Game Connectivity and Adaptive Dynamics

Tom Johnston*† Michael Savery†‡ Alex Scott‡§ Bassel Tarbush¶

3 November 2024

Abstract

We analyse the typical structure of games in terms of the connectivity properties of their best-response graphs. Our central result shows that almost every game that is 'generic' (without indifferences) and has a pure Nash equilibrium and a 'large' number of players is *connected*, meaning that every action profile that is not a pure Nash equilibrium can reach every pure Nash equilibrium via best-response paths. This has important implications for dynamics in games. In particular, we show that there are simple, uncoupled, adaptive dynamics for which period-by-period play converges almost surely to a pure Nash equilibrium in *almost every* large generic game that has one (which contrasts with the known fact that there is no such dynamic that leads almost surely to a pure Nash equilibrium in *every* generic game that has one). We build on recent results in probabilistic combinatorics for our characterisation of game connectivity.

1 Introduction

A fundamental question at the heart of the literature on learning in games and distributed systems is whether there are adaptive dynamics that are guaranteed to lead to a Nash equilibrium in every game. Examples of adaptive dynamics include better- and best-response dynamics, fictitious play (Fudenberg and Levine, 1998), adaptive play (Young, 1993), regret matching (Hart and Mas-Colell, 2000), regret testing (Foster and

*School of Mathematics, University of Bristol, Bristol, BS8 1UG, UK.
†Heilbronn Institute for Mathematical Research, Bristol, UK.
‡Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK.
§Supported by EPSRC grant EP/X013642/1.
¶Department of Economics, University of Oxford, Oxford, OX1 3UQ, UK.
Email: tom.johnston@bristol.ac.uk, {savery,scott}@maths.ox.ac.uk, bassel.tarbush@economics.ox.ac.uk
Keywords: game connectivity, adaptive dynamics, best-response graphs, probabilistic combinatorics
Thanks: Sergiu Hart, Sam Jindani, Bary Pradelski, John Quah, Marco Scarsini, Alex Teytelboym, Peyton Young.

Young, 2006), trial-and-error learning (Young, 2009), and many more.[1] Several important results have outlined the boundary between the possible and the impossible, i.e. between classes of adaptive dynamics that are guaranteed to lead to a Nash equilibrium in every game and classes that lack such a guarantee (e.g. see Young, 2007 for an overview). A particularly influential impossibility result due to Hart and Mas-Colell (2003, 2006) establishes that there is no simple adaptive dynamic that is guaranteed to lead to a pure Nash equilibrium in every game that has one, where 'simple' qualifies the amount of information that each player has access to.

In this paper, we show that learning pure Nash equilibria via simple adaptive dynamics is not as hopeless an endeavour as the impossibility result of Hart and Mas-Colell (2003) might suggest: we look at the space of all (ordinal and generic) n-player games that have at least one pure Nash equilibrium, and we show that simple adaptive dynamics lead to a pure Nash equilibrium in all but a small fraction of such games when n is large compared to the number of actions per player, with the fraction vanishing exponentially quickly in n. To establish this result, we study the 'connectivity' properties of games—a concept that we formalize below—and we show that all but a quantifiably small fraction of games with relatively large n have a connectivity property that is conducive to equilibrium convergence.

What we do can be seen as a 'beyond the worst-case' analysis of learning pure Nash equilibrium in games. If one interprets simple adaptive dynamics as algorithms whose inputs are games, then Hart and Mas-Colell (2003) have established that simple adaptive dynamics perform very poorly on their worst-case inputs: they have shown that there are games (inputs) on which such dynamics (algorithms) do not lead to a pure Nash equilibrium. But, while algorithms are often assessed in terms of their worst-case performance, in practice we are often interested in how they perform on 'typical' problem instances, and this has given rise to 'beyond the worst-case' analysis of algorithms (Roughgarden, 2021). The simplex algorithm, for example, performs poorly on worst-case inputs but tends to perform well on typical problem instances. In this spirit, what we do in our paper is to characterize the structure of 'typical' games and, in contrast with Hart and Mas-Colell (2003), we establish that simple adaptive dynamics perform very well on typical inputs, i.e. are guaranteed to lead to a pure Nash equilibrium in 'typical' games.

The central argument of our paper is outlined in points (i)-(iv) below.

(i) We classify games according to the connectivity properties of their best-response graphs. Our interest in such a classification stems from the fact that the behaviors of many game dynamics are determined by such connectivity properties. A game's best-response graph is a directed graph whose vertex set is the set of pure action profiles and whose directed edges correspond to best-responses (Young, 1993). An

[1] Hart (2005) distinguishes between three types of dynamics in games: learning, evolutionary, and adaptive. Learning requires high levels of rationality (e.g. Kalai and Lehrer, 1993) whereas players in evolutionary dynamics instead mechanically inherit traits (e.g. Weibull, 1997; Hofbauer and Sigmund, 1998; Sandholm, 2010). Adaptive agents fall somewhere in between: they use relatively little information and take actions that respond to their environment according to basic decision heuristics in a generally improving way.

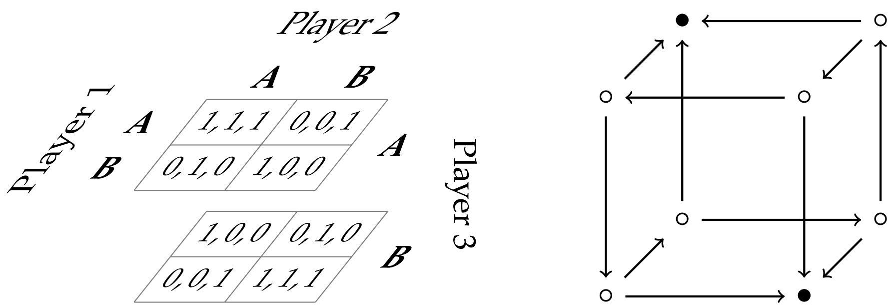


Figure 1: A 3-player 2-action game (left) and its corresponding best-response graph (right). The • vertices are sinks. The ∘ vertices form a cycle.

example is shown in Figure 1. A game's pure Nash equilibria correspond to the sinks of its best-response graph. Well-known classes of games categorized by the connectivity properties of their best-response graphs include *weakly acyclic* games, i.e. those for which every vertex of the best-response graph can reach a sink along a directed best-response path, and *acyclic* games, i.e. those whose best-response graphs contain no cycles. We introduce two new classes of games: *connected* and *super-connected* games. We say that a game is connected (super-connected) if its best-response graph has a sink and the property that every non-sink can reach every sink (non-source) via best-response paths. The logical relationships between these game classes are shown below:

$$\begin{array}{ccc} \text{acyclic} & \Longrightarrow & \text{weakly acyclic} \\ & & \Uparrow \\ & & \text{connected} \\ & & \Uparrow \\ & & \text{super-connected} \end{array}$$

The game in Figure 1, for example, is not acyclic but it is super-connected (and therefore connected and weakly acyclic).

(ii) We build on recent results in probabilistic combinatorics to enumerate games and thereby quantify the relative sizes of the game classes shown above. Throughout, we take players' preferences to be encoded ordinally via a preference relation rather than cardinally via a utility function, thus ensuring that the space of (finite) games is countable (and finite for a fixed number of players and number of actions per player). Our central result on game connectivity (Theorem 5) is that:

Almost every large generic game that has a pure Nash equilibrium is connected.

By 'generic' we mean that there are no indifferences, by 'large' we mean that the number of players is much bigger than the maximum number of actions available to each player, and by 'almost every' we mean all but an exponentially small proportion (in terms of the number of players). Our result even allows the number of actions per player to grow with the number of players, provided that the latter remains sufficiently large relative to the former.

Since connectedness is very common among large generic games that have a pure Nash equilibrium, so is weak acyclicty. In contrast, acyclicity is very rare: we show that the fraction of generic games with a pure Nash equilibrium that are acyclic vanishes super-exponentially in the number of players (Proposition 6).[2] Finally, we show that while super-connectedness is very common among large generic 2-action and 3-action games that have a pure Nash equilibrium, it is vanishingly rare among large generic k-action games when $k \geq 4$ (Proposition 7). We also derive analogous results for better-response graphs. Our characterization of the above game classes gives us insights into the 'typical' structure of large games, and this has important implications for adaptive dynamics, as we discuss below.

The results above follow from stronger results regarding the connectivity properties of random subgraphs of directed Hamming graphs, which we present in the appendix, and these are the main technical contribution of our paper. The proof of our main technical result (Theorem 15) adapts new work in probabilistic combinatorics regarding the component structure of random subgraphs of the hypercube (McDiarmid et al., 2021). Methodologically, our paper contributes to the well-established and growing literature on random games.[3] Within that literature, Amiet et al. (2021) contains results that are related to our work, and we discuss the relationship of our work to those results in Sections 4.3 and 4.4.

(iii) Turning to game dynamics, we show that there is a simple adaptive dynamic that is guaranteed to lead to a pure Nash equilibrium in every connected game. By 'simple dynamic' we mean a dynamic that is uncoupled (i.e. a player's strategy depends only on the actions of other players and on their own preferences), stationary (i.e. time-independent), and 1-recall (i.e. no more than the last period's play is available to each player), and when we say that a dynamic is 'guaranteed to lead to a pure Nash equilibrium' we mean that, starting at any action profile, the period-by-period play reaches a pure Nash equilibrium in finite time and, once there, never leaves it. Young (2004) shows that the best-response dynamic with inertia, which is a simple adaptive

[2]Because potential games have acyclic best-response graphs (Monderer and Shapley, 1996), our result on the prevalence of acyclic games implies that potential games with large numbers of players are very rare. We are not suggesting that the widely studied class of potential games is somehow unimportant: they are an appropriate model for certain types of strategic interaction such as congestion (Rosenthal, 1973), and many dynamics are guaranteed to converge to a pure Nash equilibrium in such games (e.g. Hofbauer and Sandholm, 2002; Roughgarden, 2016). Rather, we are highlighting that our implications for the convergence of adaptive dynamics are a consequence of the prevalence of connectedness rather than of acyclicity in large games.

[3]The distribution of pure Nash equilibria in random games was studied in Goldberg et al. (1968), Dresher (1970), Powers (1990), and Stanford (1995). Further results relating to the number of Nash equilibria also appear in McLennan (1997, 2005), Bárány et al. (2007), and Pei and Takahashi (2023). Alon et al. (2021) show that dominance-solvable games are rare (when the number of actions gets large for at least one player). Mimun et al. (2024) and Collevecchio et al. (2024) study best-response dynamics in two-player random games with correlated payoffs and in two-player random potential games, respectively.

dynamic, is guaranteed to lead to a pure Nash equilibrium in every weakly acyclic game.[4] It follows immediately that this is also true in connected games.

(iv) Combining our results above, we conclude (Theorem 13) that

> *There is a simple adaptive dynamic that is guaranteed to lead to a pure Nash equilibrium in almost every large generic game that has one.*

This contrasts with the aforementioned impossibility result of Hart and Mas-Colell (2003, 2006) which states that there is no simple adaptive dynamic that is guaranteed to lead to a pure Nash equilibrium in every (generic) game that has one. Our result does not overturn this impossibility, but it does limit its scope.[5]

Returning to our analogy of 'beyond the worst-case' analysis of algorithms, simple adaptive dynamics perform poorly (i.e. are not guaranteed to lead to a pure Nash equilibrium) on worst-case inputs of the type given in the impossibility result of Hart and Mas-Colell (2003), but they perform very well on 'typical' inputs because almost every large generic game has a property (namely, connectedness) which is conducive to convergent dynamics. Naturally, the strength of this analogy depends on what we understand to be a 'typical' input in the context of game dynamics. Adaptive dynamics

[4]Under the best-response dynamic with inertia, in each period, each player i independently best-responds to the current environment with probability $p_i \in (0, 1)$ and does not update their action with probability $1 - p_i$.

[5]There is a large literature on possibility and impossibility results for dynamics in games and the impossibility of Hart and Mas-Colell (2003) is by no means the only one; see Milionis et al. (2023) and Schipper (2022) for recent examples. Establishing impossibility for a class of dynamics typically consists of finding collections of games such that no dynamic in the class is guaranteed to lead to equilibrium in all of them. Naturally, this hinges on the parameters of the problem, namely, (i) the information that is allowed to determine players' decisions in the dynamic, (ii) the notion of convergence that is required, (iii) the type of equilibrium to which the dynamic converges, and (iv) the class of games to which the dynamic is applied.

There are several possibility results for dynamics that lead to mixed Nash or correlated equilibria. For example, uncoupled and completely uncoupled dynamics for which the empirical distribution of play converges almost surely to the set of correlated (or coarse correlated) equilibria in all games (Foster and Vohra, 1997; Fudenberg and Levine, 1999; Hart and Mas-Colell, 2000, 2001). There are also uncoupled and completely uncoupled dynamics for which the behavior probabilities converge almost surely to a mixed Nash equilibrium in all generic games (Foster and Young, 2006; Germano and Lugosi, 2007). Vlatakis-Gkaragkounis et al. (2020), however, show that a commonly studied regret-based dynamic does not lead to mixed Nash equilibria.

However, finding completely uncoupled, or even uncoupled, dynamics in which period-by-period play converges *almost surely* to a *pure* Nash equilibrium whenever one exists is demonstrably more challenging. There are stationary, 2-recall, uncoupled dynamics for which the period-by-period play converges almost surely to a pure Nash equilibrium in all games that have one (Hart and Mas-Colell, 2006; Cesa-Bianchi and Lugosi, 2006). Jaggard et al. (2014) identify other uncoupled dynamics with this convergence property in a bounded-recall synchronous setting. But the result of Hart and Mas-Colell (2003, 2006) shows that there is no stationary, 1-recall, uncoupled dynamic for which the period-by-period play converges almost surely to a pure Nash equilibrium in all games that have one. Babichenko (2012) shows that there is no completely uncoupled dynamic for which the period-by-period play converges almost surely to a pure Nash equilibrium in every generic game that has one, or even in every large generic game that has one. On the other hand, there is a completely uncoupled dynamic for which the period-by-period play is at a pure Nash equilibrium 'most of the time' in all generic games that have one (Young, 2009; Pradelski and Young, 2012).

are often applied in games with a large number of players and in which players select pure actions (Sandholm, 2010). Moreover, genericity is common; in fact, any utility-based game for which the utility numbers are perturbed by small random shocks independently drawn from an atomless distribution will have an ordinal representation that is almost surely generic in our sense of the term. Viewed in this way, it is reasonable to take generic games with many players, which are the focus of our paper, as 'typical' inputs for adaptive dynamics.[6]

Unlike much of the literature on learning in games, the focus of our paper is not on the dynamics (the algorithms) but on the structure of the games themselves (the inputs). Beyond having a bearing on the impossibility result of Hart and Mas-Colell (2003), this approach allows us to use our results on game connectivity to quantify the scope of existing results regarding the convergence properties of adaptive dynamics. 'Adaptive play' (Young, 1993), 'better-reply dynamics with sampling' (Friedman and Mezzetti, 2001), and regret-based dynamics (Marden et al., 2007) are guaranteed to lead to a pure Nash equilibrium in certain classes of games, and payoff-based dynamics lead to play that is at a pure Nash equilibrium 'most of the time' in certain classes of games (Marden et al., 2009). Our game connectivtiy results allow us to conclude that all of the these results apply to almost every large generic game that has a pure Nash equilibrium. More recently, building on our central result regarding the prevalence of connected games, Newton and Sawa (2024) were able to determine which Nash equilibria (according to their welfare properties) are selected by different evolutionary dynamics in large games.

2 Games

In this section we recall some standard definitions from the theory of games and introduce our notation. For $n \in \mathbb{N}$, we use $[n]$ as shorthand for the set $\{1, \ldots, n\}$. For each $a \in \mathbb{N}^n$ and $i \in [n]$, we write a_{-i} for the element of $\mathbb{N}^{n-1}$ obtained by deleting the ith coordinate of a. In an abuse of notation, for $x \in \mathbb{N}$ and $a_{-i} \in \mathbb{N}^{n-1}$, we write (x, a_{-i}) for the element of $\mathbb{N}^n$ obtained by inserting x into the ith coordinate of a_{-i}.

A *game* is a tuple

$$\Big([n], ([k_i])_{i\in[n]}, (\succsim_i)_{i\in[n]}\Big),$$

where $n \geq 2$ is an integer, $k_i \geq 2$ is an integer for each i, and for each i, $\succsim_i$ is a total preorder (i.e. a complete and transitive binary relation) on $A := \prod_{i\in[n]} [k_i]$. We say that $[n]$ is the *player set* of the game and that $[k_i]$ is the *action set* of player i. Elements of

[6]A condition that is sometimes imposed on games with many players is 'anonymity'. In anonymous games, at any action profile, a player's ranking of their actions depends only on the numbers of other players selecting particular actions but not on the identities of those other players. Certain interactive situations may be well-modeled as anonymous games but many large scale distributed interactions can be messy and idiosyncratic, in which case imposing anonymity becomes a convenient approximation. The widely studied class of weakly acyclic games, which need not be anonymous, is frequently studied in the context of adaptive dynamics with many players.

A are called *action profiles*, and $\succsim_i$ is known as i's *preference relation*.[7] For each i, let $\succ_i$ denote the asymmetric part of $\succsim_i$.

An action a_i of player i is a *best-response* to a_{-i} if $(a_i, a_{-i}) \succsim_i (x, a_{-i})$ for every $x \in [k_i]$. An action profile $a \in A$ is a *pure Nash equilibrium* if for each player $i \in [n]$, a_i is a best-response to a_{-i}.

3 Notions of game connectivity

The *best-response graph* of a game is the directed graph $(A, \rightarrow)$ whose vertex set is the set of action profiles A and whose directed edge set $\rightarrow$ is defined such that for $a, b \in A$,

> $a \rightarrow b$ if and only if there exists $i \in [n]$ such that $a_{-i} = b_{-i}$, b_i is a best-response to a_{-i}, and $b \succ_i a$.

In other words, there is a directed edge from a to b whenever b_i is a strict best-response to $a_{-i} = b_{-i}$ for some player i.

We now define various classes of games in terms of the connectivity properties of best-response graphs.[8] As part of these definitions we will use standard terminology from the theory of directed graphs which we briefly recall here. Given a directed graph $(V, \rightarrow)$ with vertex set V and edge set $\rightarrow$, a vertex $v \in V$ is a *sink* if it has no outgoing edges, and a *non-sink* otherwise. Similarly, a vertex $v \in V$ is a *source* if it has no incoming edges, and a *non-source* otherwise. For any pair of vertices $v, v' \in V$, we say that v can *reach* v' if there is a sequence $(v^1, \ldots, v^m)$ of vertices with $v^1 = v$ and $v^m = v'$ such that $v^i \rightarrow v^{i+1}$ for all $i \in [m-1]$; in this case we also say that the vertex v' can *be reached from* v. Note that every vertex can reach and be reached from itself. A *cycle* is a sequence $(v^1, \ldots, v^m)$ of distinct vertices that has length m at least 2 and that satisfies $v^m \rightarrow v^1$ and $v^i \rightarrow v^{i+1}$ for all $i \in [m-1]$.

Definition 1. A game is *acyclic* if its best-response graph has no cycles.

Definition 2. A game is *weakly acyclic* if its best-response graph has the property that every vertex can reach a sink.

Observe that, by definition, a weakly acyclic game necessarily has at least one sink. Moreover, acyclic games are weakly acyclic, but the converse need not hold.

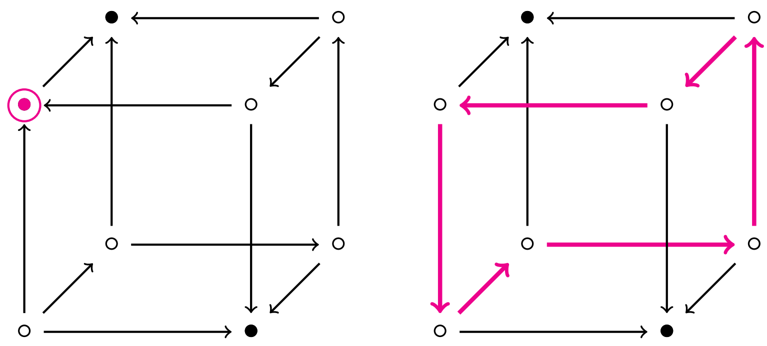
Acyclicity and weak acyclicity are standard concepts (see e.g. Fabrikant et al., 2013) though they sometimes appear under different names in the literature.[9] In our paper, the terms acyclicity and weak acyclicity follow the terminology of Young (1993), who introduced the concept of weak acyclicity to the literature on dynamics in games.

---

[7]In fact, for the dynamics that we will be interested in, the only relevant information about $\succsim_i$ is its restriction to $L(a_{-i}) := \{(x, a_{-i}) : x \in [k_i]\}$ for each $a_{-i} \in \prod_{j \in [n] \setminus \{i\}} [k_j]$.

[8]See Candogan et al. (2011) for a flow-based decomposition of games and Legacci et al. (2024) for an analysis of dynamics based on this decomposition.

[9]For example, Takahashi and Yamamori (2002) refer to weak acyclicity as quasi-acyclicity.



(a)

(b)

Figure 2: (a) Acyclic but not super-connected because some vertices, like the one shown in magenta cannot reach every non-source. (b) Super-connected but not acyclic because every non-sink can reach every non-source but there is a cycle, shown in magenta.

Acyclic games are a superset of the very widely studied class of potential games (Monderer and Shapley, 1996). Potential games have been the subject of intense research, particularly because many dynamics are guaranteed to converge to a pure Nash equilibrium in such games (e.g. Hofbauer and Sandholm, 2002; Roughgarden, 2016). Weakly acyclic games are also very widely studied because weak acyclicity is a necessary condition for the guaranteed convergence of best-response dynamics to a pure Nash equilibrium from any starting vertex (e.g. see Fabrikant et al., 2013; Apt and Simon, 2015). Yongacoglu et al. (2024) have recently studied a relaxation of weak acyclicity.

The next two notions of connectivity are ones that we introduce in this paper.

Definition 3. A game is *connected* if its best-response graph has at least one sink and the property that every non-sink can reach every sink.

Definition 4. A game is *super-connected* if its best-response graph has at least one sink and the property that every non-sink can reach every non-source.

Super-connectedness implies connectedness, and connectedness implies weak acyclicity but, in each case, the converse need not hold. Moreover, as shown in Figure 2, super-connectedness neither implies nor is implied by acyclicity.

4 Main results

We quantify the relative sizes of the game classes defined in Section 2 for large generic games, where 'large' means that the number of players is much bigger than the maximum number of actions available to each player, and 'generic' means that there are no preference ties. More specifically, a game is *generic* if for every i, and distinct action profiles a and a' that differ only in the ith index, either $a \succ_i a'$ or $a' \succ_i a$.[10]

Given an integer $n \geq 2$ and $\mathbf{k} = (k_1, \ldots, k_n) \in \{2, 3, \ldots\}^n$, we use $\mathcal{G}(n, \mathbf{k})$ to denote the set of all generic games with player set $[n]$ in which, for every $i \in [n]$, player i has action set $[k_i]$. Since we are working with ordinal games, for a fixed n and $\mathbf{k}$, the set $\mathcal{G}(n, \mathbf{k})$ is finite.

The following is our main result on game connectivity.

Theorem 5. *There exists $c > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$, if n is sufficiently large relative to $\max_i k_i$, then*

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is connected}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ has a pure Nash equilibrium}\}|} \geq 1 - e^{-cn}.$$

This result shows that, strikingly, connectedness is a ubiquitous property among large generic games that have a pure Nash equilibrium: almost every large generic game that has a pure Nash equilibrium is connected. Moreover, since connectedness implies weak acyclicity, the same is true of the latter property as well.

Our condition for n to be 'sufficiently large' is that

$$\max_i k_i \leq \delta \sqrt{n / \log(n)}$$

for a suitable constant $\delta > 0$.

While the possible dependence of k_i on n is suppressed in our notation, Theorem 5 allows for the number of actions per player to be growing with n provided that the above condition continues to be met. Of course, if the number of actions per player were fixed, our 'sufficiently large' condition would simplify to n exceeding some constant.

Theorem 5 has important implications for adaptive dynamics in games, on which we elaborate in Section 5.

Our next result, regarding acyclicity, applies to all generic games, not just those that we consider 'large'.

Proposition 6. *There exists $c > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$, we have*

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is acyclic}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ has a pure Nash equilibrium}\}|} \leq e^{-cn2^n}.$$

Together, Theorem 5 and Proposition 6 imply that there is a 'split' in large game properties: among large generic games that have a pure Nash equilibrium, acyclic

[10]Every game $([n], ([k_i])_{i\in[n]}, (\succsim_i)_{i\in[n]})$ has a utility-based representation $([n], ([k_i])_{i\in[n]}, (u_i)_{i\in[n]})$ with, for each player i, a utility function $u_i : A \to \mathbb{R}$ representing their preference relation $\succsim_i$. Genericity is equivalent to the condition that such a utility-based game has no payoff ties, i.e. that for each i and any distinct profiles a and a' that differ only in the ith index, $u_i(a) \neq u_i(a')$. Note, furthermore, that any utility-based game for which the utility numbers are perturbed by small random shocks independently drawn from an atomless distribution is almost surely generic.

games are very rare, while connected games and weakly acyclic games are very common. Note that since acyclic games are a superset of potential games, this also implies that potential games are very rare among large generic games that have a pure Nash equilibrium.

Our final main result concerns super-connectedness. Let $\mathbf{2} = (2, \ldots, 2)$, and similarly define $\mathbf{3}$ and $\mathbf{4}$.

Proposition 7. *For $\mathbf{k} = \mathbf{2}$ or $\mathbf{k} = \mathbf{3}$ there exists $c > 0$ such that for all integers $n \geq 2$,*

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{k}) \colon g \text{ is super-connected}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k}) \colon g \text{ has a pure Nash equilibrium}\}|} \geq 1 - e^{-cn}.$$

However, for each $\mathbf{k} = (k, \ldots, k) \geq \mathbf{4}$, the fraction above tends to 0 as $n \to \infty$.

This shows that almost every large generic 2-action game or 3-action game that has a Nash equilibrium is super-connected.[11] However, this is not true of k-action games for $k \geq 4$. In fact, for $k \geq 4$, the fraction of generic k-action games with a pure Nash equilibrium that are super-connected becomes vanishingly small as $n \to \infty$.

This shows two things. First, super-connectedness is too strong to be a typical property of large games. Unlike connectedness, it is not true that almost every large generic game that has a pure Nash equilibrium is super-connected. Second, connectivity properties that hold for small $\mathbf{k}$ do not necessarily extend to large $\mathbf{k}$. This is important because one cannot rely on results established for small numbers of actions as a guide for what to expect when the number of actions is large.

Our main results above are proved as corollaries of stronger results concerning the likelihood of analogous conditions holding in certain random directed graphs. The statements of these technical results on random graphs and the proofs themselves are in the appendix. The appendix also contains a section on the tightness of our main results. We provide a high-level discussion of our proof approach in Section 4.4 below.

4.1 Better-response graphs

The results above concerned best-response graphs. We now discuss their implications for better-response graphs.

An action a_i of player i is a *better-response* than a'_i to a_{-i} if $(a_i, a_{-i}) \succ_i (a'_i, a_{-i})$. The *better-response graph* of a game is the directed graph $(A, \to)$ whose vertex set is the set of action profiles A and whose directed edge set $\to$ is defined such that for $a, b \in A$,

$a \to b$ if and only if there exists $i \in [n]$ such that $a_{-i} = b_{-i}$ and b_i is a better-response to a_{-i} than a_i.

For each connectivity property $P \in$ {acyclic, weakly acyclic, connected, super-connected}, we say that a game is *globally* P if its better-response graph has that

[11] For this result we consider only games in which every player has the same number of actions.

property. For example, a game is globally connected if its better-response graph has a sink and the property that every non-sink can reach every sink. Observe that generalised ordinal potential games are precisely those that are globally acyclic (Monderer and Shapley, 1996; Fabrikant et al., 2013).[12]

Remark 8. Since a game's best-response graph is a subgraph of its better-response graph, we obtain the following relationships.

$$\begin{array}{ccccccc} \text{globally acyclic} & \Longrightarrow & \text{acyclic} & \Longrightarrow & \text{weakly acyclic} & \Longrightarrow & \text{globally weakly acyclic} \\ & & & & \Uparrow & & \Uparrow \\ & & & & \text{connected} & \Longrightarrow & \text{globally connected} \\ & & & & \Uparrow & & \Uparrow \\ & & & & \text{super-connected} & \Longrightarrow & \text{globally super-connected} \end{array}$$

The implications are now straightforward. Among large generic games that have a pure Nash equilibrium, Theorem 5 implies that connected games, weakly acyclic games and their global counterparts, are very common, while Proposition 6 implies that acyclic games and globally acyclic games are very rare. Analogous conclusions can similarly be drawn for super-connectedness.

4.2 Classes of games with positive asymptotic density

We now show that our results on game connectivity can be extended to any class of games $\mathcal{X}(n,\mathbf{k}) \subseteq \{g \in \mathcal{G}(n,\mathbf{k})\colon g \text{ has a pure Nash equilibrium}\}$ that has positive asymptotic density, by which we mean that there is a $p \in (0,1]$ such that

$$\frac{|\mathcal{X}(n,\mathbf{k})|}{|\{g \in \mathcal{G}(n,\mathbf{k})\colon g \text{ has a pure Nash equilibrium}\}|} \geq p$$

for all sufficiently large n.

The following is a corollary of Theorem 5.

Corollary 9. *If* $\mathcal{X}(n,\mathbf{k}) \subseteq \{g \in \mathcal{G}(n,\mathbf{k})\colon g \text{ has a pure Nash equilibrium}\}$ *has positive asymptotic density then, for sufficiently large* n*, almost every game in* $\mathcal{X}(n,\mathbf{k})$ *is connected.*

Here is an example. Rinott and Scarsini (2000) show that for any integer $z \geq 0$,

$$\frac{|\{g \in \mathcal{G}(n,\mathbf{k})\colon g \text{ has exactly } z \text{ pure Nash equilibria}\}|}{|\mathcal{G}(n,\mathbf{k})|} \longrightarrow \frac{e^{-1}}{z!}$$

[12] A game $g = ([n], ([k_i])_{i\in[n]}, (\succsim_i)_{i\in[n]})$ is a generalised ordinal potential game if there exists a function $\rho : A \to \mathbb{R}$ such that for each $i \in [n]$ and each pair of distinct action profiles a and a' that differ in only the ith index, $a >_i a'$ implies $\rho(a) > \rho(a')$. The game is an ordinal potential game or, simply, a potential game if, additionally, $\rho(a) > \rho(a')$ implies $a >_i a'$.

as $n \to \infty$ or as $k_i \to \infty$ for at least two players i.[13] From this we infer that for any integer $z \geq 1$, the set

$$\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ has at least } z \text{ pure Nash equilibria}\}$$

has positive asymptotic density, from which we can conclude, for example, that almost every large generic game that has say at least two pure Nash equilibria is connected.

4.3 Relationship to Amiet et al. (2021)

Our paper is related to Amiet et al. (2021). The focus of that paper is not entirely on game connectivity, but it contains results that are related to our work and we discuss this relationship here.

Consider a vertex v in the best-response graph of a game in $\mathcal{G}(n, \mathbf{k})$. We say that the game is *v-connected* if its best-response graph has at least one sink and the property that if v is a non-sink, then it can reach every sink. Similarly, we say that the game is *v-super-connected* if its best-response graph has at least one sink and the property that if v is a non-sink, then it can reach every non-source.

Expressed in the language of our paper, the arguments of Amiet et al. (2021) imply that there exists $c > 0$ such that for all integers $n \geq 2$ and any vertex v,

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{2})\colon g \text{ is } v\text{-super-connected}\}|}{|\{g \in \mathcal{G}(n, \mathbf{2})\colon g \text{ has a pure Nash equilibrium}\}|} \geq 1 - e^{-cn}.$$

This differs from our results in two main ways. First, a game is (super-)connected if it is v-(super-)connected for *every* vertex v, so (super-)connectedness is very much stronger than v-(super-)connectedness. Indeed, there are typically almost 2^n vertices that are non-sinks and connectedness requires v-connectedness to hold *simultaneously* for all of them. Second, the result of Amiet et al. (2021) applies only to two-action games, while our results apply much more generally. As we have seen, connectivity properties that hold for games with few actions per player may not extend to games with many actions per player, and these games are fundamentally different.

4.4 Proof approach

Suppose that a game G is drawn uniformly at random from $\mathcal{G}(n, \mathbf{k})$. Then

$$\Pr[G \text{ has property } P \mid G \text{ has a pure Nash equilibrium}]$$

is equal to the fraction

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ has property } P\}|}{|\{g \in \mathcal{G}(n, \mathbf{2})\colon g \text{ has a pure Nash equilibrium}\}|}.$$

[13] Rinott and Scarsini (2000) build on Arratia et al. (1989) to prove, among other things, that the distribution of the number of pure Nash equilibria in games drawn uniformly at random from among all generic games is asymptotically Poisson(1) as the number of players with at least two actions gets large or as the number of actions gets large for at least two players.

In other words, rather than directly enumerating games in $\mathcal{G}(n, \mathbf{k})$ that have certain properties, we instead draw games uniformly at random and work out the probability that such random drawn games have certain properties. The latter is simpler and yields our desired quantities of interest.

We now give a high-level explanation of some elements of the proof of Theorem 5. All the proofs for the results presented in Section 4 (and of even stronger results) are in the appendix. First, draw a game G uniformly at random from $\mathcal{G}(n, \mathbf{k})$. At the heart of our proof is an argument showing the likely existence of a large strongly connected component in the best-response graph of G; in other words, we find a large set of vertices which can all reach each other along directed paths. Indeed, we define a vertex of the best-response graph of G to be *good* if it has reasonably high in-degree and reasonably high out-degree, and we then show that with high probability every good vertex can reach all other good vertices along directed paths. Since most vertices are good, this connects up a reasonably large proportion of the graph. This part of our proof is based on work of McDiarmid et al. (2021), who study the component structure of random subgraphs of the undirected hypercube graph. Once we know that with high probability all good vertices can reach each other, we are then left with 'plugging in' the remaining, unusual, vertices. We plug in these vertices by building on arguments in Bollobás et al. (1993).

We note that our proof employs different methods from the ones used in Amiet et al. (2021). For the result that we discussed in Section 4.3, Amiet et al. (2021) also draw games uniformly at random, in their case from $\mathcal{G}(n, \mathbf{2})$. Because there are only two actions per player in every game in $\mathcal{G}(n, \mathbf{2})$, all edges in the resulting best-response graphs are independent of each other, and their proofs rely heavily on this feature. Once there are more than two actions per player, the edges are no longer independent, and this requires an alternative approach.

5 Implications for adaptive dynamics in games

We now consider the implications of our results regarding game connectivity for games played over time according to *adaptive dynamics*. We begin by recalling some standard notions.

First, a player i's *observation set* at time t, denoted o_i^t, is the set of information that i can observe at time t. Precisely what objects enter into this set varies depending on the regime under consideration, and below it will be made clear which regimes we are considering. For each integer $k \geq 2$, let O_k denote the set of all possible observation sets (under the given regime) for a player with action set $[k]$. A *strategy* for a player with action set $[k]$ is a function $f : O_k \to \Delta([k])$, where $\Delta([k])$ is the probability simplex over $[k]$. Let $n \geq 2$ and $k_1, \ldots, k_n \geq 2$ be integers, and write $\mathbf{k} = (k_1, \ldots, k_n)$. A *dynamic* on $\mathcal{G}(n, \mathbf{k})$ consists of a specification for what information enters into each player's observation set at each time, and a strategy f_i with action set $[k_i]$ for each player i.

The play of a game $g \in \mathcal{G}(n, \mathbf{k})$ under a given dynamic begins at time $t = 0$ at an initial action profile a^0 chosen arbitrarily. This informs each player's observation set o_i^1

according to the dynamic. At time $t = 1$, each player updates their action (randomly) according to $f_i(o_i^1)$, and we denote the new (random) action profile by a^1. The play continues in this manner, with each player updating their action at $t = 2$ according to $f_i(o_i^2)$ to produce an action profile a^2, and so on.

Definition 10. A dynamic is *simple* if it is uncoupled, 1-recall, and stationary. Formally, a dynamic is simple if at each time t, player i's observation set contains (at most) their own preference relation $\succsim_i$ and last period's action profile a^{t-1}.

The terms 'uncoupled', '1-recall', and 'stationary' are standard in the literature. We informally recall their definitions here: a dynamic is uncoupled if a player's observation set consists at most of their own preference relation and of the past history of play, it is 1-recall if the past history of play is restricted to only the last period, and it is stationary if their strategy is time-independent.[14]

We consider the following strong notion of convergence to a pure Nash equilibrium.

Definition 11. A dynamic on $\mathcal{G}(n, \mathbf{k})$ *converges almost surely to a pure Nash equilibrium* of a game $g \in \mathcal{G}(n, \mathbf{k})$ if when g is played according to the dynamic from any initial action profile, almost surely there exists $T < \infty$ and a pure Nash equilibrium a^* of g such that $a^t = a^*$ for all $t \geq T$.

5.1 The possibility of convergence to a pure Nash equilibrium

As mentioned in the introduction, the following impossibility result is well-known.

Theorem 12 (Hart and Mas-Colell, 2006; Jaggard et al., 2014)**.** *For all $n \geq 3$ and $\mathbf{k} \in \mathbb{N}^n$ with $k_i \geq 2$ for all i (or $k_i \geq 3$ for all i if $n = 3$), there is no simple dynamic on $\mathcal{G}(n, \mathbf{k})$ for which play converges almost surely to a pure Nash equilibrium in every game in $\mathcal{G}(n, \mathbf{k})$ that has one.*

It is instructive to revisit a proof of this result. Figure 3 shows the best-response graph of a game considered in Hart and Mas-Colell (2006). The graph's key feature is that each vertex in magenta has exactly one out-going edge. Any simple dynamic on this game initiated at one of the magenta vertices cannot get to any vertex other than the

[14] The definitions are given more formally here. A dynamic is *uncoupled* if at each time t, each player i's observation set contains (at most) their own preference relation $\succsim_i$ and the ordered history of play $a^0, \ldots, a^{t-1}$. For an integer $m \geq 1$, an uncoupled dynamic is *m-recall* if at each time t, each player i's observation set contains (at most) the current time t, their own preference relation $\succsim_i$, and the ordered history of play $a^{t-m}, \ldots, a^{t-1}$ for the past m steps, or the full history of play if $t < m$. An uncoupled and m-recall dynamic is *stationary* if at each time t each player i's observation set consists of their own preference relation $\succsim_i$ and the ordered history of play $a^{t-m}, \ldots, a^{t-1}$ for the past m steps, or the full history of play if $t < m$. Crucially, for $t \geq m$ the only information about the current time t available to the players is that $t \geq m$, so their strategies become time-independent after this point.

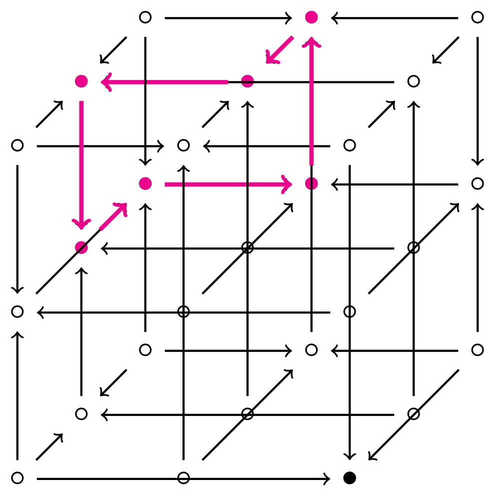


Figure 3: Best-response graph of a generic game with a pure Nash equilibrium. Any simple dynamic initiated at one of the magenta vertices cycles through those vertices forever. *Note*: some edges of the graph are omitted to keep the illustration legible.

magenta vertices.[15] In particular, any simple dynamic initiated at one of the magenta vertices cannot reach the pure Nash equilibrium. A game with a feature like this one can be embedded into a larger game (whether in terms of the number of players or the number of actions).

In contrast with the above impossibility result, we show below that if n grows (much more quickly than $\max_i k_i$), there is simple dynamic on $\mathcal{G}(n, \mathbf{k})$ that converges almost surely to a pure Nash equilibrium on all but a vanishingly small fraction of generic games in the class that have one.

Theorem 13. *For n sufficiently large relative to $\max_i k_i$, there is a simple dynamic on $\mathcal{G}(n, \mathbf{k})$ for which play converges almost surely to a pure Nash equilibrium in almost every game in $\mathcal{G}(n, \mathbf{k})$ that has one.*

The proof is straightforward . The *best-response dynamic with inertia* is defined as follows:[16] at each step t, independently of the other players, each player i sets a_i^t to be a best-response to a_{-i}^{t-1} with some fixed probability $p_i \in (0, 1)$ and sets $a_i^t = a_i^{t-1}$ with complementary probability $1 - p_i$. Young (2004) showed that, for any choice of parameters $p_i \in (0, 1)$, this dynamic converges almost surely to a pure Nash equilibrium in every weakly acyclic game. (We include a proof of Young's result in the case of generic weakly acyclic games in Appendix H in order to shed light on the link between

[15] The reason is this. Take a vertex v in the magenta cycle whose only outgoing edge is in direction j. By uncoupledness, changing player j's preferences should not affect what $i \neq j$ does at v. So change j's preferences so that v becomes the unique Nash equilibrium. If the dynamic must converge to a pure Nash equilibrium, $i \neq j$ must not move at v in this modified game. But this implies that $i \neq j$ must not move at v in the original game either.

[16] This dynamic is well-known and versions of it appear in, for example, Young (2009) and Swenson et al. (2018). The manner in which ties might be broken among multiple best-responses in non-generic games is immaterial for our purposes.

the connectivity properties of games and the convergence of dynamics.) Since every connected game is weakly acyclic, from Theorem 5 we obtain that there exists $c > 0$ such that for integers $n \geq 2$ and $\mathbf{k} \in \{2, 3, \dots\}^n$, if n is sufficiently large relative to $\max_i k_i$, then the proportion of games in

$$\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ has a pure Nash equilibrium}\}$$

for which the best-response dynamic with inertia converges almost surely to a pure Nash equilibrium is at least $1 - e^{-cn}$.[17] Noting that the best-response dynamic with inertia is simple completes the proof.

5.2 The scope of existing results on adaptive dynamics

Our results on game connectivity allow us to quantify the scope of existing results regarding the convergence properties of adaptive dynamics.

Consider the following (far from exhaustive) list of results on adaptive dynamics:[18]

- Young (1993) shows that 'adaptive play' (a class of dynamics involving inertia and finite memory) converges almost surely to a pure Nash equilibrium in all weakly acyclic games.
- Friedman and Mezzetti (2001) shows that the 'better-reply dynamic with sampling' converges almost surely to a pure Nash equilibrium in all globally weakly acyclic games.
- Marden et al. (2007) shows that a regret-based dynamic converges almost surely to a pure Nash equilibrium in all (generic) globally weakly acyclic games.
- Marden et al. (2009) show that a purely payoff-based dynamic leads to play that is at a pure Nash equilibrium in every globally weakly acyclic game 'most of the time'.

In all cases, since connected games are (globally) weakly acyclic, it follows from Theorem 5 that all the above results apply to almost every large generic game that has a pure Nash equilibrium. Observe that this is true even though the notions of convergence that are used in the aforementioned papers are sometimes different from almost sure convergence of period-by-period play.[19] For example, the notion of convergence in

[17] This result also holds for variants of the best-response dynamic with inertia. For example, it straightforwardly holds for the better-response dynamic with inertia. It also holds for a one-at-a-time version of the best-response dynamic in which, at each step t, exactly one player i is selected at random from among all players to update their action, and this player plays a best-response to a_{-i}^{t-1}. Convergence properties of this one-at-a-time version were investigated by Heinrich et al. (2023) via simulation.

[18] E.g. see Newton (2018) for further examples.

[19] Almost-sure convergence is a strong notion of convergence since it requires period-by-period play to eventually settle on, and never leave, a pure Nash equilibrium. We have focused on it, in part, because this is the notion that the impossibility results of Hart and Mas-Colell (2006) and Jaggard et al.

Marden et al. (2009) requires the dynamic to be at a pure Nash equilibrium 'most of the time', which is different from almost sure convergence.

Our results also have implications for equilibrium selection in games. When there are multiple equilibria, it is natural to ask which of these equilibria will be played. An approach taken in evolutionary game theory is to determine at which of the equilibria perturbed dynamics will spend most of their time; these are adaptive dynamics in which players' choices are subject to random errors parametrised by some $\varepsilon > 0$.[20] The 'stochastically stable' states of such a dynamic – the states at which the dynamic spends most of its time – are the action profiles that are assigned positive probability as $\varepsilon \to 0$ in the invariant distribution of the Markov process induced by the dynamic. Applying commonly used methods for determining stochastic stability such as the minimum-cost tree technique or the radius-coradius technique (Kandori et al., 1993; Kandori and Rob, 1995; Young, 1993; Freidlin et al., 2012; Ellison, 2000) can be complicated because they require checking global properties: a stochastically stable state must be 'hard' to leave (requiring many mutants to exit) and 'easy' to enter (requiring few mutants from other points in the game). Newton and Sawa (2024) observe that, in connected games, because it is very easy to move from a non-equilibrium point to *any* equilibrium point, the problem reduces to checking only 'one-shot deviations' from equilibrium, which is a local property regarding the likelihood only of exiting an equilibrium point. Combined with our results, Newton and Sawa (2024) conclude that in almost every large generic game that has a pure Nash equilibrium, stochastic stability is determined by a very simple 'one-shot' property. This implication for equilibrium selection requires our notion of connectivity. It does not follow from the connectivity result of Amiet et al. (2021), and it does not apply in weakly acyclic games that are not also connected. Using the 'one-shot' property, Newton and Sawa (2024) are able to determine which Nash equilibria (according to their welfare properties) are selected by different evolutionary dynamics in large games.

6 Open questions

We have approached the problem of finding adaptive dynamics that converge to pure Nash equilibria by studying connectivity properties of games rather than studying the properties of the dynamics themselves. We hope to have demonstrated that this

(2014) pertain to. Weaker notions of convergence (for examples, see Young, 2004) allow for possibility results that are different from ours. For example, Young (2009) shows that so-called 'trial-and-error learning' is an uncoupled (in fact, completely uncoupled) dynamic that, for any $\varepsilon > 0$, is at a pure Nash equilibrium for a $1 - \varepsilon$ proportion of time steps in any generic game that has one. This is a powerful result because it applies to *every* generic game (rather than *almost every* large generic game) but the notion of convergence there is weaker than almost-sure convergence of period-by-period play. With the latter, once a pure Nash equilibrium is reached, it is never left, whereas trial-and-error learning requires constant experimentation so there is always a positive probability of leaving a pure Nash equilibrium and wandering before settling on one again.

[20] A perturbed version of the best-response dynamic with inertia might specify that, at each time, any updating player plays a best-response with probability $1 - \varepsilon$ and, with complementary probability $\varepsilon > 0$, selects an action uniformly at random.

approach can deliver interesting conclusions, but many open questions still remain. We outline some of them below.

(i) We have worked in the regime in which the number of players is much bigger than the maximum number of actions per player. One could alternatively consider the setting where the number of players is fixed (or growing slowly) and the number of actions gets large. Results in this regime are, so far, limited. One implication of Heinrich et al. (2023) is that almost no generic two-player game that has a pure Nash equilibrium is weakly acyclic as the number of actions gets large (for both players). On the other hand, Amiet et al. (2021) show that almost every generic two-player game that has a pure Nash equilibrium is globally weakly acyclic as the number of actions gets large (again, for both players). However there are, to our knowledge, no results where the number of players is fixed above two. This 'large number of actions' regime behaves very differently from the 'large number of players' regime that we consider here. In the latter, each vertex of the best-response graph is incident to at least n edges, so n being large is likely to contribute to greater connectivity. If instead it is the number of actions that is large, then most vertices remain incident to a fixed number of edges. Characterising the 'large number of actions' regime remains an open question, which will require different arguments from those employed in our proof of Theorem 15. We will return to this question in a future paper.

(ii) General interest in uncoupled dynamics stems from the question of whether there are informationally undemanding dynamics that are guaranteed to lead to a pure Nash equilibrium when there is one. Completely uncoupled dynamics have even lower informational requirements than uncoupled ones: a dynamic is *completely uncoupled* if, at each time t, each player's observation set contains only their own realised utility payoffs and their own past actions. Babichenko (2012) showed that there is no completely uncoupled dynamic for which the period-by-period play converges almost surely to a pure Nash equilibrium in every generic game that has one, and moreover that there exist obstructions with arbitrarily large numbers of players and numbers of actions per player. We have not considered completely uncoupled dynamics here since they rely intrinsically on games with utility functions, which is outside the scope of this paper. Nevertheless, further analysis of game connectivity properties, either those we have studied here or others, may yield positive results on completely uncoupled dynamics.

(iii) Since our focus is not on any specific dynamic, we have not addressed the question of the speed of convergence to equilibrium (Arieli and Young, 2016).[21] It is known that adaptive dynamics can take a very long time to converge (Hart and Mansour, 2010), but greater knowledge of connectivity properties may help to establish general results on the speed of convergence in some classes of games.[22]

(iv) It would be fruitful to investigate the connectivity properties of specific classes of games featuring local interactions, such as graphical games (Kearns, 2007; Kearns et al., 2013) or action-graph games (Jiang et al., 2011).

[21] We expect convergence to take exponential time for the best-response dynamic with inertia.

[22] Convergence can be fast in potential games (e.g. see Awerbuch et al., 2008) and in anonymous games (e.g. see Babichenko, 2013).

(v) Amiet et al. (2021) and Collevecchio et al. (2024) derive some connectivity properties of non-generic games with two actions per player but there are, to our knowledge, no results for non-generic games with more than two actions per player.

(vi) Our analysis of convergence to pure Nash equilibrium did not consider the 'quality' of these equilibria. Pradelski and Young (2012), for example, describe a completely uncoupled dynamic that leads to a *Pareto optimal* equilibrium most of the time. Again, further analysis of game connectivity properties may help to more broadly address the question of convergence to efficient equilibria.

(vii) It might, more generally, be fruitful to investigate the prevalence of games with other types of connectivity properties. For example, it may be worth considering graphs that correspond to deviations by non-singleton coalitions of players.

A Connectivity of directed grids

Theorem 15, the main result of this section and the central technical contribution of our paper, is about the connectivity properties of random subgraphs of directed Hamming graphs. We first introduce our notation and then state the theorem, before explaining how Theorem 5 follows.

For $n \in \mathbb{N}$ and $\mathbf{k} = (k_1, \ldots, k_n) \in \{2, 3, \ldots\}^n$, the *Hamming graph* $H(n, \mathbf{k})$ is the graph with vertex set $V(n, \mathbf{k}) := \prod_{i=1}^n [k_i]$ and edges between n-tuples precisely when they differ in exactly one coordinate. For $i \in [n]$, a *line of* $V(n, \mathbf{k})$ *in coordinate* i is a subset of $V(n, \mathbf{k})$ of size k_i whose elements pairwise differ in exactly the ith coordinate. A *line* of $V(n, \mathbf{k})$ is a subset which is a line of $V(n, \mathbf{k})$ in coordinate i for some i. Note that a line induces a complete subgraph of $H(n, \mathbf{k})$. The *directed Hamming graph* $\vec{H}(n, \mathbf{k})$ is the simple directed graph formed by replacing each edge uv of $H(n, \mathbf{k})$ with directed edges $u \to v$ and $v \to u$.

Let $\vec{L}(n, \mathbf{k})$ be the random subgraph of the directed Hamming graph defined by independently and uniformly at random choosing a *winner* among the vertices of each line of $\vec{H}(n, \mathbf{k})$, and within that line keeping only those edges $u \to v$ whose endpoint, v, is the winner. Observe that in this random subgraph, each line induces a directed star in which all edges are oriented towards the winner. Our interest in $\vec{L}(n, \mathbf{k})$ stems from the following.

Remark 14. The graph $\vec{L}(n, \mathbf{k})$ has the same distribution as the best-response graph of a game drawn uniformly at random from amongst all games in $\mathcal{G}(n, \mathbf{k})$.

As in Theorem 5, we will study these objects when n is large relative to $\max_i(k_i)$; our proof breaks down when this is not the case. We now state our main theorem.

Theorem 15. *For all $\varepsilon > 0$ there exist $c, \delta > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$, if $K := \max_i(k_i)$ satisfies $K \leq \delta\sqrt{n/\log(n)}$, then with failure probability at most $\prod_{i=1}^n {k_i}^{-c}$, every vertex of $\vec{L}(n, \mathbf{k})$ can either be reached from at most $N := (1+\varepsilon)K\log(K)$ vertices, or from every non-sink.*

The full proof of Theorem 15 is postponed to Appendices B, C, D, and E. In Appendix F we examine the tightness (or lack thereof) of various aspects of Theorem 15. In particular, we show that neither the failure probability nor the value of N can be significantly improved in general.

We use the remainder of this subsection to explain how Theorem 5 follows from Theorem 15. To this end, we highlight the following corollary of Theorem 15, in which we denote by $R_{n,\mathbf{k}}$ the event that every non-sink in $\vec{L}(n, \mathbf{k})$ can reach every sink, and by $S_{n,\mathbf{k}}$ the event that $\vec{L}(n, \mathbf{k})$ has at least one sink.

Corollary 16. *There exist $c_0, c_1 > 0$ and $\delta \in (0, 1]$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$, if $K := \max_i(k_i)$ is such that $K \leq \delta\sqrt{n/\log(n)}$, then*

(a) $\mathbb{P}(R_{n,\mathbf{k}}) \geq 1 - e^{-c_0 n}$,

(b) $\mathbb{P}(R_{n,\mathbf{k}} \mid S_{n,\mathbf{k}}) \geq 1 - e^{-c_1 n}$.

Proof. Let c and δ' be as given by Theorem 15 in the case $\varepsilon = 1$, and let δ be the minimum of 1 and δ'. Then for n, $\mathbf{k}$, and K as in the statement of the corollary, we have that with failure probability at most $\prod_{i=1}^{n} k_i^{-c}$ every vertex of $\vec{L}(n, \mathbf{k})$ can either be reached from at most $2K \log(K)$ vertices or from every non-sink. However, we have $2K \log(K) \leq K^2 \log(K) \leq n$, so if this event holds then every non-sink can reach every sink, because all sinks can be reached from at least $n + 1$ vertices. Finally, note that $\prod_{i=1}^{n} k_i^{-c} \leq 2^{-cn} \leq e^{-c_0 n}$ for some $c_0 > 0$, which proves part (a).

Next,

$$\mathbb{P}(R_{n,\mathbf{k}} \mid S_{n,\mathbf{k}}) = \frac{\mathbb{P}(R_{n,\mathbf{k}} \cap S_{n,\mathbf{k}})}{\mathbb{P}(S_{n,\mathbf{k}})} \geq \frac{\mathbb{P}(R_{n,\mathbf{k}}) - (1 - \mathbb{P}(S_{n,\mathbf{k}}))}{\mathbb{P}(S_{n,\mathbf{k}})} \geq 1 - \frac{e^{-c_0 n}}{\mathbb{P}(S_{n,\mathbf{k}})},$$

where we used part (a) in the final step. It follows from work in Rinott and Scarsini (2000) that there exists a positive universal constant which lower bounds $\mathbb{P}(S_{n,\mathbf{k}})$ for all $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$, completing the proof of part (b). □

As remarked above, $\vec{L}(n, \mathbf{k})$ has the same distribution as the best-response graph of a game drawn uniformly at random from among all games in $\mathcal{G}(n, \mathbf{k})$. Because our draws are uniform, we have that

$$\mathbb{P}(R_{n,\mathbf{k}} \mid S_{n,\mathbf{k}}) = \frac{|\{g \in \mathcal{G}(n, \mathbf{k}) \colon g \text{ is connected}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k}) \colon g \text{ has a pure Nash equilibrium}\}|},$$

so Theorem 5 follows immediately from part (b) of Corollary 16.

Proof of Corollary 9. Replace $S_{n,\mathbf{k}}$ in the proof of Corollary 16 part (b) by the event that $\vec{L}(n, \mathbf{k})$ is in $\mathcal{X}(n, \mathbf{k})$. □

A.1 Outline of the proof of Theorem 15

We detail the proof of Theorem 15 in Appendices B, C, D, and E. We first start by providing a high-level overview of our arguments that is a bit more technical than the one we provided in the main text.

At the heart of our proof is an argument showing the likely existence of a certain large strongly connected component in $\vec{L}(n, \mathbf{k})$; in other words, we find a large set of vertices which can all reach each other along directed paths. Indeed, we define a vertex of $\vec{L}(n, \mathbf{k})$ to be *good* if the number of lines that it wins is close to the expected number, and then show that with high probability every good vertex can reach all other good vertices along directed paths. Since most vertices win close to the expected number of lines, this connects up a good proportion of the graph and it remains to 'plug in' the remaining, unusual, vertices. This part of our proof is based on work of McDiarmid et al. (2021), who study the component structure of random subgraphs of the undirected hypercube graph.

Once we know that all good vertices are in the same strongly connected component, it is sufficient to show that every non-sink x can reach some good vertex u, and every vertex y which can be reached from more than $N = (1+\varepsilon)K\log(K)$ vertices can be reached from some good vertex v. Indeed, this yields a directed path from x to y via u and v, where we utilise the strongly connected component to get from u to v. The first step towards this is to 'establish a foothold' by showing that with high probability (a) every non-sink can in fact reach at least $n/2$ vertices, and (b) every vertex that can be reached from more than N vertices can in fact be reached from at least $n/2$ vertices. The inspiration for considering such an event comes from the work of Bollobás et al. (1993). The final step in the proof is then to show that if a vertex can reach or be reached from at least $n/2$ vertices, then it is very unlikely that none of these vertices is good.

B Proof of Theorem 15: preliminaries

Throughout the proof of Theorem 15, i.e. throughout Appendices B, C, D, and E, we will take $n \geq 2$ to be an integer, we will take $\mathbf{k} \in \{2, 3, \dots\}^n$, and we will let $K := \max_i(k_i)$. We will describe a probability $p_\varepsilon(n, \mathbf{k})$ with parameters $\varepsilon > 0$, n, and $\mathbf{k}$ as being *very small* if for all ε there exist $c_\varepsilon, \delta_\varepsilon > 0$ depending only on ε such that $p_\varepsilon(n, \mathbf{k}) \leq \prod_{i=1}^n k_i^{-c_\varepsilon}$ for all $K \leq \delta_\varepsilon\sqrt{n/\log(n)}$. For a probability $p(n, \mathbf{k})$ with no dependence on ε, the constants c_ε and δ_ε should be replaced by universal constants. Furthermore, we will say that $p(n, \mathbf{k})$ is *extremely small* if there exist $c, \delta_0 > 0$ such that for all $\delta \in (0, \delta_0)$, if $K \leq \delta\sqrt{n/\log(n)}$, then $p(n, \mathbf{k}) \leq e^{-cn\log(K)/\delta}$. Observe that every extremely small probability is also very small.

If $p_\varepsilon(n, \mathbf{k})$ or $p(n, \mathbf{k})$ is very or extremely small, we will say that the complementary probability is very or extremely high, respectively. We say that an event $F_\varepsilon(n, \mathbf{k})$ or $F(n, \mathbf{k})$ occurs *with very high probability* (wvhp) or *with extremely high probability* (wehp) if the probability that it occurs is very or extremely high respectively.

Given this terminology, Theorem 15 is equivalent to the statement that wvhp, every vertex of $\vec{L}(n, \mathbf{k})$ can either be reached from at most $(1+\varepsilon)K\log(K)$ vertices or from every non-sink. This motivates our definition of very small probabilities. Our definition of extremely small probabilities is motivated by the following lemma, which demonstrates that such probabilities are amenable to union bounds over $V(n, \mathbf{k})$.

Lemma 17. *If $p(n, \mathbf{k})$ is an extremely small probability, then for all fixed $a > 0$ the probability $K^{an} \cdot p(n, \mathbf{k})$ is also extremely small.*

Proof. Let c and δ_0 witness the fact that $p(n, \mathbf{k})$ is extremely small. Then for all $\delta \in (0, \delta_0)$, if $K \leq \delta\sqrt{n/\log(n)}$, then

$$K^{an} \cdot p(n, \mathbf{k}) \leq e^{an\log(K) - cn\log(K)/\delta} = e^{n\log(K)(a - c/\delta)}.$$

Let $\delta_0' \in (0, \delta_0)$ be small enough that $c/(2\delta_0') > a$, then for all $\delta \in (0, \delta_0')$ we have $a - c/\delta < -c/(2\delta)$, so letting $c' = c/2$ we see that c', δ_0' witness the fact that $K^{an} \cdot p(n, \mathbf{k})$ is extremely small. □

We will also make frequent use of the following simple result which follows from elementary analyses of the various cases.

Lemma 18. *The sum of two very small probabilities is very small and the sum of two extremely small probabilities is extremely small.*

Before starting the proof Theorem 15 in earnest, we record the following two standard results which will be useful at various points. For a discussion of these results (and much more), we refer the reader to Frieze and Karoński (2015).

Lemma 19 (Chernoff bound)**.** *Let $X_1, \ldots, X_n$ be independent Bernoulli random variables, let $X = \sum_{i=1}^{n} X_i$, and let $\mu = \mathbb{E}[X]$. Then for all $\varepsilon \geq 0$ we have*

$$\mathbb{P}(X \leq (1-\varepsilon)\mu) \leq e^{-\varepsilon^2 \mu/2}.$$

Lemma 20 (Application of Markov's inequality)**.** *Let $Z_1, \ldots, Z_m$ be non-negative integer valued random variables, and suppose that $\sum_{i=1}^{m} \mathbb{E}[Z_i] \leq p$ for some $p \in [0,1]$. Then the probability that $Z_1 = \cdots = Z_m = 0$ is at least $1-p$.*

Proof. Let $Z = \sum_{i=1}^{m} Z_i$ and note that $\mathbb{E}[Z] = \sum_{i=1}^{m} \mathbb{E}[Z_i] \leq p$. By Markov's inequality, $\mathbb{P}(Z \geq 1) \leq \mathbb{E}[Z] \leq p$, and the complement of the event $\{Z \geq 1\}$ is $\{Z_1 = \cdots = Z_m = 0\}$. □

C Proof of Theorem 15: establishing a foothold

Throughout this section we let $\varepsilon > 0$ and set $N = (1+\varepsilon)K\log(K)$, as in the statement of Theorem 15. We will also assume (without loss of generality) that $k_1 \leq k_2 \leq \cdots \leq k_n$. Recall from Section A.1 that one step in our proof of Theorem 15 will be to show that, wvhp, in $\vec{L}(n, \mathbf{k})$ all non-sinks can reach more than $n/2$ vertices, and all vertices which can be reached from more than N vertices can be reached from more than $n/2$ vertices. We write A_ε for the event that this condition holds in $\vec{L}(n, \mathbf{k})$.

Definition 21 (Event A_ε)**.** Let A_ε be the event that the following two conditions are satisfied:

- all non-sinks can reach more than $n/2$ vertices; and
- every vertex that can be reached from more than N vertices can be reached from more than $n/2$ vertices.

This section is devoted to establishing the following lemma.

Lemma 22. *The event A_ε occurs with very high probability.*

Our proof of Lemma 22 follows the approach used by Bollobás et al. (1993) to study a random subgraph of $\vec{H}(n, \mathbf{2})$ with a similar distribution to that of $\vec{L}(n, \mathbf{2})$. Imitating those authors, for each $1 \leq m \leq \prod_{i=1}^{n} k_i$, define the random variable X_m to be the number of vertices of $\vec{L}(n, \mathbf{k})$ which can reach exactly m vertices (recall that every vertex can reach and be reached from itself). Analogously, let Y_m be the number of vertices which can be reached from exactly m vertices. Observe that A_ε can equivalently be defined as the event that $X_m = 0$ for all $2 \leq m \leq n/2$ and $Y_m = 0$ for all $N < m \leq n/2$.

Given a set $S \subseteq V(n, \mathbf{k})$, we say that a line of $V(n, \mathbf{k})$ in coordinate i is an *incomplete line of S* if its intersection with S has size other than 0 or k_i. If v can reach exactly m vertices, then running a depth first search from v gives a tree T with m vertices, in which all edges are oriented away from v and the winner of every incomplete line of T is in T. Similarly, if v can be reached from exactly m vertices, then we may build a tree T with m vertices where all the edges are oriented towards v and the winner of every incomplete line of T is outside of T. It follows that X_m and Y_m are bounded above by the number of pairs (v, T), where T is an appropriate tree with m vertices rooted at v.

We will use the following folklore result to upper bound the numbers of such trees in $H(n, \mathbf{k})$. A short combinatorial proof is given in McDiarmid et al. (2021).

Lemma 23. *If G is a graph with maximum degree Δ, then for each $m \in \mathbb{N}$ there are at most $(e\Delta)^{m-1}$ trees of order m in G that contain a given vertex.*

When applied to $H(n, \mathbf{k})$, Lemma 23 gives that there are at most $(enK)^{m-1}$ trees of order m in $H(n, \mathbf{k})$ that contain a given vertex. Lemma 22 follows from the next two lemmas, which handle the X_m and Y_m parts of the statement respectively.

Lemma 24. *With very high probability, $X_m = 0$ for all $2 \leq m \leq n/2$.*

Proof. We need to show that there exist universal $c, \delta > 0$ such that if $K \leq \delta\sqrt{n/\log(n)}$, then $X_m = 0$ for all $2 \leq m \leq n/2$ with probability at least $1 - \prod_{i=1}^{n} k_i^{-c}$. Thus, let $\delta > 0$ be small and assume that $K \leq \delta\sqrt{n \log(n)}$.

Fix $2 \leq m \leq n/2$ and let T be a tree of order m in $H(n, \mathbf{k})$. Given the discussion preceding Lemma 23, we wish to upper bound the probability that the winner of every incomplete line of T is in T. To this end, it will be helpful to lower bound the number of lines containing exactly one vertex of T. Each vertex of T is in n lines, so there are mn pairs (u, l) consisting of a vertex u in T and a line l containing it. For each pair of distinct vertices u and v in T, if u and v are contained in some common line l, then delete the pairs (u, l) and (v, l) from this set. Since any pair of vertices have at most one common line, this process removes at most $2\binom{m}{2}$ pairs from the set, and we deduce that there are at least $mn - m^2$ lines of $V(n, \mathbf{k})$ which contain exactly one vertex of T.

The winner of each of these lines is in T independently. Since we want to upper bound the probability that the winner of all of these lines is in T, we may assume that they are all in as low a coordinate direction as possible (recall that $k_1 \leq \cdots \leq k_n$ by assumption). At most m incomplete lines are in any given coordinate direction, so the probability that the winner of every incomplete line of T is in T is at most $\prod_{i=1}^{n-m} k_i^{-m}$.

By Lemma 23, the number of pairs (v, T) where $v \in V(n, \mathbf{k})$ and T is a tree of order m in $H(n, \mathbf{k})$ containing v is at most $(enK)^{m-1} \cdot \prod_{i=1}^{n} k_i$, so by the discussion before that lemma we have

$$\mathbb{E}[X_m] \leq \frac{(enK)^{m-1} \cdot \prod_{i=1}^{n} k_i}{\prod_{i=1}^{n-m} k_i^m} \leq \frac{K^m (enK)^{m-1} \cdot \prod_{i=1}^{n-m} k_i}{\prod_{i=1}^{n-m} k_i^m} \leq \frac{(enK^2)^m}{\prod_{i=1}^{n-m} k_i^{m-1}}.$$

Applying the fact that $m - 1 \geq m/2$ (since $m \geq 2$), we obtain

$$\mathbb{E}[X_m] \leq \frac{(enK^2)^m}{\prod_{i=1}^{n-m} k_i^{m/2}} \leq \left(\frac{enK^2}{\prod_{i=1}^{n-m} k_i^{1/2}} \right)^m \leq \prod_{i=1}^{n-m} k_i^{-m/3}$$

where the final inequality follows by taking δ small enough that $enK^2 \leq 2^{n/12}$, which is at most $\prod_{i=1}^{n-m} k_i^{1/6}$ since $m \leq n/2$.

Claim 1. *If δ is small enough, then $\prod_{i=1}^{n-m} k_i^{-m/3} \leq \prod_{i=1}^{n} k_i^{-1/2}$ for all $2 \leq m \leq n/2$.*

Proof. After rearranging, we need to show that $\prod_{i=n-m+1}^{n} k_i^{1/2} \leq \prod_{i=1}^{n-m} k_i^{m/3-1/2}$ for all $2 \leq m \leq n/2$. The left-hand side of this inequality is at most $K^{m/2}$ and the right-hand side is at least $2^{(n-m)(m/3-1/2)}$. Raising both sides to the power of $2/m$, it is sufficient that $K \leq 2^{(n-m)(2/3-1/m)}$. The right-hand side of this inequality is at least $2^{n/12}$, and we can take δ small enough that $K \leq 2^{n/12}$, so the claim is proved. □

Applying the claim, we have

$$\sum_{m=2}^{n/2} \mathbb{E}[X_m] \leq \frac{n}{2} \cdot \prod_{i=1}^{n} k_i^{-1/2}.$$

By taking δ to be sufficiently small we can ensure that this is at most $\prod_{i=1}^{n} k_i^{-c}$ for some $c > 0$. Lemma 20 now yields that $X_m = 0$ for all $2 \leq m \leq n/2$ with failure probability at most $\prod_{i=1}^{n} k_i^{-c}$, as required. □

The next lemma deals with the Y_m part of Lemma 22. Note that Lemma 22 follows immediately from Lemma 18, Lemma 24, and Lemma 25.

Lemma 25. *With very high probability, $Y_m = 0$ for all $N < m \leq n/2$.*

Proof. We need to show that there exist $c_\varepsilon, \delta_\varepsilon > 0$ depending only on ε such that if $K \leq \delta_\varepsilon \sqrt{n/\log(n)}$, then $Y_m = 0$ for all $N < m \leq n/2$ with failure probability at most $\prod_{i=1}^{n} k_i^{-c_\varepsilon}$. In fact, we will show a stronger failure probability of at most $e^{-c_\varepsilon n \log(K)}$. Thus, let $\delta_\varepsilon > 0$ be small and assume that $K \leq \delta_\varepsilon \sqrt{n/\log(n)}$.

We will employ a similar strategy to that used to prove Lemma 24. Fix $N < m \leq n/2$ and let T be a tree of order m in $H(n, \mathbf{k})$. We will upper bound the probability that the winner of every incomplete line of T is not in T using the lower bound of $mn - m^2$ on the number of incomplete lines of T (from the proof of Lemma 24). The winner of each of these lines is in T independently, so the probability that all the winners are outside T is at most $(1-1/K)^{m(n-m)}$.

Hence, by Lemma 23 and the discussion preceding it, we have

$$\mathbb{E}[Y_m] \le K^n \cdot (enK)^{m-1} \cdot \left(1-\frac{1}{K}\right)^{m(n-m)}$$
$$\le \left[enK^2 \cdot \left(K^{1/m}\left(1-\frac{1}{K}\right)\right)^{n-m}\right]^m .$$

Using that $m > N = (1+\varepsilon)K\log(K)$ and $1+x \le e^x$ for all x we have

$$K^{1/m}\left(1-\frac{1}{K}\right) \le K^{1/(1+\varepsilon)K\log(K)}e^{-1/K} = \exp\left(\frac{-\varepsilon}{(1+\varepsilon)K}\right). \tag{1}$$

Assuming that $\delta_\varepsilon \le 1/2$, we have that $K^2 \le n$. Applying this and $m \le n/2$ yields

$$\mathbb{E}[Y_m] \le \left[en^2 \cdot \exp\left(\frac{-\varepsilon(n-m)}{(1+\varepsilon)K}\right)\right]^m \le \left[en^2 \cdot \exp\left(\frac{-\varepsilon n}{2(1+\varepsilon)K}\right)\right]^m .$$

By making δ_ε small enough that $\varepsilon n/(4(1+\varepsilon)K) \ge \log(en^2)$ and using the fact that $m \ge K\log(K)$, we have

$$\mathbb{E}[Y_m] \le \exp\left(\frac{-\varepsilon mn}{4(1+\varepsilon)K}\right) \le \exp\left(\frac{-\varepsilon n\log(K)}{4(1+\varepsilon)}\right).$$

Thus,

$$\sum_{N<m\le n/2} \mathbb{E}[Y_m] \le \frac{n}{2}\cdot\exp\left(\frac{-\varepsilon n\log(K)}{4(1+\varepsilon)}\right) \le \exp\left(-\frac{\varepsilon}{8(1+\varepsilon)}n\log(K)\right)$$

for sufficiently large (depending only on ε) n. By making δ_ε sufficiently small relative to ε, we can ensure that we only need to consider values for n which are sufficiently large, and we find that $\sum_{N<m\le n/2}\mathbb{E}[Y_m]$ is at most $e^{-c_\varepsilon n\log(K)}$ for some $c_\varepsilon > 0$ depending only on ε. Lemma 20 now yields that $Y_m = 0$ for all $N < m \le n/2$ with failure probability at most $e^{-c_\varepsilon n\log(K)}$, as required. □

D Proof of Theorem 15: a strongly connected component

The next step in the proof of Theorem 15 will be to find a large strongly connected component in $\vec{L}(n,\mathbf{k})$. To this end, we make the following two definitions.

Definition 26 (Good vertex)**.** A vertex of $\vec{L}(n,\mathbf{k})$ is *good* if it wins at least $n/3K$ but at most $3n/4$ of its lines.

Definition 27 (Event B)**.** Let B be the event that every good vertex of $\vec{L}(n, \mathbf{k})$ can reach all other good vertices along directed paths, that is, all good vertices are in the same strongly connected component of $\vec{L}(n, \mathbf{k})$.

The main result of this section is that B occurs wehp.

Lemma 28. *The event B occurs with extremely high probability.*

Our proof of Lemma 28 is based on an approach taken by McDiarmid et al. (2021), and will proceed via the following two auxiliary lemmas.

Lemma 29. *With extremely high probability, every vertex in $V(n, \mathbf{k})$ has an $H(n, \mathbf{k})$-neighbour which is good in $\vec{L}(n, \mathbf{k})$.*

Lemma 30. *With extremely high probability, every good vertex in $\vec{L}(n, \mathbf{k})$ can reach all good vertices within $H(n, \mathbf{k})$-distance 3 of it.*

Given Lemmas 29 and 30, Lemma 28 follows easily. This proof and the proof of Lemma 29 are straightforward generalisations of work in McDiarmid et al. (2021), but we include them for completeness.

Proof of Lemma 28. Let $\vec{F}$ be any realisation of $\vec{L}(n, \mathbf{k})$ for which the conclusions of Lemma 29 and Lemma 30 both hold, and note that this happens wehp by those two lemmas and Lemma 18. That is, let $\vec{F}$ be any outcome of $\vec{L}(n, \mathbf{k})$ in which every good vertex can reach all other good vertices at distance at most 3 from it in $H(n, \mathbf{k})$, and in which every vertex has a neighbour in $H(n, \mathbf{k})$ which is good. To prove Lemma 28, it suffices to show that for all good vertices x and y, there exists a directed path in $\vec{F}$ from x to y.

Let x and y be good vertices of $\vec{F}$ and choose a path $P = p_0 p_1 \dots p_t$ in $H(n, \mathbf{k})$ where $p_0 = x$ and $p_t = y$. If $t \le 3$, then there exists a directed path from x to y in $\vec{F}$. Otherwise $p_2, p_3, \dots, p_{t-2}$ have good $H(n, \mathbf{k})$-neighbours $q_2, \dots, q_{t-2}$ respectively. Then x and q_2 are at distance at most 3, q_i and q_{i+1} are at distance at most 3 for all $2 \le i \le t-3$, and q_{t-2} and y are at distance at most 3, so $\vec{F}$ contains a path from x to y via $q_2, \dots, q_{t-2}$. □

Proof of Lemma 29. For a vertex $v \in V(n, \mathbf{k})$, let X be the number of lines that v wins in $\vec{L}(n, \mathbf{k})$ and note that v is good if and only if $n/3K \le X \le 3n/4$. Since X is a sum of n independent Bernoulli random variables and has mean $\mu = \sum_{i=1}^{n} 1/k_i \ge n/K$, by Lemma 19 we have

$$\mathbb{P}\Big(X < \frac{n}{3K}\Big) \le \mathbb{P}\Big(X \le \frac{\mu}{3}\Big) \le e^{-2\mu/9} \le e^{-2n/(9K)}.$$

Similarly, $n - X$ is a sum of n independent Bernoulli random variables and has mean $\mu' \ge n/2$, so by Lemma 19 again we have

$$\mathbb{P}\Big(X > 3n/4\Big) = \mathbb{P}\Big(n - X < n/4\Big) \le \mathbb{P}\Big(n - X \le \frac{\mu'}{2}\Big) \le e^{-\mu'/8} \le e^{-n/16}.$$

It follows (using $K \geq 2$) that v is good with failure probability at most $e^{-n/(10K)}$.

Now fix $u \in V(n, \mathbf{k})$ and pick one vertex other than u from each of the n lines containing it, say $v_1, \ldots, v_n$. The v_i are distinct and no two of them share a line, so they are good independently of one another. Hence, the probability that u has no good $H(n, \mathbf{k})$-neighbour is at most $e^{-n^2/(10K)}$, so by a union bound over u, the probability that there exists a vertex with no good $H(n, \mathbf{k})$-neighbour is at most $K^n \cdot e^{-n^2/(10K)}$. For $0 < \delta \leq 1$, if $K \leq \delta\sqrt{n/\log(n)}$, then $K^2 \log(K) \leq \delta n$ and $n/K \geq K \log(K)/\delta \geq \log(K)/\delta$, so $e^{-n^2/(10K)}$ is extremely small. By Lemma 17, the same is true of $K^n \cdot e^{-n^2/(10K)}$, which completes the proof of the lemma. □

It remains to give the (slightly more involved) proof of Lemma 30.

Proof of Lemma 30. In this proof we will relabel $V(n, \mathbf{k})$ as $\prod_{i=1}^{n}\{0, \ldots, k_i - 1\}$ in the natural way and will consider these vertices as elements of the vector space $\mathbb{R}^n$. We will write $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for the standard basis of this space.

We need to show that if $\delta > 0$ is sufficiently small, then whenever $K \leq \delta\sqrt{n/\log(n)}$, the probability that there exists a good vertex that cannot reach some other good vertex within $H(n, \mathbf{k})$-distance 3 of it is at most $e^{-cn\log(K)/\delta}$ for some universal $c > 0$. Thus, let $\delta > 0$ be small and assume that $K \leq \delta\sqrt{n/\log(n)}$. Note that since $n \geq 2$, by choosing δ small enough we may assume that $n/K^2 \geq \log(n)/\delta^2$ (and hence also n and n/K) is large in absolute terms.

Our proof will focus on pairs of vertices at distance exactly 3 from one another, and it will be clear how to adapt the argument to pairs at distance 1 or 2. Let u and v be vertices of $H(n, \mathbf{k})$ at distance 3 from each other. After relabelling, we may assume that $u = \mathbf{0}$ and $v = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$. Fix subsets $A', B' \subseteq [n]$ of sizes at least $n/4$ and $n/3K$ respectively; later we will assume that u and v are good vertices and take these to be the sets of coordinate directions in which u and v do not win and, respectively, win their lines. For each $i \in A'$ fix some $\alpha_i \in [k_i - 1]$; later we will take α_i to be the ith coordinate of the winner of the line through u in direction i. Now pick any $A \subseteq A' \setminus \{1, 2, 3\}$ and $B \subseteq B' \setminus \{1, 2, 3\}$ such that $|A| = \lceil n/5 \rceil$ and $|B| = \lceil n/4K \rceil$. Relabelling again, we may assume that $A, B \subseteq [\lfloor n/2 \rfloor]$.

Having fixed A, B, and the α_i, we will now define a certain type of path in $\vec{H}(n, \mathbf{k})$. First, let $i, j \in [n]$ be distinct with $i \in A$ and $j \in B$, then let $\beta \in [k_j - 1]$. A directed path in $\vec{H}(n, \mathbf{k})$ from $\alpha_i\mathbf{e}_i + \beta\mathbf{e}_j$ to $\alpha_i\mathbf{e}_i + \beta\mathbf{e}_j + v$ will be called an *(i, j, β)-path* if it has the following form: the path starts at $\alpha_i\mathbf{e}_i + \beta\mathbf{e}_j$ then follows a path of length 3 to $\alpha_i\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1$ in which the first and third edges are in a coordinate direction taken from the interval of integers $[\lfloor n/2 \rfloor + 1, \lfloor 2n/3 \rfloor]$. That is, the path starts

$$\alpha_i\mathbf{e}_i + \beta\mathbf{e}_j, \alpha_i\mathbf{e}_i + \beta\mathbf{e}_j + \gamma\mathbf{e}_\ell, \alpha_i\mathbf{e}_i + \beta\mathbf{e}_j + \gamma\mathbf{e}_\ell + \mathbf{e}_1, \alpha_i\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1,$$

for some $\ell \in [\lfloor n/2 \rfloor + 1, \lfloor 2n/3 \rfloor]$ and $\gamma \in [k_\ell - 1]$. Next, the path follows a path of length 3 to $\alpha_i\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1 + \mathbf{e}_2$ in which the first and third edges are in a coordinate direction taken from $[\lfloor 2n/3 \rfloor + 1, \lfloor 5n/6 \rfloor]$, before finally following a path of length 3 to $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 = \alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + v$ in which the first and third edges are in a coordinate direction taken from $[\lfloor 5n/6 \rfloor + 1, n]$.

Let $E^{(1)}_{(i,j,\beta)}$ be the event that there exists in $\vec{L}(n, \mathbf{k})$ a path of length 3 from $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j$ to $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \mathbf{e}_1$ in which the first and third edges are in a coordinate direction taken from $[\lfloor n/2 \rfloor + 1, \lfloor 2n/3 \rfloor]$. That is, $E^{(1)}_{(i,j,\beta)}$ is the event that there is some $\ell \in [\lfloor n/2 \rfloor + 1, \lfloor 2n/3 \rfloor]$ and some $\gamma \in [k_\ell - 1]$ such that all of the edges in the directed path

$$\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j, \alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \gamma \mathbf{e}_\ell, \alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \gamma \mathbf{e}_\ell + \mathbf{e}_1, \alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \mathbf{e}_1$$

are present in $\vec{L}(n, \mathbf{k})$. Define $E^{(2)}_{(i,j,\beta)}$ and $E^{(3)}_{(i,j,\beta)}$ analogously for the second and third parts of the (i, j, β)-path. Note that there exists an (i, j, β)-path in $\vec{L}(n, \mathbf{k})$ if and only if all three of these events occur.

The event that $\vec{L}(n, \mathbf{k})$ contains a path of length 3 from $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j$ to $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \mathbf{e}_1$ in which the first and third edges are in a given coordinate direction ℓ has probability

$$\left(1 - \frac{1}{k_\ell}\right) \frac{1}{k_1} \frac{1}{k_\ell} \geq \frac{1}{2K^2}.$$

Indeed, the probability that $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j$ does not win its line in direction ℓ is $1 - 1/k_\ell$; the probability that $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + \mathbf{e}_1$ wins its line in direction ℓ is $1/k_\ell$, the probability that the required edge in the direction 1 is present is $1/k_1$, and these three events occur independently. Since the existence of such a path is independent for different ℓ, the failure probability of $E^{(1)}_{(i,j,\beta)}$ is at most

$$\prod_{\ell = \lfloor n/2 \rfloor + 1}^{\lfloor 2n/3 \rfloor} \left(1 - \frac{1}{2K^2}\right) \leq \left(1 - \frac{1}{2K^2}\right)^{n/7} \leq e^{-n/(14K^2)} < \frac{1}{2},$$

where we have used that $1 + x \leq e^x$ for all $x \in \mathbb{R}$ and that n/K^2 is large.

Similarly, $E^{(2)}_{(i,j,\beta)}$ and $E^{(3)}_{(i,j,\beta)}$ each occur with probability at least 1/2. Moreover, it is not difficult to see that the sets of lines on whose presence each of these three events depend are pairwise disjoint, from which it follows that the events are independent. We deduce that $\vec{L}(n, \mathbf{k})$ contains an (i, j, β)-path with probability at least 1/8.

Next, we will call a path in $\vec{H}(n, \mathbf{k})$ an *extended* (i, j, β)*-path* if it is an (i, j, β)-path extended by one vertex at the end to $\beta \mathbf{e}_j + v$. The line containing $\beta \mathbf{e}_j + v$ and $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j + v$ could never be used in any (i, j, β)-path, so the probability that $\vec{L}(n, \mathbf{k})$ contains an extended (i, j, β)-path is at least $1/(8k_i)$.

To conclude our definitions, a path in $\vec{H}(n, \mathbf{k})$ will be called an (i, j)*-path* if there exists $\beta \in [k_j - 1]$ for which it is an extended (i, j, β)-path extended by one vertex at the start to $\alpha_i \mathbf{e}_i$. Note that the line through $\alpha_i \mathbf{e}_i$ in coordinate j cannot be used in an extended (i, j, β)-path for any $\beta \in [k_j - 1]$. Hence, the probability that $\vec{L}(n, \mathbf{k})$ contains an (i, j)-path is the probability that the winner of the line through $\alpha_i \mathbf{e}_i$ in coordinate j

is some $\alpha_i \mathbf{e}_i + \beta \mathbf{e}_j$ with $\beta \in [k_j - 1]$, multiplied by the probability that $\vec{L}(n, \mathbf{k})$ contains an extended (i, j, β)-path for this β. By the above, this probability is at least

$$\frac{1}{8k_i}\left(1 - \frac{1}{k_j}\right) \geq \frac{1}{16k_i}.$$

Next, observe that every line that could possibly be used in an (i, j)-path identifies the set $\{i, j\}$. There are at least $|A|\,(|B| - 1)/2 \geq n^2/(50K)$ ways to choose $\{i, j\}$, so the probability that $\vec{L}(n, \mathbf{k})$ does not contain an (i, j)-path for any (i, j) is at most

$$\left(1 - \frac{1}{16k_i}\right)^{n^2/(50K)} \leq e^{-n^2/(800K^2)}$$

Observe also that the event that there exists an (i, j)-path in $\vec{L}(n, \mathbf{k})$ for some (i, j) is independent of the behaviour of any lines of $V(n, \mathbf{k})$ containing u or v.

This analysis holds for any choice of A' and B' containing at least $n/4$ and $n/3K$ directions respectively, so if u and v are good vertices, then we may take A' to be the set of coordinate directions in which u does not win its line, and B' to be the set of directions in which v does win its line. For each $i \in A'$, let $\alpha_i \in [k_i - 1]$ be the ith coordinate of the winner of the line through u in direction i. This means that the edges from u to $\alpha_i \mathbf{e}_i$ and from $\beta \mathbf{e}_j + v$ to v are both present for any choice of $i \in A$, $j \in B$, and $\beta \in [k_j - 1]$, so if there is no path from u to v in $\vec{L}(n, \mathbf{k})$, then there is no (i, j)-path in $\vec{L}(n, \mathbf{k})$ for any (i, j). We have shown that this happens with probability at most $e^{-n^2/(800K^2)}$.

By a similar argument, the same holds for every pair of vertices at distance 1 or 2 from each other. A union bound yields that the probability that there exists a good vertex which cannot reach in $\vec{L}(n, \mathbf{k})$ some other good vertex within $H(n, \mathbf{k})$-distance 3 of it is at most $K^{2n} \cdot e^{-n^2/(800K^2)}$. Clearly $e^{-n^2/(800K^2)}$ is extremely small, so Lemma 17 implies that $K^{2n} \cdot e^{-n^2/(800K^2)}$ is also extremely small, which completes the proof of the lemma. □

E Proof of Theorem 15: accessing good vertices

Recall that in Section A.1 and Appendix D we defined a vertex of $\vec{L}(n, \mathbf{k})$ to be *good* if it wins at least $n/3K$ but at most $3n/4$ of its lines. We will start by considering the following event.

Definition 31 (Event C)**.** Let C be the event that every vertex in $V(n, \mathbf{k})$ that can be reached from more than $n/2$ vertices in $\vec{L}(n, \mathbf{k})$ can be reached from a good vertex.

We show that C is very likely.

Lemma 32. *The event C occurs with extremely high probability.*

Proof. We need to show that if $\delta > 0$ is sufficiently small, then whenever $K \leq \delta\sqrt{n/\log(n)}$, we have $\mathbb{P}(C^c) \leq e^{-cn\log(K)/\delta}$ for some $c > 0$. To this end, let $\delta > 0$ be small and assume that $K \leq \delta\sqrt{n/\log(n)}$. For each $y \in V(n, \mathbf{k})$, define C_y to be the event that either y can be reached from at most $n/2$ vertices in $\vec{L}(n, \mathbf{k})$, or y can be reached from a good vertex in $\vec{L}(n, \mathbf{k})$

Let $M = \lfloor n/3 \rfloor$. Fix y and let S be the set of trees of order M in $H(n, \mathbf{k})$ that contain y. For a tree $T \in S$, let C_T be the event that all the edges of T are oriented towards y in $\vec{L}(n, \mathbf{k})$ (so in particular, if C_T holds, then all vertices of T can reach y). If $(C_y)^c$ holds, then y can be reached from more than $n/2 \geq M$ vertices, so C_T occurs for some $T \in S$. Hence, by a union bound

$$\mathbb{P}\big((C_y)^c\big) = \mathbb{P}\left((C_y)^c \cap \left(\bigcup_{T \in S} C_T\right)\right) \leq \sum_{T \in S} \mathbb{P}\big((C_y)^c \cap C_T\big). \tag{2}$$

For fixed $T \in S$, if $(C_y)^c$ and C_T both hold, then no vertex in T is good. Since T contains exactly M vertices, each of its vertices can be assigned a set of $n - M = \lceil 2n/3 \rceil$ lines which contain that vertex and no other vertex of T (so that the number of lines that each vertex of T wins out of the $\lceil 2n/3 \rceil$ assigned to it is independent). For a fixed vertex v of T, let X be the number of the $\lceil 2n/3 \rceil$ lines assigned to v in which v is the winner, so that X is a sum of $\lceil 2n/3 \rceil$ independent Bernoulli random variables and has mean $\mu \geq 2n/3K$.

By Lemma 19,

$$\mathbb{P}\left(X \leq \frac{n}{3K}\right) \leq \mathbb{P}\left(X \leq \frac{\mu}{2}\right) \leq e^{-\mu/8} \leq e^{-n/(12K)}.$$

Meanwhile, the number of the lines assigned to v in which v is not the winner, $\lceil 2n/3 \rceil - X$, is a sum of $\lceil 2n/3 \rceil$ independent Bernoulli random variables with mean $\mu' \geq n/3$. Hence, by Lemma 19 again,

$$\mathbb{P}\Big(X > \frac{5n}{12}\Big) \leq \mathbb{P}\Big(\Big\lceil \frac{2n}{3} \Big\rceil - X \leq \frac{n}{4}\Big) \leq \mathbb{P}\Big(\Big\lceil \frac{2n}{3} \Big\rceil - X \leq \frac{3\mu'}{4}\Big) \leq e^{-\mu'/32} \leq e^{-n/96}.$$

If $n/3K < X \leq 5n/12$, then in total v wins at least $n/3K$ and at most $5n/12 + \lfloor n/3 \rfloor \leq 3n/4$ of its lines, and hence is a good vertex. It follows that each vertex of T is good with failure probability at most $e^{-n/(50K)}$ (for $n \geq 101$), so at least one of these vertices is good with failure probability at most $e^{-Mn/(50K)} \leq e^{-n^2/(200K)}$ (by choosing δ small enough, we can ensure $M \geq n/4$). This is therefore an upper bound on $\mathbb{P}((C_y)^c \cap C_T)$. It follows from Lemma 23 that $|S| \leq (enK)^{M-1}$, so by (2) we have

$$\begin{aligned}
\mathbb{P}\big((C_y)^c\big) &\leq (enK)^{M-1} \cdot e^{-n^2/(200K)} \\
&\leq \exp\left(\frac{n\log(enK)}{3} - \frac{n^2}{200K}\right) \\
&= \exp\left(\frac{n^2}{K}\left(\frac{K\log(enK)}{3n} - \frac{1}{200}\right)\right).
\end{aligned}$$

If δ is small enough, then n is large relative to $K\log(enK)$, so this probability is at most $e^{-n^2/(300K)}$.

Finally, since $C = \bigcap_y C_y$, by a union bound we have

$$\mathbb{P}(C^c) \leq K^n \cdot e^{-n^2/(300K)}.$$

Clearly $e^{-n^2/(300K)}$ is extremely small, so by Lemma 17 the same is true of $K^n \cdot e^{-n^2/(300K)}$, and the lemma follows. □

Next, we prove a very similar result for vertices that can *reach* more than $n/2$ vertices rather than can be *reached from* more than $n/2$ vertices.

Definition 33 (Event D)**.** Let D be the event that every vertex in $\vec{L}(n,\mathbf{k})$ that can reach more than $n/2$ vertices can reach a good vertex.

By an argument similar to that used to prove Lemma 32, we obtain the following.

Lemma 34. *The event D occurs with extremely high probability.*

We are now ready to put everything together to prove the main theorem.

Proof of Theorem 15. Let $n \geq 2$ be an integer, let $\mathbf{k} \in \{2,3,\dots\}^n$, and let $\varepsilon > 0$. Define $K = \max_i(k_i)$ and $N = (1+\varepsilon)K\log(K)$. Let E_ε be the event that every vertex of $\vec{L}(n,\mathbf{k})$ can either be reached from at most N vertices or can be reached from every non-sink; we want to show that E_ε occurs wvhp.

Let events A_ε, B, C, and D be as above, and suppose that they all occur simultaneously in $\vec{L}(n,\mathbf{k})$. Let $x, y \in V(n,\mathbf{k})$ where x is a non-sink and y can be reached from more than N vertices. Since A_ε occurs, x and y can reach and be reached from more than $n/2$ vertices respectively. Thus, since C occurs, there exists a good vertex $v \in V(n,\mathbf{k})$ which can reach y in $\vec{L}(n,\mathbf{k})$. Next, since D occurs, there exists a good vertex $u \in V(n,\mathbf{k})$ which can be reached from x in $\vec{L}(n,\mathbf{k})$. Since B occurs, all good vertices of $\vec{L}(n,\mathbf{k})$ are in the same strongly connected component, so in particular u can reach v. It follows that there is a directed walk from x to y in $\vec{L}(n,\mathbf{k})$ via u and v, that is, x can reach y. In other words, if A_ε, B, C, and D occur, then so does E_ε. By Lemmas 28, 32, and 34, each of B, C, and D occurs wehp, so in particular wvhp, and A_ε occurs wvhp by Lemma 22. Hence, by (repeated applications of) Lemma 18, we conclude that E_ε occurs wvhp, as required. □

F On the tightness of Theorems 5 and 15

In this section we discuss to what extent various aspects of Theorem 5 and Theorem 15 are tight. First, with regards to the relationship between n and K, it is entirely possible that this condition could be weakened considerably while still allowing results in the spirit of Theorem 5 and Theorem 15. However, the condition given in Theorem 15 seems to be at the limit of our methods, and a new approach would be needed to improve it. See also point (i) in Section 6.

Next, with regards to the tightness of the failure probability in the theorems, first note that the failure probability in Theorem 15 carries through to Theorem 5 by the proof of Corollary 16. The following theorem shows that this slightly stronger lower bound is tight up to the value of the exponent c. In fact, the theorem shows that even if Theorem 5 were weakened to only consider weakly acyclic games rather than connected games, the probability inherited from Theorem 15 would still be tight up to the value of the constant in the exponent.

Theorem 35. *There is a constant $c' > 0$ such that*

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is weakly acyclic}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is has a pure Nash equilibrium}\}|} \leq 1 - \prod_{i=1}^{n} k_i^{-c'}$$

for all integers $n \geq 4$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$.

Proof. Let n and $\mathbf{k}$ be as in the statement of the theorem. Define a 4-cycle in a subgraph of $\vec{H}(n, \mathbf{k})$ to be *sticky* if each of its vertices can only reach the other vertices of the 4-cycle. The probability that a given sticky 4-cycle (where the edges in the cycle are the first and second coordinate directions, say) appears in $\vec{L}(n, \mathbf{k})$ is $k_1^{-2} k_2^{-2} \prod_{i=3}^{n} k_i^{-4} \geq \prod_{i=1}^{n} k_i^{-4}$. Since $n \geq 4$, there is a vertex none of whose lines intersect this sticky 4-cycle. The event that this vertex is a sink, which occurs with probability $\prod_{i=1}^{n} k_i^{-1}$, is therefore independent of whether or not the sticky 4-cycle appears, so with probability at least $\prod_{i=1}^{n} k_i^{-5}$ there is both a sink and a sticky 4-cycle in $\vec{L}(n, \mathbf{k})$.

The vertices of a sticky 4-cycle cannot reach a sink, so the result now follows from arguments similar to those used to prove Corollary 16. □

One might ask whether taking a larger value for N in the statement of Theorem 15 would allow a significantly smaller failure probability, but a similar argument to the proof of Theorem 35 shows that it would not. Indeed, there is some $c' > 0$ such that if n is large relative to $\max_i(k_i)$, then with probability at least $1 - \prod_{i=1}^{n} k_i^{-c'}$ there is a vertex in $\vec{L}(n, \mathbf{k})$ which can be reached from $\prod_{i=1}^{n-1} k_i$ vertices but not from every non-sink.

This follows from an argument similar to the above: suppose that the desired sticky 4-cycle has vertices $(1, 1, 1, \ldots, 1), (1, 2, 1, \ldots, 1), (2, 1, 1, \ldots, 1)$, and $(2, 2, 1, \ldots, 1)$, then the subgraph G of $\vec{L}(n, \mathbf{k})$ induced on $\prod_{i=1}^{n-1}[k_i] \times \{2\}$ has the same distribution as $\vec{L}(n-1, (k_1, \ldots, k_{n-1}))$, and behaves independently of whether the desired sticky 4-cycle appears or not. Applying Theorem 15 to G and using work of Rinott and Scarsini (2000), one can show that there exists $p > 0$ such that if n is large enough relative to $\max_i(k_i)$, then with probability at least p, G contains exactly one sink and this can be reached from every vertex in G. It follows that there exists $c' > 0$ such that if n is large relative to $\max_i(k_i)$, then with probability at least $\prod_{i=1}^{n} k_i^{-c'}$ there is a vertex in $\vec{L}(n, \mathbf{k})$ which can be reached from $\prod_{i=1}^{n-1} k_i$ vertices but not from every non-sink.

Finally, to what extent is it possible to take a smaller N in the statement of Theorem 15? It turns out that, for large K, the value of N cannot be substantially improved as a function of K: it is possible to take r not much smaller than $\log(K-1)$ in the following theorem,[23] so one cannot hope for a value of N any better than $K\log(K) - O(\log(K))$. Here we let $\mathbf{K} = (K, \ldots, K)$ denote the all K's vector of the appropriate length.

Theorem 36. *There is a constant $c > 0$ such that for all integers $n \geq 2$, $2 \leq K \leq \sqrt{n}$, and*

$$1 \leq r \leq \frac{\log(K-1)}{(K-1)(\log(K) - \log(K-1))},$$

the probability that there is a vertex in $\vec{L}(n, \mathbf{K})$ which can be reached from exactly $r(K-1)+1$ vertices is at least $1 - c/n$.

Proof. Let $f(K)$ denote the expression upper bounding r in the theorem. One can show that this is increasing for $K \geq 2$ and that $f(\sqrt{n}) \leq n$ for $n \geq 2$, so letting n, K, and r be as in the statement, we have $r \leq n$. Let X_a be the indicator random variable of the event that $a \in V(n, \mathbf{K})$ wins exactly r of its lines and every vertex on those r lines except a is a source, and write $X = \sum_{a \in [K]^n} X_a$. We wish to upper bound the probability that $X = 0$, for which we will use a second moment calculation.

First, note that X_a and X_b are independent if the Hamming distance between a and b (i.e. the number of coordinates on which a and b differ), denoted by $d(a,b)$, is at least four. It follows that

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{a \in [K]^n} \sum_{b \in [K]^n} \mathbb{P}(X_a X_b = 1) \\
&\leq \sum_{a \in [K]^n} \sum_{b \in [K]^n} \mathbb{P}(X_a = 1)\mathbb{P}(X_b = 1) + \sum_{a \in [K]^n} \sum_{b \in [K]^n : d(a,b) \leq 3} \mathbb{P}(X_a = 1) \\
&\leq \mathbb{E}[X]^2 + K^3 n^3 \,\mathbb{E}[X] \\
&\leq \mathbb{E}[X]^2 + n^{9/2}\,\mathbb{E}[X].
\end{aligned}$$

Hence, to apply Chebyshev's inequality, we need to show that $\mathbb{E}[X]$ grows more quickly than $n^{9/2}$. We have

$$\begin{aligned}
\mathbb{E}[X] &= K^n \binom{n}{r} \frac{1}{K^r} \left(1 - \frac{1}{K}\right)^{n-r+r(n-1)(K-1)} \\
&\geq \left(\frac{n}{Kr}\right)^r \left[K\left(1 - \frac{1}{K}\right)^{1+(K-1)r}\right]^n \\
&\geq \left(\frac{\sqrt{n}}{r}\right)^r \left[K\left(1 - \frac{1}{K}\right)^{1+(K-1)r}\right]^n,
\end{aligned}$$

where we have used $\binom{n}{r} \geq (n/r)^r$ in the second line and $K \leq \sqrt{n}$ in the last line.

[23] By applying the mean value theorem to log one can show that $\log(K) - \log(K-1) = 1/K + O(1/K^2)$.

We will analyse the two terms in this product separately. For fixed n, the first term, $(\sqrt{n}/r)^r$, is increasing for $r \in [0, \sqrt{n}/e]$. Since $f(\sqrt{n}) \leq \sqrt{n}/e$ for all $n \geq 2$, it follows that this term is always at least $\sqrt{n} \geq 1$, and if $r \geq 11$, then it is at least $n^{11/2}/11^{11}$.

Turning to the second term, it is straightforward to check that if $f(K) = r$, then the expression in square brackets is equal to 1, and that for fixed $K \geq 2$ this expression is strictly decreasing for $r \in [1, f(K)]$. It follows that the second term is always at least 1. There are no integer solutions K to $f(K) = r$ for any $r \in [10]$, so for $r \leq 10$ the expression in square brackets is always strictly greater than 1. Moreover, for fixed r we have $K(1 - 1/K)^{1+r(K-1)} \to \infty$ as $K \to \infty$, so in fact there exists some universal $\varepsilon > 0$ such that the second term is at least $(1 + \varepsilon)^n$ whenever $r \leq 10$.

Combining, we have $\mathbb{E}[X] \geq \min\left\{(1+\varepsilon)^n, n^{11/2}/11^{11}\right\}$ for all admissible n, K, and r. By Chebyshev's inequality, this yields

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq \tfrac{1}{2}\mathbb{E}[X]\right) \leq 4 \cdot \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{\mathbb{E}[X]^2} \leq \frac{4n^{9/2}}{\min\left\{(1+\varepsilon)^n, n^{11/2}/11^{11}\right\}},$$

and since $\mathbb{E}[X] > 0$ this is, in turn, an upper bound on $\mathbb{P}(X = 0)$. There is a constant $c' > 0$ such that this upper bound is at most c'/n for all sufficiently large n, and we can choose $c > 0$ large enough to accommodate the (finitely many) remaining cases. □

If, however, we are prepared to allow the exponent in the failure probability to depend on K, then we can adapt the proof of Theorem 15 to slightly improve the value of N.

Theorem 37. *For all integers $K \geq 2$, there exists $c_K > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots K\}^n$, every vertex of $\vec{L}(n, \mathbf{k})$ can either be reached from at most*

$$N' = \frac{\log(K)}{\log(K) - \log(K-1)}$$

vertices or from every non-sink, with failure probability at most $e^{-c_K n}$.

Proof. Note first that it is sufficient to show that the result holds when n is large relative to K, since this covers all but finitely many cases for each K, and c_K can be chosen to handle these.

By the $\varepsilon = 1$ case of Theorem 15, there exists $c > 0$ such that if n is large relative to K, then for all $\mathbf{k} \in \{2, \ldots, K\}^n$, the failure probability of the event that every vertex of $\vec{L}(n, \mathbf{k})$ can either be reached from at most $2K\log(K)$ vertices, or from every non-sink, is at most e^{-cn}. Hence, to prove the theorem it is enough to show that for all K there exists $c'_K > 0$ such that if n is large enough relative to K, then for all $\mathbf{k} \in \{2, \ldots, K\}^n$, with failure probability at most $e^{-c'_K n}$, no vertices of $\vec{L}(n, \mathbf{k})$ can be reached from more than N' vertices but at most $2K\log(K)$ vertices.

This can be achieved by modifying the proof of Lemma 25. Defining Y_m as in Appendix C, we need to show that $Y_m = 0$ for all $N' < m \leq 2K\log(K)$ with failure probability $e^{-c'_K n}$. Fix such an m, then as in the proof of Lemma 25 we have

$$\mathbb{E}[Y_m] \le \left[4nK^2 \cdot \left(K^{1/m}\left(1-\frac{1}{K}\right)\right)^{n-m}\right]^m.$$

In place of (1), it is not difficult to check that $m > N'$ ensures that $K^{1/m}(1-1/K) < 1-\eta_K$ for some $\eta_K \in (0,1)$. Thus, for n large in terms of K (uniformly in m), we have

$$\mathbb{E}[Y_m] \le \left(4nK^2(1-\eta_K)^{n-m}\right)^m \le \left(1-\eta_K^2\right)^{m(n-m)} \le e^{-c''_K n},$$

for some $c''_K > 0$. Then, if n is large relative to K,

$$\sum_{N<m\le 2K\log(K)} \mathbb{E}[Y_m] \le 2K\log(K)\cdot e^{-c''_K n} \le e^{-c'_K n},$$

for some $c'_K > 0$, so by Lemma 20 we have that $Y_m = 0$ for all $N' < m \le 2K\log(K)$ with failure probability at most $e^{-c'_K n}$, as required. □

Together, Theorem 36 and Theorem 37 essentially determine the 'correct' value for N as a function of K. Indeed, if we ignore the fact that r must be an integer in Theorem 36, then that theorem implies that when n is much larger than K, $\vec{L}(n,\mathbf{K})$ typically contains a vertex that can be reached from 'exactly' $\log(K)/(\log(K)-\log(K-1))$ vertices. Meanwhile Theorem 37 implies that typically every vertex which can be reached from more than this many vertices can be reached from every non-sink.

Although the improvement to the value of N represented by Theorem 37 is modest, it has the following consequence for $\vec{L}(n,\mathbf{2})$ and $\vec{L}(n,\mathbf{3})$ which is of independent interest.

Corollary 38. *There exists a constant $c > 0$ such that with failure probability at most e^{-cn}, every non-sink in $\vec{L}(n,\mathbf{2})$ can reach every non-source. If $n \ge 2$, the same is true for $\vec{L}(n,\mathbf{3})$. Conversely, for each $K \ge 4$, the probability that there is a non-sink in $\vec{L}(n,\mathbf{K})$ which cannot reach every non-source tends to 1 as $n \to \infty$.*

Corollary 38 is restated as Proposition 7 in the main text.

The positive direction of the corollary follows straightforwardly from Theorem 37 and the observation that every non-source in $\vec{L}(n,\mathbf{k})$ can be reached from at least $\min_i(k_i)$ vertices, and the negative direction follows immediately from setting $r = 1$ in Theorem 36. While the $\mathbf{k} = \mathbf{2}$ case of the corollary follows from Theorem 15, the $\mathbf{k} = \mathbf{3}$ case does not.

G Acyclicity of directed grids

Using the terminology and notation of Section A we state the following strengthening of Proposition 6.

Proposition 39. *There exists $c > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \dots\}^n$, the probability that $\vec{L}(n, \mathbf{k})$ is acyclic is at most $\exp(-cnk^{n-2})$, where $k := \min_i(k_i)$.*

The proof below actually yields a slightly better upper bound on the probability that $\vec{L}(n, \mathbf{k})$ is acyclic, but for clarity we have not included this in the statement.

For distinct $i, j \in [n]$, we define an $\{i, j\}$-*plane* of $V(n, \mathbf{k})$ to be a subset of $V(n, \mathbf{k})$ of size $k_i k_j$ whose elements pairwise differ in at most their ith and jth coordinates. A subset of $V(n, \mathbf{k})$ will be called a *plane* of $V(n, \mathbf{k})$ if it is an $\{i, j\}$-plane for some i and j.

Proof of Proposition 39. We begin with the following claim.

Claim 2. *Let $k_1, k_2 \geq 2$ be integers, then $\vec{L}(2, (k_1, k_2))$ contains a cycle with probability at least $1/8$.*

Proof. We will define a random process $\mathbf{X} = (X_0, X_1, X_2, \dots)$ coupled to $\vec{L}(2, (k_1, k_2))$. Let $X_0 = (1, 1)$ and for each $t \geq 1$, given $X_{t-1} \in [k_1] \times [k_2]$, if t is odd, let X_t be the winner of the line in coordinate 1 which contains X_{t-1}. If t is even, let X_t be the winner of the line in coordinate 2 which contains X_{t-1}. Thus, $\mathbf{X}$ is a random walk on $[k_1] \times [k_2]$ starting at $(1, 1)$, which at odd time steps traverses the available edge of $\vec{L}(2, (k_1, k_2))$ in the first coordinate direction (if there is one), and at even time steps traverses the available edge in the second.

Let T be the least time t at which there exists $1 \leq i < t$ such that X_i and X_t have the same first coordinate if t is odd, or the same second coordinate if t is even. It is not difficult to check (bearing in mind that we do not explore from X_0 in the second dimension at the start of the process) that for each $t < T$, when we explore from X_t we do so in an unexplored line.

Once we reach X_T, we already know the winner in the line we want to explore: it is the X_i with $1 \leq i < T$ that has the same first or second coordinate (depending on the parity of T) as X_T. Hence, $X_{T+1} = X_i$ for this i and the process is deterministic from here, with $X_{T+2} = X_{i+1}$ and so on. The process either becomes stationary at this point or enters a (non-trivial) cycle. It is straightforward to see that the process is stationary exactly when $X_T = X_{T-1}$, so if $X_T \neq X_{T-1}$, then $\vec{L}(2, (k_1, k_2))$ contains a cycle. Hence, let A be the event that $X_T \neq X_{T-1}$. We will show that $\mathbb{P}(A) \geq 1/8$.

Given T and X_{T-1}, X_T is chosen uniformly at random from among those vertices in the unexplored line through X_{T-1} that also belong to a previously explored line. There are $\lfloor T/2 \rfloor$ such vertices to choose from. Writing $\tau = \tau(k_1, k_2)$ for the maximum possible value of T, it follows that, for each $t \in \{2, \dots, \tau\}$, we have $\mathbb{P}(A \,|\, T = t) = (\lfloor t/2 \rfloor - 1)/\lfloor t/2 \rfloor$. This is at least $1/2$ for $t \geq 4$, so

$$\mathbb{P}(A) = \sum_{t=2}^{\tau} \mathbb{P}(A \,|\, T = t)\mathbb{P}(T = t) \geq \frac{\mathbb{P}(T \geq 4)}{2}.$$

We have $T \geq 4$ exactly when X_1 does not win its line in the second dimension and X_2 does not win its line in the first dimension. This event occurs with probability $(1 - 1/k_1)(1 - 1/k_2) \geq 1/4$, so $\mathbb{P}(A) \geq 1/8$, as required. ◻

Let $n \geq 2$ be an integer and let $\mathbf{k} \in \{2, 3, \ldots\}^n$. By the claim, any given plane of $V(n, \mathbf{k})$ induces a cyclic subgraph of $\vec{L}(n, \mathbf{k})$ with probability at least $1/8$. In a family of planes which pairwise intersect in at most one vertex, each plane induces a cyclic subgraph of $\vec{L}(n, \mathbf{k})$ independently. The collection consisting of all $\{1, 2\}$-planes, all $\{3, 4\}$-planes, and so on, up to the $\{2\lfloor n/2 \rfloor - 1, 2\lfloor n/2 \rfloor\}$-planes, is such a family. For distinct $i, j \in [n]$, the number of $\{i, j\}$-planes in $V(n, \mathbf{k})$ is $\prod_{a \in [n] \setminus \{i,j\}} k_a$, so this family has size at least $\lfloor n/2 \rfloor \min(k_i)^{n-2}$, and the proposition follows. □

Proposition 6 follows from Proposition 39 by the same reasoning with which we deduced Theorem 5 from Theorem 15: since $\vec{L}(n, \mathbf{k})$ has the same distribution as the best-response graph of a game drawn uniformly at random from all games in $\mathcal{G}(n, \mathbf{k})$, we have that $\mathbb{P}\big(\vec{L}(n, \mathbf{k}) \text{ is acyclic}\big) / \mathbb{P}(S_{n,\mathbf{k}})$ is equal to

$$\frac{|\{g \in \mathcal{G}(n, \mathbf{k}) \colon g \text{ is acyclic}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k}) \colon g \text{ has a pure Nash equilibrium}\}|},$$

where $S_{n,\mathbf{k}}$ is the event that $\vec{L}(n, \mathbf{k})$ contains a sink, as in Section A. Thus, since $k_i \geq 2$ for all i, Proposition 6 follows from Proposition 39 and the fact that $\mathbb{P}(S_{n,\mathbf{k}})$ is at least a positive constant for all n and $\mathbf{k}$ under consideration, as noted in Section A.

H Best-response dynamics with inertia

In this section we recap the proof by Young (2004) to show that the best-response dynamic with inertia converges almost surely to a pure Nash equilibrium in every generic weakly acyclic game.

Theorem 40 (Young, 2004)**.** *For any choice of parameters $p_i \in (0, 1)$, the best-response dynamic with inertia converges almost surely to a pure Nash equilibrium in every generic weakly acyclic game.*

Proof. Let g be a generic weakly acyclic game and let A be its set of action profiles. Fix some $a \in A$ which is not a pure Nash equilibrium, let t be an arbitrary time, and condition on the event that $a^t = a$. The vertex corresponding to a in the best-response graph of g is not a sink, so let i be a coordinate direction in which it has an outgoing edge. Then there exists some $\varepsilon > 0$ which is independent of a, t and i, such that with probability at least ε, at time $t + 1$ player i changes their action by playing the best-response to a_{-i}, while all other players j repeat their existing action. In other words, after conditioning on $a^t = a$ for some non-sink a, the probability that at time step $t + 1$ we move along any given out-edge of a in the best-response graph is at least ε.

Since g is weakly acyclic, for each $a \in A$ there is a path of length at most $|A|$ from a to a sink in the best-response graph. After fixing such a path for each a, we can repeatedly apply the above to lower bound the probability that at each step we move along the next edge in that path. Using the fact that the dynamic never leaves a pure Nash equilibrium once it arrives at one, this yields that for all times t and all

$a \in A$, conditioned on $a^t = a$, the probability that $a^{t+|A|}$ is a pure Nash equilibrium is at least $\varepsilon^{|A|}$. For each $m \in \{0, 1, 2, \dots\}$, denote by B_m the event that $a^{m|A|}$ is not a sink. It follows from the above that $\mathbb{P}(B_m \mid B_{m-1}) \leq 1 - \varepsilon^{|A|}$ for each $m \geq 1$, and so an inductive argument gives $\mathbb{P}(B_m) \leq (1 - \varepsilon^{|A|})^m$. If the dynamic does not eventually settle at a pure Nash equilibrium, then B_m occurs for all m, but this has probability 0 since $(1 - \varepsilon^{|A|})^m \to 0$ as $m \to \infty$. This completes the proof of the theorem. □

References

Alon, N., K. Rudov, and L. Yariv (2021). Dominance solvability in random games. *arXiv:2105.10743* preprint.

Amiet, B., A. Collevecchio, and K. Hamza (2021). When better is better than best. *Operations Research Letters 49*(2), 260 – 264.

Amiet, B., A. Collevecchio, M. Scarsini, and Z. Zhong (2021). Pure Nash equilibria and best-response dynamics in random games. *Mathematics of Operations Research 46*(4), 1552 – 1572.

Apt, K. R. and S. Simon (2015). A classification of weakly acyclic games. *Theory and Decision 78*(4), 501 – 524.

Arieli, I. and H. P. Young (2016). Stochastic learning dynamics and speed of convergence in population games. *Econometrica 84*(2), 627 – 676.

Arratia, R., L. Goldstein, and L. Gordon (1989). Two moments suffice for Poisson approximations: the Chen-Stein method. *The Annals of Probability 17*(1), 9 – 25.

Awerbuch, B., Y. Azar, A. Epstein, V. S. Mirrokni, and A. Skopalik (2008). Fast convergence to nearly optimal solutions in potential games. In *Proceedings of the 9th ACM conference on Electronic Commerce*, pp. 264 – 273.

Babichenko, Y. (2012). Completely uncoupled dynamics and Nash equilibria. *Games and Economic Behavior 76*(1), 1 – 14.

Babichenko, Y. (2013). Best-reply dynamics in large binary-choice anonymous games. *Games and Economic Behavior 81*, 130–144.

Bárány, I., S. Vempala, and A. Vetta (2007). Nash equilibria in random games. *Random Structures & Algorithms 31*(4), 391–405.

Bollobás, B., C. Gotsman, and E. Shamir (1993). Connectivity and dynamics for random subgraphs of the directed cube. *Israel Journal of Mathematics 83*(3), 321 – 328.

Candogan, O., I. Menache, A. Ozdaglar, and P. A. Parrilo (2011). Flows and decompositions of games: Harmonic and potential games. *Mathematics of Operations Research 36*(3), 474 – 503.

Cesa-Bianchi, N. and G. Lugosi (2006). *Prediction, Learning, and Games*. Cambridge University Press.

Collevecchio, A., H. A. Mimun, M. Quattropani, and M. Scarsini (2024). Basins of attraction in two-player random ordinal potential games. *arXiv preprint arXiv:2407.05460*.

Collevecchio, A., T.-M. Nguyen, and Z. Zhong (2024). Finding pure nash equilibria in large random games. *arXiv preprint arXiv:2406.09732*.

Dresher, M. (1970). Probability of a pure equilibrium point in n-person games. *Journal of Combinatorial Theory 8*(1), 134 – 145.

Ellison, G. (2000). Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *The Review of Economic Studies 67*(1), 17–45.

Fabrikant, A., A. D. Jaggard, and M. Schapira (2013). On the structure of weakly acyclic games. *Theory of Computing Systems 53*(1), 107 – 122.

Foster, D. P. and R. V. Vohra (1997). Calibrated learning and correlated equilibrium. *Games and Economic Behavior 21*(1 – 2), 40 – 55.

Foster, D. P. and H. P. Young (2006). Regret testing: Learning to play Nash equilibrium without knowing you have an opponent. *Theoretical Economics 1*(3), 341 – 367.

Freidlin, M., J. Szucs, and A. Wentzell (2012). *Random Perturbations of Dynamical Systems*. Grundlehren der mathematischen Wissenschaften. Springer New York.

Friedman, J. W. and C. Mezzetti (2001). Learning in games by random sampling. *Journal of Economic Theory 98*(1), 55 – 84.

Frieze, A. and M. Karoński (2015). *Introduction to Random Graphs*. Cambridge University Press.

Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games*. MIT Press.

Fudenberg, D. and D. K. Levine (1999). Conditional universal consistency. *Games and Economic Behavior 29*(1 – 2), 104 – 130.

Germano, F. and G. Lugosi (2007). Global Nash convergence of Foster and Young's regret testing. *Games and Economic Behavior 60*(1), 135 – 154.

Goldberg, K., A. Goldman, and M. Newman (1968). The probability of an equilibrium point. *Journal of Research of the National Bureau of Standards 72*(2), 93 – 101.

Hart, S. (2005). Adaptive heuristics. *Econometrica 73*(5), 1401 – 1430.

Hart, S. and Y. Mansour (2010). How long to equilibrium? The communication complexity of uncoupled equilibrium procedures. *Games and Economic Behavior 69*(1), 107 – 126.

Hart, S. and A. Mas-Colell (2000). A simple adaptive procedure leading to correlated equilibrium. *Econometrica 68*(5), 1127 – 1150.

Hart, S. and A. Mas-Colell (2001). A reinforcement procedure leading to correlated equilibrium. In *Economics essays*, pp. 181 – 200. Springer.

Hart, S. and A. Mas-Colell (2003). Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review 93*(5), 1830 – 1836.

Hart, S. and A. Mas-Colell (2006). Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economic Behavior 57*(2), 286 – 303.

Heinrich, T., Y. Jang, L. Mungo, M. Pangallo, A. Scott, B. Tarbush, and S. Wiese (2023). Best-response dynamics, playing sequences, and convergence to equilibrium in random games. *International Journal of Game Theory 52*, 703 – 735.

Hofbauer, J. and W. H. Sandholm (2002). On the global convergence of stochastic fictitious play. *Econometrica 70*(6), 2265 – 2294.

Hofbauer, J. and K. Sigmund (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

Jaggard, A. D., N. Lutz, M. Schapira, and R. N. Wright (2014). Self-stabilizing uncoupled dynamics. In *Algorithmic Game Theory: 7th International Symposium, SAGT 2014, Proceedings 7*, pp. 74 – 85. Springer.

Jiang, A. X., K. Leyton-Brown, and N. A. Bhat (2011). Action-graph games. *Games and Economic Behavior 71*(1), 141–173.

Kalai, E. and E. Lehrer (1993). Rational learning leads to Nash equilibrium. *Econometrica 61*(5), 1019 – 1045.

Kandori, M., G. J. Mailath, and R. Rob (1993). Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, 29–56.

Kandori, M. and R. Rob (1995). Evolution of equilibria in the long run: A general theory and applications. *Journal of Economic Theory 65*(2), 383–414.

Kearns, M. (2007). Graphical games. In *Algorithmic game theory*. Cambridge University Press.

Kearns, M., M. L. Littman, and S. Singh (2013). Graphical models for game theory. *arXiv preprint arXiv:1301.2281*.

Legacci, D., P. Mertikopoulos, and B. Pradelski (2024). A geometric decomposition of finite games: Convergence vs. recurrence under exponential weights. In *ICML 2024-41st International Conference on Machine Learning*.

Marden, J. R., G. Arslan, and J. S. Shamma (2007). Regret based dynamics: convergence in weakly acyclic games. In *Proceedings of the sixth international joint conference on Autonomous Agents and Multiagent Systems*, pp. 1 – 8.

Marden, J. R., H. P. Young, G. Arslan, and J. S. Shamma (2009). Payoff-based dynamics for multiplayer weakly acyclic games. *SIAM Journal on Control and Optimization 48*(1), 373 – 396.

McDiarmid, C., A. Scott, and P. Withers (2021). The component structure of dense random subgraphs of the hypercube. *Random Structures & Algorithms 59*(1), 3–24.

McLennan, A. (1997). The maximal generic number of pure Nash equilibria. *Journal of Economic Theory 72*(2), 408 – 410.

McLennan, A. (2005). The expected number of Nash equilibria of a normal form game. *Econometrica 73*(1), 141 – 174.

Milionis, J., C. Papadimitriou, G. Piliouras, and K. Spendlove (2023). An impossibility theorem in game dynamics. *Proceedings of the National Academy of Sciences 120*(41), e2305349120.

Mimun, H. A., M. Quattropani, and M. Scarsini (2024). Best-response dynamics in two-person random games with correlated payoffs. *Games and Economic Behavior 145*, 239–262.

Monderer, D. and L. S. Shapley (1996). Potential games. *Games and Economic Behavior 14*(1), 124 – 143.

Newton, J. (2018). Evolutionary game theory: A renaissance. *Games 9*(2), 31.

Newton, J. and R. Sawa (2024). Conventions and social choice in large games. *Available at SSRN*.

Pei, T. and S. Takahashi (2023). Nash equilibria in random games with right fat-tailed distributions. *International Journal of Game Theory 52*(4), 1153–1177.

Powers, I. Y. (1990). Limiting distributions of the number of pure strategy Nash equilibria in n-person games. *International Journal of Game Theory 19*(3), 277 – 286.

Pradelski, B. S. and H. P. Young (2012). Learning efficient Nash equilibria in distributed systems. *Games and Economic Behavior 75*(2), 882 – 897.

Rinott, Y. and M. Scarsini (2000). On the number of pure strategy Nash equilibria in random games. *Games and Economic Behavior 33*(2), 274 – 293.

Rosenthal, R. W. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory 2*(1), 65 – 67.

Roughgarden, T. (2016). *Twenty Lectures on Algorithmic Game Theory*. Cambridge University Press.

Roughgarden, T. (2021). *Beyond the worst-case analysis of algorithms*. Cambridge University Press.

Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. MIT Press.

Schipper, B. C. (2022). Strategic teaching and learning in games. *American Economic Journal: Microeconomics 14*(3), 321–352.

Stanford, W. (1995). A note on the probability of k pure Nash equilibria in matrix games. *Games and Economic Behavior 9*(2), 238 – 246.

Swenson, B., C. Eksin, S. Kar, and A. Ribeiro (2018). Distributed inertial best-response dynamics. *IEEE Transactions on Automatic Control 63*(12), 4294 – 4300.

Takahashi, S. and T. Yamamori (2002). The pure Nash equilibrium property and the quasi-acyclic condition. *Economics Bulletin 3*(22), 1 – 6.

Vlatakis-Gkaragkounis, E.-V., L. Flokas, T. Lianeas, P. Mertikopoulos, and G. Piliouras (2020). No-regret learning and mixed nash equilibria: They do not mix. *Advances in Neural Information Processing Systems 33*, 1380–1391.

Weibull, J. W. (1997). *Evolutionary Game Theory*. MIT Press.

Yongacoglu, B., G. Arslan, L. Pavel, and S. Yüksel (2024). Paths to equilibrium in normal-form games. *arXiv preprint arXiv:2403.18079*.

Young, H. P. (1993). The evolution of conventions. *Econometrica 61*(1), 57 – 84.

Young, H. P. (2004). *Strategic Learning and its Limits*. Oxford University Press.

Young, H. P. (2007). The possible and the impossible in multi-agent learning. *Artificial Intelligence 171*(7), 429 – 433.

Young, H. P. (2009). Learning by trial and error. *Games and Economic Behavior 65*(2), 626 – 643.