Approximate Nash Equilibria in Anonymous Games

Constantinos Daskalakis[*]
EECS, MIT
costis@mit.edu

Christos H. Papadimitriou[†]
EECS, UC Berkeley
christos@cs.berkeley.edu

January 26, 2014

Abstract

We study from an algorithmic viewpoint *anonymous games* [Mil96, Blo99, Blo05, Kal05]. In these games a large population of players shares the same strategy set and, while players may have different payoff functions, the payoff of each depends on her own choice of strategy and the number of the other players playing each strategy (*not* the identity of these players). We show that, the intractability results of [DGP09a] and [Das11] for general games notwithstanding, approximate mixed Nash equilibria in anonymous games can be computed in polynomial time, for any desired quality of the approximation, as long as the number of strategies is bounded by some constant. In addition, if the payoff functions have a Lipschitz continuity property, we show that an approximate pure Nash equilibrium exists, whose quality depends on the number of strategies and the Lipschitz constant of the payoff functions; this equilibrium can also be computed in polynomial time. Finally, if the game has two strategies, we establish that there always exists an approximate Nash equilibrium in which either only a small number of players randomize, or of those who do, they all randomize the same way. Our results make extensive use of certain novel Central Limit-type theorems for discrete approximations of the distributions of multinonial sums.

JEL Classification: C72 (Noncooperative Games)
Keywords: Anonymous Games, Nash Equilibrium, Approximation Algorithms

[*]Supported by NSF Awards CCF-0953960 (CAREER) and CCF-1101491, a Sloan Foundation Fellowship, and a Microsoft Research Faculty Fellowship.

[†]Supported by NSF Award CCF-0964033 and a Google University Research Award.

1 Introduction

The concept of Nash equilibrium is paramount in Game Theory. It is not the only solution concept, or the most sophisticated, or even the most widely accepted one, but it is the golden standard against which all other solution concepts are measured. In his 1999 paper [Mye99], Myerson argues that the Nash equilibrium lies at the foundations not just of Game Theory, but of all modern economic thought. This is because the exacting brand of rationality that it exudes became the standard way of looking at economic problems. Another reason is because Nash's proof inspired the seminal price equilibrium results of Arrow and Debreu [AD54]. Nash's existence theorem is crucial in this regard because it establishes the concept's *universality*—the fact that it is guaranteed to exist in every conceivable situation—in the absence of which no solution concept can be taken seriously.

Seen in this light, the recent result that finding a Nash equilibrium in a game is a computationally intractable problem [DGP06, DGP09b, DGP09a] is quite disturbing. It suggests that, assuming that the world of computation is the way it is broadly believed to be, the coveted universality of the Nash equilibrium is suspect, indeed that it may be false in practice, since there are games whose Nash equilibria, even though they exist, are inaccessible within any conceivable time scale.

Computer scientists have over the past decades learned to live with this kind of complexity. Once a problem of interest is shown to be intractable, more modest goals are pursued: One seeks to solve important special cases of the problem, or to solve them approximately. That is what we accomplish in this paper: we identify a very broad and significant special class of games, the *anonymous games* studied in the past (see [Mil96, Blo99, Blo05, Kal05] and the definition below), and develop algorithms for finding *approximate Nash equilibria* in such games. By "approximate Nash equilibrium" we mean a strategy profile in which each player cannot improve her lot by more than a fixed amount denoted ϵ.

But applying such approaches to problems in Game Theory is rife with subtlety. For hard optimization problems such as the traveling salesman problem, algorithms achieving some approximation ϵ is meaningful, no matter how large ϵ may be—because doing better may be difficult. In contrast, unless ϵ is minuscule (so small that no reasonable person would lift a finger, so to speak, to acquire it), an approximate Nash equilibrium is not interesting; the reason is, playing best response is always easy in games, and thus approximate Nash equilibria with significant ϵ lack stability—that is to say, they are no equilibria at all. Consequently, approximations of the Nash equilibrium are of interest only if the approximation gap ϵ can be made arbitrarily small. An efficient approximation algorithm that can be made to work for arbitrarily small ϵ is called a *polynomial-time approximation scheme* or *PTAS* for short (see Section 4 for the formal definition). Whether there is a PTAS for finding Nash equilbria is the central open problem in this area. A positive answer would go a long way towards moderating the negative implications of the complexity result [DGP06]. Here we do not settle this important question, but we do develop a PTAS for a large and important class of games, namely the anonymous games.

A second difficulty in applying complexity concepts to games has to do with *input size and representation*. Games are useful modeling tools in Economics, and, especially, in Computer Science in the context of the Internet, often only to the extent that they may involve large numbers of players. However, multi-player games are a challenge to deal with computationally. An algorithm works on an input containing all relevant information about the problem being solved. To describe an n-player game, in which each player has $\xi \geq 2$ strategies, takes $n\xi^n$ numbers—an astronomical amount of information when n is reasonably large. Consequently, generic multiplayer games are of no computational interest; one must focus on broad and well-motivated classes of games that can be represented *succinctly*—that is, by an amount of information polynomial in n, the number of players. Indeed, over the past decade several classes of such "succinct games" have been identified.

One such class are *graphical* or *network games,* in which players are nodes in a sparsely connected network, and the payoff of each player depends on the actions only of its neighbors in the network. However, approximate Nash equilibria in network games seem very hard to find [Das11].

Anonymous games comprise a second interesting class of games which can be represented succinctly. In anonymous games all players have the same action set $\{1, 2, \ldots, \xi\}$, and our working hypothesis is that ξ is finite, often two. In an anonymous game with n players and ξ strategies, the payoff of a player depends only on two things: (a) the strategy played by the player; and (b) the number of other players who play each of the strategies. Imagine, for example, a population of commuters living in a suburb and deciding each morning between driving and taking the train. Each player has an individual way of evaluating highway congestion and crowded train cars, but only the numbers of other players matter in this evaluation, not their identities. An anonymous game can be represented by fewer than ξn^{ξ} numbers, which, for fixed ξ, is a polynomial in n; compare with $n\xi^n$ for general games. Notice, incidentally, that anonymous games generalize the more familiar class of *symmetric* games (i.e. games invariant under all player permutations) for which it had been known for some time that Nash equilibria can be computed exactly in polynomial time when ξ is a constant [PR05].

In this paper we prove several results regarding approximate Nash equilibria in anonymous games:

1. We develop a PTAS for finding approximate mixed Nash equilibria in any anonymous game, as long as the number ξ of strategies is constant (Theorem 2). The running time is polynomial in the game description but exponential in the approximation $1/\epsilon$ and the number of pure strategies ξ. Our algorithm is derived by a novel and rather comprehensive methodology that may be of interest in its own right: Any mixed strategy profile generates a distribution over the set of partitions of the players to strategies. We establish that, essentially, if two mixed strategy profiles generate distributions that are close (in one particular metric over distributions called variational distance), and one of them is a Nash equilibrium, then the other is an approximate Nash equilibrium (Lemma 4). We then invoke a theorem from probability theory (Theorem 3) stating that any mixed strategy profile can be discretized, in the sense that the probabilities are multiples of a given fraction $\frac{1}{k}$, so that the generated distribution does not move much in variational distance. The conclusion is that an approximate Nash equilibrium can be found by exhaustively enumerating all strategy profiles with probabilities that are multiples of $\frac{1}{k}$. These are exponentially many, but using maximum flow techniques they can be searched over in polynomial time.

2. For the case of anonymous games with two strategies, we show that there is always an approximate Nash equilibrium with a dichotomous structure: Either very few players randomize, or of those who do they all randomize the same way (Theorem 7). Our structural result is reminiscent of Nash's theorem that symmetric games always have a Nash equilibrium where all players play the same mixed strategy [Nas51]. It is surprising that a similar result to Nash's can be proven for approximate Nash equilibria in anonymous games, despite the weaker symmetry structure present in these games. Again, such an equilibrium can be found in polynomial time using maximum flow techniques. Indeed, exploiting the structural result we show that we can significantly improve the running time of our PTAS (Theorem 6).

3. Both algorithms discussed so far involve exhaustive enumeration. As a result, they are quite robust: We argue that they can be generalized to more complex situations, in which, for example, players are divided into *types,* and utilities depend on the number of players *of each type* using each strategy.

4. At the same time, these algorithms are *oblivious*. They enumerate over a set of (permuted) candidate equilibria with the property that, for every anonymous game, there exists an element in this set that, appropriately permuted, is an ϵ-Nash equilibrium of the game. The algorithms are called "oblivious" because they look at the input not in order to guide the search for an approximate Nash equilibrium, but only to check whether the enumerated profiles are approximate equilibria. Nevertheless obliviousness should come at a cost. Indeed, we show that any oblivious algorithm must pay exponential in $1/\epsilon$ running time (Theorem 8).

5. In view of this lower bound for oblivious algorithms we devise a non-oblivious algorithm for 2-strategy anonymous games whose running time is polynomial in the game-description times a factor of $(1/\epsilon)^{O(\log^2 1/\epsilon)}$ (Theorem 10). The algorithm is based on a strong probabilistic theorem for sums of independent indicators (Theorem 9) and it enumerates over a set of moments of mixed strategy profiles rather than of mixed strategy profiles.

6. Finally, we address the question of whether pure approximate Nash equilibria exist in anonymous games, inspired by positive answers for special kinds of anonymous games [Mil96]. In a large anonymous game, it is natural to assume that the payoff functions have some kind of *continuity* property, namely that the payoffs don't jump by more than a small parameter λ (the Lipschitz constant of the game) if only one other player changes their strategy. We point out through a fixed point argument that such games always have an approximate pure Nash equilibrium with $\epsilon = O(\lambda\xi)$ (Theorem 1 in Section 3), and this equilibrium can be found efficiently. Since it might be expected that λ is small, about $\frac{1}{n}$ (we assume that payoffs have been normalized to lie in $[0,1]$), this may be considered a reasonable approximation.

Related Work

For the long history of algorithms for computing exact and approximate Nash equilibria in general games see the survey in [MM96], and see [DGP09b] for the recently established intractability of the problem. An approximation scheme for Nash equilibria in general games is provided in [LMM03]. For 2 player games, or a constant (non-scaling) number of players, their algorithm is sub-exponential (but not polynomial), and is based on an existence theorem establishing, through sampling the Nash equilibrium, that there always exists an approximate Nash equilibrium with small (logarithmic) support, i.e. in which the players' mixed strategies are distributions over a logarithmic size subset of the pure strategies. Their algorithm then enumerates all such strategy profiles until it finds an approximate Nash equilibrium, guaranteed to exist by the existence theorem. So their algorithm is oblivious. In comparison to [LMM03], our approximation schemes are computationally efficient, and apply to different games: anonymous, with a large (scaling) number of players, and a constant (non-scaling) number of strategies. Moreover, they perform a different kind of enumeration, based on probabilistic cover constructions (Theorems 3, 5, and 9) and exploiting the symmetry of anonymous games, and they are not necessarily oblivious (Theorem 10).

Anonymous games have been studied quite extensively in the Economics literature, see [Mil96, Blo99, Blo05, Kal05]. One very special genre of anonymous games are the so-called *congestion games*, first considered by Rosenthal [Ros73], and more recently studied extensively from the points of view of the quality of the Nash equilibria (see the survey [Rou05]) and the complexity of computing [FPT04, ARV08], or approximating them [CS11, CFGS12].

A weaker form of Theorem 1, Corollary 1, and Theorem 6 of this paper were given in [DP07]. Theorem 2 was given in [DP08], Theorems 6 and 7 in [Das08a], and Theorems 8 and 10 in [DP09]. The probabilistic approximation theorems employed in our proofs are, as indicated when they are stated, from [DP08], [Das08a] and [DP09]. See also [DP13] for a survey of these probabilistic theorems, with minor improvements (which can also be used to slightly improve our running times).

2 Notation

Games. An *anonymous game* is a triple $G = (n, \xi, \{u_\ell^i\}_{i\in[n],\ell\in[\xi]})$ where $[n] = \{1, \dots, n\}$, $n \geq 2$, is the set of players, $[\xi] = \{1, \dots, \xi\}$, $\xi \geq 2$, a common set of strategies available to all players, and u_ℓ^i the payoff (or utility) function of player i when she plays strategy ℓ. This maps the set of partitions $\Pi_{n-1}^{\xi} = \{(x_1, \dots, x_\xi) \mid x_\ell \in \mathbb{N}_0 \text{ for all } \ell \in [\xi] \ \wedge \ \sum_{\ell=1}^{\xi} x_\ell = n-1\}$ to the interval $[0, 1]$. That is, it is assumed that the payoff of each player depends on her own strategy and only the number of the other players choosing each of the ξ strategies. Restricting the payoffs to $[0, 1]$ is, of course, no loss of generality. But, once this has been assumed, additive approximations, described next, become meaningful. Games whose payoffs lie in $[0, 1]$ are sometimes called *normalized.*

Our working assumptions are that n is large and ξ is fixed. Hence, anonymous games are *succinct games* in the sense that their representation requires specifying at most ξn^ξ (i.e., a fixed polynomial in the number of players) numbers, as opposed to the $n\xi^n$ (i.e. exponential in the number of players) numbers required for general games of n players and ξ strategies. Arguably, succinct games are the only multiplayer games that are computationally meaningful; see [PR05] for an extensive discussion of this point.

An anonymous game will be called λ-*Lipschitz* for some real $\lambda > 0$ iff, for all i, ℓ, $|u_\ell^i(x) - u_\ell^i(y)| \leq \lambda \cdot ||x - y||_1$ for all $x, y \in \Pi_{n-1}^{\xi}$, where $||x - y||_1$ is the ℓ_1 distance between x and y as vectors in $\mathbb{R}^\xi$. We will generally not require our anonymous games to be λ-Lipschitz except for our results of Section 3 on pure Nash equilibria.

Finally, we denote by Δ_{n-1}^{ξ} the convex hull of the set Π_{n-1}^{ξ}. That is, $\Delta_{n-1}^{\xi} = \{(x_1, \dots, x_\xi) \mid x_\ell \geq 0 \text{ for all } \ell \in [\xi] \wedge \sum_{\ell=1}^{\xi} x_\ell = n-1\}$. We also use the shorthand Δ^ξ for the set Δ_1^ξ of distributions over $[\xi]$. A *mixed strategy* is an element of Δ^ξ.

Approximate Equilibria. A *pure strategy profile* is a mapping S from $[n]$ to $[\xi]$. A pure strategy profile S is an ϵ-*approximate pure Nash equilibrium* for some $\epsilon \geq 0$ iff

$$u_{S(i)}^i(x[S, i]) + \epsilon \geq u_\ell^i(x[S, i]), \text{for all } i \in [n] \text{ and } \ell \in [\xi],$$

where $x[S, i] \in \Pi_{n-1}^{\xi}$ is the partition $(x_1, \dots, x_\xi)$ where x_t is the number of players $j \in [n] - \{i\}$ such that $S(j) = t$.

Similarly a *mixed strategy profile* is a mapping δ from $[n]$ to Δ^ξ, and we denote by δ_i the mixed strategy of player i in this profile, and by δ_{-i} the collection of all mixed strategies but i's in δ. A mixed strategy profile δ is an ϵ-*approximate mixed Nash equilibrium* for some $\epsilon \geq 0$ iff

$$\mathbb{E}_{x\sim\delta_{-i},\ell\sim\delta_i} u_\ell^i(x) + \epsilon \geq \mathbb{E}_{x\sim\delta_{-i}} u_t^i(x), \text{for all } i \in [n] \text{ and } t \in [\xi],$$

where for the purposes of the above expectations ℓ is drawn from $[\xi]$ according to δ_i and x is drawn from Π_{n-1}^{ξ} by drawing $n-1$ random samples from $[\xi]$ independently according to the distributions $\delta_j, j \neq i$, and forming the induced partition.

A stronger notion of approximation is that of an ϵ-*Nash equilibrium*, or ϵ-*well-supported Nash equilibrium.* This is a mixed strategy profile δ satisfying the following condition.

$$\text{For all } i \in [n] \text{ and } \ell, \ell' \in [\xi]\text{:} \ \mathbb{E}_{x\sim\delta_{-i}} u_\ell^i(x) > \mathbb{E}_{x\sim\delta_{-i}} u_{\ell'}^i(x) + \epsilon \Longrightarrow \delta_i(\ell') = 0,$$

where, as above, for the purposes of the expectations x is drawn from Π_{n-1}^{ξ} by drawing $n-1$ random samples from $[\xi]$ independently according to the distributions $\delta_j, j \neq i$, and forming the induced partition. It is easy to verify that an ϵ-Nash equilibrium is also an ϵ-approximate mixed Nash equilibrium but the converse need not be true. In this paper we focus on ϵ-Nash equilibria. Hence our algorithms for ϵ-Nash equilibria readily compute also ϵ-approximate mixed Nash equilibria.

Dynamic Programming. It is worth noting that given a mixed strategy profile δ it is not immediate how to compute a player's expected payoff computationally efficiently, since there are exponentially many pure strategy profiles contributing to the expectation. However, we can do this computation efficiently using dynamic programming. We explain the idea for $\xi = 2$, but the computation is similar for general ξ. Suppose that, given δ, we want to compute player n's expected payoff from strategy 1. The expected payoff can be written as:

$$u_1^n((0, n-1)) \cdot \Pr[X = 0] + u_1^n((1, n-2)) \cdot \Pr[X = 1] + \ldots + u_1^n((n-1, 0)) \cdot \Pr[X = n-1],$$

where X is the random variable representing how many of the players $1, \ldots, n-1$ play strategy 1 under mixed strategy profile δ_{-n}.

Hence, it suffices to compute the probabilities $\Pr[X = 0], \ldots, \Pr[X = n-1]$ efficiently. Notice that X can be written as the sum $X = \sum_{i=1}^{n-1} X_i$, where X_i is a Bernoulli random variable indicating whether player i plays strategy 1, which happens with probability $\delta_i(1)$ independently of the values of the other random variables. To compute the probabilities $\Pr[X = 0], \ldots, \Pr[X = n-1]$ efficiently, we fill the entries of a $(n-1) \times n$ table $T(i, \ell)$, where $i \in \{1, \ldots, n-1\}$ and $\ell \in \{0, \ldots, n-1\}$ as follows. Entry $T(i, \ell)$ is supposed to contain the value $\Pr[\sum_{j \le i} X_j = \ell]$. Thus we first set $T(1, 1) = \delta_1(1)$, $T(1, 0) = \delta_1(2)$ and $T(i, \ell) = 0$, for all i and $\ell > i$. To complete the rest of the table we work bottom up, filling layer $T(2, \cdot)$ first, then layer $T(3, \cdot)$, etc. To fill layer $T(i, \cdot)$, $i \in \{2, \ldots, n-1\}$, we use the formula:

$$T(i, \ell) = \begin{cases} \delta_i(1) \cdot T(i-1, \ell-1) + \delta_i(2) \cdot T(i-1, \ell), & \text{if } 0 < \ell < i; \\ \delta_i(1) \cdot T(i-1, i-1), & \text{if } \ell = i; \\ \delta_i(2) \cdot T(i-1, 0), & \text{if } \ell = 0; \\ 0, & \text{if } \ell > i. \end{cases}$$

Clearly, filling in table $T(\cdot, \cdot)$ as prescribed above takes polynomial time. Once this is complete, we can read off the values $\Pr[X = 0], \ldots, \Pr[X = n-1]$ by looking at the last layer of the table. Namely, $\Pr[X = \ell] \equiv T(n-1, \ell)$, for all $\ell \in \{0, \ldots, n-1\}$.

Probability Tools. The mixed strategy δ_i of player $i \in [n]$ defines a random unit vector $\mathcal{X}_i$ ranging in $\{e_1, \ldots, e_\xi\}$, where e_ℓ is the unit vector along dimension ℓ of $\mathbb{R}^\xi$. We let $\mathcal{X}_i$ take values according to the measure $\Pr[\mathcal{X}_i = e_\ell] = \delta_i(\ell)$, for all ℓ. With this convention, if $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ is a mixed strategy profile, then the expected payoff of player $i \in [n]$ for using pure strategy $\ell \in [\xi]$ is just

$$\mathbb{E} u_\ell^i \left(\sum_{j \neq i} \mathcal{X}_j \right),$$

where for the purposes of the expectation the variables $\mathcal{X}_1, \ldots, \mathcal{X}_n$ are taken to be independent.

We will also define the *total variation distance* between two distributions $\mathbb{P}$ and $\mathbb{Q}$ over a finite set $\mathcal{A}$ as follows.

$$||\mathbb{P} - \mathbb{Q}||_{\text{TV}} = \frac{1}{2} \sum_{\alpha \in \mathcal{A}} |\mathbb{P}(\alpha) - \mathbb{Q}(\alpha)|.$$

Similarly, if X and Y are two random variables ranging over a finite set, their total variation distance, denoted

$$||X \; - \; Y||_{\text{TV}},$$

is defined to be the total variation distance between their distributions.

We will use the following simple lemma.

Lemma 1. *Let $\mathcal{X}_1, \ldots, \mathcal{X}_n$ be mutually independent random vectors, and $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$ also be mutually independent random vectors. Then*

$$\left\| \sum_{i=1}^{n} \mathcal{X}_i \;-\; \sum_{i=1}^{n} \mathcal{Y}_i \right\|_{\mathrm{TV}} \le \sum_{i=1}^{n} \|\mathcal{X}_i \;-\; \mathcal{Y}_i\|_{\mathrm{TV}}.$$

Proof of Lemma 1: The coupling lemma says that for any coupling of $\mathcal{X}_1, \ldots, \mathcal{X}_n, \mathcal{Y}_1, \ldots, \mathcal{Y}_n$:

$$\begin{aligned} \left\| \sum_{i=1}^{n} \mathcal{X}_i \;-\; \sum_{i=1}^{n} \mathcal{Y}_i \right\|_{\mathrm{TV}} &\le \Pr\left[\sum_{i=1}^{n} \mathcal{X}_i \neq \sum_{i=1}^{n} \mathcal{Y}_i \right] \\ &\le \sum_{i=1}^{n} \Pr[\mathcal{X}_i \neq \mathcal{Y}_i]. \end{aligned} \tag{1}$$

We proceed to fix a specific coupling. For all i, the optimal coupling theorem says that there exists a coupling of $\mathcal{X}_i$ and $\mathcal{Y}_i$ such that $\Pr[\mathcal{X}_i \neq \mathcal{Y}_i] = \|\mathcal{X}_i \;-\; \mathcal{Y}_i\|_{\mathrm{TV}}$. Using these individual couplings for each i we define a grand coupling of $\mathcal{X}_1, \ldots, \mathcal{X}_n, \mathcal{Y}_1, \ldots, \mathcal{Y}_n$ such that $\Pr[\mathcal{X}_i \neq \mathcal{Y}_i] = \|\mathcal{X}_i \;-\; \mathcal{Y}_i\|_{\mathrm{TV}}$, for all i. This coupling is faithful because $\mathcal{X}_1, \ldots, \mathcal{X}_n$ are mutually independent and $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$ are also mutually independent. Under this coupling Eq (1) implies:

$$\left\| \sum_{i=1}^{n} \mathcal{X}_i \;-\; \sum_{i=1}^{n} \mathcal{Y}_i \right\|_{\mathrm{TV}} \le \sum_{i=1}^{n} \Pr[\mathcal{X}_i \neq \mathcal{Y}_i] \equiv \sum_{i=1}^{n} \|\mathcal{X}_i - \mathcal{Y}_i\|_{\mathrm{TV}}. \tag{2}$$

□

Discussion. *Range of Utility Functions:* Throughout this paper we restrict our attention to normalized anonymous games, i.e. anonymous games whose payoffs lie in $[0, 1]$. We noted above that this restriction is without loss of generality and is made so that the approximation size is meaningful. All our algorithms can be readily generalized to anonymous games whose payoffs lie in some general $[u_{\min}, u_{\max}]$ by applying an affine transformation to the utility functions of the latter to bring their range into $[0, 1]$. Using the simple transformation that subtracts $u_{\min}$ and divides every payoff by $\frac{1}{u_{\max} - u_{\min}}$ and we obtain the following:

Proposition 1. *Any game $\mathcal{G}$ whose payoffs lie in $[u_{\min}, u_{\max}]$ can be affinely transformed in polynomial time into a game $\mathcal{G}'$ with the same players whose payoffs lie in $[0, 1]$ and such that any ϵ-Nash equilibrium of $\mathcal{G}'$ is a $\epsilon \cdot (u_{\max} - u_{\min})$-Nash equilibrium of $\mathcal{G}$.*

Types: Anonymous games can be generalized to *typed anonymous games* in the obvious way. In a typed anonymous game, every player $i \in [n]$ has a type t_i from a set of types $\mathcal{T}$, and each player's payoff depends on (a) his/her own strategy; and (b) the number of other players of each type playing each strategy; but it is not assumed that the players of the same type have the same payoff function. Clearly anonymous games are a special case of typed anonymous games where the set of types contains a single element. All algorithms presented in this paper can be extended to typed anonymous games.

Order Notation: Let $f(x)$ and $g(x)$ be two positive functions defined on some infinite subset of $\mathbb{R}_+$. One writes $f(x) = O(g(x))$ if and only if, for sufficiently large values of x, $f(x)$ is at most a constant times $g(x)$. That is, $f(x) = O(g(x))$ if and only if there exist positive real numbers M and x_0 such that

$$f(x) \leq \ Mg(x), \text{ for all } x > x_0.$$

Similarly, we write $f(x) = \Omega(g(x))$ if and only if there exist positive reals M and x_0 such that

$$f(x) \geq \ Mg(x), \text{ for all } x > x_0.$$

Finally, we write $f(x) = \Theta(g(x))$ if and only if $f(x) = O(g(x))$ and $g(x) = O(f(x))$.

We are casual in our use of order notation throughout the paper in that we do not specify which (among the possibly several) variables in our bounds are assumed to scale (and how they scale), as long as it is clear from context. Whenever we write $O(f(n))$ or $\Omega(f(n))$ in some bound where n ranges over the integers, we mean that there exists a constant $c > 0$ such that the bound holds true for sufficiently large n if we replace the $O(f(n))$ or $\Omega(f(n))$ in the bound by $c \cdot f(n)$. On the other hand, whenever we write $O(f(1/\epsilon))$ or $\Omega(f(1/\epsilon))$ in some bound where ϵ ranges over $(0, 1]$, we mean that there exists a constant $c > 0$ such that the bound holds true for sufficiently *small* ϵ if we replace the $O(f(1/\epsilon))$ or $\Omega(f(1/\epsilon))$ in the bound with $c \cdot f(1/\epsilon)$.

3 Approximation in Pure Strategies

We show that, if an anonymous game is Lipschitz, there always exists an approximate pure Nash equilibrium whose approximation quality depends on the number of strategies and the Lipschitz constant of the game, as described by the following theorem. Our result recalls Milchtaich's existence proof of a pure Nash equilibrium in a special case of anonymous games called congestion games with player-specific payoff functions [Mil96], even though our proof is different and more involved.

Theorem 1. *Every λ-Lipschitz anonymous game with ξ strategies has an ϵ-approximate pure Nash equilibrium, where $\epsilon = O(\lambda\xi)$. In particular, the number of players does not affect the approximation guarantee.*

Proof of Theorem 1: We outline the proof first. There are three steps. In the first, we ignore the nth player and construct a "best response" map ϕ from Π_{n-1}^{ξ} to itself. ϕ maps a partition x into a partition y that arises if the first $n-1$ players best respond to x. By interpolating ϕ we obtain an "interpolated best response map" $\hat{\phi}$, mapping Δ_{n-1}^{ξ} to itself. This map is continuous, and thus has a fixed point. If we were dealing with a continuum of players, we would already be done, as we would have a partition of players into strategies that is a best response to itself. But we have atomic players, and the fixed point will in general not have integral coordinates. Moreover, we have excluded player n from the definition of $\hat{\phi}$. In the second step, we use the Shapley-Folkman lemma [Sta69] to "round" the fixed point of $\hat{\phi}$ into a pure strategy profile that is an approximate best response to itself, utilizing the Lipschitz condition. Finally, we show that the last player can be incorporated in this picture without affecting the utilities much further, again by Lipschitzness.

We proceed with the details of our argument. We start by defining a function $\phi : \Pi_{n-1}^{\xi} \to \Pi_{n-1}^{\xi}$ as follows: For any $x \in \Pi_{n-1}^{\xi}$, $\phi(x)$ is defined to be $(y_1, \ldots, y_\xi) \in \Pi_{n-1}^{\xi}$ such that, for all $i \in [\xi]$, y_i is the number of players $p \in \{1, \ldots, n-1\}$ (notice that player n is excluded) such that, for all $j < i$, $u_i^p[x] > u_j^p[x]$, and, for all $j > i$, $u_i^p[x] \geq u_j^p[x]$. In other words, $\phi(x)$ is the partition of $n-1$ into $[\xi]$ induced if players $[n-1]$ best respond to x, where ties are broken lexicographically.

We next interpolate ϕ to obtain a continuous function $\hat{\phi} : \Delta_{n-1}^{\xi} \to \Delta_{n-1}^{\xi}$. To do so, we first choose a simplicization of Δ_{n-1}^{ξ} whose vertices are the points of Π_{n-1}^{ξ}. There are several ways to obtain such a simplicization—we use the method described in [LT82] so that all simplices in the simplicization have L1 diameter $O(\xi)$. Next, we define the interpolated function $\hat{\phi}$ in terms of the simplicization as follows. For all $x \in \Delta^{\xi}_{n-1}$, find the simplex of the simplicization where x belongs; if x belongs to multiple simplices pick an arbitrary one. Let $x^1, \ldots, x^{\xi}$ be the vertices of that simplex, and suppose that $x = \sum_{j \in [\xi]} \alpha_j x^j$, for some $\alpha_1, \ldots, \alpha_{\xi} \geq 0$ such that $\sum_j \alpha_j = 1$. We define $\hat{\phi}$ at x to be

$$\hat{\phi}(x) \equiv \sum_{j \in [\xi]} \alpha_j \phi(x^j).$$

Notice that $\hat{\phi}$ (as defined above) is continuous. Since Δ^{ξ}_{n-1} is convex and compact, Brouwer's fixed point theorem implies that $\hat{\phi}$ has a fixed point $x^* = \hat{\phi}(x^*)$. Suppose that x^* belongs to a simplex of our simplicization with vertices $x^1, \ldots, x^{\xi}$. Then it must be that

$$x^* = \sum_{j=1}^{\xi} \gamma_j x^j,$$

for some $\gamma_1, \ldots, \gamma_{\xi} \geq 0$ and $\sum_j \gamma_j = 1$, and because x^* is fixed it must be that:

$$x^* = \sum_{j=1}^{\xi} \gamma_j \phi(x^j). \tag{3}$$

The term $\phi(x^j)$ in Equation (3) is induced by some pure strategy profile where each of the first $n-1$ players chooses a strategy that is her best response to x^j (with the aforementioned tie-breaking rule). In this strategy profile all first $n-1$ players would be happy if the aggregate behavior of the other players were x^j. A useful observation is this: *Suppose that a player is playing her best response to x^j but is facing some other $x \in \Pi^{\xi}_{n-1}$ instead. Then, because the game is λ-Lipschitz the regret experienced by the player for not playing her best response to x is at most* $2\lambda ||x^j - x||_1$. Inspired by this observation, we show that there exists some $x \in \Pi^{\xi}_{n-1}$ such that: (i) $||x - x^*||_1 = O(\xi)$; and (ii) x is induced by a pure strategy profile where each of the first $n-1$ players chooses a strategy that is her best response to one of $x^1, \ldots, x^{\xi}$. This is almost what we need to get the desired approximate pure Nash equilibrium, as we will see next.

But let us first establish the existence of some x satisfying (i) and (ii). For all $i \in [n-1]$, let

$$T_i = \{e_\ell \mid \text{strategy } \ell \text{ is a best response of player } i \text{ to } x^j \text{ for some } j \in \{1 \ldots, \xi\}\} \subseteq \Pi^{\xi}_1,$$

where e_ℓ is the unit vector along dimension ℓ of $\mathbb{R}^{\xi}$. I.e. T_i is the set of pure best responses of player i to at least one of $x^1, \ldots, x^{\xi}$. We establish the following.

Lemma 2. *There exists some $x \in \Pi^{\xi}_{n-1}$ such that: (i) $||x - x^*||_1 \leq 2\xi$; and (ii) $x = \sum_{i \in [n-1]} v_i$, where $v_i \in T_i$ for all $i \in [n-1]$.*

Proof of Lemma 2: Equation (3) implies that x^* is in the convex hull of the Minkowski sum $T_1 + T_2 + \cdots + T_{n-1}$. (Indeed, for all j, $\phi(x^j)$ belongs to the Minkowski sum.) It follows from the Shapley-Folkman lemma that there exists some $\mathcal{P} \subset [n-1]$ with $|\mathcal{P}| = \xi$ such that

$$x^* = z + w, \quad \text{where } z \in \sum_{i \in \mathcal{P}} \text{Conv}(T_i) \text{ and } w \in \sum_{i \in [n-1] \setminus \mathcal{P}} T_i,$$

where '$\sum$' represents Minkowski addition, and $\text{Conv}(\cdot)$ is taking the convex hull of its argument.

Let then $x = z' + w$, where z' is an arbitrary point in $\sum_{i \in \mathcal{P}} T_i$. We notice that $||z - z'||_1 \leq 2\xi$. Hence, $||x - x^*||_1 \leq 2\xi$. Moreover, by construction $x \in \sum_i T_i$, concluding the proof of the lemma.$\Box$

Let $v_1, \ldots, v_{n-1}$ and $x = \sum_{i \in [n-1]} v_i$ be as in Lemma 2. Also, let $v_n \in \Pi_1^\xi$ be player n's best response to x. We argue that the pure strategy profile $\tilde{P}$ defined by the v_i's is a $O(\lambda\xi)$-approximate pure Nash equilibrium. Indeed, notice first that $||x - x^j||_1 = O(\xi)$ for all j, since $||x - x^*||_1 \leq 2\xi$ and x^* is in the simplex of our simplicization with vertices $x^1, \ldots, x^\xi$ (and the simplices in our simplicization have L1 diameter $O(\xi)$). Now, fix some player $i \in [n-1]$. Player i plays a best response to one of $x^1, x^2, \ldots, x^\xi$, but is facing instead $x - v_i + v_n$. Given that $x - v_i + v_n$ is within $O(\xi)$ from all of $x^1, x^2, \ldots, x^\xi$ and our earlier observation, it follows that v_i is an $O(\lambda\xi)$-approximate best response to $x - v_i + v_n$. Finally, player n is playing a best response to x and he is facing x so he is happy. This concludes the proof.$\Box$

Remark 1. *Theorem 1 appeared first in [DP07] with a weaker approximation guarantee of $O(\lambda\xi^2)$. The approach was essentially the same as the proof given above, except that the analysis was wasteful by an extra factor of ξ. Indeed it was conjectured in [DP07] that an approximation of $O(\lambda\xi)$ should be possible. This was recently proved by Azrieli and Shmaya [AS13] with a different construction that invoked the Shapley-Folkman lemma. Inspired by their use of that lemma, here we notice that the analysis in [DP07] can be tightened to recover the approximation guarantee of $O(\lambda\xi)$. The two proofs are otherwise quite different.*

The immediate implication of Theorem 1 is the following algorithmic result.

Corollary 1. *In a λ-Lipschitz anonymous game with ξ strategies and n players an ϵ-approximate pure Nash equilibrium, where $\epsilon = O(\lambda\xi)$ as in Theorem 1, can be found in total number of bit operations of $U\xi n^\xi \times (n + \xi)\text{poly}(\log n)$, where U is the number of bits required to represent a payoff value of the game, i.e. in total number of bit operations that is $(n + \xi)\text{poly}(\log n)$ times the description complexity of the game.*

Proof of Corollary 1: Theorem 1 guarantees that some pure strategy profile S is an ϵ-approximate pure Nash equilibrium. However we cannot afford to exhaustively search over pure strategy profiles as there are too many of them. We search instead over partitions in Π_n^ξ. For each partition $x = (x_1, \ldots, x_\xi) \in \Pi_n^\xi$ we need to answer the following algorithmic question: Is there an ϵ-approximate pure Nash equilibrium S such that the number of players playing strategy i in S is exactly x_i for all i? We answer this question by computing a maximum flow on a related flow network. The nodes of the network are the elements of the set $\{s\} \cup \{t\} \cup [n] \cup [\xi]$, where s is the source node of the network, t is the sink node of the network, there is a directed edge from s to every node in $[n]$ of capacity 1, an edge from every node $i \in [\xi]$ to t of capacity x_i, and an edge of capacity 1 from $p \in [n]$ to $i \in [\xi]$ iff playing i is an ϵ-pure best response for player p against $x - e_i$, where e_i is the unit vector along dimension i of $\mathbb{R}^\xi$. If the maximum flow in this network is n then the answer to the above question is 'yes', and any integral maximum flow gives an assignment of players p to strategies i that comprises an ϵ-approximate pure Nash equilibrium.

The total number of bit operations required is $n^{\xi-1} \times Un\xi(n + \xi)\text{poly}(\log n)$, as we need to perform at most $|\Pi_n^\xi| \leq n^{\xi-1}$ max-flow computations on graphs of $n + \xi + 2$ vertices, $O(n\xi)$ edges, and capacities of $\lceil \log n \rceil$ bits. We use Orlin's algorithm [Orl13] for max-flow computations.$\Box$

4 Approximation in Mixed Strategies

The approximation in pure strategies achieved in Section 3 is interesting only if the game is λ-Lipschitz for a reasonably small constant λ. In games with significant discontinuities in the utilities, one needs to resort to mixed strategies. Our main algorithmic result is a polynomial-time approximation scheme (PTAS) for anonymous games with a large number n of players and a constant number ξ of strategies. More precisely, we obtain a family of algorithms $(\mathcal{A}_\epsilon)_\epsilon$, indexed by the approximation parameter ϵ, such that for all $\epsilon > 0$ Algorithm $\mathcal{A}_\epsilon$ computes an ϵ-Nash equilibrium of a given anonymous game of n players and ξ strategies in time $n^{g(\xi,1/\epsilon)}$, where g is some function of ξ and ϵ, which does not depend on n. The algorithm is called a polynomial-time approximation scheme, since whenever ξ and ϵ are absolute constants the running time of the algorithm is polynomial in the description of the game.

Theorem 2. *There is a PTAS for the mixed Nash equilibrium problem for normalized anonymous games with a constant number of strategies. More precisely, there exists some function g such that, for all $\epsilon \geq 0$, an ϵ-Nash equilibrium of a normalized anonymous game of n players and ξ strategies can be computed in time $n^{g(\xi,1/\epsilon)} \cdot U$, where U is the number of bits required to represent a payoff value of the game.*

We provide the proof of Theorem 2 in Section 4.2, after presenting the main technical lemma needed for the proof in Section 4.1. Before proceeding let us give some intuition. The basic idea of our algorithm is extremely simple and intuitive: Instead of performing the search for an approximate mixed Nash equilibrium over the full set of mixed strategy profiles, we restrict our attention to mixed strategies assigning to each strategy in their support probability mass which is an integer multiple of $\frac{1}{z}$, where z is a large enough natural number. We call this process *discretization*. Searching the space of discretized mixed strategy profiles can be done efficiently. Indeed, there are less than $(z+1)^{\xi-1}$ discretized mixed strategies available to each player, so at most $n^{(z+1)^{\xi-1}-1}$ partitions of the number n of players into these discretized mixed strategies. And checking if there is an approximate Nash equilibrium consistent with such a partition can be done efficiently using a max-flow argument similar to the one used in the proof of Corollary 1 (see full details in the proof of Theorem 2 in Section 4.2).

The challenge, however, lies somewhere else: We need to establish that restricting our search to some moderate z (independent of n) suffices. We do this in two steps. The first is to relate the approximation performance of two mixed strategy profiles to some measure of their probabilistic distance. This is achieved by the following lemma.

Lemma 3. *Let $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ and $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ be two mixed strategy profiles of a normalized anonymous game of n players and ξ strategies. Then for all $i \in [n]$ and $\ell \in [\xi]$:*

$$\left| \mathbb{E}u_\ell^i \left(\sum_{j \neq i} \mathcal{X}_j \right) - \mathbb{E}u_\ell^i \left(\sum_{j \neq i} \mathcal{Y}_j \right) \right| \leq 2 \left\| \sum_{j \neq i} \mathcal{X}_j - \sum_{j \neq i} \mathcal{Y}_j \right\|_{\mathrm{TV}}. \tag{4}$$

Proof of Lemma 3: Observe that

$$\mathbb{E}u_\ell^i \left(\sum_{j \neq i} \mathcal{X}_j \right) = \sum_{x \in \Pi_{n-1}^\xi} u_\ell^i(x) \Pr \left[\sum_{j \neq i} \mathcal{X}_j = x \right],$$

and a similar equality holds for $\mathbb{E}u_\ell^i\left(\sum_{j\neq i}\mathcal{Y}_j\right)$. Hence, the LHS of (4) equals

$$\left|\sum_{x\in\Pi_{n-1}^{\xi}} u_\ell^i(x)\left(\Pr\left[\sum_{j\neq i}\mathcal{X}_j = x\right] - \Pr\left[\sum_{j\neq i}\mathcal{Y}_j = x\right]\right)\right|$$
$$\leq \sum_{x\in\Pi_{n-1}^{\xi}} |u_\ell^i(x)| \cdot \left|\Pr\left[\sum_{j\neq i}\mathcal{X}_j = x\right] - \Pr\left[\sum_{j\neq i}\mathcal{Y}_j = x\right]\right|$$
$$\leq 2\left|\left|\sum_{j\neq i}\mathcal{X}_j - \sum_{j\neq i}\mathcal{Y}_j\right|\right|_{\mathrm{TV}},$$

where in the last line we used that $|u_\ell^i(x)| \leq 1$ and the definition of total variation distance. □

The immediate corollary of Lemma 3 is the following.

Lemma 4. *Suppose $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ is a Nash equilibrium of a normalized anonymous game of n players and ξ strategies and $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ a mixed strategy profile satisfying:*

1. *the support of $\mathcal{Y}_i$ is a subset of the support of $\mathcal{X}_i$, for all i; and*
2. *for some $\epsilon \geq 0$, $\left|\left|\sum_{j\neq i}\mathcal{X}_j - \sum_{j\neq i}\mathcal{Y}_j\right|\right|_{\mathrm{TV}} \leq \epsilon$, for all i.*

Then $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ is a 4ϵ-Nash equilibrium.

Proof of Lemma 4: Fix some player i. We want to show that every strategy ℓ in the support of $\mathcal{Y}_i$ is a 4ϵ-approximate best response to $(\mathcal{Y}_j)_{j\neq i}$. Since the support of $\mathcal{Y}_i$ is a subset of the support of $\mathcal{X}_i$ and $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ is a Nash equilibrium we know that ℓ is a best response to $(\mathcal{X}_j)_{j\neq i}$. Hence, for all ℓ', we have:

$$\mathbb{E}u_\ell^i\left(\sum_{j\neq i}\mathcal{X}_j\right) \geq \mathbb{E}u_{\ell'}^i\left(\sum_{j\neq i}\mathcal{X}_j\right).$$

But Lemma 3 implies

$$\mathbb{E}u_\ell^i\left(\sum_{j\neq i}\mathcal{Y}_j\right) \geq \mathbb{E}u_\ell^i\left(\sum_{j\neq i}\mathcal{X}_j\right) - 2\epsilon; \quad \text{and}$$
$$\mathbb{E}u_{\ell'}^i\left(\sum_{j\neq i}\mathcal{X}_j\right) \geq \mathbb{E}u_{\ell'}^i\left(\sum_{j\neq i}\mathcal{Y}_j\right) - 2\epsilon.$$

Putting the above together, we get the expected utility that i gets for playing ℓ against $(\mathcal{Y}_j)_{j\neq i}$ is within an additive 4ϵ from the expected utility obtained by playing ℓ'. Since ℓ' was arbitrary, this concludes the proof. □

Given Lemmas 3 and 4, the route to arguing that a moderate discretization size z, independent of n, suffices for our search for ϵ-Nash equilibria, is establishing this: Let e_ℓ be the unit vector along dimension ℓ of the ξ-dimensional Euclidean space. Then the distribution of the sum of n independent random unit vectors with values ranging over $\{e_1, \ldots, e_\xi\}$ can be approximated by the distribution of the sum of another set of independent unit vectors whose probabilities of obtaining each value are multiples of $\frac{1}{z}$, and so that the total variation distance of the two distributions depends only on z (in fact, a decreasing function of z) and the dimension ξ, but not on the number n of vectors. The intention is for the original random vectors to correspond to the strategies of the players in an (unknown) Nash equilibrium, and the discretized ones to a discretized mixed strategy profile. In the next section we present a probabilistic approximation theorem achieving the afore-mentioned approximation, and show how it can be used to obtain a PTAS for anonymous games in Section 4.2.

4.1 An Approximation Theorem for Sums of Multinomial Distributions

We present the main probabilistic tool required for our PTAS. Its proof can be found in [DP08].

Theorem 3 ([DP08]). *Let $\{p_i \in \Delta^\xi\}_{i\in[n]}$, and let $\{\mathcal{X}_i \in \mathbb{R}^\xi\}_{i\in[n]}$ be a set of independent ξ-dimensional random unit vectors such that, for all $i \in [n]$, $\ell \in [\xi]$, $\Pr[\mathcal{X}_i = e_\ell] = p_{i,\ell}$, where e_ℓ is the unit vector along dimension ℓ; also, let $z > 0$ be an integer. Then there exists another set of probability vectors $\{\widehat{p}_i \in \Delta^\xi\}_{i\in[n]}$ such that*

1. $|\widehat{p}_{i,\ell} - p_{i,\ell}| = O\left(\frac{1}{z}\right)$, *for all* $i \in [n], \ell \in [\xi]$;
2. $\widehat{p}_{i,\ell}$ *is an integer multiple of* $\frac{1}{2^\xi}\frac{1}{z}$, *for all* $i \in [n], \ell \in [\xi]$;
3. *if* $p_{i,\ell} = 0$, *then* $\widehat{p}_{i,\ell} = 0$, *for all* $i \in [n], \ell \in [\xi]$;
4. *if* $\{\widehat{\mathcal{X}}_i \in \mathbb{R}^\xi\}_{i\in[n]}$ *is a set of independent random unit vectors such that* $\Pr[\widehat{\mathcal{X}}_i = e_\ell] = \widehat{p}_{i,\ell}$, *for all* $i \in [n]$, $\ell \in [\xi]$, *then*

$$\left|\left|\sum_i \mathcal{X}_i - \sum_i \widehat{\mathcal{X}}_i\right|\right|_{\mathrm{TV}} = O\left(f(\xi)\frac{\log z}{z^{1/5}}\right) \tag{5}$$

and, moreover, for all $j \in [n]$,

$$\left|\left|\sum_{i\neq j} \mathcal{X}_i - \sum_{i\neq j} \widehat{\mathcal{X}}_i\right|\right|_{\mathrm{TV}} = O\left(f(\xi)\frac{\log z}{z^{1/5}}\right), \tag{6}$$

where $f(\xi)$ *some function of* ξ.

The theorem states that, for all finite $\epsilon \in [0, 1]$, there is a way to quantize any set of n independent random vectors into another set of n independent random vectors, whose probabilities of obtaining each value are integer multiples of ϵ, so that the total variation distance between the distribution of the sum of the vectors before and after the quantization is bounded by $O(f(\xi)2^{\xi/6}\epsilon^{1/6})$. The crucial, and perhaps surprising, property of this bound is the lack of dependence on the number n of random vectors. In particular, notice that a naive rounding of the probabilities into integer multiples of ϵ would result in total variation distance of $\Theta(n \cdot \xi \cdot \epsilon)$, which fails to give a polynomial time approximation scheme (as will become clear in the proof of Theorem 2 in the next section). It is crucial to obtain a bound that is independent of n such as the one guaranteed by Theorem 3.

4.2 Proof of Theorem 2

The proof follows from the technical tools presented earlier using flow arguments.

Proof of Theorem 2: Take an arbitrary mixed Nash equilibrium $(p_1, \ldots, p_n)$ of an n-player ξ-strategy anonymous game, where $p_i \in \Delta^\xi$ for all i. Theorem 3 guarantees the existence of a mixed strategy profile $(\widehat{p}_1, \ldots, \widehat{p}_n)$ satisfying Properties 1 through 4 of the theorem. Then Lemma 4 guarantees that this mixed strategy profile is an $O\left(f(\xi) z^{-1/6}\right)$-Nash equilibrium of the game. And, if $z = (f(\xi)/\epsilon)^6$, it is a δ-Nash equilibrium, for $\delta = O(\epsilon)$.

The above discussion shows that there is a discretized mixed strategy profile $\{\widehat{p}_i\}_i$ that satisfies Property 2 of Theorem 3 and is a δ-Nash equilibrium for z chosen as above. The algorithmic challenge is, however, that there is an exponential number of such discretized mixed strategy profiles so we cannot afford to do exhaustive search over them. We do instead the following.

Notice first that there is at most $K := \xi + (2^\xi z)^{\xi-1} \leq 2^{\xi^2} (f(\xi)/\epsilon)^{6\xi}$ "quantized" mixed strategies with each probability being an integer multiple of $\frac{1}{2^\xi}\frac{1}{z}$. Let $\mathcal{K}$ be the set of such quantized mixed strategies. We start our algorithm by guessing the partition of the number n of players into quantized mixed strategies; let $\theta = \{\theta_\sigma\}_{\sigma \in \mathcal{K}}$ be the guessed partition, where θ_σ represents the number of players choosing the discretized mixed strategy $\sigma \in \mathcal{K}$. Now we need to determine if there exists an assignment of quantized mixed strategies to the players in $[n]$, with θ_σ of them being assigned $\sigma \in \mathcal{K}$, so that the resulting mixed strategy profile is a δ-Nash equilibrium. To answer this question it is enough to solve the following *max-flow* problem. Let us consider the bipartite graph $([n], \mathcal{K}, E)$ with edge set E defined as follows: $(i, \sigma) \in E$, for $i \in [n]$ and $\sigma \in \mathcal{K}$, if $\theta_\sigma > 0$ and σ is a δ-best response for player i, if the partition of the other players into the mixed strategies in $\mathcal{K}$ is the partition θ, with one unit subtracted from θ_σ.[1] Note that to define E expected payoff computations are required. By straightforward dynamic programming (see Section 2), the expected utility of player i for playing pure strategy $s \in [\xi]$ given the mixed strategies of the other players can be computed with $O(\xi n^{\xi+1})$ operations on numbers with at most $b(n, z, \xi) := O(n(\xi + \log_2 z) + U)$ bits, where U is the number of bits required to specify the values in the range of the payoff functions.[2] To conclude the construction of the max-flow instance we add a source node u connected to all the left hand side nodes and a sink node v connected to all the right hand side nodes. We set the capacity of the edge (σ, v) equal to θ_σ, for all $\sigma \in \mathcal{K}$, and the capacity of all other edges equal to 1. If the max-flow from u to v has value n then there is a way to assign discretized mixed strategies to the players so that θ_σ of them play mixed strategy $\sigma \in \mathcal{K}$ and the resulting mixed strategy profile is a δ-Nash equilibrium. There are at most $(n+1)^{K-1}$ possible guesses for θ; hence, the search takes overall time

$$O\left((nK\xi^2 n^{\xi+1} b(n, z, \xi) + p(n + K + 2)) \cdot (n+1)^{K-1} \right),$$

where $p(n + K + 2)$ is the time needed to find an integral maximum flow in a graph with $n + K + 2$ nodes and edge capacities encoded with at most $\lceil \log_2 n \rceil$ bits. Hence, the overall running time is

$$n^{O\left(2^{\xi^2} \left(\frac{f(\xi)}{\epsilon}\right)^{6\xi}\right)} \cdot U.$$

[1] For our discussion, a mixed strategy σ of player i is a δ*-best response* to a set of mixed strategies for the other players iff the expected payoff of player i for playing any pure strategy in the support of σ is no more than δ worse than her expected payoff for playing any other pure strategy.

[2] To compute a bound on the number of bits required for the expected utility computations, note that every non-zero probability value that is computed along the execution of the algorithm must be an integer multiple of $(\frac{1}{2^\xi}\frac{1}{z})^j$ for some $j \leq n-1$, since the mixed strategies of all players are from the set $\mathcal{K}$. Further note that the expected utility is a weighted sum of at most n^ξ payoff values, with U bits required to represent each value, and all weights being probabilities.

□

4.3 Extension to Typed Anonymous Games

We can extend Theorem 2 to typed anonymous games as follows.

Theorem 4. *There is a PTAS for the mixed Nash equilibrium problem for normalized typed anonymous games with a constant number of strategies and types. More precisely, there exists some function g such that, for all $\epsilon \geq 0$, an ϵ-Nash equilibrium of a normalized anonymous game of n players, ξ strategies and t types can be computed in time $n^{g(\xi,1/(\epsilon),t)} \cdot U$, where U is the number of bits required to represent a payoff value of the game.*

Proof of Theorem 4: The proof is similar to that of Theorem 2 except that we need to treat every type separately, exploiting the following straightforward generalizations of Lemmas 3 and 4.

Lemma 5. *Let $\mathcal{X} = (\mathcal{X}_1, \ldots, \mathcal{X}_n)$ and $\mathcal{Y} = (\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ be two mixed strategy profiles of a normalized anonymous game of n players, ξ strategies and type-set $\mathcal{T}$. Then for all $i \in [n]$ and $\ell \in [\xi]$:*

$$\left|\mathbb{E}u_\ell^i(\mathcal{X}) - \mathbb{E}u_\ell^i(\mathcal{Y})\right| \leq \sum_{t \in \mathcal{T}} 2 \left|\left| \sum_{j \neq i,\ t_j = t} \mathcal{X}_j - \sum_{j \neq i,,\ t_j = t} \mathcal{Y}_j \right|\right|_{\mathrm{TV}}. \tag{7}$$

Lemma 6. *Suppose $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ is a Nash equilibrium of a normalized anonymous game of n players, ξ strategies and type-set $\mathcal{T}$, and $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ a mixed strategy profile satisfying:*

1. *the support of $\mathcal{Y}_i$ is a subset of the support of $\mathcal{X}_i$, for all i; and*
2. *for some $\epsilon \geq 0$, for all $t \in \mathcal{T}$: $\left|\left|\sum_{j \neq i,\ t_j = t} \mathcal{X}_j - \sum_{j \neq i,\ t_j = t} \mathcal{Y}_j\right|\right|_{\mathrm{TV}} \leq \epsilon$, for all i.*

Then $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ is a $4\epsilon|\mathcal{T}|$-Nash equilibrium.

Now take an arbitrary mixed Nash equilibrium $(p_1, \ldots, p_n)$ of a normalized n-player ξ-strategy anonymous game with type-set $\mathcal{T}$, where $p_i \in \Delta^\xi$ for all i. Theorem 3 (applied to the mixed strategy profile restricted to each type separately) guarantees the existence of a mixed strategy profile $(\widehat{p}_1, \ldots, \widehat{p}_n)$ such that:

1. $|\widehat{p}_{i,\ell} - p_{i,\ell}| = O\left(\frac{1}{z}\right)$, for all $i \in [n], \ell \in [\xi]$;
2. $\widehat{p}_{i,\ell}$ is an integer multiple of $\frac{1}{2^\xi}\frac{1}{z}$, for all $i \in [n], \ell \in [\xi]$;
3. if $p_{i,\ell} = 0$, then $\widehat{p}_{i,\ell} = 0$, for all $i \in [n], \ell \in [\xi]$;
4. if $\{\widehat{\mathcal{X}}_i \in \mathbb{R}^\xi\}_{i \in [n]}$ is a set of independent random unit vectors such that $\Pr[\widehat{\mathcal{X}}_i = e_\ell] = \widehat{p}_{i,\ell}$, for all $i \in [n]$, $\ell \in [\xi]$, then for all $t \in \mathcal{T}$:

$$\left|\left| \sum_{i, t_i = t} \mathcal{X}_i - \sum_{i, t_i = t} \widehat{\mathcal{X}}_i \right|\right|_{\mathrm{TV}} = O\left(f(\xi)\frac{\log z}{z^{1/5}}\right) \tag{8}$$

and, moreover, for all $j \in [n]$,

$$\left|\left| \sum_{i \neq j, t_i = t} \mathcal{X}_i - \sum_{i \neq j, t_i = t} \widehat{\mathcal{X}}_i \right|\right|_{\mathrm{TV}} = O\left(f(\xi)\frac{\log z}{z^{1/5}}\right). \tag{9}$$

Through Lemma 6 the above conditions guarantee that this mixed strategy profile is a $O\left(|\mathcal{T}|f(\xi)z^{-1/6}\right)$-Nash equilibrium of the game. And, if we choose $z = (f(\xi)|\mathcal{T}|/\epsilon)^6$, it is a δ-Nash equilibrium, for $\delta = O(\epsilon)$.

The above discussion shows that there is a mixed strategy profile $\{\widehat{p_i}\}_i$ that is discretized as above and is a δ-Nash equilibrium of the game. To find such a mixed strategy profile we proceed in a similar fashion as in the proof of Theorem 2. Briefly:

- There is at most $K := \xi + (2^\xi z)^{\xi-1} \leq 2^{\xi^2}\left(f(\xi)|\mathcal{T}|/\epsilon\right)^{6\xi}$ "quantized" mixed strategies with each probability being an integer multiple of $\frac{1}{2^\xi}\frac{1}{z}$, for the afore-given choice of z; call $\mathcal{K}$ their set.

- We guess how many players of each type use each of these discretized mixed strategies; for all $t \in \mathcal{T}$ and $\sigma \in \mathcal{K}$, θ^t_σ represents the number of players of type t using the discretized mixed strategy σ.

- To decide if these guesses are consistent with a δ-Nash equilibrium, we set up a maximum flow problem for each type. The flow network for type $t \in \mathcal{T}$ is set up as follows:

  – Consider a bipartite graph whose left side has one node for every player of type t and right side has one node for every element of $\mathcal{K}$.

  – There is an edge from player i of type t to discretized mixed strategy $\sigma \in \mathcal{K}$ iff σ is a δ-best response for player i when the players of type $t' \neq t$ play as specified by $\theta^{t'}$, for all $t' \neq t$, and the players of type t that are different than i play according to θ^t, with one unit subtracted from θ^t_σ.

  – Add a source node u connected to all the left hand side nodes and a sink node v connected to all the right hand side nodes.

  – Set the capacity of edge (σ, v) equal to θ^t_σ, for all $\sigma \in \mathcal{K}$, and the capacity of all other edges equal to 1.

- If, for all $t \in \mathcal{T}$, the maximum flow in the flow network for type t has value $|i \mid t_i = t|$, then there is a way to assign discretized mixed strategies to the players so that θ^t_σ players of type t play mixed strategy $\sigma \in \mathcal{K}$ and the resulting mixed strategy profile is a δ-Nash equilibrium.

The running time analysis proceeds in the same way as in the proof of Theorem 2. □

5 The Structure of Approximate Equilibria

The PTAS of Theorem 2 exploits the probabilistic approximation theorem for sums of multinomial distributions presented in Theorem 3. Indeed the probabilistic approximation results in a small enough discretized set that we can search over in time polynomial in the number of players (albeit exponential in the approximation parameter ϵ). Is it possible to improve the running time by obtaining a stronger probabilistic approximation? In this section we show that the answer is "yes" for 2-strategy anonymous games. We obtain a significantly faster algorithm by exploiting a stronger probabilistic approximation. As a corollary we also get an interesting structural theorem for approximate equilibria in these games, which is reminiscent of Nash's theorem on the existence of symmetric equilibria in symmetric games [Nas51].

5.1 A Stronger Approximation, and a Faster Algorithm

We present a probabilistic approximation theorem for sums of independent indicators that is tighter than Theorem 3 (which nevertheless applies to the more general case of sums of independent categorical random variables). Roughly speaking the theorem guarantees that the sum of a collection of independent indicators can be in one of two modes: (i) a *dense mode* whereby the sum is ϵ-close to a shifted Binomial distribution; or (ii) a *sparse mode* whereby the sum is ϵ-close to the shifted sum of just $O(1/\epsilon^3)$ independent indicators, for all ϵ. The benefit is that both cases correspond to qualitatively simpler types of distributions. We exploit their simplicity to search over them quickly.

Theorem 5 ([Das08a, Das08b]). *Let $\{p_i\}_{i=1}^n$ be arbitrary probability values, $p_i \in [0,1]$ for $i = 1,\dots,n$; $\{X_i\}_{i=1}^n$ be independent indicator random variables such that X_i has expectation $\mathbb{E}[X_i] = p_i$; and k be a positive integer. Then there exists another set of probability values $\{q_i\}_{i=1}^n$, $q_i \in [0,1]$, $i = 1,\dots,n$, which satisfy the following properties:*

1. *if $\{Y_i\}_{i=1}^n$ are independent indicator random variables such that Y_i has expectation $\mathbb{E}[Y_i] = q_i$, then,*

$$\left|\left|\sum_i X_i \;-\; \sum_i Y_i\right|\right|_{\mathrm{TV}} = O(1/k), \tag{10}$$

$$\text{and, for all } j = 1,\dots,n, \left|\left|\sum_{i\neq j} X_i \;-\; \sum_{i\neq j} Y_i\right|\right|_{\mathrm{TV}} = O(1/k), \tag{11}$$

   where the constant in the $O(\cdot)$ notation does not depend on n or $p_1,\dots,p_n$.

2. *the set $\{q_i\}_{i=1}^n$ is such that:*

   (a) if $p_i = 0$ then $q_i = 0$, and if $p_i = 1$ then $q_i = 1$;

   (b) one of the following is true:

      i. either there exists some $S \subseteq [n]$ and some value q which is an integer multiple of $\frac{1}{kn}$, such that, for all $i \notin S$, $q_i \in \{0,1\}$, and, for all $i \in S$, $q_i = q$;

      ii. or, there exists some $S \subset [n]$, $|S| < k^3$ such that, for all $i \notin S$, $q_i \in \{0,1\}$, and, for all $i \in S$, q_i is an integer multiple of $\frac{1}{k^2}$.

As promised earlier, we exploit Theorem 5 to improve the running time of our approximation scheme. Compared to Theorem 2, the approximation scheme presented below has an important qualitative difference: in the running time the approximation quality ϵ does not affect the degree of the polynomial that depends on the input size. Such PTAS's are called in the algorithmic literature *efficient PTAS's*. The proof of Theorem 6 follows the same approach used for Theorem 2, namely searching for approximate Nash equilibria in the set of strategy profiles of the form 2(b)i or 2(b)ii of Theorem 5.

Theorem 6. *For all $\epsilon \geq 0$, an ϵ-Nash equilibrium of a normalized anonymous game of n players and 2 strategies can be computed in time $(1/\epsilon)^{O(1/\epsilon^2)} \cdot \mathrm{poly}(n) \cdot U$, where U is the number of bits required to represent a payoff value of the game.*

Proof of Theorem 6: Consider a mixed Nash equilibrium $(p_1,\dots,p_n)$, where p_i is the probability that player i plays strategy 2. It follows from Lemma 4 that a mixed strategy profile $(q_1,\dots,q_n)$ satisfying Properties 1 and 2 of Theorem 5 is a $O(1/k)$-Nash equilibrium. Hence there is an ϵ-Nash equilibrium $\{q_i\}_i$ satisfying Property 2b in the statement of Theorem 5 for $k = O(1/\epsilon)$. The problem is, of course, that we do not know such a mixed strategy profile and also do not know whether it is of the kind specified by Property 2(b)i or the kind specified by Property 2(b)ii. Moreover, we cannot afford to do exhaustive search over all mixed strategy profiles satisfying Property 2(b)i or 2(b)ii, since there is an exponential number of those. We do instead the following two searches, corresponding to each of the two cases; one of them is guaranteed to find an ϵ-Nash equilibrium.

Search corresponding to 2(b)i: We can first guess the cardinality m of the set S (at most $n+1$ choices), the value q ($kn+1$ choices), and the number m' of q_i's in $[n] \setminus S$ which are equal to 1 (at most $n+1$ choices). Then we only need to determine if there is a set of players $S \subseteq [n]$ and another set of players $S' \subseteq [n] \setminus S$ such that, if all players in S are assigned mixed strategy q, all players in S' mixed strategy 1 and all players in $[n] \setminus S \setminus S'$ mixed strategy 0, then the resulting mixed strategy profile is an ϵ-Nash equilibrium. To answer this question it is enough to solve a *max-flow* problem. Let us define the constants $\theta_0 = n - m - m'$, $\theta_q = m$ and $\theta_1 = m'$, and consider a flow network with node-set $\{s\} \cup \{t\} \cup [n] \cup \{0, q, 1\}$ and edge-set as follows. There is a directed edge of capacity 1 from node $i \in [n]$ to node $\sigma \in \{0, q, 1\}$ iff $\theta_\sigma > 0$ and mixed strategy σ is an ϵ-best response for player i, when the partition of the other players into the mixed strategies 0, q and 1 is the partition θ, with one unit subtracted from θ_σ.[3] Moreover, there is a directed edge of capacity 1 from s to all nodes in $[n]$ and a directed edge of capacity θ_σ from σ to t for all $\sigma \in \{0, q, 1\}$. If the max-flow from s to t in this network has value n then there is a way to select S and S' by looking at the edges used by an integral maximum flow. As far as the running time goes, notice that to define the edge-set of the flow network expected payoff computations are required. By straightforward dynamic programming (see Section 2), the expected utility of a player for playing some pure strategy given the mixed strategies of the other players can be computed with $O(n^3)$ operations on numbers of $b(n,k) = O(n \log_2(kn) + U)$ bits, where U is the number of bits required to specify a payoff value of the game.[4] Since there are $O(n^3 k)$ possible guesses for (m, q, m'), the search takes overall time

$$O\left((n^4 b(n,k) + p_1(n)) \cdot n^3 k\right), \tag{12}$$

where $p_1(n)$ is the time needed to solve a max-flow problem on a graph with $n+5$ nodes, $O(n)$ edges, and edge capacities with at most $\lceil \log_2 n \rceil$ bits.

Search corresponding to 2(b)ii: We can guess the cardinality m of the set S (there are k^3 choices), the number m' of q_i's in $[n] \setminus S$ which are equal to 1 (at most $n+1$ choices), and a partition of $m \equiv |S|$ into the integer multiples of $\frac{1}{k^2}$ in $[0,1]$; let $\{\phi_{i/k^2}\}_{i \in [k^2] \cup \{0\}}$ be the partition. Then we only need to determine if there is a set of players $S \subseteq [n]$ of cardinality m, a set of players $S' \subseteq [n] \setminus S$ of cardinality m', and an assignment of mixed strategies to the players in S with ϕ_{i/k^2} of them being assigned i/k^2 so that, if additionally the players in S' are assigned mixed strategy 1 and the players in $[n] \setminus S \setminus S'$ are assigned mixed strategy 0, then the corresponding mixed strategy profile is an ϵ-Nash equilibrium. We answer this question in the same way we did in the previous case, i.e., by reducing the problem to a max-flow problem. Let us define the vector $\{\theta_{i/k^2}\}_{i \in [k^2] \cup \{0\}}$ by setting $\theta_{i/k^2} = \phi_{i/k^2}$ for all $i \neq 0, 1$, $\theta_{0/k^2} = \phi_{0/k^2} + n - m - m'$ and $\theta_{k^2/k^2} = \phi_{k^2/k^2} + m'$. The

[3] As above, a mixed strategy σ of a player i is an *ϵ-best response* to a set of mixed strategies for the other players iff the expected payoff of player i for playing any pure strategy in the support of σ is no more than ϵ worse than her expected payoff for playing any other pure strategy.

[4] To compute a bound on the number of bits required for the expected utility computations, note that every non-zero probability value that is computed along the execution of the algorithm must be an integer multiple of $(\frac{1}{kn})^j$ for some $j \leq n-1$, since the mixed strategies of all players are from the set $\{0, q, 1\}$. Further note that the expected utility is a weighted sum of n payoff values, with U bits required to represent each value, and all weights being probabilities.

flow network has vertex set $\{s, t\} \cup [n] \cup \{i/k^2\}_{i=0}^n$ and edge set as follows: There is a directed edge from $j \in [n]$ to $\sigma \in \{i/k^2\}_{i=0}^n$ iff $\theta_\sigma > 0$ and mixed strategy σ is an ϵ-best response for player j when the partition of the other players into the mixed strategies i/k^2 is the partition θ, with one unit subtracted from θ_σ. Moreover, there is a directed edge of capacity 1 from s to all nodes in $[n]$ and a directed edge of capacity θ_σ from node σ to t, for all $\sigma \in \{i/k^2\}_{i \in [k^2] \cup \{0\}}$. If the maximum flow from s to t has value n then there is a way to select S, S' and ϕ by looking at the edges used by an integral maximum flow. As far as the running time goes, to define the edge-set of the flow network we need expected payoff computations. Every such computation can be carried out with $O(n^3)$ operations on numbers of at most $b'(n, k) = O(n \log(k^2) + U)$ bits.[5] Since there are at most $k^3 \cdot (n+1) \cdot ((k+1)e)^{k^2}$ choices for (m, m', ϕ), this step takes overall time

$$O\left(\left(n(k^2+1)n^3 b'(n,k) + p_2(n,k)\right) \cdot (n+1)k^3((k+1)e)^{k^2}\right), \tag{13}$$

where $p_2(n, k)$ is the time needed to solve a max-flow problem on a graph with $n + k^2 + 3$ nodes, at most $(n + 1)(k^2 + 2) - 1$ edges, and edge capacities of at most $\lceil \log_2 n \rceil$ bits.

From (12) and (13), it follows that the overall running time is at most

$$\text{poly}(n) \cdot U \cdot (1/\epsilon)^{O(1/\epsilon^2)}.$$

□

Remark 2. *Theorem 6 generalizes to typed anonymous games with* 2 *strategies per player improving the running time of the algorithm of Theorem 4. The generalization amounts to guessing the number of players of each type that uses each of the discretized mixed strategies and setting up a maximum flow problem for each type separately, as discussed in the proof of Theorem 4.*

5.2 The Structural Result

Recall that the algorithm of Theorem 6 effectively searches for an ϵ-Nash equilibrium in the set of mixed strategy profiles satisfying Property 2(b)i or 2(b)ii of Theorem 5. The immediate corollary of this realization is the following structural result for approximate Nash equilibria. The result is reminiscent of Nash's theorem for symmetric games, namely that every game that is symmetric with respect to all player permutations has a Nash equilibrium where every player plays the same mixed strategy. In the case of anonymous games the symmetry of the game is much weaker since players have different utility functions. Despite this, we show that there always exists an approximate equilibrium where either a small number of players randomize or every randomizing player plays the same mixed strategy.

Theorem 7. *For all $\epsilon > 0$, a normalized n-player 2-strategy anonymous game has an ϵ-Nash equilibrium such that for some integer $k = O(1/\epsilon)$:*

1. *either at most k^3 players randomize, and their mixed strategies use probabilities that are integer multiples of $1/k^2$; or*
2. *all players who randomize use the same mixed strategy, and the probabilities used by this mixed strategy are integer multiples of $\frac{1}{kn}$.*

[5]The bound on the number of bits follows from the fact that every non-zero probability value that is computed along the execution of the algorithm must be an integer multiple of $(\frac{1}{k^2})^j$ for some $j \leq n-1$, since the mixed strategies of all players are from the set $\{i/k^2\}_{i \in [k^2] \cup \{0\}}$. Further, note that the expected utility is a weighted sum of n payoff values, with U bits required to represent each value, and all weights being probabilities.

6 Oblivious Algorithms: A Computational Lower Bound

The algorithms presented in Theorems 2 and 6 share the following properties. First, the approximation guarantee ϵ appears in the exponent of the running time. Second, they can both be extended to anonymous games with types (see Theorem 4 and Remark 2). Finally, they both enumerate over a predetermined set of unordered n-tuples of mixed strategies, and only use the input game's description to find out which tuple can be "permuted" to a Nash equilibrium. So both algorithms belong to the following, more general class of algorithms, called *oblivious approximation algorithms.*

Definition 1. An oblivious ϵ-approximation algorithm for anonymous games *is defined in terms of a sequence of distributions, indexed by the number of players n. The n-th distribution in the sequence is a distribution over* unordered *n-tuples of mixed strategies, i.e. collections of mixed strategies that are unassigned to players. For a given n-player (possibly typed) anonymous game, the algorithm samples from the n-th distribution an unordered n-tuple of mixed strategies and determines whether there is an assignment of these strategies to the players of the game so that the resulting mixed strategy profile (with each player using her assigned mixed strategy) is an ϵ-Nash equilibrium. If not, it continues sampling from the n-th distribution until an ϵ-Nash equilibrium is found.*

The expected running time of the algorithm is the inverse of its probability of success times the time needed to determine whether an unordered n-tuple of mixed strategies can be permuted to an ϵ-Nash equilibrium.

Our algorithms from Theorems 2 and 6 can be viewed as oblivious algorithms with the following components: (i) The distribution over unordered n-tuples of mixed strategies is the uniform distribution over unordered n-tuples of mixed strategies discretized as specified by Theorems 3 and 5 respectively (in particular, satisfying Property 2 of Theorem 3 for $z = (f(\xi)/\epsilon)^6$ and, respectively, Property 2b of Theorem 5 for $k = O(1/\epsilon)$); (ii) determining whether an unordered n-tuple of mixed strategies can be permuted to an ϵ-Nash equilibrium is carried out with max-flow computations.

Note that the running time of both algorithms, as well as their generalizations to typed anonymous games, contains $(1/\epsilon)^{\Theta(\xi)}$, i.e. some polynomial function of $1/\epsilon$, in the exponent. The question is this: *Can we remove this polynomial dependence on the approximation from the exponent*? We show that the answer is "no". In particular, we show that any oblivious ϵ-approximation algorithm for 3 type, 2 strategy anonymous games whose expected running time is polynomial in the number of players must have expected running time exponential in $(\frac{1}{\epsilon})^{1/3}$.

Theorem 8. *For any constants $c, \epsilon \geq 0$, no oblivious ϵ-approximation algorithm for normalized anonymous games with 2 strategies and 3 player types has probability of success larger than $n^{-c} \cdot 2^{-\Omega(1/\epsilon^{1/3})}$.*

We only sketch the proof here and postpone further details to Appendix A. We first establish the following (Theorem 11 in Appendix A.1): given any ordered n-tuple $(p_1, \ldots, p_n)$ of probabilities, we can construct a 3 type, 2 strategy anonymous game with n players of type A and two players of their own type such that, in any ϵ-Nash equilibrium, the i-th player of type A plays strategy 2 with probability very close (depending on ϵ and n) to the prescribed p_i. To obtain this game, we need to understand how to exploit the difference in the payoff functions of the players of type A to enforce different behaviors at equilibrium, despite the fact that in all other aspects of the game the players of group A are indistinguishable.

The construction is based on the following idea: For all i, let us denote by $\mu_{-i} := \sum_{j \neq i} p_j$ the *target* expected number of type-A players different than i who play strategy 2; and let us give this payoff to player i if she plays strategy 1, regardless of what the other players are doing. If i chooses

2 instead, we give her expected payoff equal to the *realized* expected number of players different than i who play 2 (which depends on the mixed strategies of the other players). By setting the payoffs in this way we can ensure that $(p_1, \ldots, p_n)$ is an equilibrium of the game, since for every player the payoff she gets from strategy 1 matches the expected payoff she gets from strategy 2, if all players play according to the prescribed probabilities. However, enforcing that $(p_1, \ldots, p_n)$ is also the unique equilibrium is a more challenging task. To do this we include two other players of their own type: we use these players to ensure that the sum of the mixed stategies of the players of type A matches the target $\sum p_i$ at equilibrium, so that a player i deviating from her prescribed strategy p_i is pushed back towards p_i. We show how this can be done in Appendix A.1. We also provide guarantees not only for exact but also for ϵ-Nash equilibria of the resulting game.

The construction outlined above enables one to define a family of $2^{\Omega(1/\epsilon^{1/3})}$ anonymous games of $O((1/\epsilon)^{1/3})$ players with the property that no two games in the family share an ϵ-Nash equilibrium, even as an unordered tuple of mixed strategies (Claims 4 and 5 in Appendix A.2). Then, by an averaging argument, we can deduce that for any oblivious algorithm there is a game in the family for which the probability of success is at most $2^{-\Omega(1/\epsilon^{1/3})}$. The detailed proof of Theorem 8 is given in Appendix A.2.

Remark 3. *We can show an analog of Theorem 8 for oblivious ϵ-approximation algorithms for anonymous games with* 2*-player types and* 3 *strategies per player. The details are omitted.*

7 Beyond Oblivion: Moments of Mixed Strategy Profiles

In the previous section we identified a barrier for improving the running time of our approximation algorithms. In this section we overcome this barrier, providing a PTAS for 2-strategy anonymous games whose running time is some fixed polynomial in the number of players, n, times a factor of $(\frac{1}{\epsilon})^{O\left(\log^2 \frac{1}{\epsilon}\right)}$, where ϵ is the desired approximation. Our new PTAS is of course non-oblivious. Indeed it is so in the following interesting way: Instead of sampling from a distribution over unordered n-tuples of mixed strategies, our algorithm samples from a distribution over $\log(1/\epsilon)$-*tuples, representing the first* $\log(1/\epsilon)$ *moments of these n-tuples.* We can think of the these moment vectors as *more succinct aggregates* of mixed strategy profiles than the unordered n-tuples of mixed strategies considered earlier in this paper, as several of these unordered tuples may share the same moments.

But what is the intuition for considering moment vectors of mixed strategy profiles? Our algorithm of Theorem 6 is founded on the probabilistic approximation presented in Theorem 5. In its heart this approximation quantifies the following intuitive (albeit not quantitatively precise) fact about sums of independent indicators: *If two sums of independent indicators have close means and variances, then their total variation distance should be small.* However, the precise bound obtained by quantifying this intuition is weak enough that the set of unordered n-tuples of mixed strategies that our algorithm has to sample from is exponential in $1/\epsilon$; and our lower bound from the previous section supports that this cannot be improved. Given this realization it seems intuitive that a tighter probabilistic approximation and a faster algorithm may be found by considering higher moments of mixed strategies. Indeed, our new PTAS is founded on the following theorem, which provides a rather strong quantification of how the total variation distance between two sums of indicators depends on the number of their first moments that are equal.

Theorem 9 ([DP09])**.** *Let $\mathcal{P} := (p_i)_{i=1}^n \in (0, 1/2]^n$ and $\mathcal{Q} := (q_i)_{i=1}^n \in (0, 1/2]^n$ be two collections of probability values in $(0, 1/2]$. Let also $\mathcal{X} := (X_i)_{i=1}^n$ and $\mathcal{Y} := (Y_i)_{i=1}^n$ be two collections of* *independent indicators with* $\mathbb{E}[X_i] = p_i$ *and* $\mathbb{E}[Y_i] = q_i$, *for all* $i \in [n]$. *If for some* $d \in [n]$ *the following condition is satisfied:*

$$(C_d): \quad \sum_{i=1}^{n} p_i^\ell = \sum_{i=1}^{n} q_i^\ell, \quad \textit{for all } \ell = 1, \ldots, d,$$

$$\textit{then} \quad \left|\left|\sum_i X_i \; - \; \sum_i Y_i\right|\right|_{\mathrm{TV}} \leq 20(d+1)^{1/4} 2^{-(d+1)/2}. \tag{14}$$

The exact same conclusion holds if instead of the interval $(0, 1/2]$ *we use the interval* $[1/2, 1)$ *in the hypothesis of the theorem.*

Remark 4. *Condition* (C_d) *considers the power sums of the expectations of the indicators. Using the theory of symmetric polynomials it can be shown that* (C_d) *is equivalent to the following condition on the moments of the sums of the indicators (for the proof see [DP09]):*

$$(V_d): \quad \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^\ell\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^\ell\right], \; \textit{for all } \ell \in [d].$$

Theorem 9 provides the following strong approximation guarantee for two sums of indicator random variables: *If two sums of independent indicators with expectations bounded by 1/2 have equal first d moments, then their total variation distance is* $2^{-\Omega(d)}$. It is important to note that this bound applies for all n and, in particular, does not rely upon summing up a large number of indicators. This is very crucial as we explain in the next sections.

7.1 The Oblivious PTAS of Section 5, Revisited

The algorithm presented in the proof of Theorem 6 can be summarized as follows.

1. Choose $k = O(1/\epsilon)$, according to Theorem 7.
2. Guess the number t of players who randomize, the number t_0 of players who use mixed strategy 0 (i.e. play pure strategy 1), and the number $t_1 = n - t - t_0$ of players who use mixed strategy 1 (i.e. play pure strategy 2).
3. Depending on the number t of players who mix do one of the following:
   (a) If $t > k^3$, guess an integer multiple i/kn of $1/kn$ and, solving a maximum flow problem whose details are given in the proof of Theorem 6, check if there is an ϵ-Nash equilibrium in which t players use mixed strategy i/kn, t_0 players use mixed strategy 0, and t_1 players use mixed strategy 1.
   (b) If $t \leq k^3$, guess the number of players ψ_i whose mixed strategy is i/k^2, for all $i \in \{1, \ldots, k^2 - 1\}$, and, solving a maximum flow problem whose details are given in the proof of Theorem 6, check if there is a ϵ-Nash equilibrium in which ψ_i players use mixed strategy i/k^2, for all i, t_0 players use mixed strategy 0, and t_1 players use mixed strategy 1.

Figure 1: The oblivious PTAS of Section 5

There are clearly $O(n^2)$ possible choices for Step 2 of the algorithm. Moreover, the search of Step 3a can be completed in time (see the proof of Theorem 6)

$$U \cdot \text{poly}(n) \cdot (1/\epsilon) \log_2(1/\epsilon),$$

which is quasi-linear in $1/\epsilon$.

On the other hand, Step 3b involves searching over all partitions of t balls into $k^2 - 1$ bins. The resulting running time for this step (see the proof of Theorem 6) is

$$U \cdot \text{poly}(n) \cdot (1/\epsilon)^{O(1/\epsilon^2)},$$

which is exponential in $1/\epsilon$.

7.2 Moment Search

It is clear that the exponential dependence of the running time of the algorithm of Figure 1 on $1/\epsilon$ is due to Step 3b of the algorithm. Moment Search, presented below, utilizes Theorem 9 to improve precisely this step of the algorithm. Indeed, rather than guessing the number of players whose mixed strategy is every multiple of $1/k^2$ at Nash equilibrium, the algorithm only guesses the first $O(\log 1/\epsilon)$ moments of their mixed strategy profile at Nash equilibrium (Step 3c below). Then it uses the players' payoff functions to try to disentangle this moment vector into possible mixed strategies that each player could be using at a Nash equilibrium resulting in the guessed moment vector (Step 3d below).

Algorithm Moment Search
Input: A 2-strategy anonymous game $\mathcal{G}$, the desired approximation ϵ.
Output: An ϵ-Nash equilibrium of $\mathcal{G}$.

1. */*Replaces Step 1 of Figure 1*/* For technical reasons (that will be clear in the proof of correctness of this algorithm), we choose a value of $k = \lceil \frac{c}{\epsilon} \rceil$ (where c is some universal constant) that is by a factor of 2 larger than the value of k required by Theorem 7. This is the value k that guarantees the existence of an $\epsilon/2$-Nash equilibrium having the form 1 or 2 of the theorem.

2. */*Same as Step 2 of Figure 1*/* Guess the number t of players who randomize, the number t_0 of players who use mixed strategy 0 (i.e. play pure strategy 1), and the number $t_1 = n - t - t_0$ of players who use mixed strategy 1 (i.e. play pure strategy 2);

3. Depending on the number t of players who mix do one of the following:

   (a) */*Same as Step 3a of Figure 1*/* If $t > k^3$, guess an integer multiple i/kn of $1/kn$ and, solving a maximum flow problem whose details are given in the proof of Theorem 6, check if there is an ϵ-Nash equilibrium in which t players use mixed strategy i/kn, t_0 players use mixed strategy 0, and t_1 players use mixed strategy 1.

   (b) */*The following steps replace Step 3b of Figure 1. If the control of the algorithm is here the algorithm has guessed that there exists an $\epsilon/2$-Nash equilibrium in which at most t players randomize in integer multiples of $1/k^2$.*/*

   If $t \leq k^3$, guess positive integers t_s, t_b such that $t_s + t_b = t$, where t_s is the number of players who mix with probability $\leq \frac{1}{2}$ (i.e. play pure strategy 2 with probability at most 1/2), and $t_b = n - t_0 - t_1 - t_s$ is the number of players who mix with probability $> \frac{1}{2}$ (i.e.

play pure strategy 2 with probability larger than 1/2). (Note that we have to distinguish between these groups of players because our approximation theorem (Theorem 9) is making a distinction between the cases of probabilities lying in $(0, 1/2]$ and $[1/2, 1)$.)

(c) For $d = \lceil 3 \log_2(320/\epsilon) \rceil$, guess $\mu_1, \mu_2, \ldots, \mu_d, \mu'_1, \mu'_2, \ldots, \mu'_d$, where, for all $\ell \in [d]$:

$$\mu_\ell \in \left\{ j \left(\frac{1}{k^2} \right)^\ell : \; t_s \leq j \leq t_s \left(\frac{k^2}{2} \right)^\ell \right\},$$

and

$$\mu'_\ell \in \left\{ j \left(\frac{1}{k^2} \right)^\ell : t_b \left(\frac{k^2}{2} + 1 \right)^\ell \leq j \leq t_b (k^2 - 1)^\ell \right\}.$$

For all ℓ, μ_ℓ represents the ℓ-power sum of the mixed strategies of the players who mix and choose mixed strategies from the set $\{1/k^2, \ldots, 1/2\}$. Similarly, μ'_ℓ represents the ℓ-power sum of the mixed strategies of the players who mix and choose mixed strategies from the set $\{1/2 + 1/k^2, \ldots, (k^2-1)/k^2\}$. **Remark:** Whether there actually exist probability values $\pi_1, \ldots, \pi_{t_s} \in \{1/k^2, \ldots, 1/2\}$ and $\theta_1, \ldots, \theta_{t_b} \in \{1/2 + 1/k^2, \ldots, (k^2-1)/k^2\}$ such that $\mu_\ell = \sum_{i=1}^{t_s} \pi_i^\ell$ and $\mu'_\ell = \sum_{i=1}^{t_b} \theta_i^\ell$, for all $\ell = 1, 2, \ldots, d$, will be determined later.

(d) For each player $i = 1, \ldots, n$, find a subset

$$\mathcal{S}_i \subseteq \left\{ 0, \frac{1}{k^2}, \ldots, \frac{k^2 - 1}{k^2}, 1 \right\}$$

of permitted mixed strategies for that player in an $\frac{\epsilon}{2}$-Nash equilibrium, conditioned on the guesses in the previous steps. By this, we mean determining the answer to the following: "Given our guesses for the aggregates t_0, t_1, t_s, t_b, μ_ℓ, μ'_ℓ, for all $\ell \in [d]$, what multiples of $1/k^2$ could player i be playing in an $\epsilon/2$-Nash equilibrium?" Our test exploits the anonymity of the game and uses Theorem 9 to achieve the following:

- if a multiple of $1/k^2$ can be assigned to player i and complemented by choices of multiples for the other players, so that the aggregate conditions are satisfied and player i is at $3\epsilon/4$-best response (that is, she experiences at most $3\epsilon/4$ regret), then this multiple of $1/k^2$ is included in the set $\mathcal{S}_i$;
- if, given a multiple of $1/k^2$ to player i, there exists no assignment of multiples to the other players so that the aggregate conditions are satisfied and player i is at $3\epsilon/4$-best response, the multiple is rejected from set $\mathcal{S}_i$.

Observe that the value of $3\epsilon/4$ used in our classifier is intentionally chosen midway between $\epsilon/2$ and ϵ. The reason for this value is that, if we only match the first d moments of a mixed strategy profile, our estimation of the real approximation of that strategy profile is distorted by an additive error $\epsilon/4$ (coming from (14) and the choice of d). Hence, with a threshold at $3\epsilon/4$ we make sure that: a. we are not going to "miss" the mixed strategy of player i in the $\epsilon/2$-Nash equilibrium that we know exists in multiples of $1/k^2$ by virtue of our choice of k and Theorem 7; and b. *any* strategy profile that is consistent with the aggregate conditions and the sets $\mathcal{S}_i$ found in this step is going to be an approximate Nash equilibrium of approximation $3\epsilon/4 + \epsilon/4 = \epsilon$. The fairly involved details of our test are given in Appendix B.1, and the way its analysis ties in with the search for an ϵ-Nash equilibrium is given in the proofs of Claims 6 and 7 of Appendix B.2.

(e) Find an assignment of mixed strategies $v_1 \in \mathcal{S}_1, \ldots, v_n \in \mathcal{S}_n$ to players, such that:

- t_0 players are assigned mixed strategy 0 and t_1 players mixed strategy 1;
- t_s players are assigned a mixed strategy in $(0, 1/2]$, and we satisfy $\sum_{i:v_i \in (0,1/2]} v_i^\ell = \mu_\ell$, for all $\ell \in [d]$;
- t_b players are assigned a mixed strategy in $(1/2, 1)$, and we satisfy $\sum_{i:v_i \in (1/2,1)} v_i^\ell = \mu'_\ell$, for all $\ell \in [d]$.

Solving this assignment problem is non-trivial, but can be done by dynamic programming in time

$$O(n^3) \cdot \left(\frac{1}{\epsilon}\right)^{O(\log^2(1/\epsilon))},$$

because the sets $\mathcal{S}_i$ are subsets of $\{0, 1/k^2, \ldots, 1\}$. The algorithm is given in the proof of Claim 8 in Appendix B.2.

(f) If an assignment is found, then the vector $(v_1, \ldots, v_n)$ constitutes an ϵ-Nash equilibrium.

Theorem 10. Moment Search *is a PTAS for n-player 2-strategy anonymous games with running time $U \cdot \text{poly}(n) \cdot (1/\epsilon)^{O(\log^2(1/\epsilon))}$, where U is the number of bits required to represent a payoff value of the game.*

Proof. **(Sketch; complete proof in Appendix B.2)** Correctness follows from the following observation. Theorem 7 and the choice of k guarantee that there exists a $\frac{\epsilon}{2}$-approximate Nash equilibrium satisfying 1 or 2 of that theorem. Therefore, if t_0, t_1 and t are guessed correctly and the control of the algorithm goes to Step 3a correctness follows as in the proof of Theorem 6. Otherwise, if control goes to Step 3b and t_s, t_b are guessed correctly, Step 3d will find non-empty $\mathcal{S}_i$'s for all players for the correct guesses in Step 3c (since in particular the $\epsilon/2$-Nash equilibrium will survive the tests of Step 3d—by Theorem 9 and the choice of d, at most $\epsilon/4$ accuracy is lost if the correct values for the moments are guessed); and thus Step 3e will find an ϵ-approximate Nash equilibrium (ϵ instead of $3\epsilon/4$, because another $\epsilon/4$ may be lost in this step). The full proof and the running time analysis are provided in Appendix B.2. □

Remark 5. *Theorem 10 generalizes to typed anonymous games with 2 strategies per player. The generalization amounts to guessing the moment vector of the mixed strategy profile of the players of each type separately.*

8 Conclusion and Open Problems

Multiplayer games are of great interest to the interface between Game Theory and Computer Science, and yet they are problematic when seen from the computational standpoint because of issues of representation. Anonymous games are an important class of multiplayer games that can be represented very succinctly (the other important genre being graphical games). In this paper, we develop a comprehensive methodology, both analytical and algorithmic, for equilibrium problems, and use it to prove a number of positive results relating to the Nash equilibrium problem for anonymous games:

- A Lipschitz continuity assumption on the utilities and a fixed point argument yields an approximate pure Nash equilibrium (Theorem 1 and Corollary 1).

- A polynomial-time approximation scheme for mixed Nash equilibrium can be obtained (Theorem 2, and Theorem 6 for two-strategy anonymous games) by relating the quality of approximation to the total variation distance between mixed strategy profile distributions (Lemmas 3 and 4), and using novel probabilistic approximation theorems for Poisson Multinomial and Poisson Binomial distributions (Theorems 3 and 5 respectively), which may have broader application in equilibrium computations.

- We show that anonymous games with two strategies always possess approximate Nash equilibria in which either very few players randomize, or those who do randomize the same way (Theorem 7); this leads to improved algorithms.

- Finally, also for two-strategy anonymous games, by using a more sophisticated argument about moments of the strategy profile distribution (Theorem 9), we obtain a non-oblivious algorithm with quasi-polynomial dependence on the approximation parameter (Theorem 10)—no oblivious algorithm can so perform (Theorem 8).

Several problems arise:

1. Devise better approximation algorithms. More elaborate probabilistic approximation results, or stronger results about the structure of equilibria of anonymous games are possibilities. Are Nash equilibria of two-strategy anonymous games even guaranteed to be rational?

2. Is there a *fully* polynomial-time approximation scheme, that is one whose running time depends polynomially on the approximation parameter?

3. Notice that it is not known whether finding an exact Nash equilibrium in an anonymous game (with a scaling number of players and a non-scaling number of strategies) is PPAD-complete. We conjecture that it is.

APPENDIX

A The Oblivious Lower Bound

For the purposes of the lower bound it is more convenient to work with anonymous games whose payoffs lie in $[-1, 1]$. These games can be normalized using the affine transformation discussed before Proposition 1, with a factor 2 increase in the approximation.

A.1 Constructing Anonymous Games with Prescribed Equilibria

Theorem 11. *For all $\delta > 0$ and $k \in \mathbb{N}$ such that $3\delta k < 1$, and for any collection $\mathcal{P} := (p_i)_{i\in[k]}$, where $p_i \in [3\delta k, 1]$ for all i, there exists a $(k+2)$-player 2-strategy anonymous game $\mathcal{G}_{\mathcal{P}}$ with payoffs in $[-1, 1]$ and three player types, A,B and C such that: (i) k players, $1, \ldots, k$, belong to type A, 1 player to type B, and 1 player to type C; and (ii) for all $\delta' < \delta$, in every δ'-Nash equilibrium of the game the following is satisfied: For every i, player i's mixed strategy belongs to the set $[p_i - 7k^2\delta, p_i + 7k^2\delta]$; moreover, at least one of the players belonging to types B and C play strategy 2 with probability 0.*

Proof. Let us call B the player of type B and C the player of type C. Let us also use the notation: $\mu = \sum_{i\in[k]} p_i$, and $\mu_{-i} = \sum_{j\in[k]\setminus\{i\}} p_j$, for all i. Now, let us assign the following payoffs to the players B and C:

- $u_2^B = \frac{1}{k} \cdot (t_A - \mu)$, where t_A is the number of players of type A who play strategy 2;
- $u_1^B = 2\delta$;
- $u_2^C = \frac{1}{k} \cdot (\mu - t_A)$, where t_A is the number of players of type A who play strategy 2;
- $u_1^C = 2\delta$;

The payoff functions of the players of type A are defined as follows. For all $i \in [k]$:

- $u_1^i = \frac{1}{k}(\mu_{-i} \cdot \mathcal{X}_{\text{B plays 1}} \cdot \mathcal{X}_{\text{C plays 1}} - \delta k \cdot \mathcal{X}_{\text{C plays 2}})$, where $\mathcal{X}_{\text{B plays 1}}$, $\mathcal{X}_{\text{C plays 1}}$ and $\mathcal{X}_{\text{C plays 2}}$ are the indicators of the events 'B plays 1', 'C plays 1' and 'C plays 2' respectively.
- $u_2^i = \frac{1}{k}(t_{A,-i} \cdot \mathcal{X}_{\text{B plays 1}} \cdot \mathcal{X}_{\text{C plays 1}} - \delta k \cdot \mathcal{X}_{\text{B plays 2}})$, where $t_{A,-i}$ is the number of players of type A who are different than i and play 2, and $\mathcal{X}_{\text{B plays 1}}$, $\mathcal{X}_{\text{C plays 1}}$ and $\mathcal{X}_{\text{B plays 2}}$ are the indicators of the events 'B plays 1', 'C plays 1' and 'B plays 2' respectively.

Note that the range of all payoffs of the game thus defined is $[-1, 1]$. We claim the following:

Claim 1. *For all $\delta' < \delta$, in every δ'-Nash equilibrium of the game it must be that*

$$\sum_{i \in [k]} q_i = \mu \pm 3\delta k,$$

where $q_1, \ldots, q_k$ are the probabilities that players $1, \ldots, k$ play strategy 2.

Proof of Claim 1: Let $\mu' = \sum_{i \in [k]} q_i$. Suppose for a contradiction that in a δ'-Nash equilibrium $\mu' > \mu + 3\delta k$; then

$$\frac{1}{k}(\mu' - \mu) > 3\delta.$$

Note however that $\mathbb{E}[u_2^B] = \frac{1}{k}(\mu' - \mu)$ and $\mathbb{E}[u_2^C] = -\frac{1}{k}(\mu' - \mu)$. Hence, the above implies

$$\mathbb{E}[u_2^B] > \mathbb{E}[u_1^B] + \delta, \tag{15}$$
$$\mathbb{E}[u_2^C] < \mathbb{E}[u_1^C] - \delta. \tag{16}$$

Since $\delta' < \delta$, it must be that $\Pr[\text{B plays 2}] = 1$ and $\Pr[\text{C plays 2}] = 0$. It follows then that for all $i \in [k]$:

$$\mathbb{E}[u_1^i] = 0,$$
$$\mathbb{E}[u_2^i] = -\delta.$$

Hence, in a δ'-Nash equilibrium with $\delta' < \delta$, it must be that $\Pr[i \text{ plays } 2] = q_i = 0$, for all $i \in [k]$. This is a contradiction since we assumed that $\mu' = \sum_{i \in [k]} q_i > \mu + 3\delta k$, and μ is non-negative. Via similar arguments we show that the assumption $\mu' < \mu - 3\delta k$ also leads to a contradiction. Hence, in every δ'-Nash equilibrium with $\delta' < \delta$, it must be that

$$\mu' = \mu \pm 3\delta k.$$

□

We next show that in every δ'-Nash equilibrium with $\delta' < \delta$, at least one of the players B and C will not include strategy 2 in her support.

Claim 2. *For all $\delta' < \delta$, in every δ'-Nash equilibrium of the game it must be that*

$$\Pr[B \textit{ plays } 2] = 0 \textit{ or } \Pr[C \textit{ plays } 2] = 0.$$

Proof of Claim 2: Let $q_1, \dots, q_k$ be the probabilities that players $1, \dots, k$ play strategy 2 in some δ'-Nash equilibrium of the game with $\delta' < \delta$. Let us consider the quantity $\mathcal{M} = \frac{1}{k}(\mu' - \mu)$, where $\mu' = \sum_{i \in [k]} q_i$. We distinguish the following cases:

- $\mathcal{M} \leq \delta$: In this case, $\mathbb{E}[u_2^B] = \frac{1}{k}(\mu' - \mu) \leq \delta \leq 2\delta - \delta = \mathbb{E}[u_1^B] - \delta$. Since $\delta' < \delta$, Pr[B plays 2] = 0.
- $\mathcal{M} \geq \delta$: In this case, $\mathbb{E}[u_2^C] = -\frac{1}{k}(\mu' - \mu) \leq -\delta \leq 2\delta - \delta = \mathbb{E}[u_1^C] - \delta$. Since $\delta' < \delta$, Pr[C plays 2] = 0.

□

Finally, we establish the following.

Claim 3. *For all $\delta' < \delta$, in every δ'-Nash equilibrium of the game it must be that for all $i \in [k]$:*

$$\mu'_{-i} := \sum_{j \in [k] \setminus \{i\}} q_j = \mu_{-i} \pm 4\delta k^2,$$

where $q_1, \dots, q_k$ are the probabilities that players $1, \dots, k$ play strategy 2.

Proof of Claim 3: Let us fix an arbitrary δ'-Nash equilibrium. From Claim 2 it follows that either player B or C plays strategy 2 with probability 0. Without loss of generality, we will assume that Pr[C plays 2] = 0 (the argument for the case Pr[B plays 2] = 0 is identical to the one that follows).

Let us now fix an arbitary player $i \in [k]$. We show first that under the assumption Pr[C plays 2] = 0, Pr[C plays 1] = 1, it must be that

$$\mu_{-i} \leq \mu'_{-i} + \delta k. \tag{17}$$

Assume for a contradiction that $\mu_{-i} > \mu'_{-i} + \delta k$. It follows then that

$$\begin{aligned}
&\mu_{-i}(1 - \Pr[\text{B plays 2}]) \geq \mu'_{-i}(1 - \Pr[\text{B plays 2}]) \\
&\qquad\qquad\qquad\qquad\qquad + \delta k(1 - \Pr[\text{B plays 2}]) \\
\Rightarrow \quad &\mu_{-i} \Pr[\text{B plays 1}] \geq \mu'_{-i} \Pr[\text{B plays 1}] \\
&\qquad\qquad\qquad\qquad\qquad + \delta k(1 - \Pr[\text{B plays 2}]) \\
\Rightarrow \quad &\mu_{-i} \Pr[\text{B plays 1}] \Pr[\text{C plays 1}] \geq \\
&\qquad\qquad\qquad \mu'_{-i} \Pr[\text{B plays 1}] \Pr[\text{C plays 1}] \\
&\qquad\qquad\qquad\qquad - \delta k \Pr[\text{B plays 2}] + \delta k \\
\Rightarrow \quad &\mathbb{E}[u_1^i] \geq \mathbb{E}[u_2^i] + \delta.
\end{aligned}$$

But we have fixed a δ'-Nash equilibrium with $\delta' < \delta$; hence the last equation implies that $q_i = 0$. But this quickly leads to a contradiction since, if $q_i = 0$, then using Claim 1 we have

$$\mu'_{-i} = \mu' \geq \mu - 3\delta k \geq \mu - p_i = \mu_{-i},$$

where we also used that $p_i \geq 3\delta k$. The above inequality contradicts our assumption that $\mu_{-i} > \mu'_{-i} + \delta k$. Hence, (17) must be satisfied. Using then $\mu' \leq \mu + 3\delta k$ (which is implied by Claim 1) we get

$$q_i \leq p_i + 4\delta k.$$

As i was arbitrary in the above discussion, it follows that

$$q_j \leq p_j + 4\delta k, \text{for all } j. \tag{18}$$

Now fix $i \in [k]$ again. Summing (18) over all $j \neq i$, we get that

$$\mu'_{-i} \leq \mu_{-i} + 4\delta k^2. \tag{19}$$

Combining (17) and (19) we get

$$\mu'_{-i} = \mu_{-i} \pm 4\delta k^2.$$

□

To conclude the proof of Theorem 11, we combine Claims 1 and 3, as follows. For every player $i \in [k]$, we have from Claims 1 and 3 that in every δ'-Nash equilibrium with $\delta' < \delta$,

$$\mu'_{-i} = \mu_{-i} \pm 4\delta k^2 \text{ and } \mu' = \mu \pm 3\delta k.$$

By combining these equations we get

$$q_i = p_i \pm 7\delta k^2.$$

□

A.2 The Lower Bound

Given Theorem 11, we can establish our lower bound.

Proof of Theorem 8: Let us fix any oblivious ϵ-approximation algorithm for anonymous games with 2 strategies and 3 player types. The algorithm comes together with a distribution over unordered sets of mixed strategies—parametrized by the number of players n—which we denote by D_n.

We will consider the performance of the algorithm on the family of games specified in the statement of Theorem 11 for the following setting of parameters:

$$k = \lfloor (1/\epsilon)^{1/3} \rfloor, \ \delta = 1.01\epsilon, \ \mathcal{P} \in \mathcal{T}_\epsilon^k$$
$$\text{where } \mathcal{T}_\epsilon := \left\{ j \cdot 15\epsilon^{1/3} \,\middle|\, j = 1, \ldots, t_\epsilon \right\}, \quad t_\epsilon = \left\lfloor \frac{1}{15}\epsilon^{-1/3} \right\rfloor.$$

For technical reasons, let us define the following notion of distance between $\mathcal{P}, \mathbb{Q} \in \mathcal{T}_\epsilon^k$.

$$d(\mathcal{P}, \mathbb{Q}) := \sum_{j=1}^{t_\epsilon} \left| v_j^{\mathcal{P}} - v_j^{\mathbb{Q}} \right|.$$

where $v^{\mathcal{P}} = (v_1^{\mathcal{P}}, v_2^{\mathcal{P}}, \ldots, v_{t_\epsilon}^{\mathcal{P}})$ is a vector storing the frequencies of various elements of the set $\mathcal{T}_\epsilon$ in the collection $\mathcal{P}$, i.e. $v_j^{\mathcal{P}} := |\{i \,|\, i \in [k], p_i = j \cdot 15\epsilon^{1/3}\}|$. To find the distance between two collections $\mathcal{P}, \mathbb{Q}$ we compute the ℓ_1 distance of their frequency vectors. Notice in particular that this distance must be an even number. We also need the following definition.

Definition 2. *We say that two anonymous games $\mathcal{G}$ and $\mathcal{G}'$ share an ϵ-Nash equilibrium in unordered form if there exists an ϵ-Nash equilibrium $\sigma_{\mathcal{G}}$ of game $\mathcal{G}$ and an ϵ-Nash equilibrium $\sigma_{\mathcal{G}'}$ of game $\mathcal{G}'$ such that $\sigma_{\mathcal{G}}$ and $\sigma_{\mathcal{G}'}$ are equal as unordered sets of mixed strategies.*

We show first the following about the shareability of ϵ-Nash equilibria among the games $\mathcal{G}_{\mathcal{P}}$, $\mathcal{P} \in T_\epsilon^k$.

Claim 4. *If, for $\mathcal{P}, \mathbb{Q} \in T_\epsilon^k$, $d(\mathcal{P}, \mathbb{Q}) > 0$, then there is no ϵ-Nash equilibrium that is shared between the games $\mathcal{G}_\mathcal{P}$ and $\mathcal{G}_\mathbb{Q}$ in unordered form.*

Proof of Claim 4: For all j, let us define the $7.07k^2\epsilon$ ball around probability $j \cdot 15\epsilon^{1/3}$ in the natural way:

$$B_j := [j \cdot 15\epsilon^{1/3} - 7.07k^2\epsilon, j \cdot 15\epsilon^{1/3} + 7.07k^2\epsilon].$$

Observe that for all $j \geq 2$:

$$(j+1) \cdot 15\epsilon^{1/3} - j \cdot 15\epsilon^{1/3} = 15\epsilon^{1/3} > 2 \cdot 7.07k^2\epsilon.$$

Hence, for all j, j': $B_j \cap B_j' = \emptyset$.

Now, let us consider any pair of ϵ-Nash equilibria $\sigma_{\mathcal{G}_\mathcal{P}}$, $\sigma_{\mathcal{G}_\mathbb{Q}}$ of the games $\mathcal{G}_\mathcal{P}$ and $\mathcal{G}_\mathbb{Q}$ and let us consider the vectors $v^{\sigma_{\mathcal{G}_\mathcal{P}}} = (v_1^{\sigma_{\mathcal{G}_\mathcal{P}}}, \ldots, v_{t_\epsilon}^{\sigma_{\mathcal{G}_\mathcal{P}}})$ and $v^{\sigma_{\mathcal{G}_\mathbb{Q}}} = (v_1^{\sigma_{\mathcal{G}_\mathbb{Q}}}, \ldots, v_{t_\epsilon}^{\sigma_{\mathcal{G}_\mathbb{Q}}})$ whose j-th components are defined as follows:

$$v_j^{\sigma_{\mathcal{G}_\mathcal{P}}} = \left(\begin{array}{c} \text{number of players who are} \\ \text{assigned a mixed strategy from} \\ \text{the set } B_j \text{ in } \sigma_{\mathcal{G}_\mathcal{P}} \end{array} \right),$$

$$v_j^{\sigma_{\mathcal{G}_\mathbb{Q}}} = \left(\begin{array}{c} \text{number of players who are} \\ \text{assigned a mixed strategy from} \\ \text{the set } B_j \text{ in } \sigma_{\mathcal{G}_\mathbb{Q}} \end{array} \right).$$

It is not hard to see that Theorem 11 and our assumption $d(\mathcal{P}, \mathbb{Q}) > 0$ imply that $\|v^{\sigma_{\mathcal{G}_\mathcal{P}}} - v^{\sigma_{\mathcal{G}_\mathbb{Q}}}\|_1 > 0$, hence $\sigma_{\mathcal{G}_\mathcal{P}}$ and $\sigma_{\mathcal{G}_\mathbb{Q}}$ cannot be permutations of each other. This concludes the proof. □

Next, we show that there exists a large family of games such that no two members of the family share an ϵ-Nash equilibrium.

Claim 5. *There exists a subset $T \subseteq \mathcal{T}_\epsilon^k$ such that:*

1. *for every $\mathcal{P}, \mathbb{Q} \in T$: $d(\mathcal{P}, \mathbb{Q}) > 0$;*
2. *$|T| \geq 2^{\Omega\left(\left(\frac{1}{\epsilon}\right)^{1/3}\right)}$;*

Proof of claim 5: The total number of distinct multi-sets of cardinality k with elements from $\mathcal{T}_\epsilon$ is

$$\binom{t_\epsilon + k - 1}{k}.$$

Hence, it is easy to create a subset $T \subseteq \mathcal{T}_\epsilon^k$ such that:

- for every $\mathcal{P}, \mathbb{Q} \in T$: $d(\mathcal{P}, \mathbb{Q}) > 0$;
- $|T| = \binom{t_\epsilon + k - 1}{k}$.

Clearly, the set T satisfies Property 1 in the statement. For the cardinality bound we have:

$$|T| \geq \binom{t_\epsilon + k - 1}{k} \geq \left(\frac{t_\epsilon + k - 1}{k}\right)^k$$
$$\geq \left(1 + \frac{1}{15} - \frac{2}{k}\right)^k \geq 2^{\Omega\left(\left(\frac{1}{\epsilon}\right)^{1/3}\right)}.$$

□

Now let us consider the performance of the distribution D_k on the family of anonymous games $\{\mathcal{G}_{\mathcal{P}}\}_{\mathcal{P}\in T}$, where T is the set defined in Claim 5. By Claims 4 and 5, no two games in the family share an ϵ-Nash equilibrium in unordered form. Hence, no matter what D_k is, there will be some game in our family for which the probability that D_k samples an ϵ-Nash equilibrium of that game is at most

$$1/|T| \leq 2^{-\Omega\left(\left(\frac{1}{\epsilon}\right)^{1/3}\right)}.$$

This concludes the proof of Theorem 8. □

B The non-oblivious PTAS for Anonymous Games

B.1 Moment Search : Missing Details

We first describe in detail Step 3d of Moment Search.

3. (d) For each player $i = 1, \ldots, n$, find a subset

$$\mathcal{S}_i \subseteq \left\{0, \frac{1}{k^2}, \ldots, \frac{k^2-1}{k^2}, 1\right\}$$

of permitted mixed strategies for that player in an $\epsilon/2$-Nash equilibrium, "conditioning" on the total number of players playing mixed strategy 0 being t_0, the total number of players playing mixed strategy 1 being t_1, and the mixed strategies of the players who mix resulting in the power-sums $\mu_1, \ldots, \mu_d$ and $\mu'_1, \ldots, \mu'_d$. The way we compute the set $\mathcal{S}_i$ is as follows:

i. To determine whether $0 \in \mathcal{S}_i$:

A. Find *any* set of mixed strategies $q_1, \ldots, q_{t_s} \subseteq \{\frac{1}{k^2}, \frac{2}{k^2}, \ldots, \frac{1}{2}\}$ such that $\sum_{\iota=1}^{t_s} q_\iota^\ell = \mu_\ell$, for all $\ell = 1, \ldots, d$. Find *any* set of mixed strategies $r_1, \ldots, r_{t_b} \subseteq \{\frac{1}{2}+\frac{1}{k^2}, \frac{1}{2}+\frac{2}{k^2}, \ldots, 1-\frac{1}{k^2}\}$ such that $\sum_{\iota=1}^{t_b} r_\iota^\ell = \mu'_\ell$, for all $\ell = 1, \ldots, d$. If such values do not exist Fail.

Remark: An efficient algorithm to solve this optimization problem is given by Claim 8.

B. Define the random variable

$$Y = (t_0 - 1)\cdot 0 + \sum_{\iota=1}^{t_s} S_\iota + \sum_{\iota=1}^{t_b} B_\iota + t_1 \cdot 1,$$

where the variables $S_1, \ldots, S_{t_s}, B_1, \ldots, B_{t_b}$ are mutually independent with expectations $\mathbb{E}[S_\iota] = q_\iota$, for all $\iota = 1, \ldots, t_s$, and $\mathbb{E}[B_\iota] = r_\iota$, for all $\iota = 1, \ldots, t_b$.

C. Compute the expected payoff $\mathcal{U}_1^i = \mathbb{E}[u_1^i(Y)]$ and $\mathcal{U}_2^i = \mathbb{E}[u_2^i(Y)]$ of player i for playing pure strategy 1 and 2 respectively, if the number of the other players playing 2 is distributed identically to Y.

D. if $\mathcal{U}_1^i \geq \mathcal{U}_2^i - 3\epsilon/4$, then include 0 to the set $\mathcal{S}_i$, otherwise do not.

ii. To determine whether $1 \in \mathcal{S}_i$, follow the same procedure except now Y is defined as follows

$$Y = t_0 \cdot 0 + \sum_{\iota=1}^{t_s} S_\iota + \sum_{\iota=1}^{t_b} B_\iota + (t_1 - 1) \cdot 1,$$

to account for the fact that we are testing for the candidate mixed strategy 1 for player i. Also, the test that determines whether $1 \in \mathcal{S}_i$ is now whether $\mathcal{U}_2^i \geq \mathcal{U}_1^i - 3\epsilon/4$.

iii. For all $j \in \{1, \ldots, k^2/2\}$, to determine whether $j/k^2 \in \mathcal{S}_i$ do the following slightly modified test:

A. Find any set of mixed strategies $q_1, \ldots, q_{t_s-1} \subseteq \{\frac{1}{k^2}, \frac{2}{k^2}, \ldots, 1/2\}$ such that $\sum_{\iota=1}^{t_s-1} q_\iota^\ell = \mu_\ell - (j/k^2)^\ell$, for all $\ell = 1, \ldots, d$. Find any set of mixed strategies

$$r_1, \ldots, r_{t_b} \subseteq \left\{ \frac{1}{2} + \frac{1}{k^2}, \frac{1}{2} + \frac{2}{k^2}, \ldots, 1 - \frac{1}{k^2} \right\}$$

such that $\sum_{\iota=1}^{t_b} r_\iota^\ell = \mu'_\ell$, for all $\ell = 1, \ldots, d$. If such values do not exist FAIL.

B. Define the random variable

$$Y = t_0 \cdot 0 + \sum_{\iota=1}^{t_s-1} S_\iota + \sum_{\iota=1}^{t_b} B_\iota + t_1 \cdot 1,$$

where the variables $S_1, \ldots, S_{t_s-1}, B_1, \ldots, B_{t_b}$ are mutually independent with $\mathbb{E}[S_\iota] = q_\iota$, for all $\iota = 1, \ldots, t_s - 1$, and $\mathbb{E}[B_\iota] = r_\iota$, for all $\iota = 1, \ldots, t_b$.

C. Compute the expected payoff $\mathcal{U}_1^i = \mathbb{E}[u_1^i(Y)]$ and $\mathcal{U}_2^i = \mathbb{E}[u_2^i(Y)]$ of player i for playing pure strategy 1 and 2 respectively, if the number of the other players playing 2 is distributed identically to Y.

D. if $\mathcal{U}_1^i \in [\mathcal{U}_2^i - 3\epsilon/4, \mathcal{U}_2^i + 3\epsilon/4]$, then include j/k^2 to the set $\mathcal{S}_i$, otherwise do not.

iv. For all $j \in \{(k^2+2)/2, \ldots, k^2-1\}$, to determine whether $j/k^2 \in \mathcal{S}_i$ do the appropriate modifications to the method described in Step 3(d)iii.

B.2 The Analysis of MOMENT SEARCH

Correctness The correctness of MOMENT SEARCH follows from the following two claims.

Claim 6. *If there exists an $\epsilon/2$-Nash equilibrium in which $t \leq k^3$ players mix, and their mixed strategies are integer multiples of $1/k^2$, then* MOMENT SEARCH *will not fail, i.e. it will output a set of mixed strategies $(v_1, \ldots, v_n)$.*

Claim 7. *If* MOMENT SEARCH *outputs a set of mixed strategies $(v_1, \ldots, v_n)$, then these strategies constitute an ϵ-Nash equilibrium.*

Proof of Claim 6: Let $(p_1, \ldots, p_n)$ be an $\epsilon/2$-Nash equilibrium in which t_0 players use mixed strategy 0, t_1 players use mixed strategy 1, and $t \leq k^3$ players mix, and their mixed strategies are integer multiples of $1/k^2$. It suffices to show that there exist guesses for t_0, t_1, t_s, t_b, $\mu_1, \ldots, \mu_d$, $\mu'_1, \ldots, \mu'_d$, such that $p_1 \in \mathcal{S}_1$, $p_2 \in \mathcal{S}_2$,..., $p_n \in \mathcal{S}_n$. Indeed, let

$$\mathcal{I}_0 := \{i | p_i = 0\},\ \mathcal{I}_s := \{i | p_i \in (0, 1/2]\},$$

$$\mathcal{I}_b := \{i | p_i \in (1/2, 1)\},\ \mathcal{I}_1 := \{i | p_i = 1\},$$

and let us choose the following values for our guesses

$$t_0 := |\mathcal{I}_0|, t_s = |\mathcal{I}_s|,\ t_b = |\mathcal{I}_b|, t_1 := |\mathcal{I}_1|$$

and, for all $\ell \in [d]$,

$$\mu_\ell = \sum_{i \in \mathcal{I}_s} p_i^\ell, \quad \mu'_\ell = \sum_{i \in \mathcal{I}_b} p_i^\ell.$$

We will show that for the guesses defined above $p_i \in \mathcal{S}_i$, for all i. We distinguish the following cases: $i \in \mathcal{I}_0$, $i \in \mathcal{I}_s$, $i \in \mathcal{I}_b$, $i \in \mathcal{I}_1$. The proof for all cases proceeds in the same fashion, so we will only argue the case $i \in I_s$. In particular, we will show that in Step 3(d)iii of Moment Search the test succeeds for $j/k^2 = p_i$.

At the equilibrium point $(p_1, \ldots, p_n)$, the number of players different than i who choose pure strategy 2 is distributed identically to the random variable:

$$Z := \sum_{\iota \in \mathcal{I}_s \setminus \{i\}} X_\iota + \sum_{\iota \in \mathcal{I}_b} X_\iota + t_1 \cdot 1,$$

where $(X_\iota)_{\iota \in \mathcal{I}_b \cup \mathcal{I}_s \setminus \{i\}}$ are mutually independent random indicators with expectations $\mathbb{E}[X_\iota] = p_\iota$ for all ι. Since $(p_1, \ldots, p_n)$ is an $\epsilon/2$-Nash equilibrium where player i mixes it must be the case that

$$|\mathbb{E}[u_1^i(Z)] - \mathbb{E}[u_2^i(Z)]| \leq \epsilon/2. \tag{20}$$

We will argue that, if in the above equation, we replace Z by Y, where Y is the random variable defined in Step 3(d)iiiB of Moment Search, the inequality still holds with slightly updated upper bound:

$$|\mathbb{E}[u_1^i(Y)] - \mathbb{E}[u_2^i(Y)]| \leq 3\epsilon/4. \tag{21}$$

If (21) is established, the proof is completed since Step 3(d)iiiD will include j/k^2 into the set $\mathcal{S}_i$.

Let $S_1, \ldots, S_{t_s-1}, B_1, \ldots, B_{t_b}$ be the random variables with expectations $q_1, \ldots, q_{t_s-1}, r_1, \ldots, r_{t_b}$ defined in Step 3(d)iiiB of Moment Search. Observe that, for all $\ell = 1, \ldots, d$,

$$\sum_{\iota=1}^{t_s-1} q_\iota^\ell = \mu_\ell - (j/k^2)^\ell = \sum_{\iota \in \mathcal{I}_s \setminus \{i\}} p_\iota^\ell,$$

since $p_i = j/k^2$. Hence, by Theorem 9,

$$\left\| \sum_{\iota=1}^{t_s-1} S_\iota - \sum_{\iota \in \mathcal{I}_s \setminus \{i\}} X_\iota \right\|_{\mathrm{TV}} \leq 20(d+1)^{1/4} 2^{-(d+1)/2} \leq \epsilon/32. \tag{22}$$

Via similar arguments and Theorem 9, we get

$$\left\| \sum_{\iota=1}^{t_b} B_\iota - \sum_{\iota \in \mathcal{I}_b} X_\iota \right\|_{\mathrm{TV}} \leq \epsilon/32. \tag{23}$$

(22) and (23) imply using Lemma 1 that

$$\|Y - Z\|_{\mathrm{TV}} \leq \frac{\epsilon}{16}. \tag{24}$$

Lemma 3 implies then that

$$|\mathbb{E}[u_1^i(Y)] - \mathbb{E}[u_1^i(Z)]| \leq 2\|Y - Z\|_{\rm TV} \leq \frac{\epsilon}{8},$$

where we used (24). Similarly,

$$|\mathbb{E}[u_2^i(Y)] - \mathbb{E}[u_2^i(Z)]| \leq 2\|Y - Z\|_{\rm TV} \leq \frac{\epsilon}{8}.$$

Combining the above with (20) we get (21). This concludes the proof. □

Proof of Claim 7: Let

$$\mathcal{I}_0 := \{i | v_i = 0\},\ \mathcal{I}_s := \{i | v_i \in (0, 1/2]\},$$
$$\mathcal{I}_b := \{i | v_i \in (1/2, 1)\},\ \mathcal{I}_1 := \{i | v_i = 1\},$$
$$t_s = |\mathcal{I}_s|, \text{and } t_b = |\mathcal{I}_b|.$$

Observe that the moment values that were guessed in Step 3c of Moment Search satisfy

$$\mu_\ell = \sum_{i \in \mathcal{I}_s} v_i^\ell,\quad \mu'_\ell = \sum_{i \in \mathcal{I}_b} v_i^\ell,\quad \text{for all } \ell = 1, \ldots, d.$$

We will argue that $(v_1, \ldots, v_n)$ is an ϵ-Nash equilibrium. To do this we need to argue that, for each player i, v_i is an ϵ-best response to the strategies of her opponents. We distinguish the following cases: $i \in \mathcal{I}_0$, $i \in \mathcal{I}_s$, $i \in \mathcal{I}_b$ and $i \in \mathcal{I}_1$. The proof for all cases proceeds in a similar fashion, so we only present the argument for the case $i \in I_s$.

Let $v_i = j/k^2$ for some $j \in \{1, \ldots, \frac{k^2}{2}\}$. From the perspective of player i, the number of other players who play pure strategy 2 in the mixed strategy profile $(v_1, \ldots, v_n)$ is distributed identically to the random variable

$$Z := \sum_{\iota \in [n] \setminus \{i\}} X_\iota,$$

where $(X_\iota)_\iota$ is a collection of mutually independent random indicators with expectations $\mathbb{E}[X_\iota] = v_\iota$ for all ι. To argue that v_i is an ϵ-best response against the strategies of i's opponents, we need to show that

$$|\mathbb{E}[u_1^i(Z)] - \mathbb{E}[u_2^i(Z)]| \leq \epsilon. \tag{25}$$

Let us go back to the execution of Step 3(d)iii in which the probability value j/k^2 was inserted into the set $\mathcal{S}_i$. Let $q_1, \ldots, q_{t_s-1}$, $r_1, \ldots, r_{t_b}$ be the values that were selected in Step 3(d)iiiA of that execution, and let

$$Y = \sum_{\iota=1}^{t_s-1} S_\iota + \sum_{\iota=1}^{t_b} B_\iota + t_1 \cdot 1,$$

be the random variable defined in Step 3(d)iiiB, where the variables $S_1, \ldots, S_{t_s-1}, B_1, \ldots, B_{t_b}$ are mutually independent with expectations $\mathbb{E}[S_\iota] = q_\iota$, for all $\iota = 1, \ldots, t_s - 1$, and $\mathbb{E}[B_\iota] = r_\iota$, for all $\iota = 1, \ldots, t_b$. Observe that the q_ι's and r_ι's where chosen by Step 3(d)iiiA so that the following are satisfied

$$\sum_{\iota=1}^{t_s-1} q_\iota^\ell = \mu_\ell - (j/k^2)^\ell = \mu_\ell - v_i^\ell = \sum_{\iota \in \mathcal{I}_s \setminus \{i\}} v_\iota^\ell,\ \text{for all } \ell \in [d], \tag{26}$$

$$\text{and}\quad \sum_{\iota=1}^{t_b} r_\iota^\ell = \mu'_\ell = \sum_{\iota \in \mathcal{I}_b} v_\iota^\ell,\quad \text{for all } \ell = 1, \ldots, d. \tag{27}$$

Equation (26) implies via Theorem 9 that

$$\left\|\sum_{\iota=1}^{t_s-1} S_\iota - \sum_{\iota\in\mathcal{I}_s\setminus\{i\}} X_\iota\right\|_{\mathrm{TV}} \le 20(d+1)^{1/4}2^{-(d+1)/2} \le \epsilon/32. \tag{28}$$

Equation (27) and Theorem 9 imply

$$\left\|\sum_{\iota=1}^{t_b} B_\iota - \sum_{\iota\in\mathcal{I}_b} X_\iota\right\|_{\mathrm{TV}} \le \epsilon/32. \tag{29}$$

(28) and (29) imply using Lemma 1 that

$$\|Y-Z\|_{\mathrm{TV}} \le \frac{\epsilon}{16}. \tag{30}$$

Lemma 3 implies then that

$$|\mathbb{E}[u_1^i(Y)] - \mathbb{E}[u_1^i(Z)]| \le 2\|Y-Z\|_{\mathrm{TV}} \le \frac{\epsilon}{8}, \tag{31}$$

where we used (30). Similarly,

$$|\mathbb{E}[u_2^i(Y)] - \mathbb{E}[u_2^i(Z)]| \le \|Y-Z\|_{\mathrm{TV}} \le \frac{\epsilon}{8}. \tag{32}$$

Moreover, notice that the random variable Y satisfies the following condition

$$|\mathbb{E}[u_1^i(Y)] - \mathbb{E}[u_2^i(Y)]| \le 3\epsilon/4, \tag{33}$$

since, in order for v_i to be included into $\mathcal{S}_i$, the test in Step 3(d)iiiD of Moment Search must have succeeded. Combining (31), (32) and (33) we get (25). This concludes the proof. □

Computational Complexity We will argue that there is an implementation of Moment Search that runs in time

$$U\cdot \mathrm{poly}(n)\cdot (1/\epsilon)^{O(\log^2(1/\epsilon))},$$

where U is the number of bits required to represent a payoff value of the game. We already argued in Section 7.1 that the number of guesses needed in Step 2 is $O(n^2)$ and that Step 3a can be completed in time $U\cdot \mathrm{poly}(n)\cdot(1/\epsilon)\log_2(1/\epsilon)$. So we only need to pay attention to Steps 3b-3f.

Observe first that the number of possible guesses for t_s, t_b in Step 3b is at most $O((1/\epsilon)^3)$. Observe further that the number of possible guesses for μ_ℓ in Step 3c is at most $t\left(\frac{k^2}{2}\right)^\ell$ (where $t\le k^3$ is the number of players who mix), so jointly the number of possible guesses for $\mu_1,\dots,\mu_d$ is at most

$$\prod_{\ell=1}^{d} t\left(\frac{k^2}{2}\right)^\ell = t^d\left(\frac{k^2}{2}\right)^{d(d+1)/2} = \left(\frac{1}{\epsilon}\right)^{O\left(\log^2\frac{1}{\epsilon}\right)}.$$

The same asymptotic upper bound applies to the total number of guesses for $\mu_1',\dots,\mu_d'$. Given the above the total number of guesses that Moment Search has to do in Steps 3b and 3c is

$$\left(\frac{1}{\epsilon}\right)^{O\left(\log^2\frac{1}{\epsilon}\right)}.$$

We next argue that the running time required to complete Steps 3d, 3e, and 3f is

$$\text{poly}(n) \cdot U \cdot \left(\frac{1}{\epsilon}\right)^{O(\log^2(1/\epsilon))}.$$

For this we establish the following. We give its proof in the end of this section.

Claim 8. *Given a collection of values $\mu_1, \ldots, \mu_d, \mu'_1, \ldots, \mu'_d$, where, for all $\ell = 1, \ldots, d$,*

$$\mu_\ell, \mu'_\ell \in \left\{0, \left(\frac{1}{k^2}\right)^\ell, 2\left(\frac{1}{k^2}\right)^\ell, \ldots, B\left(\frac{k^2}{k^2}\right)^\ell\right\},$$

for some $B \in \mathbb{N}$, discrete sets $\mathcal{T}_1, \ldots, \mathcal{T}_m \subseteq \left\{0, \frac{1}{k^2}, \frac{2}{k^2}, \ldots, 1\right\}$, and four integers $m_0, m_1 \leq m$, $m_s, m_b \leq B$, it is possible to solve the system of equations:

$$\begin{aligned}
(\Sigma): \qquad \sum_{p_i \in (0,1/2]} p_i^\ell &= \mu_\ell, \text{ for all } \ell = 1, \ldots, d, \\
\sum_{p_i \in (1/2,1)} p_i^\ell &= \mu'_\ell, \text{ for all } \ell = 1, \ldots, d, \\
|\{i | p_i = 0\}| &= m_0 \\
|\{i | p_i = 1\}| &= m_1 \\
|\{i | p_i \in (0, 1/2]\}| &= m_s \\
|\{i | p_i \in (1/2, 1)\}| &= m_b
\end{aligned}$$

with respect to unknowns $p_1 \in \mathcal{T}_1, \ldots, p_m \in \mathcal{T}_m$, or to determine that no solution exists, in time

$$O(m^3) B^{O(d)} k^{O(d^2)}.$$

Applying Claim 8 with $m \leq t$, $B \leq t$ (where $t \leq k^3$ is the number of players who mix), $m_0 = 0$, $m_1 = 0$, shows that Steps 3(d)iA, 3(d)iiiA can be completed in time

$$O(t^3) t^{O(d)} k^{O(d^2)} = \left(\frac{1}{\epsilon}\right)^{O(\log^2(1/\epsilon))}.$$

Another application of Claim 8 with $m = n$, $B \leq t$, $m_0 \leq n$, $m_1 \leq n$ shows that Step 3e of Moment Search can be completed in time

$$O(n^3) t^{O(d)} k^{O(d^2)} = O(n^3) \cdot \left(\frac{1}{\epsilon}\right)^{O(\log^2(1/\epsilon))}.$$

Finally, as we argued in the proof of Theorem 6, the computation of the expected utilities $\mathcal{U}_1^i$ and $\mathcal{U}_2^i$ required in Steps 3(d)iC, 3(d)iiiC of Moment Search can be carried out with $O(n^3)$ operations on numbers of at most $O(n \log(k^2) + U)$ bits.

Therefore, the overall time required for the execution of Steps 3b-3f of Moment Search is

$$\text{poly}(n) \cdot U \cdot \left(\frac{1}{\epsilon}\right)^{O(\log^2(1/\epsilon))}.$$

Proof of Claim 8: We use dynamic programming. Let us consider the following tensor of dimension $2d+5$:

$$A(i, z_0, z_1, z_s, z_b; \nu_1, \ldots, \nu_d; \nu'_1, \ldots, \nu'_d),$$

where $i \in [m]$, $z_0, z_1 \in \{0, \ldots, m\}$, $z_s, z_b \in \{0, \ldots, B\}$ and

$$\nu_\ell, \nu'_\ell \in \left\{0, \left(\frac{1}{k^2}\right)^\ell, 2\left(\frac{1}{k^2}\right)^\ell, \ldots, B\right\},$$

for $\ell = 1, \ldots, d$. The total number of cells in A is

$$m \cdot (m+1)^2 \cdot (B+1)^2 \cdot \left(\prod_{\ell=1}^{d}(Bk^{2\ell}+1)\right)^2$$
$$\leq O(m^3)B^{O(d)}k^{2d(d+1)}.$$

Every cell of A is assigned value 0 or 1, as follows:

$$A(i, z_0, z_1, z_s, z_b; \nu_1, \ldots, \nu_d, \nu'_1, \ldots, \nu'_d) = 1$$
$$\Leftrightarrow \left(\begin{array}{c} \text{There exist } p_1 \in \mathcal{T}_1, \ldots, p_i \in \mathcal{T}_i \text{ such} \\ \text{that } |\{j \leq i | p_j = 0\}| = z_0, \\ |\{j \leq i | p_j = 1\}| = z_1, \\ |\{j \leq i | p_j \in (0, 1/2]\}| = z_s, \\ |\{j \leq i | p_j \in (1/2, 1)\}| = z_b, \\ \sum_{j \leq i: p_j \in (0,1/2]} p_j^\ell = \nu_\ell, \text{ for all} \\ \ell = 1, \ldots, d, \sum_{j \leq i: p_j \in (1/2,1)} p_j^\ell = \nu'_\ell, \text{ for} \\ \text{all } \ell = 1, \ldots, d. \end{array}\right).$$

It is easy to complete A working in layers of increasing i. We initialize all entries to value 0. Then, the first layer $A(1, \cdot, \cdot, \cdot, \cdot\ ;\ \cdot, \ldots, \cdot\ ;\ \cdot, \ldots, \cdot)$ can be completed easily as follows:

$$A(1, 1, 0, 0, 0; 0, 0, \ldots, 0; 0, 0, \ldots, 0) = 1 \Leftrightarrow 0 \in \mathcal{T}_1$$
$$A(1, 0, 1, 0, 0; 0, 0, \ldots, 0; 0, 0 \ldots, 0) = 1 \Leftrightarrow 1 \in \mathcal{T}_1$$
$$A(1, 0, 0, 1, 0; p, p^2, \ldots, p^d; 0, \ldots, 0) = 1 \Leftrightarrow p \in \mathcal{T}_1 \cap (0, 1/2]$$
$$A(1, 0, 0, 0, 1; 0, \ldots, 0; p, p^2, \ldots, p^d) = 1 \Leftrightarrow p \in \mathcal{T}_1 \cap (1/2, 1)$$

Inductively, to complete layer $i+1$, we consider all the non-zero entries of layer i and for every such non-zero entry and for every $v_{i+1} \in \mathcal{T}_{i+1}$, we find which entry of layer $i+1$ we would transition to if we chose $p_{i+1} = v_{i+1}$. We set that entry equal to 1 and we also save a pointer to this entry from the corresponding entry of layer i, labeling that pointer with the value v_{i+1}. The time we need to complete layer $i+1$ is bounded by

$$|\mathcal{T}_{i+1}|.(m+1)^2 B^{O(d)} k^{2d(d+1)} \leq O(m^2) B^{O(d)} k^{O(d^2)}.$$

Therefore, the overall time needed to complete A is

$$O(m^3) B^{O(d)} k^{O(d^2)}.$$

After completing tensor A, it is easy to check if there exists a solution to (Σ). A solution exists if and only if

$$A(m, m_0, m_1, m_s, m_b; \mu_1, \ldots, \mu_d; \mu'_1, \ldots, \mu'_d) = 1,$$

and it can be found by tracing back the pointers from this cell of A. The overall running time is dominated by the time needed to fill in A. □

References

[AD54] Kenneth J. Arrow and Gerard Debreu. Existence of an Equilibrium for a Competitive Economy. *Econometrica*, 22(3):265–290, 1954.

[ARV08] Heiner Ackermann, Heiko Röglin, and Berthold Vöcking. On the Impact of Combinatorial Structure on Congestion Games. *Journal of the ACM (JACM)*, 55(6):25, 2008.

[AS13] Yaron Azrieli and Eran Shmaya. Lipschitz Games. *Mathematics of Operations Research*, 38(2), 2013.

[Blo99] Matthias Blonski. Anonymous Games with Binary Actions. *Games and Economic Behavior*, 28(2):171–180, 1999.

[Blo05] Matthias Blonski. The Women of Cairo: Equilibria in Large Anonymous Games. *Journal of Mathematical Economics*, 41(3):253–264, 2005.

[CFGS12] Ioannis Caragiannis, Angelo Fanelli, Nick Gravin, and Alexander Skopalik. Computing Approximate Pure Nash Equilibria in Congestion Games. *ACM SIGecom Exchanges*, 11(1):26–29, 2012.

[CS11] Steve Chien and Alistair Sinclair. Convergence to Approximate Nash Equilibria in Congestion Games. *Games and Economic Behavior*, 71(2):315–327, 2011.

[Das08a] Constantinos Daskalakis. An Efficient PTAS for Two-Strategy Anonymous Games. In *the 4th International Workshop on Internet and Network Economics (WINE)*, 2008.

[Das08b] Constantinos Daskalakis. An Efficient PTAS for Two-Strategy Anonymous Games. *CoRR*, abs/0812.2277, 2008.

[Das11] Constantinos Daskalakis. On the Complexity of Approximating a Nash Equilibrium. In *the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2011.

[DGP06] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The Complexity of Computing a Nash Equilibrium. In *the 38th Annual ACM Symposium on Theory of Computing (STOC)*, 2006.

[DGP09a] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The Complexity of Computing a Nash Equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.

[DGP09b] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The Complexity of Computing a Nash Equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

[DP07] Constantinos Daskalakis and Christos H. Papadimitriou. Computing Equilibria in Anonymous Games. In *the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2007.

[DP08] Constantinos Daskalakis and Christos H. Papadimitriou. Discretized Multinomial Distributions and Nash Equilibria in Anonymous Games. In *the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2008.

[DP09] Constantinos Daskalakis and Christos H. Papadimitriou. On Oblivious PTAS's for Nash Equilibrium. In *the 41st Annual ACM Symposium on Theory of Computing (STOC)*, 2009.

[DP13] Constantinos Daskalakis and Christos H. Papadimitriou. Sparse Covers for Sums of Indicators. *Arxiv report*, 2013.

[FPT04] Alex Fabrikant, Christos H. Papadimitriou, and Kunal Talwar. The Complexity of Pure Nash Equilibria. In *the 36th Annual ACM Symposium on Theory of Computing (STOC)*, 2004.

[Kal05] Ehud Kalai. Partially-Specified Large Games. In *the 1st International Workshop on Internet and Network Economics (WINE)*, 2005.

[LMM03] Richard J. Lipton, Evangelos Markakis, and Aranyak Mehta. Playing Large Games Using Simple Strategies. In *the 4th ACM Conference on Electronic Commerce (EC)*, 2003.

[LT82] Gerard van der Laan and Adolphus Johannes Jan Talman. On the Computation of Fixed Points in the Product Space of Unit Simplices and an Application to Noncooperative N Person Games. *Mathematics of Operations Research*, 7(1):1–13, 1982.

[Mil96] Igal Milchtaich. Congestion Games with Player-Specific Payoff Functions. *Games and Economic Behavior*, 13:111–124, 1996.

[MM96] Richard D. McKelvey and Andrew McLennan. Computation of Equilibria in Finite Games. *Handbook of Computational Economics*, 1:87–142, 1996.

[Mye99] Roger B. Myerson. Nash Equilibrium and the History of Economic Theory. *Journal of Economic Literature*, 37(3):1067–1082, 1999.

[Nas51] John F. Nash. Non-Cooperative Games. *Annals of Mathematics*, 54(2):286–295, 1951.

[Orl13] James B. Orlin. Max Flows in $O(nm)$ Time or Better. In *the 45th Annual ACM Symposium on Theory of Computing (STOC)*, 2013.

[PR05] Christos H. Papadimitriou and Tim Roughgarden. Computing Equilibria in Multi-Player Games. In *the 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 82–91, 2005.

[Ros73] Robert W. Rosenthal. A Class of Games Possessing Pure-Strategy Nash Equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[Rou05] Tim Roughgarden. *Selfish Routing and the Price of Anarchy*. MIT press, 2005.

[Sta69] Ross M. Starr. Quasi-Equilibria in Markets with Non-Convex Preferences. *Econometrica*, 37(1):25–38, 1969.