Game Connectivity and Adaptive Dynamics

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

19 September 2023

Abstract

We analyse the typical structure of games in terms of the connectivity properties of their best-response graphs. In particular, we show that almost every 'large' generic game that has a pure Nash equilibrium is *connected*, meaning that every non-equilibrium action profile can reach every pure Nash equilibrium via best-response paths. This has implications for dynamics in games: many adaptive dynamics, such as the best-response dynamic with inertia, lead to equilibrium in connected games. It follows 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. 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 simple adaptive dynamics that are guaranteed to lead to a Nash equilibrium in all games. Several influential results have outlined the boundary between the possible and the impossible, i.e. between classes of dynamics that are guaranteed to lead to a Nash equilibrium in all games and classes that lack such a guarantee (Young, 2007; Hart and Mas-Colell, 2013). For example, it has been demonstrated that there is no dynamic with low informational requirements on the players (i.e. no dynamic that is uncoupled or completely uncoupled) that is guaranteed to converge to a pure Nash equilibrium in every game that has one (Hart and Mas-Colell, 2003; Babichenko, 2012).

*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

We consider how this boundary shifts when the requirement of convergence in *all* games is relaxed to *almost all* games. Our approach is to quantitatively classify games based on the connectivity properties of their best-response graphs. We say that a game is *connected* if it possesses a pure Nash equilibrium and every non-equilibrium action profile can reach every pure Nash equilibrium via best-response paths. Our key result establishes that almost every 'large' generic game that has a pure Nash equilibrium is connected. Connectedness is a strong property that is conducive to convergent dynamics. For example, it follows from Young (2004) that there is a simple, uncoupled, adaptive dynamic (namely, the best-response dynamic with inertia) that leads almost surely to a pure Nash equilibrium in every connected generic game. Consequently, there is a simple, uncoupled, adaptive dynamic that is guaranteed to converge almost surely to a pure Nash equilibrium in *almost every* large generic game that has one, even though it has been demonstrated that no such guarantee can be given for *every* game that has one.

Our approach, leveraging recent results in probabilistic combinatorics (McDiarmid et al., 2021) to quantify the proportions of games that have certain connectivity properties, differs significantly from much of the literature on adaptive dynamics, where the properties of the dynamics themselves are the more usual focus.

We now provide a more detailed overview of our results regarding game connectivity and the consequences for adaptive dynamics.

Game connectivity Key strategic features of a game can be captured by its *best-response graph*. This is the graph whose set of vertices is the set of action profiles and whose directed edges correspond to best-responses (Young, 1993). Throughout, we focus on games in which each player's preferences are encoded ordinally (via a preference relation) rather than cardinally (via a utility function).

The behavior of many game dynamics are crucially determined by the connectivity properties of a game's best-response graph. Weakly acyclic games, for example, are those whose best-response graphs have the property that every vertex can reach a sink (i.e. a pure Nash equilibrium) along a directed path. It is well-known that weak acyclicity is a necessary condition for the convergence of best-response dynamics to a pure Nash equilibrium from any starting vertex (see e.g. Fabrikant et al., 2013; Apt and Simon, 2015). Young (1993), Friedman and Mezzetti (2001), Marden et al. (2007), and Marden et al. (2009) all study the convergence properties of different dynamics on the class of weakly acyclic games. Potential games (Monderer and Shapley, 1996), a subset of games that have acyclic best-response graphs, are another widely studied class of games. 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 (see e.g. Hofbauer and Sandholm, 2002; Roughgarden, 2016).

We introduce a new class of games which we refer to as *connected games*. These are those games whose best-response graphs contain a sink and have the property that every non-sink can reach *every* sink along a directed path. It is easy to see that a connected game must be weakly acyclic, but the converse need not hold.

We focus on generic games, i.e. those without indifferences. Our main result on game connectivity (Theorem 5) is that:

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

Here, 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). This result is striking given how strong the property of connectedness is.

We also show, by contrast, that the fraction of generic games with a pure Nash equilibrium that are acyclic vanishes super-exponentially in the number of players (Proposition 6). A consequence of this is that potential games with large numbers of players are very rare.

Our results on game connectivity (Theorem 5 and Proposition 6) are, in fact, corollaries of stronger results regarding the connectivity properties of random subgraphs of directed Hamming graphs (Theorem 8 and Proposition 14). Theorem 8 is the main technical contribution of our paper. It extends recent work of Amiet et al. (2021), and our proof of it adapts new results in probabilistic combinatorics regarding the component structure of random subgraphs of the hypercube (McDiarmid et al., 2021).

Adaptive dynamics Our results on game connectivity have implications for dynamics in games (Fudenberg and Levine, 1998; Young, 2004; Hart and Mas-Colell, 2000).

There are well-known possibility and impossibility results regarding dynamics in games. Establishing impossibility for a class of dynamics consists in 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.[1] With regards to (i), some of the most commonly studied classes of dynamics are those that are *uncoupled*, meaning that a player's strategy depends only on the actions of other players and on their own payoff function, or that are *completely uncoupled*, meaning that a player's strategy depends only on their own realised payoffs and actions.

There are several possibility results for dynamics that lead to mixed or correlated Nash equilibria,[2] but finding completely uncoupled, or even uncoupled, dynamics

[1]Examples of restrictions in each category include: (i) the length of a player's recall, whether the player's strategy is stationary, and whether it is uncoupled or completely uncoupled, (ii) almost sure convergence vs. being near an equilibrium 'most of the time', (iii) pure, mixed, correlated, or ε-Nash equilibrium, (iv) generic, dominance-solvable, potential, weakly acyclic, and so on.

[2]There are, for example, uncoupled and completely uncoupled dynamics for which the empirical distribution of play converges almost surely to the set of correlated Nash 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).

in which period-by-period play converges *almost surely* to a *pure* Nash equilibrium whenever one exists is demonstrably more challenging. The following impossibility result is well-known.[3,4]

> *Impossibility for uncoupled dynamics* (Hart and Mas-Colell, 2006; Jaggard et al., 2014) Given a stationary, 1-recall, uncoupled dynamic for a game that has at least three players,[5] there is a generic game with a pure Nash equilibrium that has the same action profile space, but for which the period-by-period play under the dynamic does not almost surely converge to a pure Nash equilibrium.

Our results on game connectivity imply that this impossibility result doesn't hold when we consider the 'typical' structure of many-player games. In particular, we show that

> *There exists a stationary, 1-recall, uncoupled dynamic for which the period-by-period play converges almost surely to a pure Nash equilibrium in* almost every *large generic game that has one.*

This follows from combining our result that almost every large generic game that has a pure Nash equilibrium is connected (and hence weakly acyclic) with a result by Young (2004) that shows that period-by-period play under the *best-response dynamic with inertia* (which is stationary, 1-recall, and uncoupled) converges almost surely to a pure Nash equilibrium in every weakly acyclic game.[6] This, of course, does not overturn the above impossibility result, but it limits its scope: the best-response dynamic with inertia is not guaranteed to converge to a pure Nash equilibrium in *all* large generic games that have one, but it does have such a guarantee in *almost all* large generic games that have one.

The best-response dynamic with inertia is, according to a classification by Hart (2005), an *adaptive* dynamic. 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

[3]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.

[4]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).

[5]We require that every player has at least two actions (or at least three actions if there are only three players).

[6]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 $1 - p_i$.

information and take actions that respond to their environment according to simple decision heuristics in a generally improving way. 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 Young, 2006), trial-and-error learning (Young, 2009), and many more. As argued in Section 6, our results imply that several of these simple adaptive dynamics are guaranteed to lead to a pure Nash equilibrium in almost every large generic game that has one.

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 order on $A := \prod_{i\in[n]} [k_i]$. (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 $a_{-i} \in \prod_{j\in[n]\setminus\{i\}}[k_j]$.) We say that $[n]$ is the *player set* of the game and that $[k_i]$ is the *action set* of player i. Elements of A are called *action profiles*, and $\succsim_i$ is known as i's *preference relation*. For each i, let $\succ_i$ denote the asymmetric part of $\succsim_i$.

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 games with player set $[n]$ in which, for every $i \in [n]$, player i has action set $[k_i]$.

An action a_i of player i is a *better-response* than a'_i to a_{-i} if $(a_i, a_{-i}) \succ_i (a'_i, a_{-i})$. An action a_i of player i is a *best-response* to a_{-i} if (a_i, a_{-i}) is $\succsim_i$-maximal in $L(a_{-i})$, that is, 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}. A game is *generic* if for every i, a_{-i}, and distinct $(a_i, a_{-i}), (a'_i, a_{-i}) \in L(a_{-i})$, either $(a_i, a_{-i}) \succ_i (a'_i, a_{-i})$ or $(a'_i, a_{-i}) \succ_i (a_i, a_{-i})$.[7]

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 $a, b \in A$,

[7]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.

$a \longrightarrow b$ 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.

The *best-response graph* of a game is defined in the same way, except that the edge condition is now that $a, b \in A$ satisfy

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

Clearly, the best-response graph of a game is a subgraph of its better-response graph.

3 A graph-based classification of games

We now define various classes of games in terms of the connectivity properties of best- and better-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 $(V, \longrightarrow)$ on vertex set V and edge set $\longrightarrow$, 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 $v^i \longrightarrow v^{i+1}$ 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 at least 2 and that satisfies $v^m \longrightarrow v^1$ and $v^i \longrightarrow v^{i+1}$ for all $i \in [m-1]$.

Definition 1. A game is *acyclic* if its best-response graph has no cycles. A game is *globally acyclic* if its better-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. A game is *globally weakly acyclic* if its better-response graph has the property that every vertex can reach a sink.

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

Acyclicity, weak acyclicity, and their 'global' counterparts, have become 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.

[8] See Candogan et al. (2011) for a flow-based decomposition of games.
[9] For example, Takahashi and Yamamori (2002) refer to weak acyclicity as quasi-acyclicity.

Acyclic games are a superset of the very widely studied class of potential games (Monderer and Shapley, 1996).[10] 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). Coordination games, dominance-solvable games (Alon et al., 2021), and congestion games (Rosenthal, 1973) are all special cases of potential games.

Weakly acyclic games are also very widely studied because weak acyclicity is a necessary condition for the guaranteed convergence of better- and best-response dynamics to a pure Nash equilibrium (e.g. see Fabrikant et al., 2013; Apt and Simon, 2015).

The notion of connectedness, which is stronger than weak acyclicity, is one that we introduce in this paper.

Remark 4. The following diagram summarises the relationships between the game classes above.

$$\begin{tikzcd}
\text{globally acyclic} \arrow[r, Rightarrow] & \text{acyclic} \arrow[r, Rightarrow] & \text{weakly acyclic} \arrow[r, Rightarrow] & \text{globally weakly acyclic} \\
& & \text{connected} \arrow[u, Rightarrow] \arrow[r, Rightarrow] & \text{globally connected} \arrow[u, Rightarrow]
\end{tikzcd}$$

4 Results on game connectivity

Remark 4 is useful for qualitatively characterising the relationships between game classes, but it does not offer a *quantitative* description of the relationships. For example, we know that acyclic games are a subset of weakly acyclic games, but do they constitute a small or large subset of weakly acyclic games?

The results we present in this section quantify the relative sizes of the game classes defined in Section 3 for large games, where 'large' here means that the number of players is much bigger than the maximum number of actions available to each player. 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*

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

This result shows that, strikingly, connectedness is a ubiquitous property among

[10]A (not necessarily generic) 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, $a \succ_i a'$ implies $\rho(a) > \rho(a')$. The game is an ordinal potential game or, simply, a potential game if, additionally, $\rho(a) > \rho(a')$ implies $a \succ_i a'$. Generalised ordinal potential games are precisely those that are globally acyclic (Monderer and Shapley, 1996; Fabrikant et al., 2013).

large generic games that have a pure Nash equilibrium.[11] Theorem 5 has important implications for adaptive dynamics in games, on which we elaborate in Section 6.

Our quantitative characterisation of game classes is completed by the following result, which applies to all 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*

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

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 and globally acyclic games are very rare, while connected games, weakly acyclic games, and their global counterparts, 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.[12]

Theorem 5 and Proposition 6 will be proved as corollaries of stronger results (Theorem 8 and Proposition 14 respectively) concerning the likelihood of analogous conditions holding in certain random directed graphs. We now turn to these results.

5 Results on directed grids

5.1 Connectivity of directed grids

Theorem 8, the main result of this section and the central 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.

[11] Rinott and Scarsini (2000) show that

$$\lim_{n\to\infty} \frac{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is generic and has a pure Nash equilibrium}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is generic}\}|} = 1 - e^{-1},$$

where this limit is uniform over all $\mathbf{k} = \mathbf{k}(n) \in \mathbb{N}^n$ with $k_i \geq 2$ for all i. That is, about $1 - e^{-1} \approx 63\%$ of large generic games have a pure Nash equilibrium. Combined with Theorem 5, this implies that approximately 63% of all large generic games are connected.

In fact, building on Arratia et al. (1989), Rinott and Scarsini (2000) prove the much stronger result 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. The distribution of pure Nash equilibria in random games was previously also 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).

[12] This does not, of course, imply that potential games are somehow unimportant: many types of strategic interaction, such as congestion, are appropriately modeled as potential games. Our result merely quantifies the prevalence of such games.

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 7. 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 generic 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 8. *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 8 will be postponed to the appendices but we provide an outline of the proof in Section 5.3. In Section 5.2 we examine the tightness (or lack thereof) of various aspects of Theorem 8. First, however, we will use the remainder of this subsection to explain how Theorem 5 follows from Theorem 8 and to discuss how these results relate to recent work of Amiet et al. (2021). To these ends, we highlight the following corollary of Theorem 8, 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 9. *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 8 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

$2K\log(K) \le K^2\log(K) \le n$, 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 $\prod_{i=1}^n k_i^{-c} \le 2^{-cn} \le e^{-c_0 n}$ for some $c_0 > 0$, which proves part (a).

Next,

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

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 $n \ge 2$ and all $\mathbf{k} \in \{2,3,\dots\}^n$, completing the proof of part (b). □

We now (i) explain how Theorem 5 follows from Corollary 9 part (b), and (ii) discuss how Corollary 9 part (a) is related to the work of Amiet et al. (2021).

(i) 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 generic games in $\mathcal{G}(n,\mathbf{k})$. Because our draws are uniform, we have that

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

so Theorem 5 is immediate from part (b) of Corollary 9.

(ii) Amiet et al. (2021) derive connectivity properties of the graph $\vec{L}(n,\mathbf{2})$, where $\mathbf{2} = (2,\dots,2)$. For each vertex x of $\vec{L}(n,\mathbf{k})$, define $R^x_{n,\mathbf{k}}$ to be the event that if x is a non-sink, then it can reach every sink. Amiet et al. (2021) show that for a fixed vertex x, the event $R^x_{n,\mathbf{2}}$ occurs with failure probability $e^{-\Omega(n)}$ (i.e. there exists $c > 0$ such that the failure probability is at most e^{-cn} for all n). Part (a) of Corollary 9 is a stronger result for two reasons. Firstly, our result implies the stronger statement that $R_{n,\mathbf{2}} = \bigcap_x R^x_{n,\mathbf{2}}$ occurs with with failure probability $e^{-\Omega(n)}$; in other words, Amiet et al. (2021) show that any fixed non-sink can reach every sink whereas we show that every non-sink can *simultaneously* reach every sink. In particular, their results do not imply that almost all generic games with a pure Nash equilibrium are connected in the $\mathbf{k} = \mathbf{2}$ case. Secondly, our result allows for the case $k_i > 2$, and indeed k_i growing with n. This is a non-trivial extension: the edges of $\vec{L}(n,\mathbf{2})$ are oriented independently of one another, and the proof of Amiet et al. (2021) relies heavily on this feature. In the more general case, edges in a line of $\vec{L}(n,\mathbf{k})$ are no longer independent of each other and for this reason our approach to the proof of Theorem 8 is different.

5.2 On the tightness of Theorem 5 and Theorem 8

In this section we discuss to what extent various aspects of Theorem 5 and Theorem 8 are tight. The proofs of all the theorems in this section can be found in Appendix E.

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 8. However, the condition given in Theorem 8 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 7.

Next, with regards to the tightness of the failure probability in the theorems, first note that the failure probability in Theorem 8 carries through to Theorem 5 by the arguments of the previous subsection. 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 8 would still be tight up to the value of the constant in the exponent.

Theorem 10. *There is a constant $c' > 0$ such that*

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

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

One might ask whether taking a larger value for N in the statement of Theorem 8 would allow a significantly smaller failure probability, but a similar argument to the proof of Theorem 10 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. See Appendix E for a proof of this claim.

Finally, to what extent is it possible to take a smaller N in the statement of Theorem 8? 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 m not much smaller than $\log(K-1)$ in the following theorem,[13] 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 11. *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 $1 - c/n$.

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 8 to slightly improve the value of N.

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

Theorem 12. *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 $\overrightarrow{L}(n, \mathbf{k})$ 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 $e^{-c_K n}$.

Together, Theorem 11 and Theorem 12 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 11, then that theorem implies that when n is much larger than K, $\overrightarrow{L}(n, \mathbf{K})$ typically contains a vertex that can be reached from 'exactly' $\log(K)/(\log(K) - \log(K-1))$ vertices. Meanwhile Theorem 12 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 12 is modest, it has the following consequence for $\overrightarrow{L}(n, \mathbf{2})$ and $\overrightarrow{L}(n, \mathbf{3})$ which may be of independent interest.

Corollary 13. *With failure probability $e^{-\Omega(n)}$, every non-sink in $\overrightarrow{L}(n, \mathbf{2})$ can reach every non-source. The same is true for $\overrightarrow{L}(n, \mathbf{3})$. Conversely, for each $K \geq 4$, the probability that there is a non-sink in $\overrightarrow{L}(n, \mathbf{K})$ which cannot reach every non-source tends to 1 as $n \to \infty$.*

Note that the game property corresponding to the condition with which Corollary 13 is concerned is much stronger even than being connected. The positive direction of the corollary follows straightforwardly from Theorem 12 and the observation that every non-source in $\overrightarrow{L}(n, \mathbf{k})$ can be reached from at least $\min_i(k_i)$ vertices, and the negative direction follows immediately from setting $r = 1$ in Theorem 11. While the $\mathbf{k} = \mathbf{2}$ case of the corollary also follows from Theorem 8, the $\mathbf{k} = \mathbf{3}$ case does not. We also note that in the $\mathbf{k} = \mathbf{2}$ case the arguments of Amiet et al. (2021) show that for any vertex x, with failure probability $e^{-\Omega(n)}$, if x is a non-sink then it can reach every non-source.

5.3 Outline of the proof of Theorem 8

We now provide a high-level explanation of the key elements of our proof of Theorem 8, the full version of which can be found in Appendices A, B, C, and D. Very loosely, we will say that an event occurs in $\overrightarrow{L}(n, \mathbf{k})$ *with very high probability* (wvhp) if its failure probability is of the form $\prod_{i=1}^n k_i^{-c}$ for some $c > 0$ independent of n and $\mathbf{k}$, and *with extremely high probability* (wehp) if its failure probability is small enough to facilitate a bounded number of union bounds over $V(n, \mathbf{k})$ while remaining of the form $\prod_{i=1}^n k_i^{-c}$. See Appendix A for rigorous definitions.

Let $H'(n, \mathbf{k})$ be the random subgraph of $H(n, \mathbf{k})$ in which, for each line in $H(n, \mathbf{k})$ in coordinate i, we independently keep the edges induced by the line with probability $1/k_i$, and delete them otherwise. We will say that a line is present in $H'(n, \mathbf{k})$ if we have kept the edges in the line. The proof of Theorem 8 is based on a coupling between $\overrightarrow{L}(n, \mathbf{k})$ and $H'(n, \mathbf{k})$ in which, given a 'root' vertex $z \in V(n, \mathbf{k})$, we build a random subgraph G of $H(n, \mathbf{k})$ which has the same distribution as $H'(n, \mathbf{k})$, and has the property that every vertex in the same connected component of G as z can reach z in $\overrightarrow{L}(n, \mathbf{k})$. This is achieved via a 'backwards exploration process' from z, with the condition that we cannot re-explore any line which has already been revealed.

McDiarmid et al. (2021) study the component structure of $H'(n, \mathbf{k})$ in the case $\mathbf{k} = (2, \ldots, 2)$. In Appendix B we adapt their work to $H'(n, \mathbf{k})$ for $\mathbf{k}$ as in Theorem 8. In particular, we follow them in defining a vertex of $H'(n, \mathbf{k})$ to be *good* if the number of present lines which contain it is at least half the expected number. We then show that wehp all good vertices of $H'(n, \mathbf{k})$ are in the same connected component. It follows by a union bound over $V(n, \mathbf{k})$ that wehp, for any choice of root vertex z the coupled graph G has all its good vertices in the same component.

Let $x, y \in V(n, \mathbf{k})$ be distinct. To prove Theorem 8, we would like to show that the probability that x is not a sink and y can be reached from more than N vertices, yet there is no directed path from x to y, is very small. More precisely we would like to show that the complement of this event occurs wehp, so that we could take a union bound over x and y to complete the proof. However, for example, the probability that x can only reach one other vertex is not this small, so our argument needs an extra ingredient.

To this end, we first consider the event that every non-sink can reach more than $n/2$ vertices, and every vertex which can be reached from more than N vertices can be reached from more than $n/2$ vertices. We show in Appendix C that this event occurs wvhp. The inspiration for considering such an event comes from work of Bollobás et al. (1993). Returning to x and y, we may now assume that they can (respectively) reach and be reached from at least $n/2$ vertices. This is sufficient to allow us to show that wehp, for every $z \in V(n, \mathbf{k})$, there is a vertex u which can be reached from x and which is good in the coupled graph G built from root z. Similarly, we show that wehp y can be reached from a vertex v which is good in the coupled graph built from root v.

Recalling that wehp, for any choice of root vertex the coupled graph G has all its good vertices in the same component, we deduce that wehp there is a directed walk from x to y in $\overrightarrow{L}(n, \mathbf{k})$, via u and v. Finally, we apply a union bound over x and y, and account for the probability that some non-sink can reach at most $n/2$ vertices or some vertex can be reached from more than N but at most $n/2$ vertices, to complete the proof of the theorem.

5.4 Acyclicity of directed grids

Using the terminology and notation of Section 5.1 we can state the following strengthening of Proposition 6.

Proposition 14. *There exists $c > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \ldots\}^n$, the probability that $\overrightarrow{L}(n, \mathbf{k})$ is acyclic is at most $\exp(-cnk^{n-2})$, where $k := \min_i(k_i)$.*

The proof of Proposition 14 can be found in Appendix F. Proposition 6 follows from Proposition 14 by the same reasoning with which we deduced Theorem 5 from Theorem 8: since $\vec{L}(n, \mathbf{k})$ has the same distribution as the best-response graph of a game drawn uniformly at random from all generic games in $\mathcal{G}(n, \mathbf{k})$, we have

$$\frac{\mathbb{P}(\vec{L}(n, \mathbf{k}) \text{ is acyclic})}{\mathbb{P}(S_{n,\mathbf{k}})} = \frac{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is generic and acyclic}\}|}{|\{g \in \mathcal{G}(n, \mathbf{k})\colon g \text{ is generic with a pure Nash equilibrium}\}|},$$

where $S_{n,\mathbf{k}}$ is the event that $\vec{L}(n, \mathbf{k})$ contains a sink, as in Section 5.1. Thus, since $k_i \geq 2$ for all i, Proposition 6 follows from Proposition 14 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 5.1.

6 Results on adaptive dynamics

We now consider 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 $f\colon O_k \to \Delta([k])$, 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. We now define various types of dynamic.

Definition 15. 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 $a^0, \ldots, a^{t-1}$.

Definition 16. 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 $t < m$.

Definition 17. 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.

In what follows, we consider the following strong notion of convergence to a pure Nash equilibrium.

Definition 18. 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 $t \geq T$.

As mentioned in the introduction, the following impossibility result due to Hart and Mas-Colell (2006) and Jaggard et al. (2014) is well-known: 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 stationary, 1-recall, uncoupled dynamic on $\mathcal{G}(n, \mathbf{k})$ for which play converges almost surely to a pure Nash equilibrium in every generic game in the class that has one. However, we show below that as n grows (much more quickly than $\max_i(k_i)$), there are dynamics on $\mathcal{G}(n, \mathbf{k})$ of this type that converge almost surely to a pure Nash equilibrium on all but a vanishingly small fraction of generic games in the class that have one.

Consider the following family of stationary, 1-recall, uncoupled dynamics.[14]

Definition 19 (Best-response dynamic with inertia)**.** 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 this result in the case of generic weakly acyclic games in order to shed light on the link between the connectivity properties of games and the convergence of dynamics.

Theorem 20 (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,

[14]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.

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. □

Since every connected game is weakly acyclic, combining Theorem 5 and Theorem 20 yields the following.[15]

Corollary 21. *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*

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

for which the best-response dynamic with inertia converges almost surely to a pure Nash equilibrium is at least $1 - e^{-cn}$.

In contrast with the impossibility result of Hart and Mas-Colell (2006) and Jaggard et al. (2014), this result shows that there is a stationary, 1-recall, uncoupled dynamic that converges almost surely to a pure Nash equilibrium in almost every large generic game that has one.

Comment on convergence The notion of convergence considered above is strong, since it requires period-by-period play to eventually settle on, and never leave, a pure Nash equilibrium. 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,

[15]Corollary 21 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^{t-1}_{-i}. Convergence properties of this one-at-a-time version were investigated by Heinrich et al. (2023) via simulation.

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.

Comment on fragility The fact that almost every large generic game that has a pure Nash equilibrium is connected implies a certain fragility for adaptive dynamics in large games. We have seen that the best-response dynamic with inertia converges almost surely to a pure Nash equilibrium in almost every large game that has one. But now suppose the dynamic is at a pure Nash equilibrium and we shock the dynamic out of this equilibrium (e.g. some player commits an error and plays a non-best-response action). Since every non-sink can reach *every* sink in a connected game, the shocked dynamic can go on to converge to *any* pure Nash equilibrium.[16] Weakly acyclic games need not have this property.

Implications for other dynamics Since connected games are weakly acyclic, it follows from Theorem 5 that all existing results on dynamics converging in weakly acyclic games also apply in almost every large generic game that has a pure Nash equilibrium. We have focussed here on the best-response dynamic with inertia but there are many more dynamics to which this applies. For example, Young (1993) shows that 'adaptive play' converges almost surely to a pure Nash equilibrium in all globally weakly acyclic games, and Friedman and Mezzetti (2001) shows the same for 'better-reply dynamics with sampling'. Moreover Marden et al. (2007) and Marden et al. (2009) describe, respectively, regret-based and payoff-based dynamics that lead to play that is at a pure Nash equilibrium in every weakly acyclic game 'most of the time'. By Theorem 5, all of these results apply to almost every large generic game with a pure Nash equilibrium.

7 Limitations and 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 approach can deliver interesting conclusions. There are also limitations and open questions that we outline here.

(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

[16] This type of 'fragility' is investigated by Rudov (2022) in the context of matching markets.

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 8.

(ii) The 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) We have not addressed the question of speed of convergence to equilibrium (Arieli and Young, 2016). 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 results on the speed of convergence in some classes of games.[17]

(iv) Our analysis is focused entirely on generic games. Amiet et al. (2021) 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.

(v) 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.

(vi) We have quantified the proportion of games with specific types of connectivity properties. As noted in the points above, it might be fruitful to investigate the

[17]Convergence can be fast in potential games. For example, see Awerbuch et al. (2008).

prevalence of games with other types of connectivity properties. For example, it may be worth considering graphs that correspond to deviations by multiple players at a time, rather than just one.

A Proof of Theorem 8: preliminaries

In Appendices A, B, C, and D we detail the proof of Theorem 8, which is restated below for convenience. We will continue to use the notation and terminology introduced in Section 5.

Theorem 8. *For all $\varepsilon > 0$ there exist $c, \delta > 0$ such that for all integers $n \geq 2$ and all $\mathbf{k} \in \{2, 3, \dots\}^n$, 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 $\overrightarrow{L}(n, \mathbf{k})$ can either be reached from at most $N := (1+\varepsilon)K\log(K)$ vertices, or from every non-sink.*

Throughout the proof of this theorem, i.e. throughout Appendices A, B, C, and D, 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 8 is equivalent to the statement that wvhp, every vertex of $\overrightarrow{L}(n, \mathbf{k})$ 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 22. *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 23. *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 8 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 24 (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 25 (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 $\{Z_1 = \cdots = Z_m = 0\}$. $\square$

B Proof of Theorem 8: results on $H'(n, \mathbf{k})$

In this section, we state and prove a series of lemmas concerning $H'(n, \mathbf{k})$, the random subgraph of $H(n, \mathbf{k})$ defined in Section 5.3. Our results and our proof strategy are based on those of McDiarmid et al. (2021), who consider the case $\mathbf{k} = (2, \ldots, 2)$. Following those authors, we will say that a vertex of $H'(n, \mathbf{k})$ is *good* if at least $\frac{1}{2} \sum_{i=1}^{n} \frac{1}{k_i}$ of the lines containing it are present, that is, the number of present lines containing it is at least half the expected number. The main result of this section is that wehp all good vertices of $H'(n, \mathbf{k})$ are in the same connected component.

Lemma 26. *With extremely high probability, all good vertices of $H'(n, \mathbf{k})$ are in the same connected component.*

Lemma 26 will be proved using the following two auxiliary lemmas.

Lemma 27. *With extremely high probability, every vertex in $V(n, \mathbf{k})$ has an $H(n, \mathbf{k})$-neighbour which is good in $H'(n, \mathbf{k})$.*

Lemma 28. *With extremely high probability, every pair of good vertices in $H'(n, \mathbf{k})$ which are at distance at most 3 from each other in $H(n, \mathbf{k})$ are joined by a path in $H'(n, \mathbf{k})$.*

Given Lemmas 27 and 28, Lemma 26 follows easily. This proof and the proof of Lemma 27 are straightforward generalisations of work in McDiarmid et al. (2021), but we include them for completeness.

Proof of Lemma 26. Let F be any realisation of $H'(n, \mathbf{k})$ for which the conclusions of Lemma 27 and Lemma 28 both hold, and note that this happens wehp by those two lemmas and Lemma 23. That is, let F be any spanning subgraph of $H(n, \mathbf{k})$ in which there is a path in F between every pair of good vertices which are at distance at most 3 in $H(n, \mathbf{k})$, and in which every vertex has a neighbour in $H(n, \mathbf{k})$ which is good. To prove Lemma 26, it suffices to show that there exists a path in F between every pair of good vertices.

Let x and y be good vertices (of F). Choose a path $P = p_0p_1 \dots p_t$ in $H(n, \mathbf{k})$ where $p_0 = x$ and $p_t = y$. If $t \le 3$, then there exists a path from x to y in 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 F contains a path from x to y via $q_2, \dots, q_{t-2}$. □

Proof of Lemma 27. For a vertex $v \in V(n, \mathbf{k})$, let X be the number of lines containing v which are present in $H'(n, \mathbf{k})$. Then X is a sum of n independent Bernoulli random variables with parameters $1/k_1, \dots, 1/k_n$ respectively, so by Lemma 24 we have

$$\mathbb{P}(v \text{ is not good}) = \mathbb{P}\Big(X < \frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}\Big) \le e^{-\sum_{i=1}^{n} 1/8k_i} \le e^{-n/8K},$$

where $K = \max(k_i)$ as before.

Now fix $u \in V(n, \mathbf{k})$ and pick one vertex other than u from each of the n lines containing it, say $v_1, \dots, v_n$. 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 $e^{-n^2/8K}$, 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 $K^n \cdot e^{-n^2/8K}$. For $0 < \delta \le 1$, if $K \le \delta\sqrt{n/\log(n)}$, then $K^2 \log(K) \le \delta n$ and $n/K \ge K\log(K)/\delta \ge \log(K)/\delta$, so $e^{-n^2/8K}$ is extremely small. By Lemma 22, the same is true of $K^n \cdot e^{-n^2/8K}$, which completes the proof of the lemma. □

It remains to give the (slightly more involved) proof of Lemma 28.

Proof of Lemma 28. In this proof we will relabel $V(n, \mathbf{k})$ as $\prod_{i=1}^{n}\{0, \dots, k_i - 1\}$ in the natural way and will consider these vertices as elements of the vector space $\mathbb{R}^n$. We will write $\mathbf{e}_1, \dots, \mathbf{e}_n$ for the standard basis of this space.

We need to show that if $\delta > 0$ is sufficiently small, then whenever $K \le \delta\sqrt{n/\log(n)}$, the probability that there exists a pair of good vertices in $H'(n, \mathbf{k})$ which are at distance at most 3 from each other in $H(n, \mathbf{k})$ but are not in the same component of $H'(n, \mathbf{k})$ is at most $e^{-cn\log(K)/\delta}$ for some universal $c > 0$. Thus, let $\delta > 0$ be small and assume that $K \le \delta\sqrt{n/\log(n)}$. Note that since $n \ge 2$, by choosing δ small enough we may assume that $n/K^2 \ge \log(n)/\delta^2$ (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 $A', B' \subseteq [n]$ each of size at least $n/2K$; later we will assume that u and v are good vertices and take these to be the sets of coordinate directions in which the lines containing each of them (respectively) are present. Now pick any $A \subseteq A' \setminus \{1,2,3\}$ and $B \subseteq B' \setminus \{1,2,3\}$ such that $|A| = |B| = \lceil n/3K \rceil$. Relabelling again, we may assume that $A, B \subseteq [\lfloor n/2 \rfloor]$.

Having fixed A and B, we will now define a certain type of path in $H(n, \mathbf{k})$. First, let $i, j \in [n]$ be distinct with $i \in A$ and $j \in B$, then let $\alpha \in [k_i - 1]$ and $\beta \in [k_j - 1]$. A path in $H(n, \mathbf{k})$ from $\alpha\mathbf{e}_i + \beta\mathbf{e}_j$ to $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + v$ will be called an *(i, j, α, β)-path* if it has the following form: the path starts at $\alpha\mathbf{e}_i + \beta\mathbf{e}_j$ then follows a path of length 3 to $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1$ 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

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

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 $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1 + \mathbf{e}_2$ 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 $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 = \alpha\mathbf{e}_i + \beta\mathbf{e}_j + v$ in which the first and third edges are in a coordinate direction taken from $[\lfloor 5n/6 \rfloor + 1, n]$.

Let $E^{(1)}_{(i,j,\alpha,\beta)}$ be the event that there exists in $H'(n, \mathbf{k})$ a path of length 3 from $\alpha\mathbf{e}_i + \beta\mathbf{e}_j$ to $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1$ 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, $E^{(1)}_{(i,j,\alpha,\beta)}$ 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 path

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

are present in $H'(n, \mathbf{k})$. Define $E^{(2)}_{(i,j,\alpha,\beta)}$ and $E^{(3)}_{(i,j,\alpha,\beta)}$ analogously for the second and third parts of the (i, j, α, β)-path. Note that there exists an (i, j, α, β)-path in $H'(n, \mathbf{k})$ if and only if all three of these events occur. To analyse the probability of these events it will be helpful to have the following simple fact which can be proved by induction on m.

Fact 1. *Let $x \geq 1$ and $m \in \mathbb{N}$ satisfy $m - 1 \leq x$. Then*

$$\left(1 - \frac{1}{x}\right)^m \leq 1 - \frac{m}{2x}.$$

The event that $H'(n, \mathbf{k})$ contains a path of length 3 from $\alpha\mathbf{e}_i + \beta\mathbf{e}_j$ to $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1$ in which the first and third edges are in a given coordinate direction ℓ is the intersection of three independent events: the line in coordinate ℓ containing $\alpha\mathbf{e}_i + \beta\mathbf{e}_j$ must be present, the line in coordinate ℓ containing $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \mathbf{e}_1$ must be present and there must be some $\gamma \in [k_\ell - 1]$ for which the line in coordinate 1 containing $\alpha\mathbf{e}_i + \beta\mathbf{e}_j + \gamma\mathbf{e}_\ell$ is present. Hence, the probability of this event is

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

By the claim, this is at least

$$\frac{1}{k_\ell^2}\left(1-\left(1-\frac{k_\ell-1}{2K}\right)\right)=\frac{k_\ell-1}{2k_\ell^2K}\geq\frac{1}{4K^2},$$

where for the final inequality we have used the fact that $2 \leq k_\ell \leq K$.

Since the existence of such a path is independent for different ℓ, the failure probability of $E^{(1)}_{(i,j,\alpha,\beta)}$ is at most

$$\prod_{\ell=\lfloor n/2\rfloor+1}^{\lfloor 2n/3\rfloor}\left(1-\frac{1}{4K^2}\right)\leq\left(1-\frac{1}{4K^2}\right)^{n/7}\leq e^{-n/28K^2}<\frac{1}{2},$$

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

Similarly, $E^{(2)}_{(i,j,\alpha,\beta)}$ and $E^{(3)}_{(i,j,\alpha,\beta)}$ 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 $H'(n,\mathbf{k})$ contains an (i,j,α,β)-path with probability at least $1/8$.

Next, we will call a path in $H(n,\mathbf{k})$ an *extended* (i,j,α,β)*-path* if it is an (i,j,α,β)-path extended by one vertex at each end to $\alpha\mathbf{e}_i$ and to $\beta\mathbf{e}_j+v$. The line containing $\alpha\mathbf{e}_i$ and $\alpha\mathbf{e}_i+\beta\mathbf{e}_j$ and the line containing $\beta\mathbf{e}_j+v$ and $\alpha\mathbf{e}_i+\beta\mathbf{e}_j+v$ are distinct, and neither could be used in an (i,j,α,β)-path, so the probability that $H'(n,\mathbf{k})$ contains an extended (i,j,α,β)-path is at least $1/(8k_ik_j)$.

To conclude our definitions, a path in $H(n,\mathbf{k})$ will be called an (i,j)*-path* if there exist $\alpha\in[k_i-1]$ and $\beta\in[k_j-1]$ for which it is an extended (i,j,α,β)-path. Note that (once the pair (i,j) is fixed) for each choice of α and β, every line that could possibly be used in an extended (i,j,α,β)-path identifies the pair (α,β). It follows that as α and β vary, the events that $H'(n,\mathbf{k})$ contains an extended (i,j,α,β)-path are independent. Hence, the probability that $H'(n,\mathbf{k})$ contains no (i,j)-path is at most

$$\left(1-\frac{1}{8k_ik_j}\right)^{(k_i-1)(k_j-1)}\leq\exp\left(-\frac{1}{8}\left(1-\frac{1}{k_i}\right)\left(1-\frac{1}{k_j}\right)\right)\leq e^{-1/32}.$$

Next, observe that every line that could possibly be used in an (i,j)-path identifies the set $\{i,j\}$. There are at least $|A|(|B|-1)/2\geq n^2/20K^2$ ways to choose $\{i,j\}$, so the probability that $H'(n,\mathbf{k})$ does not contain an (i,j)-path for any (i,j) is at most $e^{-n^2/(640K^2)}$. Observe also that the event that there exists an (i,j)-path in $H'(n,\mathbf{k})$ 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 $n/2K$ directions, so if u and v are good vertices, then we may take A' to be the set of coordinate directions in which the line containing u is present, and let B' be the same but for v. This means that the edges from u to $\alpha\mathbf{e}_i$ and from $\beta\mathbf{e}_j+v$ to v are both present for any choice of $i\in A$ and $j\in B$, so if there is no path between u and v in $H'(n,\mathbf{k})$, then there is no (i,j)-path in $H'(n,\mathbf{k})$ for any (i,j). This happens with probability at most $e^{-n^2/(640K^2)}$.

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 exist two good vertices at distance at most 3 from one another which are not joined by a path in $H'(n, \mathbf{k})$ is at most $K^{2n} \cdot e^{-n^2/(640K^2)}$. Clearly $e^{-n^2/(640K^2)}$ is extremely small, so Lemma 22 implies that $K^{2n} \cdot e^{-n^2/(640K^2)}$ is also extremely small, which completes the proof of the lemma. □

C Proof of Theorem 8: establishing a foothold

Throughout this section we let $\varepsilon > 0$ and set $N = (1 + \varepsilon)K \log(K)$, as in the statement of Theorem 8. We will also assume (without loss of generality) that $k_1 \leq k_2 \leq \cdots \leq k_n$. Recall from Section 5.3 that one step in our proof of Theorem 8 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 29 (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 30. *The event A_ε occurs with very high probability.*

Our proof of Lemma 30 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 31. *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 31 gives that there are at most $(enK)^{m-1}$ trees of order m in $H(n, \mathbf{k})$ that contain a given vertex. Lemma 30 follows from the next two lemmas, which handle the X_m and Y_m parts of the statement respectively.

Lemma 32. *With very high probability, $X_m = 0$ for all $2 \le m \le n/2$.*

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

Fix $2 \le m \le n/2$ and let T be a tree of order m in $H(n, \mathbf{k})$. Given the discussion preceding Lemma 31, 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 line is in T, we may assume that they are all in as low a coordinate direction as possible (recall that $k_1 \le \cdots \le k_n$ 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 31, 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

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

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

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

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

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 25 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 30. Note that Lemma 30 follows immediately from Lemma 23, Lemma 32, and Lemma 33.

Lemma 33. *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 32. 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 32). 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 31 and the discussion preceding it, we have

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

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

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

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

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

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

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

Thus,

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

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 $\sum_{N<m\leq n/2} \mathbb{E}[Y_m]$ is at most $e^{-c_\varepsilon n\log(K)}$ for some $c_\varepsilon > 0$ depending only on ε. Lemma 25 now yields that $Y_m = 0$ for all $N < m \leq n/2$ with failure probability at most $e^{-c_\varepsilon n\log(K)}$, as required. □

D Proof of Theorem 8: the coupling

Recall from Section 5.3 that we wish to define a process of constructing a random subgraph G of $H(n,\mathbf{k})$ given $\vec{L}(n,\mathbf{k})$ and a 'root' vertex $z \in V(n,\mathbf{k})$. We want G to have the same distribution as $H'(n,\mathbf{k})$, and have the property that any vertex in the same connected component of G as z can reach z in $\vec{L}(n,\mathbf{k})$.

We define the coupling as follows. Initialise a process with $V(G) = V(n,\mathbf{k})$ and $E(G) = \emptyset$, fix an ordering of $V(G)$ and set $P = \emptyset$. Repeat the following steps until it is no longer possible to choose the vertex v. Pick the first (according to the ordering) vertex v which is not already in P and is in the same connected component of G as z. For each line in $H(n,\mathbf{k})$ containing v whose edges are not already in G or marked as forbidden, check if v is the winner of that line in $\vec{L}(n,\mathbf{k})$. If v is the winner, add to G the edges induced by the line, and if v is not the winner, mark these edges as forbidden. Once the lines have been checked, add the vertex v to P. It is not hard to see that every vertex in the same connected component of G as z can reach z in $\vec{L}(n,\mathbf{k})$, and that each line considered has been added independently with the appropriate probability.

To complete G, go through those lines which have not been considered and add their edges to G if the winner of the line is first in the fixed ordering. This does not grow the component of G which contains z, but ensures that G has the same distribution as $H'(n,\mathbf{k})$, as promised.

We now prove some results about this coupling process. In this section, as in the previous one, we will assume without loss of generality that $k_1 \leq k_2 \leq \cdots \leq k_n$. Recall that in Section 5.3 and Appendix B we defined a vertex of $H'(n,\mathbf{k})$ to be *good* if at least $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}$ of the lines containing it are present. We will start by considering the following event.

Definition 34 (Event B)**.** Let B 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 vertex which wins at least $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}$ of its lines in $\vec{L}(n,\mathbf{k})$.

We show that B is very likely.

Lemma 35. *The event B 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 $\mathbb{P}(B^c) \leq e^{-cn\log(K)/\delta}$ 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 B_y to be the event that either y can be reached from at most $n/2$ vertices in $\vec{L}(n,\mathbf{k})$, or there exists $v \in V(n,\mathbf{k})$ which can reach y and which is the winner in at least $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}$ of its lines.

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

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

For fixed $T \in S$, if $(B_y)^c$ and B_T both hold, then every vertex in T must win fewer than $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}$ of its lines. Since T contains exactly M vertices, each of its vertices can be assigned a set of at least $n-M$ lines which contain the vertex and no other vertex of T (so that each vertex of T wins at least $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}$ of these $n-M$ lines independently). For a fixed vertex v of T, let X be the number of the $n-M$ lines assigned to v in which v is the winner, so that X is a sum of $n-M$ independent Bernoulli random variables and has mean $\mu \geq \sum_{i=M+1}^{n} k_i^{-1} \geq (n-M)/K \geq n/(2K)$, since $k_1 \leq \cdots \leq k_n$.

Using the inequalities $M \leq n/(3K)$, and $M \leq n/2$, we find that

$$\frac{1}{2}\sum_{i=1}^{M}\frac{1}{k_i} \leq \frac{M}{4} \leq \frac{n}{12K} \leq \frac{n-M}{6K} \leq \frac{1}{6}\sum_{i=M+1}^{n}\frac{1}{k_i}.$$

Adding $\frac{1}{2}\sum_{i=M+1}^{n}\frac{1}{k_i}$ to the left- and right-hand sides gives

$$\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i} \leq \frac{2}{3}\sum_{i=M+1}^{n}\frac{1}{k_i}.$$

Hence, by Lemma 24,

$$\begin{aligned}\mathbb{P}\left(X \le \frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}\right) &\le \mathbb{P}\left(X \le \frac{2}{3}\sum_{i=M+1}^{n}\frac{1}{k_i}\right)\\ &\le \mathbb{P}\left(X \le \frac{2\mu}{3}\right)\\ &\le e^{-\mu/18}\\ &\le e^{-n/(36K)}.\end{aligned}$$

It follows that at least one of the vertices in T wins at least $\frac{1}{2}\sum_{i=1}^{n}\frac{1}{k_i}$ of its lines with failure probability at most $e^{-Mn/(36K)} \le e^{-n^2/(216K^2)}$. This is therefore an upper bound on $\mathbb{P}((B_y)^c \cap B_T)$. It follows from Lemma 31 that $|S| \le (enK)^{M-1}$, so by (2) we have

$$\begin{aligned}\mathbb{P}\big((B_y)^c\big) &\le (enK)^{M-1} \cdot e^{-n^2/(216K^2)}\\ &\le \exp\left(\frac{n\log(enK)}{3K} - \frac{n^2}{216K^2}\right)\\ &= \exp\left(\frac{n^2}{K^2}\left(\frac{K\log(enK)}{3n} - \frac{1}{216}\right)\right).\end{aligned}$$

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

Finally, since $B = \bigcap_y B_y$, by a union bound we have

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

Clearly $e^{-n^2/(300K^2)}$ is extremely small, so by Lemma 22 the same is true of $K^n \cdot e^{-n^2/(300K^2)}$, and the lemma follows. □

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

Definition 36 (Event C)**.** Let C be the event that for every vertex x in $\vec{L}(n,\mathbf{k})$ that can reach more than $n/2$ vertices and every $z \in V(n,\mathbf{k})$, there exists $u \in V(n,\mathbf{k})$ that can be reached from x and that is good in the coupled process with root z.

We will show that this event occurs wehp.

Lemma 37. *The event C occurs with extremely high probability.*

Proof. For each pair $x,z \in V(n,\mathbf{k})$, let $C_{x,z}$ be the event that x can reach at most $n/2$ vertices or that there exists a vertex $u \in V(n,\mathbf{k})$ which is good in the graph produced by the coupling process with root z and can be reached from x in $\vec{L}(n,\mathbf{k})$. By an argument very similar to that used in the preceding proof to upper bound $\mathbb{P}\big((B_y)^c\big)$, we can show that if $\delta > 0$ is small enough, then for each x and z, if $K \leq \delta\sqrt{n/\log(n)}$, then $\mathbb{P}\big((C_{x,z})^c\big) \leq e^{-n^2/(300K^2)}$. The lemma follows by an application of Lemma 22 to a union bound over x and z. $\square$

Before proving Theorem 8 we state a corollary of Lemma 26 concerning the following event.

Definition 38 (Event D)**.** Let D be the event that, for every root vertex z, the good vertices in the graph produced by the coupling process with root z are all in the same component.

It is a straightforward consequence of Lemma 26 that D occurs wehp.

Corollary 39. *The event D occurs with extremely high probability.*

Proof. For each $z \in V(n, \mathbf{k})$, let D_z be the event that the good vertices in the coupling process with root z are all in the same component. Note that by symmetry, $\mathbb{P}(D_z)$ is independent of z. By Lemma 26 and the construction of the coupling process, $\mathbb{P}(D_z)$ is extremely high, so by Lemma 22 and a union bound over z, $\mathbb{P}(D)$ is extremely high too. $\square$

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

Proof of Theorem 8. 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 B occurs, there exists a vertex $v \in V(n, \mathbf{k})$ which can reach y in $\vec{L}(n, \mathbf{k})$, and which is a good vertex in the graph G produced by the coupling process with root v. Next, since C occurs, there exists a vertex $u \in V(n, \mathbf{k})$ which can be reached from x in $\vec{L}(n, \mathbf{k})$ and which is a good vertex in G. Since D occurs, all good vertices of G are in the same connected component and thus, by the design of the coupling process, u can reach v in $\vec{L}(n, \mathbf{k})$.

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 Lemma 35, Lemma 37, and Corollary 39, each of B, C, and D occurs wehp, so in particular wvhp, and A_ε occurs wvhp by Lemma 30. Hence, by (repeated applications of) Lemma 23, we conclude that E_ε occurs wvhp, as required. $\square$

E On improvements to Theorem 5 and Theorem 8

In this section we give proofs for the claims made in Section 5.2. We will start with Theorem 10, which asserts that the improved failure probability for Theorem 5 given by Theorem 8 cannot be significantly improved. We will then explain how to adapt this proof to show that increasing N cannot meaningfully improve the probability of success in Theorem 8.

Theorem 10. *There is a constant $c' > 0$ such that*

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

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

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 9. □

We also claimed that there is some $c' > 0$ 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. This follows from an argument similar to the above: suppose that the desired sticky 4-cycle has vertices $(1, 1, 1, \dots, 1)$, $(1, 2, 1, \dots, 1)$, $(2, 1, 1, \dots, 1)$, and $(2, 2, 1, \dots, 1)$, 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 $\vec{L}(n-1, (k_1, \dots, k_{n-1}))$, and behaves independently of whether the desired sticky 4-cycle appears or not. Applying Theorem 8 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.

We now move on to the matter of improving the value of N. We begin with a proof of Theorem 11, which states that the value of N in Theorem 8 cannot be significantly improved. Recall that $\mathbf{K} = (K, \dots, K)$ denotes a vector of the appropriate length in which every entry is K.

Theorem 11. *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.

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 $n^{11/2}/11^{11}$. For the second term, note that for fixed r the expression in square brackets is strictly increasing in $K \in [1, \infty)$, so we need only lower bound the second term for the least integer K satisfying $f(K) \geq r$. It is straightforward to check that if $f(K) = r$, then the expression in the square brackets is equal to 1. Hence, the second term is always at least 1. Moreover, since there are no integer solutions K to $f(K) = r$ for any $r \in [10]$, 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. □

Next, we turn to the proof of Theorem 12, which gives a slight improvement to the value of N in Theorem 8 at the expense of allowing the constant in the exponent of the failure probability to depend on K. The proof is a fairly straightforward adaption of the proof of Theorem 8.

Theorem 12. *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 8, 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 33. 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 33 we have

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

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

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

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

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

for some $c'_K > 0$, so by Lemma 25 we have that $Y_m = 0$ for all $N' < m \leq 2K\log(K)$ with failure probability at most $e^{-c'_K n}$, as required. □

F Proof of Proposition 14

In this section we give the proof of Proposition 14, which we restate below for convenience. 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.

Proposition 14. *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 $k := \min_i(k_i)$.*

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_ik_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 14. 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. Without loss of generality, assume that $k_1 \leq k_2$. 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 $i < t$ 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. Since $k_1 \leq k_2$, we have $1 \leq T \leq 2k_1 - 1$. Observe that 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 $\mathbb{P}(A) \geq 1/8$.

Since the choice of winner in each line of $\vec{L}(2,(k_1,k_2))$ is independent, for each $t \in \{3,\dots,2k_1-1\}$ we have $\mathbb{P}(A \mid T = t) = (\lceil t/2\rceil - 1)/\lceil t/2\rceil \geq 1/2$. Thus,

$$\mathbb{P}(A) = \sum_{t=1}^{2k_1-1} \mathbb{P}(A \mid T = t)\mathbb{P}(T = t) \geq \frac{\mathbb{P}(T \geq 3)}{2}.$$

We have $\mathbb{P}(T = 1) = 1/k_1$ and $\mathbb{P}(T = 2) = (1 - 1/k_1)(1/k_2)$, so

$$\mathbb{P}(A) \geq \frac{1}{2}\left(1 - \frac{1}{k_1} - \frac{1}{k_2} + \frac{1}{k_1 k_2}\right).$$

The right-hand side of this inequality is increasing in both $k_1 \geq 2$ and $k_2 \geq 2$, so $\mathbb{P}(A) \geq 1/8$, as required. □

Let $n \geq 2$ be an integer and let $\mathbf{k} \in \{2, 3, \dots\}^n$. By the claim, any given plane of $V(n, \mathbf{k})$ induces a cyclic subgraph of $\overrightarrow{L}(n, \mathbf{k})$ 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 $\overrightarrow{L}(n, \mathbf{k})$ 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. □

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.

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.

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

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.

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.

Hart, S. and A. Mas-Colell (2013). *Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics*, Volume 4. World Scientific.

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.

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

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.

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

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.

Rudov, K. (2022). Fragile stable matchings. https://krudov.com/papers/JMP_Kirill_Rudov.pdf. Accessed on 12 September 2023.

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

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.

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

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.