **15**, pp287-299, 1988.

[22] R. Lipton, E. Markakis and A. Mehta, “Playing large games using simple strategies”, *Proceedings of E-Commerce*, pp36-41, 2003.

[23] A. McLennan, “The expected number of Nash equilibria in a normal form game”, to appear in *Econometrica*.

[24] A. McLennan and J. Berg, ”The asymptotic expected number of Nash equilibria of two player normal form games”, to appear in *Games and Economic Behavior*.

[25] N. Megiddo, “Improved asymptotic analysis of the average number of steps performed by the self-dual simplex algorithm”, *Mathematical Programming*, **35(2)**, pp140-172, 1986.

[26] N. Megiddo and C. H. Papadimitriou, ”On total functions, existence theorems and computational complexity”, *Theoret. Comput. Sci.*, **81(2)**, pp317-324, 1991.

[27] C. H. Papadimitriou, ”On the complexity of the parity argument and other inefficient proofs of existence”, *Journal of Computer and System Sciences*, **48(3)**, pp498-532, 1994.

[28] C. H. Papadimitriou, ”Algorithms, games and the internet”, *Proceedings of 33rd STOC*, pp749-753, 2001.

[29] C. H. Papadimitriou, ”Computing correlated equilibria in multiplayer game”, *Proceedings of 37th STOC*, pp49-56, 2005.

[30] C. Papadimitriou and T. Roughgarden, “Computing equilibria in multi-player games”, *Proceedings of 16th SODA*, pp82-91, 2005.

[31] H. Raynaud, “Sur l’enveloppe convexe des nuages de points aleatiores dans $\mathbb{R}^n$. I.”, *Journal of Applied Probability*, **7**, pp35-48, 1970.

[32] R. Savani and B. von Stengel, "Exponentially Many Steps for Finding a Nash Equilibrium in a Bimatrix Game", *Proceedings of 45th FOCS*, pp258-267, 2004.

[33] R. Schneider, "Discrete aspects of stochastic geometry", in eds. J. Goodman and J. O'Rourke, *Handbook of Discrete and Computational Geometry*, Boca Raton, CRC Press, pp255-278, 2004.

[34] S. Smale, "On the average number of steps in the simplex method of linear programming", *Mathematical Programming*, **27**, pp241-262, 1983.

[35] D. A. Spielman, S. Teng, "Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time", *Journal of the ACM*, **51(3)**, pp385-463, 2004.

[36] B. von Stengel, "New maximal numbers of equilibria in bimatrix games", *Discrete and Computational Geometry*, **21**, pp557-568, 1999.

[37] B. von Stengel, "Computing equilibria for two-person games", in eds. R. Aumann and S. Hart, *Handbook of Game Theory, Vol. 3*, North-Holland, Amsterdam, pp1723-1759, 2002.

[38] B. van Wel, "The convex hull of a uniform sample from the interior of a simple d-polytope", *Journal of Applied Probability*, **26**, pp259-273, 1989.

[39] W. Weil and J. Wieacker, "Stochastic geometry", In *Handbook of convex geometry*, North-Holland, Amsterdam, pp1391-1438, 1993.