The Complexity of Nash Equilibria as Revealed by Data

Siddharth Barman* Umang Bhaskar † Federico Echenique ‡ Adam Wierman §

Abstract

In this paper we initiate the study of the computational complexity of Nash equilibria in bimatrix games that are specified via data. This direction is motivated by an attempt to connect the emerging work on the computational complexity of Nash equilibria with the perspective of revealed preference theory, where inputs are data about observed behavior, rather than explicit payoffs. Our results draw such connections for large classes of data sets, and provide a formal basis for studying these connections more generally. In particular, we derive three structural conditions that are sufficient to ensure that a data set is both consistent with Nash equilibria and that the observed equilibria could have been computed efficiently: (i) small *dimensionality* of the observed strategies, (ii) small *support size* of the observed strategies, and (iii) small *chromatic number* of the data set. Key to these results is a connection between data sets and the player rank of a game, defined to be the minimum rank of the payoff matrices of the players. We complement our results by constructing data sets that require rationalizing games to have high player rank, which suggests that computational constraints may be important empirically as well.

1 Introduction

At the core of the intersection of computer science and economics lie questions about the computational complexity of finding Nash equilibria, Walrasian equilibria, and other economic solution concepts. Over the last decade, significant progress has been made on these topics; and the basic message that has emerged is that finding equilibria of economic models is, in general, computationally hard. For example, computing Nash equilibria is PPAD-complete even for 2-player games [5] and, similarly, computing Walrasian equilibria in general is PPAD-hard [4, 19].

However, economists and computer scientists approach games and game theory in different ways. The results mentioned above all take the model as given. For example, in the case of Nash equilibrium computation, the assumption is that the payoff matrices are explicitly specified. In contrast, in economics, game theory is typically viewed as a mostly positive science, where *observed*

*Center for the Mathematics of Information, California Institute of Technology. barman@caltech.edu
†Center for the Mathematics of Information, California Institute of Technology. umang@caltech.edu
‡Humanities and Social Sciences, California Institute of Technology. fede@hss.caltech.edu
§Computing and Mathematical Sciences, California Institute of Technology. adamw@caltech.edu
This research was supported by NSF grants CNS-0846025 and CCF-1101470.

phenomena are modeled using games. Thus, the parameters of the model (e.g., payoffs) are not specified precisely, rather they must be inferred from *data*.

A typical situation is one where an economist observes the behavior of a group of strategic agents in a game-theoretic setting, and would like to infer a model (e.g., payoff matrices) that explains the data. The economist can observe the strategies taken by the agents, but cannot typically observe payoffs. Indeed, even if some quantifiable outcome of the players' interaction is observable, the payoffs themselves may depend arbitrarily on such a quantity, e.g., a reward of \$100 in a game does not imply a payoff of \$100. Economists view payoffs as purely subjective quantities, and therefore intrinsically unobservable.

The situation above is classically formalized via *revealed preference theory*, which was pioneered by [13] and has a long tradition in economics (see, e.g., [2, 15–18]). Revealed preference theory seeks to understand when data about observed behavior of agents fits within a model (e.g., when is observed behavior consistent with equilibria?), and what particular form the model that explains the data must take. Since we typically observe the behavior of agents and not their payoffs, the revealed preference approach is unavoidable.

However, revealed preference theory typically ignores issues of computational complexity. This presents a problem. Even though an economic explanation need not include an explicit algorithm for calculating a Nash equilibria, the explanation implicitly assumes that the agents can efficiently find the equilibrium since the agents were observed playing the equilibrium. Thus, the explanation is incomplete if equilibria cannot be efficiently computed for the payoff matrices used to explain the observations.

The above discussion highlights the importance of connecting the results that have emerged in the computer science community about the computational complexity of finding equilibria to the empirical perspective on economic models in the revealed preference community. This paper seeks to accomplish this in the context of bimatrix games. In particular, *the goal of this paper is to understand the computational complexity of bimatrix games that are specified via data.*

This goal is very related to, and builds on, the literature that seeks to understand the computational complexity of Nash equilibria. The distinction is in how the game is specified. In the current paper, the game is not specified directly via the payoff matrices, rather it is specified indirectly through empirical observations of behavior. Thus, the payoff matrices must be inferred from the set of observations, i.e., the data. However, modulo this difference, the questions of interest parallel those classically studied, e.g., in lieu of considering what classes of payoff matrices are computationally tractable, the following question is of interest in this context: what classes of inputs (data) can be explained via payoff matrices for which equilibria be computed efficiently?

1.1 Bimatrix Games, Observations, and Player Rank

Bimatrix games are two-player games where the payoffs of the two players, called the row player and the column player, are specified by matrices A and B. Throughout, we assume that both players have n strategies available to them, and hence A and B are $n \times n$ matrices, with A_{ij} and B_{ij} being the payoff of the row player and column player respectively, when the row player plays pure strategy i and the column player plays pure strategy j.

Following the revealed preference approach, we are not explicitly given the payoff matrices A and B. Instead, we assume that, on multiple occasions, two players play a fixed bimatrix game that is known to the players but unknown to us. On each occasion, each player plays a mixed strategy, given by a probability distribution over its pure strategies. The observed data, and hence the input, then consists of these pairs of mixed strategies played on each occasion. We call such an input a data set, and the pair of mixed strategies played on each occasion an observation.

Given this setting, the first-order question is about the *rationalizability* of the data. Specifically, do there exist payoff matrices A and B so that every observation in the data set is a strict mixed Nash equilibrium[1] in the inferred bimatrix game? Such data sets are called *rationalizable*, and are *rationalized* by the payoff matrices A, B. This is the fundamental problem in revealed preference theory, and the existence of such rationalizing payoff matrices validates—minus computational constraints—the model of Nash equilibrium for a given data set.

The question asked in this work is about *computational complexity*: does there exist a rationalization for the data set in which players can efficiently compute the observed Nash equilibria? This is our primary focus, and the answer to this question has implications for both revealed preference theory and computational complexity. The results in this paper derive three structural conditions that are each sufficient to ensure that a data set is both consistent with Nash equilibria and that the observed equilibria could have been computed efficiently. To do this, we focus on rationalizing data via games having low *player rank*; such games are known to have efficient algorithms that can compute all equilibria.

The player rank of a game (A, B) is defined as the minimum of the rank of A and B. Our focus on player rank stems from a number of important properties. In particular, if a game has player rank $k \in \{1, 2, \ldots, n\}$, then an equilibrium can be computed in time $O(n^{O(k)})$ [8,10]. Thus the time taken by algorithms based on player rank increase smoothly with the player rank. Further, for a game of player rank k, all extreme equilibria can be enumerated in time $O(n^{O(k)})$ [8]. Importantly, this provides an explanation for not just some arbitrary equilibrium, but a credible rationale for the specific observations in the data set.

The above provides strong motivation for the use of player rank; however there are many other properties which ensure the existence of efficient algorithms for computing Nash equilibria. Other well-known examples of such properties are: a *game rank* of zero or one[2], where the game rank is the rank of $C := A + B$; the existence of a potential function; or the existence of a pure Nash equilibrium. None of these properties is appropriate for use in the exercise here because each is binary: either a data set possesses the particular property, or it does not; and absence of the property renders algorithms based on the property useless in computing equilibria. Further, in Appendix A, we show that for each property other than player rank, simple data sets with a small number of observations necessitate rationalizations that do not satisfy the property.

[1] We require strictness in order to avoid rationalization by trivial games, such as games in which all the entries of the payoff matrices are identical.

[2] If the game rank is 4 or larger then the computation of a Nash equilibria is PPAD-hard [11]

1.2 Contributions of This Paper

The primary contribution of this paper is to identify structural properties of data sets that admit rationalizations in which Nash equilibria can be computed efficiently. We show that large and structurally complex data sets of observations — including data sets with overlapping observed strategies, each of which may be an arbitrary distribution — can be modeled by games that have low player rank, and hence allow efficient computation of Nash equilibria. Thus, such data are consistent with the theory of Nash equilibrium under a form of "bounded rationality" defined via computational constraints.

More specifically, we identify three measures of structural complexity in data sets, and show that if a data set has a low value on any one of these measures for either player, or can be partitioned in a manner such that each partition has a low value on any one of these measures, then it has a rationalization with low player rank. The three measures we study are (i) the *dimensionality* of the observed strategies, (ii) the *support size* of the observed strategies, and (iii) the *chromatic number* of the data set. We believe these are natural and complementary measures to evaluate the structural complexity of a data set, and our contribution is to show that each of these measures individually translates to the existence of rationalizing games with low player rank.

We begin by observing that our first objective—determining rationalizability—is easily met, and can be done in polynomial time by solving a linear program (Proposition 1). In contrast, determining the existence of a rationalization with low player rank is non-linear and non-convex, and so we focus our attention on this.

1. ***Dimensionality of observed mixed strategy vectors*:** Our first result shows that *if either player has at most s linearly independent observed strategies, there is a rationalization of player rank s (Theorem 2).* The basic technique of the proof is to construct a low-rank matrix that satisfies the property that the maximum entries in each column are in certain rows. For example, consider a matrix where the maximum entry in the ith column is in row i. An obvious candidate for such a matrix is the identity matrix. Of course, the identity matrix has high rank; however, there also exist matrices of rank 2 that satisfy this property. More concretely, given a rationalization $(\hat{A}, \hat{B})$, where $\hat{A}$ and $\hat{B}$ are of high rank, we give a construction that can be used to replace matrices $\hat{A}$ and $\hat{B}$ by equivalent low-rank matrices A and B.

2. ***Support size of the observed mixed strategies*:** Our second result states that *if all observed strategies of either player have support size at most s and the data set is generic*[3]*, then there exists a rationalization with player rank at most $2s + 1$ (Theorem 3).* Note that in this case the set of observed strategies may span a high-dimensional subspace and hence the previous construction is not applicable. In order to establish the result, we first prove an interesting connection between polynomials of low degree and the rank of matrices they generate. We then show that, in the case of strategies of small support, there are low-degree polynomials that are not only maximized at points in the support of these strategies, but also

[3] A data set is *generic* if for each player, the vectors of observed strategies are linearly independent.

have equal value at these supports. We use these polynomials to generate rationalizations for these data sets. The low degree of the polynomials gives us the required bound on the rank.

3. ***Chromatic number of the data set*:** Our third result introduces a novel notion of structural complexity for the data set: the *chromatic number*. The row chromatic number of a data set is the chromatic number of a graph that contains a vertex for each observation and an edge between observations that place positive probability on the same strategy for the row player. The column chromatic number is defined similarly using intersection in the column player's strategies, and the chromatic number of a data set is the minimum of the row and column chromatic numbers.

   We prove that, *for generic data sets with chromatic number s, there exists a rationalization where the player rank is at most $2s+1$ (Theorem 4).* Data sets with low chromatic number may have high dimension and large support size, and hence neither of the previous results are useful here. Instead, we first consider the simplified case where the set of observed strategies of the row player have disjoint support and the strategies for the column player are generic. In this case, we give a rank 2 rationalization. We then show how to utilize the decomposition of the data set obtained from the chromatic number to partition the data set, obtain rank 2 rationalizations for each partition, and then combine these to obtain a single low-rank rationalization.

We further extend these results to data sets that have a high value for each of these measures, but can be partitioned into three subsets such that each partition has a low value for one of these measures (Corollary 1). Thus data sets that are structurally simple modulo a few observations have low-rank rationalizations as well. Overall, our results show that large classes of data sets admit computationally-tractable rationalizations.

Finally, we also show that the bounds we obtain on the player rank for these measures are nearly tight by giving an example of a data set with $n-1$ observations that necessitates player rank $n-1$ for any rationalization (Theorem 5). This data set additionally highlights that the hypothesis of low player rank can be refuted by data.

1.3 Related Work

The complexity of economic models that are specified via data has been studied in two two previous papers [3, 6], each focusing on a different context than the current paper.

The question was first asked in the context of consumer choice theory in [6]. In this context it was shown that a data set of n observations of a consumer choosing among d indivisible goods can always be explained by a consumer utility function that can be optimized in $O(nd)$ time. Thus, despite the fact that the consumer choice problem is NP-hard, consumer choice data is not rich enough to expose computationally hard utility functions without an exponentially large data set.

The question was also asked in the context of bimatrix games in [3]; however the focus was on pure strategy equilibria and so differs from the current paper, which focuses on mixed strategy equilibria. In [3], the message is different than in [6]. In this case, the core problem is not computationally hard, and so the focus is on the structural complexity of the game, as formalized via game rank. The paper shows that there exists a measure of richness for the data, formalized via the crossing number, which connects to game rank.

The purpose of the current paper is to develop a connection between the structural complexity of the data and the computational complexity of the game in the context of mixed strategy Nash equilibria.

This problem is significantly more difficult than the ones solved in [6] and [3]. To highlight this, note that, there are no results in the revealed preference literature characterizing which data sets can be explained as mixed Nash equilibria. So, our results represent a contribution to pure economic theory, as well as to algorithmic game theory. In particular, note that most of the economic literature on revealed preference theory deals with single-person decision problems, e.g., see the survey of Varian [18]. There is much less known about revealed preference in game theory. The approach was first formulated in game theory by [14], and then extended in, among others, [9] and [7]. However, this literature deals exclusively with pure Nash equilibria.

2 Preliminaries

Bimatrix Games. Bimatrix games are two player games in normal form. Such games are specified by a pair of matrices (A, B) of size $n \times n$, which are termed the payoff matrices for the players. The first player, also called the row player, has payoff matrix A, and the second player, or the column player, has payoff matrix B. The strategy set for each player is $[n] = \{1, 2, \ldots, n\}$, and, if the row player plays strategy i and column player plays strategy j, then the payoffs of the two players are A_{ij} and B_{ij} respectively. The player rank of game (A, B) is defined to be $\min\{\text{rank}(A), \text{rank}(B)\}$.

Let Δ^n be the set of probability distributions over the set of pure strategies $[n]$. For $x \in \Delta^n$, we define $\text{Supp}(x) := \{i : x_i > 0\}$. Further, $e_i \in \mathbb{R}^n$ is the vector with 1 in the ith coordinate and 0's elsewhere. The players can randomize over their strategies by selecting any probability distribution in Δ^n, called a mixed strategy. When the row and column players play mixed strategies x and y respectively, the expected payoff of the row player is $x^T Ay$ and the expected payoff of the column player is $x^T By$.

Given a mixed strategy $y \in \Delta^n$ for the column player, the best-response set of the row player, β_r, is defined as $\beta_r(y) := \{i \in [n] \mid e_i^T Ay \geq e_k^T Ay \quad \forall k \in [n]\}$. Similarly, the best-response set, β_c, of the column player (against mixed strategy $x \in \Delta^n$ of the row player) is defined as $\beta_c(x) := \{j \in [n] \mid x^T Be_j \geq x^T Be_k \quad \forall k \in [n]\}$. The best response sets β_r and β_c are defined with respect to the payoff matrices A and B. When we want to emphasize this fact we use superscripts: β_r^A and β_c^B.

Definition 1 (Nash Equilibrium). *A pair of mixed strategies (x, y), $x, y \in \Delta^n$, is a Nash equilibrium if and only if:*

$$\begin{aligned} x^T Ay &\geq e_i^T Ay \qquad \forall i \in [n] \text{ and} \\ x^T By &\geq x^T Be_j \qquad \forall j \in [n]. \end{aligned}$$

The Nash equilibrium is strict *if additionally the support and the best-response sets are equal, i.e.,* *Supp*$(x) = \beta_r(y)$ *and Supp*$(y) = \beta_c(x)$. *Thus, for strict Nash equilibrium* $x^T Ay > e_i^T Ay$ *for all* $i \notin$ *Supp*(x) *and* $x^T By > x^T Be_j$ *for all* $j \notin$ *Supp*(y).

Observations and Data. A key component in the revealed preference setting is the specification of the observed behavior of the agents, i.e., the data. There is no prior work on rationalizing mixed behavior of agents, as we study here, and so we adapt and extend a specification from previous work in the context of pure Nash equilibria [3,7,9]. We extend the standard setup from the context of pure Nash equilibria to mixed Nash equilibria in the natural way. The one key difference is that in our formalization the strategy space for each observation is the same, whereas in the papers on pure strategy observations the strategy space is allowed to differ across observations. This is an interesting generalization to consider, but one that we leave to future work.

The setting we consider is the following. We assume that the same game is played on multiple occasions by two players and the data set consists of observations of the mixed strategies played on each occasion. Specifically, we define a data set $\mathcal{D}$ of size m as a collection of mixed-strategy pairs, $\mathcal{D} = \{(x_k, y_k) \in \Delta^n \times \Delta^n \mid k \in [m]\}$. Throughout we consider only finite data sets and we refer to mixed-strategy pairs (x_k, y_k) as observations. Crucially, no information about the payoff matrices is known, i.e., the payoffs of the players are not observed.

We denote the set of observed mixed strategies of the row and column player in a data set $\mathcal{D}$ by $\mathcal{O}_r(\mathcal{D})$ and $\mathcal{O}_c(\mathcal{D})$ respectively: $\mathcal{O}_r(\mathcal{D}) := \{x \in \Delta^n \mid \exists y \in \Delta^n \text{ such that } (x, y) \in \mathcal{D}\}$ and $\mathcal{O}_c(\mathcal{D}) := \{y \in \Delta^n \mid \exists x \in \Delta^n \text{ such that } (x, y) \in \mathcal{D}\}$. When there is a single data set under consideration, for ease of notation, we simply refer to these sets as $\mathcal{O}_r$ and $\mathcal{O}_c$.

Rationalization. Given data as described above, the first order goal of revealed preference theory is to understand whether the data is rationalizable, i.e., whether the data can be explained as Nash equilibria resulting from some bimatrix game. More formally, we have the following definition.

Definition 2 (Rationalizable Data)**.** *A data set* $\mathcal{D} = \{(x_k, y_k)\}_k$ *is said to be* rationalizable *if there exist payoff matrices* A *and* B *such that for all* k, (x_k, y_k) *is a strict Nash equilibrium in the game* (A, B).

Note that strictness is required in the above definition in order to avoid rationalization by trivial games, such as games in which all the entries of the matrices are identical.

It is also useful to talk about rationalization with respect to the row and column players. We say that a matrix A rationalizes the row player's strategies (present in the data set) if A satisfies the strict Nash requirement of the row player for all the observations in the data set. In other words, with payoffs from A, the following equality holds for all $(x, y) \in \mathcal{D}$: $\text{Supp}(x) = \beta_r^A(y)$. Similarly, B is said to rationalize the column player's strategies if $\text{Supp}(y) = \beta_c^B(x)$ for all $(x, y) \in \mathcal{D}$.

Our first result is a straightforward observation that determining if a data set is rationalizable can be done efficiently by solving a linear program. The variables in the linear program are the entries of the payoff matrices A and B, and the payoffs π_k, π'_k obtained by the players for each observation (x_k, y_k) in the data set.

Proposition 1. *A data set $\mathcal{D} = \{(x_k, y_k) \in \Delta^n \times \Delta^n \mid 1 \leq k \leq m\}$ is rationalizable if an only if the optimal value of the linear program below is strictly greater than zero.*

$$
\begin{array}{lll}
\text{maximize} & \delta & \\
\text{subject to} & (Ay_k)_i = \pi_k & \forall k, \forall i \in \mathit{Supp}(x_k) \\
& (Ay_k)_j \leq \pi_k - \delta & \forall k, \forall j \notin \mathit{Supp}(x_k) \\
& (x_k^T B)_i = \pi'_k & \forall k, \forall i \in \mathit{Supp}(y_k) \\
& (x_k^T B)_j \leq \pi'_k - \delta & \forall k, \forall j \notin \mathit{Supp}(y_k) \\
& 0 \leq A_{i,j}, B_{i,j} \leq 1 & \forall i, j \in [n] \\
& \delta \geq 0. &
\end{array}
\tag{LP}
$$

Proof. If the optimal value of the linear program is strictly greater than zero then we have matrices A and B (corresponding to an optimal solution) under which the observations (x_k, y_k) are strict Nash equilibrium. Hence $\mathcal{D}$ is rationalizable.

On the other hand, if $\mathcal{D}$ is rationalizable, say via game $(\hat{A}, \hat{B})$, then we can scale the entries of $\hat{A}$ and $\hat{B}$ (by a large enough positive constant) and add a fixed number to all of them to obtain a (normalized) game (A, B) with entries between 0 and 1. Since such an affine transformation preserves the set of strict Nash equilibria, $\mathcal{D}$ is also rationalized by (A, B). In other words, (A, B) gives us a feasible solution to the linear program with objective function value strictly greater than zero. This establishes the claim. □

3 The Empirical Implications of Player Rank

This section includes the main results of the paper, which identify structural properties of data sets that guarantee the existence of rationalizations for which Nash equilibria can be computed efficiently.

While one could perhaps approach this goal by constructing rationalizations for which very specific and ad-hoc algorithms are computationally efficient, such rationalizations would not represent convincing explanations for how the observations arose. In particular, for the explanation to be convincing, the algorithm for equilibria computation must work independently of knowledge of the observations in the data: after all, the agents did not have such information when playing the game. Thus, a more convincing approach is to construct rationalizations for which known algorithms compute Nash equilibria efficiently. We focus on *player rank* as a property of a rationalization that guarantees efficient computation of equilibria.

The key feature of player rank that makes it appealing for the revealed preference exercise in this paper is summarized in the following theorem, which follows from results in [8].

Theorem 1 ([8]). *If the player rank of a bimatrix game is k then all extreme Nash equilibria can be computed in time $O(n^{O(k)})$.*

This theorem provides a polynomial-time algorithm for computing a Nash equilibria when the player rank k is constant, and shows that the bound on computation time increases smoothly with player rank beyond constant k. Further, since the algorithm computes all extreme equilibria, it at can be used to compute the observations in the data set rather than simply some arbitrary equilibria. This fact is crucial to the exercise since the goal is to explain the specific observations in the data set.

The above highlights why player rank is an appealing measure to target when constructing rationalizations, and the results in this section validate this choice by showing that rich classes of data sets can be rationalized by games with small player rank. In particular, we identify three measures of structural complexity in data sets, and show that if a data set has a low value on any one of these measures for either player, or can be partitioned into three subsets such that each partition has a low value on any one of these measures, then it has a rationalization with low player rank.

3.1 Observations from a Low Dimensional Subspace

The first structural property we connect to player rank is the dimensionality of the observed strategies in the data set. Given a finite set $S \subset \mathbb{R}^n$ of m vectors $s_1, \ldots, s_m$, write $\dim(S)$ to denote the maximum number of linearly independent vectors in S. Observed strategies that form a low dimensional subspace are natural candidates for low player rank rationalizations and, the following theorem shows that — independent of the size of the data set — if the observations form a low dimensional subspace then they can be rationalized by a game of low player rank.

Theorem 2. *If a data set $\mathcal{D}$ is rationalizable then it can be rationalized by a game of player rank at most* $\min\{dim(\mathcal{O}_r), dim(\mathcal{O}_c)\}$.

An immediate consequence of the above theorem is that, if a data set $\mathcal{D}$ is rationalizable, then it can be rationalized by a game of player rank at most $|\mathcal{D}|$. Additionally, later, in Theorem 5, we prove a lower bound that highlights that this result is tight.

Importantly, Theorem 2 has rationalizability as one of its hypotheses, and thus implies that, for data with low-dimensional strategies, the computational constraints have no added empirical or observational content. In other words, any low dimensional data set that is rationalizable without computational constraints is also rationalizable with them.

To prove Theorem 2, a key technical piece is the following lemma about reconstructing the product of an arbitrary matrix and a low-rank matrix.

Lemma 1. *Suppose $Y \in \mathbb{R}^{n\times m}$ is a matrix of rank t. Then for every matrix $\hat{A} \in \mathbb{R}^{n\times n}$, there exists a matrix $A \in \mathbb{R}^{n\times n}$ of rank at most t that satisfies $AY = \hat{A}Y$.*

Proof. Let $\{y_1, y_2, \ldots, y_t\}$ be a set of linearly independent columns in Y, and define $\hat{Y}$ as the $n \times t$ matrix $[y_1\ y_2\ \ldots\ y_t]$. Since each of the columns of $\hat{Y}$ are linearly independent, we can obtain a matrix Γ of size $n \times t$ that satisfies $\Gamma^T \hat{Y} = I_t$. Thus if we denote the ith column of Γ by γ_i, then for all $j \in [t]$,

$$\begin{aligned} \gamma_j^T y_j &= 1 \qquad \text{and} \\ \gamma_j^T y_k &= 0 \qquad \forall k \in [t] \setminus \{j\}\,. \end{aligned}$$

We define matrix A as the following sum of t rank-1 outer products:

$$A = \sum_{j=1}^{t} \hat{A} y_j \gamma_j^T .$$

By construction, the rank of A is at most t. Further, for all y_j for $j \in [t]$:

$$A\, y_j = \sum_{k=1}^{t} \hat{A} y_k \gamma_k^T y_j \;=\; \hat{A}\, y_j . \tag{1}$$

Since any column y of matrix Y can be expressed as a linear combination of the columns of $\hat{Y}$, we can write $y = \sum_{j=1}^{t} \lambda_j y_j$. Then $\hat{A}\, y_i$ is given by

$$\begin{aligned} \hat{A}\, y = \hat{A} \left(\sum_{j=1}^{t} \lambda_j y_j \right) &= \sum_{j=1}^{t} \lambda_j \hat{A} y_j = \sum_{j=1}^{t} \lambda_j A y_j \\ = A \left(\sum_{j=1}^{t} \lambda_j y_j \right) &= Ay \,. \end{aligned}$$

where the third equality is obtained from (1). Overall we have the desired claim, $AY = A\hat{Y}$. □

Using this lemma, Theorem 2 can be established as follows.

Proof of Theorem 2. Consider a game $(\hat{A}, \hat{B})$ that rationalizes the data set $\mathcal{D} = \{(x_k, y_k)\}_{k=1}^m$. Write X (Y) to denote the matrix whose kth row (column) is equal to x_k (y_k), for all $k \in [m]$. Note that $\text{rank}(X) = \dim(\mathcal{O}_r)$ and $\text{rank}(Y) = \dim(\mathcal{O}_c)$. By Lemma 1, there exist matrices A, B so that $\text{rank}(A) \leq \text{rank}(Y)$, $\text{rank}(B) \leq \text{rank}(X)$, and $AY = \hat{A}Y$, $XB = X\hat{B}$. Then (A, B) is the rationalization required by the theorem. We have already shown that A and B are of the required rank. To see that (A, B) rationalize $\mathcal{D}$, note that since $AY = \hat{A}Y$, for all $(x, y) \in \mathcal{D}$ we have $\beta_r^A(y) = \beta_r^{\hat{A}}(y)$. Hence $\text{Supp}(x) = \beta_r^A(y)$. Similarly, since $XB = X\hat{B}$, $\text{Supp}(y) = \beta_c^B(x)$. Hence (x, y) is a strict Nash equilibrium in (A, B). □

3.2 Observations with Small Support Size

The second structural property of the data set we consider is the support size of the observations. In spirit, the following theorem complements the result of Lipton et al. [10] wherein they establish that if the rank of both the payoff matrices is low then the game contains a small-support equilibrium. The following result highlights that there is a connection in the other direction as well.

Theorem 3. *Let $\mathcal{D} = \{(x_k, y_k) \in \Delta^n \times \Delta^n \mid 1 \leq k \leq m\}$ be a data set in which $|Supp(x_k)| \leq s$ for all $k \in [m]$ or $|Supp(y_k)| \leq s$ for all $k \in [m]$. If the observed strategies $\mathcal{O}_r(\mathcal{D})$ and $\mathcal{O}_c(\mathcal{D})$ are generic then $\mathcal{D}$ can be rationalized by a game with player rank $\leq 2s + 1$.*

Note that, later, Theorem 5 highlights that the bound in Theorem 3 is tight to within a factor of 2.

Our proof of Theorem 3 uses a construction based on polynomials, and the following lemma is a key technical piece in the argument. The lemma states that if for all $j \in [m]$, the jth column of an $n \times m$ matrix M is obtained by evaluating a degree d polynomial p_j at $1, 2, \ldots n$ (i.e., the jth column of M is equal to $(p_j(1), p_j(2), \ldots, p_j(n))^T$), then the rank of M is at most $d+1$.

Lemma 2. *Let $p_1, p_2, \ldots, p_m$ be m univariate polynomials over $\mathbb{R}$, and suppose that the degree of each of them is at most d. If the (i,j)th entry of an $n \times m$ matrix M is equal to $p_j(i)$, for all $i \in [n]$ and $j \in [m]$, then the rank of M is at most $d+1$.*

Proof. Write $p_j(x) = a_d^{(j)} x^d + a_{d-1}^{(j)} x^{d-1} + \ldots + a_1^{(j)} x + a_0$ for all $j \in [m]$. M can be expressed as a sum of $d+1$ outer products:

$$M = \begin{pmatrix} 1^d \\ 2^d \\ \vdots \\ n^d \end{pmatrix} \begin{pmatrix} a_d^{(1)} & a_d^{(2)} & \cdots & a_d^{(n)} \end{pmatrix} + \begin{pmatrix} 1^{d-1} \\ 2^{d-1} \\ \vdots \\ n^{d-1} \end{pmatrix} \begin{pmatrix} a_{d-1}^{(1)} & a_{d-1}^{(2)} & \cdots & a_{d-1}^{(n)} \end{pmatrix} + \ldots$$

$$\ldots + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} a_0^{(1)} & a_0^{(2)} & \cdots & a_0^{(n)} \end{pmatrix}.$$

Note that the rank of an outer product is one and the rank of the sum of two matrices satisfies $\text{rank}(X+Y) \le \text{rank}(X) + \text{rank}(Y)$. Hence the rank of M is no more than $d+1$. □

Using this lemma, Theorem 3 can be established as follows.

Proof of Theorem 3. We prove the claim for the case in which the mixed strategies of the row player in the data set $\mathcal{D} = \{(x_k, y_k) \in \Delta^n \times \Delta^n \mid 1 \le k \le m\}$ are of support size at most s. A construction similar to the one presented below takes care of the alternate case wherein $|\text{Supp}(y_k)| \le s$ for each $k \in [m]$.

We consider a polynomial p_k that satisfies $\arg\max_x p_k(x) = \text{Supp}(x_k)$ and has degree $2|\text{Supp}(x_k)|$. In particular,

$$p_k(x) := - \prod_{i \in \text{Supp}(x_k)} (x - i)^2.$$

Say $\text{Supp}(x_k) = \{i_1, i_2, \ldots, i_s\} \subset [n]$, then the polynomial p_k vanishes exactly at $i_1, i_2, \ldots, i_s$ and is negative elsewhere. Hence, $\text{Supp}(x_k)$ is the set of points at which p_k attains its maximum value. In addition, the degree of p_k is $2|\text{Supp}(x_k)|$.

Consider the $n \times m$ matrix P in which the kth column is equal to $(p_k(1), p_k(2), \ldots, p_k(n))^T$. By construction, for all $k \in [m]$, degree of the polynomial p_k is no more than $2s$. Therefore, Lemma 2 implies that the rank of P is at most $2s+1$. Moreover, the set of the largest components of the kth column of P (i.e., $\arg\max_i P_{i,k}$) is exactly equal to $\mathrm{Supp}(x_k)$. Since,

$$\arg\max_{i\in[n]} P_{i,k} = \arg\max_{i\in[n]} p_k(i) \;\; = \mathrm{Supp}(x_k).$$

Recall that the mixed strategies in $\mathcal{O}_c(\mathcal{D})$ are generic. Therefore, we can find an $m \times n$ matrix V that satisfies the following equality for all $y_k \in \mathcal{O}_c(\mathcal{D})$: $Vy_k = e_k^{(m)}$. Here $e_k^{(m)}$ is the m-dimensional vector with a 1 in the kth coordinate and 0s elsewhere.

Set the payoff matrix of the row player $A = PV$. Rank of the product of two matrices satisfies: $\mathrm{rank}(XY) \leq \min\{\mathrm{rank}(X), \mathrm{rank}(Y)\}$. Hence $\mathrm{rank}(A) \leq 2s+1$.

For all $(x_k, y_k) \in \mathcal{D}$, we have $Ay_k = Pe_k^{(m)} = (p_k(1), p_k(2), \ldots, p_k(n))^T$. Hence, the set of the largest components of the vector Ay_k is equal to $\mathrm{Supp}(x_k)$. Overall, under the payoff matrix A, we have $\beta_r(y_k) = \mathrm{Supp}(x_k)$, for all $(x_k, y_k) \in \mathcal{D}$. That is, A rationalizes the mixed strategies of the row player. □

3.3 Observations with Low Chromatic Number

The third, and final, structural property of data sets that we consider is the chromatic number. Intuitively, the chromatic number quantifies the degree of intersection between the observed mixed strategies, and hence it is a relevant measure of the structural complexity of data.

For a data set $\mathcal{D}$, we define the row chromatic number $\kappa_r(\mathcal{D})$ and the column chromatic number $\kappa_c(\mathcal{D})$ as the chromatic numbers of graphs G_r and G_c, defined as follows. For the row chromatic number, $\kappa_r(\mathcal{D})$, construct graph G_r with a vertex corresponding to each observation in $\mathcal{O}_r$. For distinct observations (x, y) and (x', y') in $\mathcal{D}$, if $\mathrm{Supp}(x) \cap \mathrm{Supp}(x') \neq \emptyset$ then the graph G_r has an edge between the corresponding vertices. Then set $\kappa_r(\mathcal{D}) = \chi(G_r)$, i.e., the chromatic number of graph G_r. The column chromatic number is defined similarly using intersections $\mathrm{Supp}(y) \cap \mathrm{Supp}(y')$. The chromatic number of the data set, $\kappa(\mathcal{D})$, is defined to be the minimum of $\kappa_r(\mathcal{D})$ and $\kappa_c(\mathcal{D})$.

Theorem 4. *Let $\mathcal{D}$ be a data set with chromatic number equal to $\kappa(\mathcal{D})$. If the observed mixed-strategy sets $\mathcal{O}_r(\mathcal{D})$ and $\mathcal{O}_c(\mathcal{D})$ are generic then $\mathcal{D}$ can be rationalized by a game of player rank at most $2\kappa(\mathcal{D})$.*

Note that, later, Theorem 5 highlights that the bound in Theorem 3 is tight to within a factor of 2.

Importantly, like the case of support size, the bound on the player rank in Theorem 4 is not exactly dependent on the size of the data set, but rather only on the "richness" of the observations in terms of the structure of the underlying graph.

Of course, in general the chromatic number of a graph is hard to compute. However, an easy upper bound is the maximum degree of any vertex in G_r and G_c plus one, which then can be interpreted as follows: $\kappa_r(\mathcal{D}) \leq \max_{(x,y)\in\mathcal{D}} |\{(x', y') \in \mathcal{D} : \mathrm{Supp}(x) \cap \mathrm{Supp}(x') \neq \phi\}|$ and $\kappa_c(\mathcal{D}) \leq \max_{(x,y)\in\mathcal{D}} |\{(x', y') \in \mathcal{D} : \mathrm{Supp}(y) \cap \mathrm{Supp}(y') \neq \phi\}|$. Though these bounds provide intuition, it is obvious the chromatic numbers can be much less than these upper bounds, e.g., if the graph G_r is a star.

The proof of Theorem 4 starts by focusing on data sets where the row chromatic number is one, i.e., the supports of all the observations for the row player are pairwise disjoint. In this case we show, in the following Lemma, that the data set can be rationalized by a game with row player rank 2. We first prove the initial lemma and then Theorem 4.

Lemma 3. *Let $\mathcal{D}'$ be a data set with row chromatic number $\kappa_r(\mathcal{D}') = 1$. If the set of observations $\mathcal{O}_c(\mathcal{D}')$ is generic then there exists a rank 2 matrix A' that rationalizes the row player's strategies in $\mathcal{D}'$.*

Proof. Say $\mathcal{D}'$ consists of m observations $(x_1, y_1), \dots, (x_m, y_m)$. Since the support sets of the strategies of the row player are disjoint, any (pure) strategy $i \in [n]$ of the row player is in at most one such support set. For pure strategy $i \in [n]$, let $\sigma(i) := k$ such that $i \in \text{Supp}(x_k)$. If row i is not in the support of any strategy, let $\sigma(i) = m + 1$.

We construct three vectors u, w and f in $\mathbb{R}^n$. Values are assigned to the components of the vectors u and w directly: for $i \in [n]$, entry $u_i = -\sigma(i)^2$, and $w_i = \sigma(i)$. We define vector f so that for all $y_k \in \mathcal{O}_c(\mathcal{D}')$, $f^T y_k = k$. Since the set of observations $\mathcal{O}_c(\mathcal{D}')$ are generic, such a vector f exists and can be obtained. Matrix A' is then defined as follows: $A' := u\mathbf{1}_n^T + 2wf^T$, where $\mathbf{1}_n$ is the n-column vector consisting of all 1's.

Since A' is the sum of two outer products, it has rank 2. We show that with payoffs from A' the strict Nash requirement for the row player, $\text{Supp}(x_k) = \beta_r(y_k)$, is satisfied for all $(x_k, y_k) \in \mathcal{D}'$. Hence, we get the desired lemma.

For any $y_k \in \mathcal{O}_c(\mathcal{D}')$, by our construction, $A'y_k = u + 2wf^T y_k = u + 2kw$. For a fixed k, consider the jth component of the vector $A'y_k$, which by the construction is $-\sigma(j)^2 + 2k\sigma(j)$. Note that this expression is maximized when $\sigma(j) = k$ and is strictly less for other values of $\sigma(j)$. Further, by construction, if $\sigma(j) = k$ then $j \in \text{Supp}(x_k)$. Thus, the maximum components of $A'y_k$ are exactly those that correspond to the support of x_k, and hence under payoff matrix A', $\beta_r(y_k) = \text{Supp}(x_k)$. □

We show now how this lemma can then be used to construct games with low player rank for data sets with larger chromatic numbers.

Proof of Theorem 4. We constructively show that there exists a matrix A that rationalizes the mixed strategies of the row player in $\mathcal{D}$ and has rank no more than $2\kappa_r(\mathcal{D})$. Similarly, we can construct a payoff matrix B of rank at most $2\kappa_c(\mathcal{D})$ for the column player. This establishes the existence of the game (A, B) that rationalizes $\mathcal{D}$ with the required player rank.

Let t be the row chromatic number of the data set, i.e., $t = \kappa_r(\mathcal{D})$, and let $\chi(G_r)$ be the graph coloring that defines the row chromatic number. We partition the observations in data set $\mathcal{D}$ into t sets $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_t$ according to the color assigned to the vertex corresponding to each observation, so that observations with the same color are in the same partition. Note that by this technique, for all observations within the same partition, the supports of the observations for the row player are disjoint.

By Lemma 3, for each data set $\mathcal{D}_j$, we can obtain a rank-2 matrix A_j so that $\beta_r(y_i) = \text{Supp}(x_i)$ with A_j as the payoff matrix. In order to combine these matrices, for $j \in [t]$, we define matrix V_j as satisfying the following property:

$$V_j y = y \text{ for } y \in \mathcal{O}_c(\mathcal{D}_j), \text{ and}$$
$$V_j y = \mathbf{0} \text{ otherwise.}$$

Since the set of strategies of the players are generic, we can obtain such matrices. We then define the payoff matrix A as $A = \sum_{j=1}^{t} A_j V_j$. Since each A_j is of rank 2, matrix A is of rank $2t = 2\kappa_r(\mathcal{D}')$. To see that A rationalizes the data set $\mathcal{D}$, note that for $j \in [t]$ and $(x, y) \in \mathcal{D}_k$, $Ay = \sum_{j=1}^{t} A_j V_j y = A_k y$, and by construction of A_k, $\beta_r(y) = \text{Supp}(x)$. □

3.4 A Unifying Result

The previous sections have identified three structural properties for data sets that ensure the existence of low player rank rationalizations. In this section, we present a unifying result that extends the previous three theorems to provide a more robust low-rank construction and, in particular, shows that addition of a small number of observations to a data set does not have a big impact on the player rank necessary to rationalize the data. Specifically, we establish low-rank rationalizations for data sets that can be partitioned into three sets which are structurally simple in terms of dimensionality, support size, and chromatic number, respectively. For example, say we have a data set $\mathcal{D}$ in which all but t observed mixed strategies are of support size s, then it can be partitioned into a set that has support bounded by s and a set that has dimensionality bounded by t (and an empty set that has chromatic number zero). The following corollary then shows how to construct a rationalization for $\mathcal{D}$ of player rank at most $(2s + 1) + t$.[4]

To obtain such a generalization we introduce the notion of the *composite number* of a data set, which considers a 3-partition of the data set and combines the dimensionality of the first partition, the support size of the second partition, and the chromatic number of the third partition.

Definition 3 (Composite number)**.** *The row composite number $\sigma_r(\mathcal{D})$ of a data set $\mathcal{D}$ is defined to be the smallest number for which there exists a 3-partition of $\mathcal{D}$, $\{\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3\}$, that satisfies $\dim(\mathcal{O}_c(\mathcal{D}_1)) + \max_{(x,y)\in\mathcal{D}_2} |Supp(x)| + \kappa_r(\mathcal{D}_3) = \sigma_r(\mathcal{D})$. The column composite number $\sigma_c(\mathcal{D})$ is defined similarly. The composite number of a data set, $\sigma(\mathcal{D})$, is the minimum of the row and column composite number: $\sigma(\mathcal{D}) := \min\{\sigma_r(\mathcal{D}), \sigma_c(\mathcal{D})\}$.*

Corollary 1. *Let $\mathcal{D}$ be a data set with composite number $\sigma(\mathcal{D})$. If the observed mixed-strategy sets $\mathcal{O}_r(\mathcal{D})$ and $\mathcal{O}_c(\mathcal{D})$ are generic then $\mathcal{D}$ can be rationalized by a game of player rank at most $2\sigma(\mathcal{D}) + 1$.*

Proof. Below we show that there exists a payoff matrix A of rank at most $2\sigma_r(\mathcal{D})+1$ that rationalizes the row player's strategies in $\mathcal{D}$. A similar argument establishes the existence of a matrix B (which

[4]The precise bound from the corollary is $2(s + t) + 1$, but this can be strengthened to $(2s + 1) + t$. For ease of presentation, we present the corollary with slightly loose factors.

rationalizes the column player's strategies) of rank no more than $2\sigma_c(\mathcal{D})+1$, and hence we get the desired claim.

Say $\{\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3\}$ is a 3-partition that satisfies $\dim(\mathcal{O}_c(\mathcal{D}_1)) + \max_{(x,y)\in\mathcal{D}_2} |\text{Supp}(x)| + \kappa_r(\mathcal{D}_3) = \sigma_r(\mathcal{D})$. The constructions of Theorems 2, 3, and 4, imply that there exist matrices A_1, A_2, A_3 of rank $\dim(\mathcal{O}_c(\mathcal{D}_1))$, $2\max_{(x,y)\in\mathcal{D}_2} |\text{Supp}(x)|+1$, and $2\kappa_r(\mathcal{D}_3)$ respectively such that A_i rationalizes the row player's observations in $\mathcal{D}_i$ for all $i \in [3]$.

Since $\mathcal{O}_c(\mathcal{D})$ is generic, there exists matrix V_i for all $i \in [3]$ that satisfies the following equalities:

$$\begin{aligned} V_i y &= y \qquad \forall\, y \in \mathcal{O}_c(\mathcal{D}_i) \\ V_i y &= \mathbf{0} \qquad \forall y \in \mathcal{O}_c(\mathcal{D}) \setminus \mathcal{O}_c(\mathcal{D}_i) \end{aligned}$$

Note that payoff matrix $A = \sum_{i=1}^{3} A_i V_i$ rationalizes the row players observations in $\mathcal{D}$ and is of rank at most $2\sigma_r(D)+1$. □

This result serves as another illustration of why player rank is an appealing choice for the revealed preference exercise in this paper, since it allows us to merge the constructions used in Theorems 2, 3, and 4.

4 A Lower Bound on Player Rank

The results to this point of the paper have focused on constructing rationalizing games with low player rank, thus guaranteeing the observed equilibria can be computed efficiently. It is also natural to ask if there exist data sets that *require* rationalizations to have large player rank. In the following, we show that such data sets do exist. In particular, there exists a data set that requires any rationalization to have player rank at least $n-1$. Proof of Theorem 5 appears in Appendix **??**.

Theorem 5. *Any game (A,B) that rationalizes $\mathcal{D}' = \{(u_k, u_k) \mid k \in \{2,3,\dots,n\}\}$ has player rank at least $n-1$, i.e., rank$(A) \geq n-1$ and rank$(B) \geq n-1$.*

Proof. For $2 \leq k \leq n$, write $u_k \in \Delta^n$ to denote the uniform distribution over the set $\{1,2,\dots,k\}$. Consider the following data set with $n-1$ observations, $\mathcal{D} = \{(u_k,u_k) \mid k \in \{2,3,\dots,n\}\}$.

Note that the data set $\mathcal{D}$ is rationalizable. In particular, the game obtained by setting the payoff matrices of both the players to I_n (the $n \times n$ identity matrix) rationalizes $\mathcal{D}$. The player rank of the rationalization, (I_n, I_n), is n. Below we establish that in fact the player rank of any game that rationalizes $\mathcal{D}$ is at least $n-1$.

Say (A,B) is a rationalization of $\mathcal{D}$. Let $A_{(i)}$ be the ith row of the matrix A and for $2 \leq j \leq n$, we define $n-1$ vectors v_js as follows: $v_j := A_{(1)} - A_{(j)}$. Note that v_js lie in the row space of A. We will show that v_js are linearly independent and hence get that the dimension of the row space of A is at least $n-1$. This, in turn, proves that the rank of A is at least $n-1$. Since the data set $\mathcal{D}$ is symmetric, via a similar argument, we can establish that the rank of B is no less than $n-1$. This overall establishes the stated claim that the player rank of (A,B) is at least $n-1$.

Since mixed strategy pair (u_k,u_k) is a strict Nash equilibrium in (A,B), we have $e_1^T A u_k = e_j^T A u_k$ for all $j \in \{2,\dots,k\}$ and $e_1^T A u_k > e_j^T A u_k$ for all $j \in \{k+1,\dots,n\}$. That is, $A_{(1)}^T u_k = A_{(j)}^T u_k$

for all $j \in [k]$ and $A_{(1)}^T u_k > A_{(j)}^T u_k$ for all $j \notin \{k+1, \dots, n\}$. We can rewrite these equalities and inequalities using the definition of v_js as follows:

$$\begin{aligned} v_j^T u_k &= 0 \qquad \forall j \in \{2, \dots, k\} \text{ and} \\ v_j^T u_k &> 0 \qquad \forall j \in \{k+1, \dots, n\}. \end{aligned}$$

Hence, for all $2 \leq k < n$, vector $v_2, v_3, \dots, v_{k+1}$ are linearly independent. Say for contradiction that they are linearly dependent. Then we can write v_{k+1} as a linear combination of $v_2, \dots, v_k$, i.e., $v_{k+1} = \sum_{j=2}^{k} \lambda_j v_j$. Taking inner product of both sides of this equation with u_k leads to a contradiction.

Overall, we get that the vectors $v_2, \dots, v_n$ are linearly independent and this completes the proof. □

This result is important for two reasons. First, the theorem highlights that Theorems 2, 3, and 4 are (nearly) tight. Specifically, by construction, data set $\mathcal{D}'$ satisfies: (i) the observed strategies of each player lie in a subspace of dimension $n-1$; (ii) each observed strategy has support size at most n; and (iii) the chromatic number of the data set is $n-1$. It follows immediately that the bounds in Theorem 2 are exactly tight, and the bounds in Theorems 3 and 4 are tight to within a factor of 2.

Second, the lower bound strongly suggests that adding computational constraints to the theory of Nash equilibrium has testable implications, in contrast with single-person consumer theory [6]. It is still possible that rationalizing games could be simple, but it seems unlikely. Investigating this issue further is an intriguing direction for future work.

References

[1] Bharat Adsul, Jugal Garg, Ruta Mehta, and Milind A. Sohoni. Rank-1 bimatrix games: a homeomorphism and a polynomial time algorithm. In *STOC*, 2011.

[2] Sydney N. Afriat. The construction of utility functions from expenditure data. *International Economic Review*, 8(1):67–77, feb 1967.

[3] Siddharth Barman, Umang Bhaskar, Federico Echenique, and Adam Wierman. The empirical implications of rank in bimatrix games. In *ACM Conference on Electronic Commerce*, 2013.

[4] Xi Chen, Decheng Dai, Ye Du, and Shang-Hua Teng. Settling the complexity of arrow-debreu equilibria in markets with additively separable utilities. In *FOCS*, 2009.

[5] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *J. ACM*, 56(3), 2009.

[6] Federico Echenique, Daniel Golovin, and Adam Wierman. A revealed preference approach to computational complexity in economics. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 101–110. ACM, 2011.

[7] Adám Galambos. The complexity of nash rationalizability. Technical report, Tech. rep., Lawrence University, 2009.

[8] Jugal Garg, Albert Xin Jiang, and Ruta Mehta. Bilinear games: Polynomial time algorithms for rank based subclasses. In *WINE*, 2011.

[9] Sang Mok Lee. The testable implications of zero-sum games. *Journal of Mathematical Economics*, pages 39–46, 2011.

[10] Richard J Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM conference on Electronic commerce*, pages 36–41. ACM, 2003.

[11] Ruta Mehta. Constant rank bimatrix games are PPAD-hard. Personal communication.

[12] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[13] Paul A. Samuelson. A note on the pure theory of consumer's behaviour. *Economica*, 5(17):61–71, 1938.

[14] Yves Sprumont. On the testable implications of collective choice theories. *Journal of Economic Theory*, 93:205–232, 2000.

[15] Hal R. Varian. The nonparametric approach to demand analysis. *Econometrica*, 50(4):945–974, jul 1982.

[16] Hal R. Varian. Non-parametric tests of consumer behaviour. *Review of Economic Studies*, 50(1):99–110, jan 1983.

[17] Hal R. Varian. The nonparametric approach to production analysis. *Econometrica*, 52(3):579–598, may 1984.

[18] Hal R. Varian. Revealed preference. *Samuelsonian economics and the twenty-first century*, pages 99–115, 2006.

[19] Vijay V. Vazirani and Mihalis Yannakakis. Market equilibrium under separable, piecewise-linear, concave utilities. *J. ACM*, 58(3), 2011.

A Game Rank, Potential Games, and Pure Strategy Equilibria

The goal of this paper is to understand when it is possible to rationalize data via payoff matrices for which the mixed strategies observed are efficiently computable. Given the hardness of computing equilibria in general, this requires that the rationalizations we generate must have some special property that allows for efficient computation. We do this by focusing on rationalizations with small player rank; however there are a number of other properties that could be considered. For example, a small game rank, the existence of a potential function, or the existence of a pure Nash equilibrium. In the following, we highlight that these alternatives are not well-suited for use in this paper.

A.1 Game Rank

The connection between game rank and computational efficiency has only recently begun to be understood. To this point, polynomial-time algorithms to compute Nash equilibria are known when the game rank is either zero [12] or one [1], and it has recently been shown that when the game rank is four or more computing an equilibrium is PPAD-hard [11]. Though incomplete, these results are already problematic for the use of game rank in this paper. In particular, the following result highlights that only very small data sets can be guaranteed to have game rank small enough to ensure that a computationally efficient algorithm exists, e.g., there is a data set with 9 observations that necessitates game rank of at least two.

Theorem 6. *There exists a rationalizable data set $\mathcal{D}$ with $2n+1$ observations such that any game (A, B) that rationalizes $\mathcal{D}$ has game rank at least $n-2$, i.e., rank$(A+B) \geq n-2$.*

Proof. Let $u_n \in \Delta^n$ be the uniform distribution over $[n]$ and $e_k \in \Delta^n$ be the vector with a 1 in the kth coordinate and 0's elsewhere. Write v_k to denote the uniform distribution over $[n] \setminus \{k\}$. We consider the following data set with $2n+1$ observations, $\mathcal{D} = \{(e_k, e_k) \mid 1 \leq k \leq n\} \cup \{(v_k, v_k) \mid 1 \leq k \leq n\} \cup \{(u_n, u_n)\}$. Note that $\mathcal{D}$ can be rationalized by the game (I_n, I_n), where I_n is the $n \times n$ identity matrix.

Say game (A, B) rationalizes $\mathcal{D}$. First we show that in every column of A all the off-diagonal entries are equal to each other. That is, for all $k \in [n]$ and for all $i, i' \in [n] \setminus \{k\}$ we have $A_{i,k} = A_{i',k}$. A similar result holds for the rows of matrix B.

Since (u_n, u_n) is a strict Nash equilibrium in (A, B) we have $\text{Supp}(u_n) = \beta_r(u_n)$. This implies that all the components of the vector Au_n are equal, i.e., the row sums of A are equal to each other. Formally,

$$\sum_j A_{i,j} = \sum_j A_{i',j} \quad \forall i, i' \in [n]. \tag{2}$$

Similarly, the fact that (v_k, v_k) is a strict Nash equilibrium implies $\text{Supp}(v_k) = \beta_r(v_k)$. In particular, for all $i, i' \in \text{Supp}(v_k)$ the ith and the i'th component of Av_k must be equal to each other. Since the ith component of the vector Av_k is equal to $\frac{1}{n-1} \sum_{j \neq k} A_{i,j}$ and $\text{Supp}(v_k) = [n] \setminus \{k\}$, we have the following equality for all $i, i' \in [n] \setminus \{k\}$:

$$\sum_{j \neq k} A_{i,j} = \sum_{j \neq k} A_{i',j}. \tag{3}$$

Subtracting (3) from (2) gives us $A_{i,k} = A_{i',k}$ for $i, i' \in [n] \setminus \{k\}$.

Finally, using the fact that $(e_k, e_k) \in \mathcal{D}$ we get that (k, k) is a pure and strict Nash equilibrium in (A, B) for all k. Therefore, $A_{k,k} > A_{i,k}$ for all $i \neq k$. Say the off-diagonal entries of the kth column of A are equal to α_k. We have $A_{k,k} > \alpha_k$ and matrix A has the following form:

$$A = \begin{pmatrix} A_{1,1} & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_1 & A_{2,2} & \alpha_3 & \cdots & \alpha_n \\ \alpha_1 & \alpha_2 & A_{3,3} & \cdots & \alpha_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & A_{n,n} \end{pmatrix}$$

We can write A as the sum of a diagonal matrix and an outer product.

$$A = \begin{pmatrix} A_{1,1} - \alpha_1 & & & & \\ & A_{2,2} - \alpha_2 & & \text{\Large 0} & \\ & & A_{3,3} - \alpha_3 & & \\ & \text{\Large 0} & & \ddots & \\ & & & & A_{n,n} - \alpha_n \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots \alpha_n \end{pmatrix}$$

Write D to denote the above diagonal matrix and P to denote the outer product. We have $A = D + P$. Note that all the diagonal entries of D are positive, since $A_{k,k} > \alpha_k$ for all k. Similarly, we can decompose column player's payoff matrix B into a diagonal matrix D' and an outer product P', i.e., $B = D' + P'$. Like D, the all the diagonal entries of D' are positive.

Overall, we have $A + B = D + D' + P + P'$. The rank of the sum of two matrices satisfies $\text{rank}(X + Y) \leq \text{rank}(X) + \text{rank}(Y)$. Therefore, $\text{rank}(D + D') \leq \text{rank}(A + B) + \text{rank}(-(P + P'))$. Since P and P' are outer products, $\text{rank}(-(P+P')) \leq 2$. The diagonal entries of both D and D' are positive, hence matrix $D+D'$ has full rank. This gives us the desired bound, $\text{rank}(A+B) \geq n-2$. □

A.2 Potential Games

When a game has a potential function, it is termed a *potential game*, and an appealing property of such games is that a pure strategy equilibrium is guaranteed to exist (e.g., [12]). Not surprisingly, this property is limiting for the purposes of this paper. That is, if we were to use the existence of a pure strategy equilibria as a property to yield efficient computability of an equilibrium in the rationalizing game, then we would be restricted to extremely limited data sets. To see this, note that there are very simple data sets that cannot be rationalized by a game that has a pure Nash equilibrium, and consequently cannot be rationalized by a potential game.

Theorem 7. *There exists a rationalizable data set $\mathcal{D}$ with three observations such that any game (A, B) that rationalizes $\mathcal{D}$ does not possess a pure Nash equilibrium.*

Proof. We consider a game where each player has 3 strategies, and a data set consisting of the following three observations: $((1, 0, 0), (0, 1/2, 1/2))$; $((0, 1, 0), (1/2, 0, 1/2))$, and $((0, 0, 1), (1/2, 1/2, 0))$. Thus the row player plays a different pure strategy in each observation, while the column player randomizes uniformly over two strategies. To see that any rationalization by matrices A, B does not admit a pure Nash equilibrium, suppose for a contradiction that (i, j) is in fact a pure Nash equilibrium and consider the matrix B. The data set enforces that the maximum entry in each row is not unique, and hence no entry can be a strict pure Nash equilibrium. □