




Best-response dynamics, playing sequences, and convergence to equilibrium in random games

Torsten Heinrich^{1,2,3} · Yoojin Jang^{2,4} · Luca Mungo^{2,5} · Marco Pangallo⁶  · Alex Scott⁵ · Bassel Tarbush⁷ · Samuel Wiese^{2,4}

Accepted: 20 February 2023 / Published online: 14 June 2023
© The Author(s) 2023

Abstract

We analyze the performance of the best-response dynamic across all normal-form games using a random games approach. The playing sequence—the order in which players update their actions—is essentially irrelevant in determining whether the dynamic converges to a Nash equilibrium in certain classes of games (e.g. in potential games) but, when evaluated across all possible games, convergence to equilibrium depends on the playing sequence in an extreme way. Our main asymptotic result shows that the best-response dynamic converges to a pure Nash equilibrium in a vanishingly small fraction of all (large) games when players take turns according to a fixed cyclic order. By contrast, when the playing sequence is random, the dynamic converges to a pure Nash equilibrium if one exists in almost all (large) games.

Keywords Best-response dynamics · Equilibrium convergence · Random games

JEL Classification C62 · C72 · C73 · D83

1 Introduction

The best-response dynamic is a ubiquitous iterative game-playing process in which, at each time step, players myopically select actions that are a best-response to the actions last chosen by all other players. The literature at large has established the

We thank Doyné Farmer for useful comments at the early stages of this project. We acknowledge funding from Baillie Gifford (Luca Mungo), the James S Mc Donnell Foundation (Marco Pangallo) and the Foundation of German Business (Samuel Wiese). Research supported by EPSRC Grant EP/V007327/1 (Alex Scott).

✉ Marco Pangallo
marco.pangallo@centai.eu

✉ Bassel Tarbush
bassel.tarbush@economics.ox.ac.uk

Extended author information available on the last page of the article

equilibrium convergence properties of the best-response dynamic in games with specific payoff structures; particularly in potential games (Monderer and Shapley 1996), but also in weakly acyclic games (Fabrikant et al. 2013), aggregative games (Dindoš and Mezzetti 2006), and quasi-acyclic games (Friedman and Mezzetti 2001; Takahashi and Yamamori 2002). So known results are restricted to special cases. The performance of the best-response dynamic in the class of *all* games remains to be established. In this paper, we consider the question of whether the best-response dynamic converges to a pure Nash equilibrium in a small or large fraction of all possible normal-form games.

To answer our question, we take a “random games” approach: we determine whether the best-response dynamic converges to a pure Nash equilibrium in a game drawn at random from among all possible games. The random games approach has a long history in game theory (since Goldman 1957; Goldberg et al. 1968, and Dresher 1970), and has been used to address questions regarding the prevalence of Nash equilibria (Powers 1990; Stanford 1995, 1996, 1997; Cohen 1998; Stanford 1999; McLennan 2005; McLennan and Berg 2005; Takahashi 2008; Kultti et al. 2011; Daskalakis et al. 2011; Quattropiani and Scarsini 2020), the prevalence of rationalizable strategies (Pei and Takahashi 2019), convergence to equilibrium (Pangallo et al. 2019; Amiet et al. 2021a, b; Wiese and Heinrich 2022), and the prevalence of dominance solvable games (Alon et al. 2021).¹ A guiding principle of the approach is that, since the property of interest (e.g. existence of Nash equilibrium, convergence to Nash equilibrium, dominance solvability) does not hold in all games, one can at least determine how *likely* the property is to hold in the class of all games. To do so, one defines a probability distribution over all games, and computes the probability that a game drawn randomly according to this distribution has the desired property.

The *playing sequence*—the order in which players update their actions—has an important role in our analysis. We largely focus on two specific playing sequences in this paper. At one extreme, we consider the random playing sequence, where players take turns to play one at a time and the next player to play is chosen uniformly at random from among all players. At the other extreme, we consider a natural deterministic counterpart to the random sequence, which we refer to as the clockwork playing sequence, where players take turns to play one at a time according to a fixed cyclic order. The best-response dynamic under the random playing sequence is widely studied. It is often of interest in population and evolutionary games (Sandholm 2010), and its properties have been analyzed in a variety of games with specific payoff structures.² The best-response dynamic under the clockwork playing sequence appears most frequently in the algorithmic game theory literature. Its properties have inter alia been studied in auctions (Nisan et al. 2011), job scheduling (Berger et al. 2011), network formation games (Chauhan et al. 2017), and it has been

¹ The majority of the literature has focused on normal-form games. Arieli and Babichenko (2016) study random extensive form games.

² It has been analyzed in anonymous games (Babichenko 2013), near-potential games (Candogan et al. 2013), potential games (Christodoulou et al. 2012; Coucheney et al. 2014; Swenson et al. 2018; Durand et al. 2019), and games on a lattice (Blume et al. 1993). “Sink” equilibria are studied in (Goemans et al. 2005; Mirrokni and Skopalik 2009).

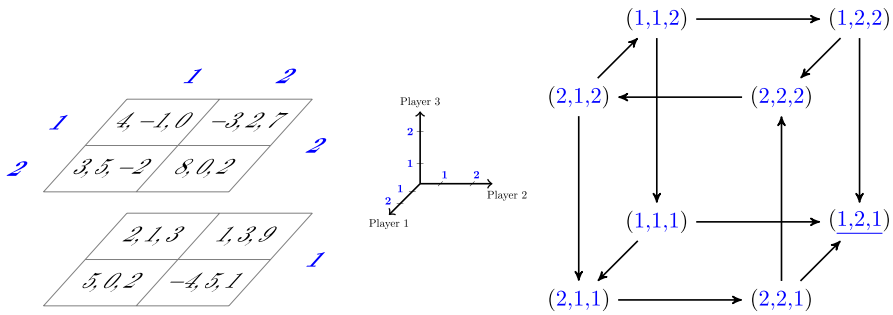


Fig. 1 Illustration of a 3-player game with 2 actions per player (left) and its associated best-response digraph (right). The axes shown in the center give us our coordinate system: player 1 selects rows (along the depth), player 2 selects columns (along the width), and player 3 selects levels (along height). In the left-hand panel, the payoffs of players 1, 2, and 3 are listed in that order. The unique pure Nash equilibrium at the profile (1, 2, 1) is a sink of the digraph and is underlined

used for equilibrium selection in potential games (Boucher 2017). Using the random games approach, Durand and Gaujal (2016) show that, in expectation, convergence to equilibrium in potential games is faster under the clockwork playing sequence than under any other playing sequence.

The playing sequence is essentially irrelevant in determining whether the best-response dynamic converges to equilibrium in potential games—which is the focus of a large part of the literature—but it is a key determinant of the dynamic’s convergence properties in non-potential games. To see this, consider the 3-player game shown in the left-hand panel of Fig. 1 and its associated *best-response digraph* shown in the right-hand panel. Best-response digraphs are a commonly used reduced-form representation of a game in which the vertices are the action profiles and the directed edges correspond to the players’ best-responses (e.g. see Young 1998, Chapter 7, or Pangallo et al. 2019). It is easy to show that potential games have acyclic best-response digraphs, which implies that the playing sequence plays almost no role: as long as each player with a remaining payoff-improving action has a chance to play—which is the case for both the random and the clockwork playing sequences—the dynamic must eventually end at a sink of the digraph, i.e. at a Nash equilibrium of the game.³ In contrast, in the non-potential game shown in Fig. 1, convergence is dependent on the playing sequence: with initial profile (1, 1, 1), the random sequence best-response dynamic must eventually converge to the Nash equilibrium, whereas the clockwork sequence best-response dynamic with cyclic player order 1-2-3-1-...will remain stuck cycling on the four profiles on the front face of the cube forever.

Since we are assessing the performance of the best-response dynamic over the class of *all* games (including non-potential games), it is necessary for us to be explicit about the details of the playing sequence. There are, of course, many possible playing

³ Any such playing sequence may affect the path taken to equilibrium but not whether the path ends at a sink.

sequences,⁴ but our focus on random vs. clockwork suffices for our main finding: whether the best-response dynamic converges to equilibrium in a small or large fraction of all games depends on the playing sequence in an extreme way. Broadly, we show that under a clockwork playing sequence, the fraction of all n -player games in which the best-response dynamic converges to a pure Nash equilibrium goes to 0 as the number of players and/or actions gets large. By contrast, under a random playing sequence, the fraction of all n -player games with a pure Nash equilibrium in which the best-response dynamic converges to a pure Nash equilibrium goes to 1 as the number of players and/or actions gets large (when $n > 2$).

That the best-response dynamic converges less often under a clockwork than under a random playing sequence is perhaps unsurprising since the clockwork sequence will have more difficulty escaping best-response cycles. We therefore expect the probability of convergence to equilibrium for the clockwork sequence to be *less* than it is for the random sequence. However, the resulting extreme jump in the asymptotic equilibrium convergence frequency from 1 to 0 is rather striking. Since most games have digraphs that contain cycles, our contribution can be seen as quantifying the fact that a clockwork playing sequence is *very* likely to become trapped in such cycles, whereas the random playing sequence is *very* likely to escape them.

We now provide a brief technical overview of our methods and results. To generate games at random, we follow the majority of papers in the ‘random games’ literature by drawing each player’s payoff at each action profile independently according to an arbitrary atomless distribution.⁵ This induces a uniform distribution over best-response digraphs, and it is in this sense that we can claim convergence in a large or small fraction of all games. The probability of convergence to a pure Nash equilibrium can be reduced to working out the probability that the best-response path initiated at a random vertex hits a sink of the randomly drawn digraph.⁶

⁴ The concept of a playing sequence is closely related to “revision functions” in Durand and Gaujal (2016) and “schedulers” in Apt and Simon (2015). Simultaneous updating by all players at each time step is studied in Quint et al. (1997) for 2-player games and in Kash et al. (2011) for anonymous games. Feldman and Tamir (2012) study the case in which the sequence of play depends on current payoffs. Feldman et al. (2017) study the dynamic inefficiency of the best-response dynamic under different playing sequences.

⁵ See Goldberg et al. (1968), Stanford (1999), Berg and Weigt (1999), Rinott and Scarsini (2000), Galla and Farmer (2013), Sanders et al. (2018) and Pangallo et al. (2019) for work on random games with payoff correlations.

⁶ Wiese and Heinrich (2022) refer to a game as being “convergent” if, from *every* initial vertex, the clockwork best-response dynamic converges to a pure Nash equilibrium. They then show, for each $k \geq 1$, that the probability that a randomly drawn game is convergent and has exactly k pure Nash equilibria is asymptotically zero in large games. Our asymptotic results imply those of Wiese and Heinrich (2022), but not vice versa. Here is why. In Theorem 1, we derive upper and lower bounds on the probability that the clockwork best-response dynamic initiated at an arbitrarily chosen vertex converges to a pure Nash equilibrium. But because equilibrium convergence from *every* starting vertex (the focus of Wiese and Heinrich 2022) implies convergence to equilibrium from an arbitrarily chosen vertex (the focus of our paper) and not conversely, the upper bound that we find in Theorem 1 is also an upper bound for the probabilities derived in Wiese and Heinrich (2022). Moreover, since we show that the upper bound in

In Sect. 3.1, we show that the probability that the clockwork best-response dynamic converges to a pure Nash equilibrium in a game with $n > 2$ players and $m_i \geq 2$ actions per player i is, up to a polynomial factor, of order $1/\sqrt{q_{n,m}}$, where $q_{n,m} := \frac{\prod_{i=1}^n m_i}{\max_i m_i}$ is the minimal number of strategic environments in the game (i.e. the minimal number of combinations of actions of all but one player). The proof relies on a coupling argument that makes it possible to deal with the path-dependence of the best-response dynamic. The result has two implications. (i) For large $q_{n,m}$, the probability of convergence is determined by the value of a single parameter, namely, the minimal number of possible strategic environments, so all games with an identical minimal number of strategic environments have similar asymptotic probabilities of convergence to equilibrium. This is also reflected in our simulations even for small values of $q_{n,m}$. (ii) When the number of players n and/or the number of actions per player is large for at least two players (implying $q_{n,m} \rightarrow \infty$), the probability that the clockwork best-response dynamic converges to a pure Nash equilibrium goes to zero. This is in stark contrast with the convergence properties of the random sequence best-response dynamic.

In Sect. 3.2, we provide more detailed theoretical results for games with $n = 2$ players. In particular, we provide results on game duration, and we derive an exact expression for the probability that the best-response dynamic converges to a (best-response) cycle of given length at a particular time. As a special case, we obtain the exact probability that the clockwork best-response dynamic converges to a pure Nash equilibrium in 2-player games with m_i actions per player. Unlike in games with $n > 2$ players in which the clockwork and random sequences behave very differently from each other, the probability of convergence to equilibrium is the same for the random and clockwork playing sequences in 2-player games. Furthermore, when $m_1 = m_2 = m$, we show that this probability is asymptotically $\sqrt{\pi/m}$ when m is large.

Section 4 present our simulation results. We investigate the extent to which our asymptotic analytical results also hold for small numbers of players and/or actions. Additionally, we investigate the behavior of playing sequences that interpolate between the extremes of clockwork and random playing sequences.

2 Best-response dynamics in games

2.1 Games

A game with $n \geq 2$ players and $m_i \geq 2$ actions per player i is a tuple

$$g_{n,m} := ([n], \{m_i\}_{i \in [n]}, \{u_i\}_{i \in [n]}),$$

Footnote 6 (continued)

Theorem 1 goes to zero in large games, our asymptotic result implies the asymptotic results of Wiese and Heinrich (2022), but not vice versa.

where $\mathbf{m} := (m_1, \dots, m_n)$, $[n] := \{1, \dots, n\}$ is the set of players, and each player $i \in [n]$ has a set of actions $[m_i] := \{1, \dots, m_i\}$ and a payoff function $u_i : \mathcal{M} \rightarrow \mathbb{R}$, where $\mathcal{M} := \times_{i \in [n]} [m_i]$.

An *action profile* is a vector of actions $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{M}$ that lists the action taken by each player. An *environment* for player i is a vector $\mathbf{a}_{-i} \in \mathcal{M}_{-i} := \times_{j \in [n] \setminus \{i\}} [m_j]$ that lists the action taken by each player but i . A *best-response correspondence* b_i for player i is a mapping from the set of environments for player i to the set of all non-empty subsets of i 's actions and is defined by

$$b_i(\mathbf{a}_{-i}) := \arg \max_{a_i \in [m_i]} u_i(a_i, \mathbf{a}_{-i}).$$

In the rest of this paper, we consider only games in which for each player i and environment \mathbf{a}_{-i} , the best-response action is unique. This is the case for games in which there are no ties in payoffs.⁷

An action profile $\mathbf{a} \in \mathcal{M}$ is a *pure Nash equilibrium* (PNE) if for all $i \in [n]$ and all $a_i \in [m_i]$, $u_i(\mathbf{a}) \geq u_i(a_i, \mathbf{a}_{-i})$. Equivalently, \mathbf{a} is a PNE if each player $i \in [n]$ is playing their (assumed unique) best-response action i.e. $a_i = b_i(\mathbf{a}_{-i})$. Denote the set of PNE of the game $g_{n,\mathbf{m}}$ by $\text{PNE}(g_{n,\mathbf{m}})$ and let $\#\text{PNE}(g_{n,\mathbf{m}})$ denote the cardinality of this set.

2.2 Best-response digraphs

The best-response structure of a game $g_{n,\mathbf{m}}$ can be represented by a *best-response digraph* $\mathcal{D}(g_{n,\mathbf{m}})$ whose vertex set is the set of action profiles \mathcal{M} and whose edges are constructed as follows: for each $i \in [n]$ and each pair of distinct vertices $\mathbf{a} = (a_i, \mathbf{a}_{-i})$ and $\mathbf{a}' = (a'_i, \mathbf{a}_{-i})$, place a directed edge from \mathbf{a} to \mathbf{a}' if and only if a'_i is player i 's best-response to environment \mathbf{a}_{-i} , i.e. $a'_i = b_i(\mathbf{a}_{-i})$. There are edges only between action profiles that differ in exactly one coordinate. A profile \mathbf{a} is a PNE of $g_{n,\mathbf{m}}$ if and only if it is a sink of the best-response digraph $\mathcal{D}(g_{n,\mathbf{m}})$. It is easy to show that potential games have acyclic best-response digraphs.⁸

2.3 Best-response dynamics

We now consider games played over time, with each player in turn myopically best-responding to their current environment.

A *playing sequence* function $s : \mathbb{N} \rightarrow [n]$ determines whose turn it is to play at each time $t \in \mathbb{N}$, where \mathbb{N} denotes the set of positive integers.⁹ We will be interested in two specific playing sequences. The *clockwork* playing sequence is defined by $s_c(t) := 1 + (t - 1) \bmod n$, so player 1 plays at time 1, followed by player 2, then 3,

⁷ There are no ties in payoffs if for all $i \in [n]$, all \mathbf{a}_{-i} , and all $a_i \neq a'_i$, $u_i(a_i, \mathbf{a}_{-i}) \neq u_i(a'_i, \mathbf{a}_{-i})$.

⁸ A game is a (generalized ordinal) potential game if there exists a function $\rho : \mathcal{M} \rightarrow \mathbb{R}$ such that for all $\mathbf{a} \in \mathcal{M}$, $i \in [n]$, and $a'_i \in [m_i]$, $u_i(a'_i, \mathbf{a}_{-i}) > u_i(\mathbf{a})$ implies $\rho(a'_i, \mathbf{a}_{-i}) > \rho(\mathbf{a})$ (Monderer and Shapley 1996). If the best-response digraph has a cycle $(\mathbf{a}_1 \dots \mathbf{a}_k)$ then the potential function would need to satisfy $\rho(\mathbf{a}_k) > \dots > \rho(\mathbf{a}_1) > \rho(\mathbf{a}_k)$, a contradiction.

⁹ Our results hold for any permutation of player labels.

and so on until player n , and then the sequence returns to player 1, and so on. The *random* playing sequence s_x is determined as follows: for each $t \in \mathbb{N}$, draw $s_x(t)$ uniformly at random from $[n]$. So, at each time, the player playing at that time is drawn uniformly at random from among all players. It is easy to see that, starting from any initial profile, the random sequence best-response dynamic must eventually converge to the PNE of the game shown in Fig. 1, but it is by no means guaranteed to converge to a PNE in all games.¹⁰ In Sects. 2 and 3 we restrict our attention to playing sequences $s \in \{s_c, s_x\}$.

A *path* $\langle \vec{\mathbf{a}}, s \rangle$ is an infinite sequence of action profiles $\vec{\mathbf{a}} = (\mathbf{a}^0, \mathbf{a}^1, \dots)$ and an associated playing sequence function $s : \mathbb{N} \rightarrow [n]$ satisfying the constraint that only one player changes her action at a time, i.e. $\mathbf{a}^t_{-s(t)} = \mathbf{a}^{t-1}_{-s(t)}$ for each $t \in \mathbb{N}$. So only the action of player $s(t)$ is allowed to differ between profiles \mathbf{a}^{t-1} and \mathbf{a}^t along a path.

The *best-response dynamic* with playing sequence $s : \mathbb{N} \rightarrow [n]$ on a game $g_{n,m}$ initiated at the action profile \mathbf{a}^0 is the following process: set the initial action profile to \mathbf{a}^0 and, at each time $t \in \mathbb{N}$, player $s(t)$ myopically plays her best-response $a^t_i = b_i(\mathbf{a}^{t-1}_{-i})$ to her current environment $\mathbf{a}^{t-1}_{-s(t)}$. The best-response dynamic effectively generates a path $\langle \vec{\mathbf{a}}, s \rangle$ by traveling along the edges of the best-response digraph $\mathcal{D}(g_{n,m})$ in direction $s(t)$ at step t starting from the initial profile \mathbf{a}^0 .¹¹

2.4 Convergence

For any path $\langle \vec{\mathbf{a}}, s \rangle$ and set of action profiles $\mathcal{A} \subseteq \mathcal{M}$ the *hitting time* $H_{\langle \vec{\mathbf{a}}, s \rangle}(\mathcal{A}) := \inf\{t \in \mathbb{N} : \mathbf{a}^t \in \mathcal{A}\}$ is the first time $t \geq 1$ at which some element of the sequence $\vec{\mathbf{a}}$ is in, or (first) hits, the set \mathcal{A} (inf is the infimum operator and we use the convention that $\inf \emptyset = \infty$).¹² We say that the s -sequence best-response dynamic on game $g_{n,m}$ initiated at \mathbf{a}^0 *converges* to a PNE if its path $\langle \vec{\mathbf{a}}, s \rangle$ hits $\text{PNE}(g_{n,m})$ in finite time. Clearly, if a path hits a PNE at some time t , it stays there forever after.

2.5 Best-response dynamics on random games

We generate random games by drawing all payoffs at random: for each $\mathbf{a} \in \mathcal{M}$ and $i \in [n]$, the payoff $U_i(\mathbf{a})$ is a random number that is drawn from an atomless distribution \mathbb{P} . The draws are independent across all $i \in [n]$ and $\mathbf{a} \in \mathcal{M}$. The distribution \mathbb{P} ensures that any ties in payoffs have zero measure, so almost surely each

¹⁰ It is, for example, easy to construct games with a PNE in which there is a cluster of non-PNE profiles that, once visited, cannot be escaped by the random sequence best-response dynamic. Amiet et al. (2021) refer to such clusters as “best-response traps”.

¹¹ More precisely, the infinite sequence of actions $\vec{\mathbf{a}}$ is determined as follows: if player $s(t)$ is already best responding then \mathbf{a}^{t-1} does not point to any vertex $(a^t_{s(t)}, \mathbf{a}^{t-1}_{-s(t)}) \neq \mathbf{a}^{t-1}$ and the next profile in the sequence is \mathbf{a}^{t-1} itself, i.e. $\mathbf{a}^t = \mathbf{a}^{t-1}$; otherwise, if player $s(t)$ is not already playing her best response then travel to the vertex that corresponds to her playing her best-response action, i.e. set $\mathbf{a}^t = (a^t_{s(t)}, \mathbf{a}^{t-1}_{-s(t)})$ where $(a^t_{s(t)}, \mathbf{a}^{t-1}_{-s(t)}) \neq \mathbf{a}^{t-1}$ is the unique vertex that \mathbf{a}^{t-1} points to.

¹² We also say that the path hits \mathcal{A} *by* t if it hits \mathcal{A} at time τ with $\tau \leq t$, and the path hits \mathcal{A} *before* (after) t if it hits \mathcal{A} at time $\tau < t$ ($\tau > t$).

environment has a unique best-response for each player. A random game drawn in this way is denoted by $G_{n,m} := ([n], \{m_i\}_{i \in [n]}, \{U_i\}_{i \in [n]})$.

The best-response dynamic on random games is described by Algorithm 1. We randomly draw a game and run the best-response dynamic on the drawn game, starting from a randomly drawn initial profile \mathbf{A}^0 .¹³ Doing so induces a distribution over paths and PNE sets.

Algorithm 1 s -sequence best-response dynamic on $G_{n,m}$

- (1) For all $i \in [n]$ and $\mathbf{a} \in \mathcal{M}$ draw $U_i(\mathbf{a})$ at random according to \mathbb{P}
 - (2) Draw \mathbf{A}^0 uniformly at random from \mathcal{M}
 - (3) For $t \in \mathbb{N}$:
 - (a) Set $i = s(t)$
 - (b) Set $\mathbf{A}_{-i}^t = \mathbf{A}_{-i}^{t-1}$
 - (c) Set $A_i^t = B_i(\mathbf{A}_{-i}^{t-1})$ where $B_i(\mathbf{A}_{-i}^{t-1}) := \arg \max_{x_i \in [m_i]} U_i(x_i, \mathbf{A}_{-i}^{t-1})$
-

The notion of convergence given in Sect. 2.4 applies here. Namely, the s -sequence best-response dynamic on game $G_{n,m}$ (and initial condition \mathbf{A}^0) converges to a PNE if its path $\langle \vec{\mathbf{A}}, s \rangle$ (generated according to Algorithm 1) hits $\text{PNE}(G_{n,m})$ in finite time.

3 Theoretical results

In this section, we present the theoretical results for best-response dynamics in random games. In Sect. 3.1 we focus on games with $n > 2$ players. In this case, we find that best-response dynamics behave very differently under clockwork vs. random playing sequences. Most of our results on the probability of convergence to equilibrium are asymptotic. In Sect. 3.2 we focus on games with $n = 2$ players. In this case, the probability of convergence to equilibrium is the same under both clockwork and random playing sequences. Furthermore, we are able to provide asymptotic as well as *exact* results for game duration and for the probability of convergence to equilibrium.

The quantity

$$q_{n,m} := \frac{\prod_{i \in [n]} m_i}{\max_{i \in [n]} m_i}$$

is central to our results and it appears frequently in the literature on random games (for example, see Dresher 1970, or Rinott and Scarsini 2000). As summarized in the proposition below, the probability that there is a pure Nash equilibrium is asymptotically $1 - \exp\{-1\} \approx 0.63$ as $q_{n,m}$ gets large.

¹³ We draw the initial profile \mathbf{A}^0 uniformly at random from among all profiles, but this is merely a stylistic choice: since the game itself is drawn at random, the choice of initial condition is actually irrelevant, i.e. our results would not change if we had arbitrarily fixed the initial profile to some specific value.

Proposition 1 (Rinott and Scarsini 2000)

$$\lim_{q_{n,m} \rightarrow \infty} \Pr [\#PNE (G_{n,m}) \geq 1] = 1 - \exp\{-1\}.$$

Since $q_{n,m} \rightarrow \infty$ if and only if $n \rightarrow \infty$ or $m_i \rightarrow \infty$ for at least two players i , the probability that there is a PNE in a randomly drawn game approaches $1 - \exp\{-1\}$ when the number of players gets large or when the number of actions per player gets large for at least two players.¹⁴

3.1 Games with $n > 2$ players

The following result shows that, in large 2-action games, the random sequence best-response dynamic converges with high probability to a PNE if there is one. Let 2 denote a n -vector of 2s.

Proposition 2 (Amiet et al. 2021)

$$\lim_{n \rightarrow \infty} \Pr [s_2\text{-best-response dynamic on } G_{n,2} \text{ converges to a PNE} \mid \#PNE(G_{n,2}) \geq 1] = 1.$$

Combined with Proposition 1, it follows that over the class of all 2-action games, the random sequence best-response dynamic converges to a PNE with probability about $(1 - \exp\{-1\})$, i.e. in approximately 63% of those games, when the number of players is large.

A generalization of Proposition 2 to games with more than 2 actions per player is non-trivial. There are currently no existing analytical results for such cases, so this area remains open for future research. However, we conjecture that for $n > 2$, the random sequence best-response dynamic converges to a PNE with high probability if there is one as $q_{n,m} \rightarrow \infty$. Consistent with this conjecture, in the simulations of Sect. 4 we show that, provided $n > 2$, the random sequence best-response dynamic does converge to a PNE with probability close to $1 - \exp\{-1\}$ when n gets large or when the number of actions gets large for at least two players.

Our main result for the clockwork sequence best-response dynamic in games with $n > 2$ players is given below.

Theorem 1

$$\frac{1}{4\sqrt{n}} \frac{1}{\sqrt{q_{n,m}}} \leq \Pr \left[\begin{array}{l} s_c\text{-best-response dynamic} \\ \text{on } G_{n,m} \text{ converges to a PNE} \end{array} \right] \leq \frac{6n\sqrt{\log(q_{n,m})}}{\sqrt{q_{n,m}}}.$$

Consequently, since the upper and lower bound both go to zero as n gets large or when the number of actions gets large for at least two players,

¹⁴ Using results from Arratia et al. (1989), Rinott and Scarsini (2000) prove the stronger result that the distribution of the number of PNE in random games is asymptotically Poisson(1) as $q_{n,m} \rightarrow \infty$. The probability that a PNE exists in a random game was previously studied by Goldberg et al. (1968) in the 2-player case and by Dresher (1970) in the n -player case as the number of actions gets large for at least two players. Powers (1990) and Stanford (1995) noted that the distribution of $\#PNE(G_{n,m})$ approaches a Poisson(1) as the number of actions gets large.

$$\lim_{q_{n,m} \rightarrow \infty} \Pr [s_c\text{-best-response dynamic on } G_{n,m} \text{ converges to a PNE}] = 0.$$

So, with high probability, the clockwork sequence best-response dynamic does not converge to a PNE as the number of players gets large or as the number of actions for at least two players gets large. This is in sharp contrast with the asymptotic behavior of the random sequence best-response dynamic. It is intuitive that the clockwork sequence converges to a PNE less often than the random sequence because it will have more difficulty escaping cycles in a best-response digraph. That said, the extreme swing in the asymptotic probability of convergence from 1 to 0 is rather striking.

We briefly comment on Theorem 1 and its implications. (i) In Algorithm 1, drawing payoffs independently at random (from an atomless distribution) induces a uniform distribution over best-response digraphs.¹⁵ It is in this sense that we can say that the best-response dynamic converges in a “large” or “small” fraction of all games. (ii) Our proof of Theorem 1 relies on a coupling argument (explained in the appendix) that makes it possible to deal with the path-dependence of the best-response dynamic (which arises from the fact that if a player encounters an environment that they had seen before, they must play the same action that they played when the environment was first encountered). The proof centers on bounding the time it takes for some player to re-encounter a previously seen environment along a best-response path and this time is fundamentally determined by $q_{n,m}$, which is the minimal number of possible environments. (iii) In fact, Theorem 1 gives us the following corollary, which shows that the asymptotic probability of convergence to equilibrium is determined primarily by the value of the parameter $q_{n,m}$.¹⁶

Corollary 1 *The asymptotic probability that the clockwork sequence best-response dynamic converges to a PNE is, up to a polynomial factor, of order $1/\sqrt{q_{n,m}}$.*

3.2 Games with $n = 2$ players

For $n = 2$ players, we provide detailed results on both game duration and on the probability of convergence to equilibrium.

If the path (\vec{a}, s_c) generated by the clockwork best-response dynamic on a 2-player game $g_{2,m}$ has the property that from t onwards, the sequence of $2k$ possibly non-distinct action profiles $\mathbf{a}^t, \dots, \mathbf{a}^{t+2k-1}$ repeats itself forever and t is the hitting time

¹⁵ This follows from the manner in which the payoffs are drawn: there is a zero probability of ties because \mathbb{P} is atomless and for each $i \in [n]$ the probability that action $a_i \in [m_i]$ is a best-response to environment \mathbf{a}_{-i} is given by

$$\Pr \left[U_i(a_i, \mathbf{a}_{-i}) = \max_{x_i \in [m_i]} U_i(x_i, \mathbf{a}_{-i}) \right] = \frac{1}{m_i}.$$

¹⁶ Since $\log(q_{n,m})$ is dominated by a polynomial in n and m , and $q_{n,m}$ grows faster than $\log(q_{n,m})$ and than n to any power, the asymptotic behavior of each bound is governed by the behavior of the term $\sqrt{q_{n,m}}$ in the denominator.

to \mathbf{a}^t , then we say that the clockwork best-response dynamic converged to a cycle of length $2k$, or a $2k$ -cycle, at time t , where $k \in \{1, \dots, m_*\}$ and $m_* := \min\{m_1, m_2\}$.

Theorem 2 For any $k \in \{1, \dots, m_*\}$ and $t \in \{1, \dots, 2(m_* - k + 1)\}$,¹⁷ the probability that the s_c -best-response dynamic on $G_{2,m}$ converges to a $2k$ -cycle at time t is given by

$$\frac{1}{m_{s_c(t+2k-1)}} \prod_{i=1}^{t+2k-2} \left(1 - \frac{1}{m_{s_c(i)}} \left\lfloor \frac{i}{2} \right\rfloor \right). \tag{1}$$

Thus we have an exact expression for the probability that the clockwork sequence best-response dynamic converges to a $2k$ -cycle at time t .¹⁸ Setting $k = 1$ in (1) yields the exact probability that the clockwork sequence best-response dynamic on $G_{2,m}$ converges to a PNE at time t .

As a straightforward corollary of Theorem 2, the probability that the clockwork sequence best-response dynamic converges to a $2k$ -cycle is obtained by summing (1) over all $t \in \{1, \dots, 2(m_* - k + 1)\}$:

Corollary 2 The probability that the s_c -best-response dynamic on $G_{2,m}$ converges to a $2k$ -cycle is given by

$$\sum_{t=1}^{2(m_*-k+1)} \frac{1}{m_{s_c(t+2k-1)}} \prod_{i=1}^{t+2k-2} \left(1 - \frac{1}{m_{s_c(i)}} \left\lfloor \frac{i}{2} \right\rfloor \right). \tag{2}$$

Setting $k = 1$ in (2) yields the exact probability that the clockwork sequence best-response dynamic on $G_{2,m}$ converges to a PNE.

To get a better sense of the behavior of (2), we now study its asymptotics, which are easiest to see when $m_1 = m_2 = m$. We maintain this restriction in the rest of this section. Let $\Phi(\cdot)$ denote the standard normal cumulative distribution function:

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{z^2}{2}\right\} dz.$$

We say that $f(n)$ is asymptotically $g(n)$ if $f(n)/g(n) \rightarrow 1$ as $n \rightarrow \infty$, and $f(n) = o(g(n))$ denotes $f(n)/g(n) \rightarrow 0$ as $n \rightarrow \infty$.

Proposition 3 Set $m_1 = m_2 = m$. If $k = o(m^{2/3})$ then, as $m \rightarrow \infty$, (2) is asymptotically

$$2\sqrt{\frac{\pi}{m}} \left(1 - \Phi\left(\frac{2k-1}{\sqrt{2m}}\right) \right).$$

If $k = o(\sqrt{m})$ then, as $m \rightarrow \infty$, (2) is asymptotically $\sqrt{\pi/m}$.

¹⁷ For any $k \in \{1, \dots, m_*\}$ the product is non-negative provided $t + 2k - 2 \leq 2m_*$.

¹⁸ See also Pangallo et al. (2019) for an exact formula giving the probability of existence of cycles of any length in 2-player games.

The asymptotics given in Proposition 3 help us to better understand the behavior of the clockwork sequence best-response dynamic in large 2-player games. (i) The probability of convergence to a PNE, which corresponds to setting $k = 1$, goes to zero when $m \rightarrow \infty$.¹⁹ (ii) Short cycles all have about the same probability. Indeed, for $k = o(\sqrt{m})$ the probability is asymptotically $\sqrt{\pi/m}$. Finally, (iii) it is very unlikely that the best-response dynamic converges to a very long cycle: if $k/\sqrt{m} \rightarrow \infty$ then the probability that the dynamic converges to a cycle of length at least $2k$ tends to 0.²⁰

Theorem 3 *Set $m_1 = m_2 = m$ and fix $x > 0$. The probability that the s_c -best-response dynamic on $G_{2,m}$ does not hit a cycle (of any length) until at least time step $x\sqrt{2m}$ is asymptotically $\exp\{-x^2/2\}$ as $m \rightarrow \infty$.*

This result shows that the clockwork sequence best-response dynamic in 2-player games is likely to converge to a $2k$ -cycle (for some $k \in \{1, \dots, m\}$) within $\sqrt{2m}$ time steps when m is large.

We now compare the behavior of the clockwork sequence best-response dynamic in 2-player games with the behavior of the random sequence best-response dynamic in 2-player games. (i) The probability of convergence to a PNE is the same for clockwork and for random playing sequences in 2-player games. The reason is that, under the random playing sequence, players' actions do not change whenever the sequence asks the same player to play several times in a row. The profiles that are therefore visited along the path are the same under both playing sequences, which induces the same probability of convergence to equilibrium. However, (ii) the expected game duration will be different since the random playing sequence introduces delays. In fact, the expected game duration for the random playing sequence should be greater than for the clockwork playing sequence by a factor of 2. The reason is that, under the clockwork playing sequence, the players alternate at the tick of each time step, whereas, under the random playing sequence, the time it takes for the playing sequence to turn to the other player is Geometric($\frac{1}{2}$). Thus the random playing sequence can be considered as a slowing down of the clockwork playing sequence in which the expected time to play the next step is 2.

4 Simulation results

In Sects. 4.1 and 4.2, we run simulations of the clockwork and random sequence best-response dynamics. Our main goal is to investigate the extent to which our asymptotic results are also valid for a small number of players and actions. In

¹⁹ In contrast, for $n = 2$, Amiet et al. (2021) find that “better”- (rather than best-) response dynamics converge to a PNE (whenever there is one) with high probability as $m \rightarrow \infty$.

²⁰ When $k = o(\sqrt{m})$, the argument of $\Phi(\cdot)$ goes to zero. Since $\Phi(0) = 1/2$ we have that the convergence probability goes to $\sqrt{\pi/m}$ which is independent of k . If, instead, $k/\sqrt{m} \rightarrow \infty$ then the argument of $\Phi(\cdot)$ grows large and since $\Phi(\infty) = 1$, the convergence probability goes zero. Our proof of Proposition 3 derives the asymptotics for $k = o(m^{2/3})$. The standard normal has small tails outside this range.

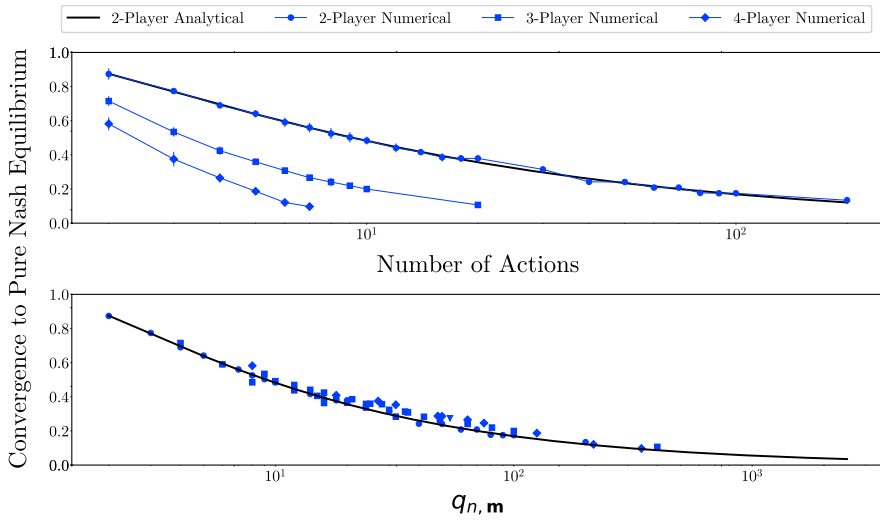


Fig. 2 Frequency of convergence to a PNE for the clockwork best-response dynamic

these simulations, for each choice of n and m , we randomly draw 10 batches of 1000 games. We run the best-response dynamic on each game and find the mean frequency of convergence to equilibrium in each batch, and then report the mean across the batches. The error bars in our figures are intervals of one empirical standard deviation (across the means for each batch).

In Sect. 4.3 we investigate the equilibrium convergence probability of playing sequences that interpolate between the extremes of clockwork and random playing sequences, and pay particular attention to the speed of convergence.

4.1 Simulations of clockwork best-response dynamics

The blue markers in Fig. 2 show the frequency of convergence to a PNE in our simulations for different values of numbers of players and actions. In both panels, the solid black line is the analytical probability of convergence to a PNE in 2-player m -action games, calculated using Eq. (2) with $m_1 = m_2 = m$.

In the top panel, we present simulation outcomes for 2, 3, and 4-player games in which all players have the same number of actions. Up to sampling noise, our analytical result for 2-player games perfectly matches the numerical simulations. We also find that convergence frequency becomes lower for a given number of actions as the number of players increases.

The blue markers in the bottom panel of Fig. 2 are the simulation means for different values of n and m , all chosen to ensure that the minimal number of environments in those games match the number of environments in a 2-player m -action game. All markers line up reasonably well along the solid black line. Corollary 1

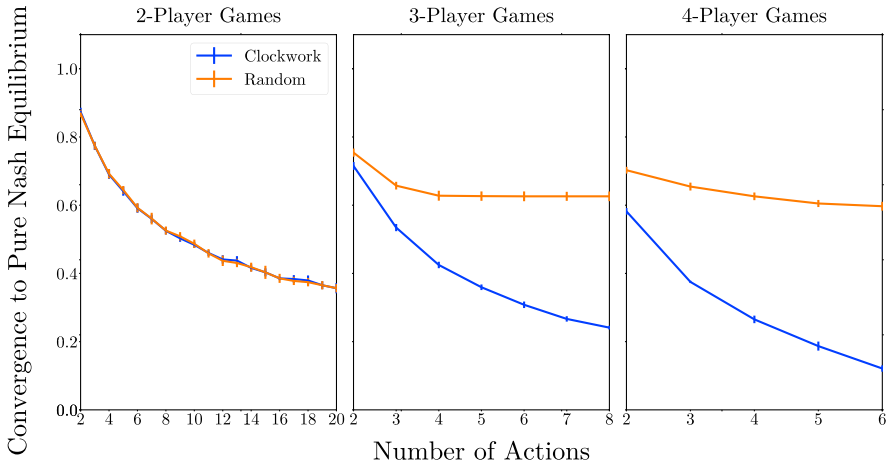


Fig. 3 Frequency of convergence to a PNE for clockwork vs. random best-response dynamics

implies that the asymptotic convergence probability in games $G_{n,m}$ and $G_{n',m'}$ is approximately the same whenever $q_{n,m} = q_{n',m'}$. Our results show that this relation holds even for relatively small games.

4.2 Simulations of random best-response dynamics

Figure 3 shows the frequency of convergence to a PNE under clockwork vs. random best-response dynamics in n -player games with m actions per player.²¹

As argued in Sect. 3.2, when there are only $n = 2$ players, the random playing sequence has the same convergence probability as the clockwork playing sequence, which can be seen in the left panel of Fig. 3.

Looking across the panels, the frequency of convergence to a PNE is decreasing in both n and m for the clockwork playing sequence, but the random playing sequence is different because its frequency of convergence rapidly settles near $1 - 1/e$ for $n > 2$. Recall, Amiet et al. (2021) proved that the random sequence best-response dynamic always converges to a PNE if there is one when $m = 2$ and $n \rightarrow \infty$. As argued in Sect. 3.1, this gives us an unconditional probability of convergence of $1 - 1/e \approx 63\%$. Our simulations show that the result of Amiet et al. (2021) also appears to hold for games with more than two actions provided $n > 2$. In fact, the random sequence best-response dynamic almost always converges to a PNE in games that have a PNE even for relatively small values of n and m .

²¹ The results also hold if we allow for different numbers of actions per player.

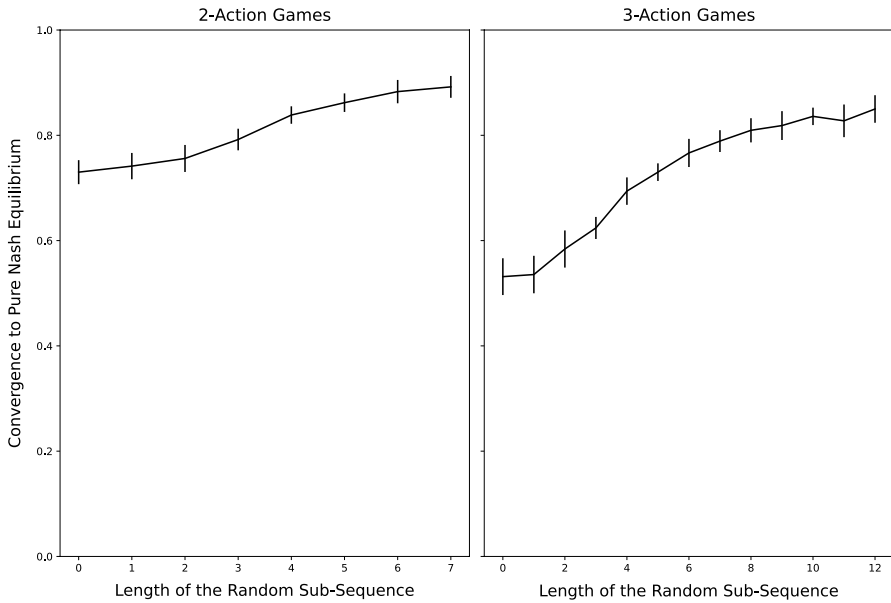


Fig. 4 Frequency of convergence to a PNE for p -periodic best-response dynamics in $n = 3$ player games

4.3 Simulations of periodic best-response dynamics

The analytical results of Sect. 3 allowed us to compare the behavior of two extreme playing sequences: clockwork and random. We now turn our attention to intermediate cases. A playing sequence is p -periodic if it consists of a sequence of players of length $p \geq n$ that is repeated forever, with the constraint that each player appears at least once in the repeated sequence. In other words,

$$i_1, i_2, \dots, i_p, i_1, i_2, \dots, i_p, i_1, i_2, \dots, i_p, \dots$$

is a p -periodic playing sequence if for each $i \in [n]$ there is some $j \in [p]$ such that $i_j = i$.

We generate p -periodic playing sequences at random as follows: construct a sequence of $p - n$ integers drawn at random from $[n]$ and append the numbers $1, \dots, n$ to this sequence. This results in a sequence of p integers. Now select a random permutation of this sequence and call it σ . Then $\sigma, \sigma, \sigma, \dots$ is a p -periodic playing sequence. Clearly, if $p = n$, we recover clockwork playing sequences. And, fixing n , we recover random playing sequences for $p \rightarrow \infty$.

Figure 4 plots the frequency of convergence to a pure Nash equilibrium against the length of the random subsequence (namely $p - n$) for p -periodic best-response dynamics in $n = 3$ player games with 2 and 3 actions per player. As might be expected, the probability of convergence to a Nash equilibrium for p -periodic playing sequences is increasing in p .

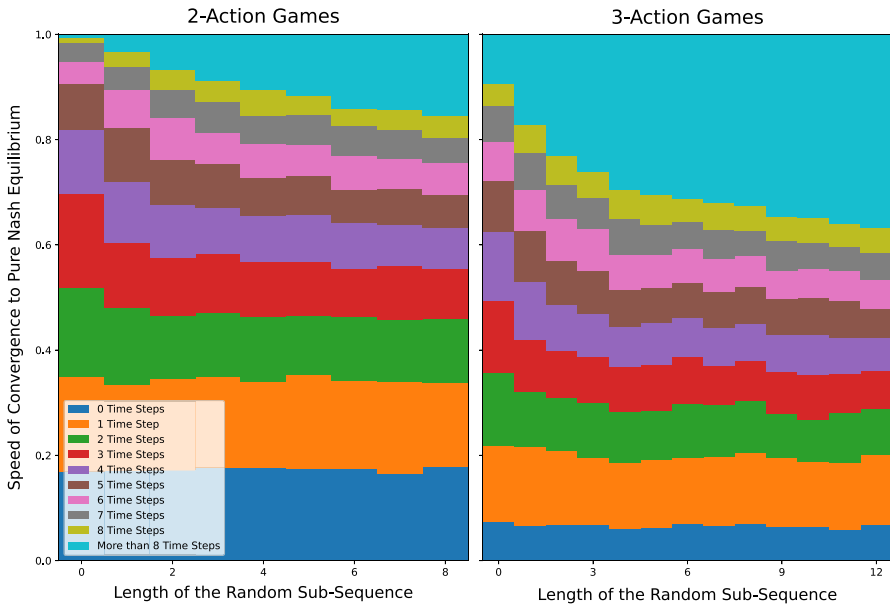


Fig. 5 Conditional speed of convergence to a PNE for p -periodic best-response dynamics in $n = 3$ player games

Figure 5 plots, for different lengths of the random subsequence, the distribution of the number of time steps until a pure Nash equilibrium is reached *conditional* on converging to a pure Nash equilibrium for p -periodic best-response dynamics in $n = 3$ player games with 2 and 3 actions per player. Interestingly, conditional on converging to a Nash equilibrium, the average number of time steps to reach equilibrium is increasing in p . This relates to the findings of Durand and Gaujal (2016) who showed that, in expectation, convergence to equilibrium in potential games is faster under the clockwork playing sequence than under any other playing sequence. Here, our results indicate that, over the space of all games, the speed of convergence (conditional on converging to equilibrium) is slower for playing sequences that have a larger share of random elements (i.e. a larger $p - n$). The apparent trade-off between the success in finding equilibria vs. the speed of convergence to equilibria is an interesting area for future research.

Appendix A: Proofs

The appendix concerns only the clockwork best-response dynamic and presents proofs for the results stated in the main body of the paper.

Proof of Theorem 1

We start by stating two lemmas that will be used to prove Theorem 1. Lemma 1 bounds the probability that the clockwork sequence best-response dynamic converges to a pure Nash equilibrium after time t . Lemma 2 bounds the probability that the clockwork sequence best-response dynamic converges to a pure Nash equilibrium by time t .

Lemma 1 *Let $\langle \vec{A}, s_c \rangle$ be generated according to Algorithm 1. For any $t \in \mathbb{N}$,*

$$\Pr \left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m}) \text{ after } t \right] \leq \exp \left\{ - \frac{\left(\left\lfloor \frac{t}{n} - 1 \right\rfloor \right)^2}{2q_{n,m}} \right\}.$$

Lemma 2 *Let $\langle \vec{A}, s_c \rangle$ be generated according to Algorithm 1. For any $t \in \mathbb{N}$,*

$$\left\lfloor \frac{t}{n} \right\rfloor \frac{1}{q_{n,m}} \left(1 - \frac{\left(\left\lfloor \frac{t}{n} \right\rfloor \right)^2}{2} \frac{n}{q_{n,m}} \right) \leq \Pr \left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m}) \text{ by } t \right] \leq \frac{t}{q_{n,m}}.$$

We now show how Theorem 1 follows from Lemmas 1 and 2 and, afterwards, we provide proofs for the lemmas themselves.

Proof of Theorem 1 Let $\langle \vec{A}, s_c \rangle$ be generated according to Algorithm 1. The probability that the s_c -best-response dynamic on $G_{n,m}$ converges to a PNE is equal to the probability that $\langle \vec{A}, s_c \rangle$ hits PNE($G_{n,m}$). Let us start with the upper bound. For any $t \in \mathbb{N}$,

$$\begin{aligned} & \Pr \left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m}) \right] \\ &= \Pr \left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m}) \text{ by } t \right] + \Pr \left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m}) \text{ after } t \right] \\ &\leq \frac{t}{q_{n,m}} + \exp \left\{ - \frac{\left(\left\lfloor \frac{t}{n} - 1 \right\rfloor \right)^2}{2q_{n,m}} \right\}. \end{aligned} \tag{3}$$

Equation (3) follows from Lemmas 1 and 2. Now, set

$$t = n \left(\left\lceil \sqrt{2q_{n,m} \log(q_{n,m})} \right\rceil + 1 \right).$$

Since $n \geq 2$ and $m_i \geq 2$ for all i , we have $\sqrt{2q_{n,m} \log(q_{n,m})} > 1$, so

$$n\left(\sqrt{2q_{n,m} \log(q_{n,m})} + 1\right) \leq t \leq n\left(\sqrt{2q_{n,m} \log(q_{n,m})} + 2\right) < 3n\sqrt{2q_{n,m} \log(q_{n,m})}.$$

It follows that

$$\frac{t}{q_{n,m}} < 3n\sqrt{\frac{2 \log(q_{n,m})}{q_{n,m}}}, \tag{4}$$

and that

$$\exp\left\{-\frac{\left(\left\lfloor \frac{t}{n} \right\rfloor - 1\right)^2}{2q_{n,m}}\right\} \leq \frac{1}{q_{n,m}} < n\sqrt{\frac{2 \log(q_{n,m})}{q_{n,m}}}. \tag{5}$$

Adding the upper bounds in (4) and (5) yields the desired result.

Let us now turn to the lower bound. For any $t \in \mathbb{N}$,

$$\begin{aligned} \Pr\left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m})\right] &\geq \Pr\left[\langle \vec{A}, s_c \rangle \text{ hits PNE}(G_{n,m}) \text{ by } t\right] \\ &\geq \left\lfloor \frac{t}{n} \right\rfloor \frac{1}{q_{n,m}} \left(1 - \frac{\left(\left\lfloor \frac{t}{n} \right\rfloor\right)^2}{2} \frac{n}{q_{n,m}}\right). \end{aligned} \tag{6}$$

Equation (6) follows from Lemma 2. Now, set

$$t = n \left\lfloor \frac{\sqrt{q_{n,m}}}{\sqrt{n}} \right\rfloor.$$

Then,

$$1 - \frac{\left(\left\lfloor \frac{t}{n} \right\rfloor\right)^2}{2} \frac{n}{q_{n,m}} \geq \frac{1}{2}. \tag{7}$$

And since $n \geq 2$ and $m_i \geq 2$ for all i , we have $t \geq \frac{1}{2}\sqrt{nq_{n,m}}$, so

$$\left\lfloor \frac{t}{n} \right\rfloor \frac{1}{q_{n,m}} \geq \frac{1}{2\sqrt{n}} \frac{1}{\sqrt{q_{n,m}}}. \tag{8}$$

Multiplying the lower bounds in (7) and (8) together yields the desired result. □

Lemmas

We now turn to the proofs of Lemmas 1 and 2. These require additional notation which we introduce here.

The notion of convergence given in Sect. 2.4 applies to all playing sequences but we can provide a more direct characterization of convergence (and non-convergence) in terms of path properties when the sequence is clockwork. We refer to one complete rotation of the clockwork sequence as a *round* of play; e.g. if a round starts at player i then each player plays once in order and the round is complete when it is once again i 's turn to play. For any $k \in \mathbb{N}$ define

$$T_{\langle \vec{a}, s_c \rangle}(k) := \inf \{ t \in \mathbb{N} : \mathbf{a}^t = \mathbf{a}^{t+nk} \text{ and } \mathbf{a}^t \neq \mathbf{a}^{t+nk'} \text{ for all } k' \in \mathbb{N} \text{ such that } k' < k \},$$

to be the first time at which an action profile is repeated k rounds later (and at no earlier round). If $T_{\langle \vec{a}, s_c \rangle}(k)$ is finite, the path $\langle \vec{a}, s_c \rangle$ has the property that from time $T_{\langle \vec{a}, s_c \rangle}(k)$ onwards, the sequence of nk possibly non-distinct action profiles $\mathbf{a}^t, \dots, \mathbf{a}^{t+nk-1}$ repeats itself forever, and we say that the path $\langle \vec{a}, s_c \rangle$ (first) hits an nk -cycle at time $T_{\langle \vec{a}, s_c \rangle}(k)$. Note that there is exactly one k such that $T_{\langle \vec{a}, s_c \rangle}(k)$ is finite.

If the action profile is \mathbf{a}^t at some time t and no one deviates from this profile in a single round (i.e. $\mathbf{a}^t = \mathbf{a}^{t+n}$), then \mathbf{a}^t must be a PNE. Therefore, if the path $\langle \vec{a}, s_c \rangle$ hits an nk -cycle at time $T_{\langle \vec{a}, s_c \rangle}(k)$ and $k = 1$ ($k > 1$) then the clockwork sequence best-response dynamic converges to a PNE (a best-response cycle of length nk) at that time.

Let

$$T_{\langle \vec{a}, s_c \rangle} := \inf \{ T_{\langle \vec{a}, s_c \rangle}(k) : k \in \mathbb{N} \},$$

denote the first time (necessarily finite) at which the path $\langle \vec{a}, s_c \rangle$ hits a PNE or a best-response cycle.

For any path $\langle \vec{a}, s_c \rangle$ and for each $t \in \mathbb{N}$ define

$$f_{\langle \vec{a}, s_c \rangle}(t) := \min \left\{ u \leq t : \mathbf{a}_{-s_c(u)}^{u-1} = \mathbf{a}_{-s_c(t)}^{t-1} \text{ and } s_c(u) = s_c(t) \right\}.$$

So $f_{\langle \vec{a}, s_c \rangle}(t)$ is the first time along the path $\langle \vec{a}, s_c \rangle$ that player $s_c(t)$ encounters the environment $\mathbf{a}_{-s_c(t)}^{t-1}$. Finally, define

$$F_{\langle \vec{a}, s_c \rangle} := \inf \{ t \in \mathbb{N} : f_{\langle \vec{a}, s_c \rangle}(t) < t \}.$$

So $F_{\langle \vec{a}, s_c \rangle}$ is the first time (necessarily finite) at which some player encounters an environment that they encountered previously along the path.

Remark 1 notes that any path generated by the clockwork best-response dynamic must hit a PNE or a best-response cycle before any player encounters an environment for the second time.

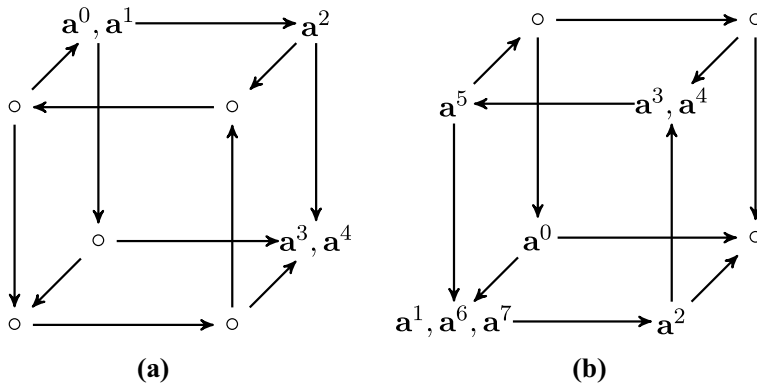


Fig. 6 The digraphs above are identical and correspond to the best-response digraph of the game shown in Fig. 1 but we now omit labels to avoid clutter. In panel (a), the initial profile is set to \mathbf{a}^0 . The first few elements of the infinite sequence $\bar{\mathbf{a}}$ are shown. Once at the profile \mathbf{a}^3 at $t = 3$, which is the unique PNE, the path remains there forever. Here, $T_{(\bar{\mathbf{a}}, s_c)} = T_{(\bar{\mathbf{a}}, s_c)}(1) = 3$ and $F_{(\bar{\mathbf{a}}, s_c)} = 6$. In panel (b), we have a different initial profile \mathbf{a}^0 . The path moves to the bottom left corner on the front face of the cube at $t = 1$ and then cycles forever among the four profiles on the front face of the cube. In fact, $T_{(\bar{\mathbf{a}}, s_c)} = T_{(\bar{\mathbf{a}}, s_c)}(2) = 1$, so the path hits a 6-cycle at time 1: once reached, the (not all distinct) action profiles in the sequence $\mathbf{a}^1, \dots, \mathbf{a}^6$ are repeated forever. Here, $F_{(\bar{\mathbf{a}}, s_c)} = 7$

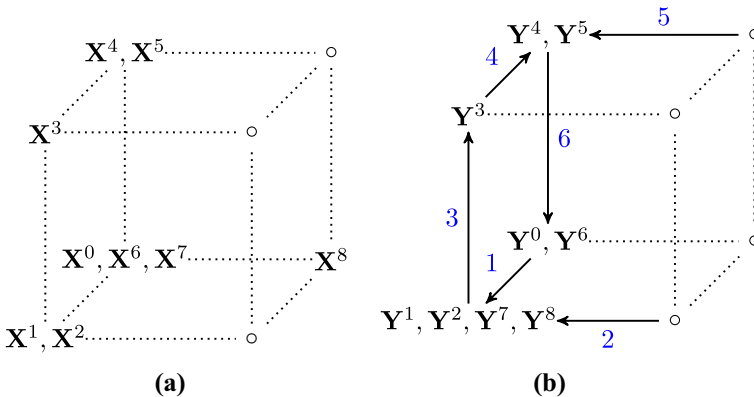


Fig. 7 Illustration of Algorithms 2 and 3. Panel (a) shows the first few elements of a possible path $(\bar{\mathbf{X}}, s_c)$ generated according to the clockwork random walk starting at the profile \mathbf{X}^0 . Panel (b) illustrates the first few elements of the corresponding path $(\bar{\mathbf{Y}}, s_c)$ generated according to Algorithm 3, starting with an empty digraph and numbering the directed edges according to the time at which they are first placed. The paths in panels (a) and (b) are identical up to and including time 6. At time step 7, however, player 1 encounters the same environment that she had encountered at time 1 ($F_{(\bar{\mathbf{X}}, s_c)} = F_{(\bar{\mathbf{Y}}, s_c)} = 7$); namely, players 2 and 3 each choosing action 1. The first time that player 1 encountered this environment, she responded by playing action 2, so she must play action 2 again at time 7. From then on, the path in panel (b) will keep cycling among the action profiles on the left-hand side of the cube forever whereas the path in panel (a) is allowed to wander freely

Remark 1 $T_{\langle \bar{\mathbf{a}}, s_c \rangle} < F_{\langle \bar{\mathbf{a}}, s_c \rangle}$.

Roughly speaking, $T_{\langle \bar{\mathbf{a}}, s_c \rangle}$ denotes the time at which the path $\bar{\mathbf{a}}$ hits an nk -cycle (for some $k \geq 1$) whereas $F_{\langle \bar{\mathbf{a}}, s_c \rangle}$ denotes the time at which the path completes its first circuit.

The quantities $T_{\langle \bar{\mathbf{a}}, s_c \rangle}(k)$, $T_{\langle \bar{\mathbf{a}}, s_c \rangle}$, and $F_{\langle \bar{\mathbf{a}}, s_c \rangle}$ are illustrated in an example in Fig. 6.

The main challenge posed by paths generated according to Algorithm 1 is that they have “memory”: whenever player $s_c(t)$ encounters an environment that she has encountered before (i.e. $\mathbf{A}_{-s_c(t)}^{t-1} = \mathbf{A}_{-s_c(u)}^{u-1}$ for some $u < t$ with $s_c(t) = s_c(u)$) then, at time t , the player must play the same action that she played when she previously encountered the environment (i.e. $A_{s_c(t)}^t = A_{s_c(u)}^u$). This path-dependence complicates the analysis of the clockwork best-response dynamic. We therefore study a simpler (random walk) process that is “memoryless” to which we couple a dynamic that induces the same distribution over paths as Algorithm 1. The coupled dynamic follows the random walk process until an environment is encountered by some player for the second time and becomes deterministic thereafter. We elaborate on our argument’s reliance on this coupling after the proof of Lemma 1.

The coupled system is described by Algorithms 2 and 3 and is illustrated in Fig. 7. $\langle \vec{\mathbf{X}}, s_c \rangle$ and $\langle \vec{\mathbf{Y}}, s_c \rangle$ denote paths generated according to Algorithms 2 and 3 respectively.

Algorithm 2 Clockwork random walk

- (1) Draw an initial profile \mathbf{X}^0 uniformly at random from \mathcal{M}
 - (2) For $t \in \mathbb{N}$:
 - (a) Set $i = s_c(t)$
 - (b) Set $\mathbf{X}_{-i}^t = \mathbf{X}_{-i}^{t-1}$
 - (c) Independently draw X_i^t uniformly at random from $[m_i]$
-

Algorithm 3 Coupled dynamic

- (1) Set $R_i(\mathbf{a}_{-i}) = 0$ for all $i \in [n]$ and $\mathbf{a}_{-i} \in \mathcal{M}_{-i}$
 - (2) Set the initial action profile to $\mathbf{Y}^0 = \mathbf{X}^0$
 - (3) For $t \in \mathbb{N}$:
 - (a) Set $i = s_c(t)$
 - (b) Set $\mathbf{Y}_{-i}^t = \mathbf{Y}_{-i}^{t-1}$
 - (c) If $R_i(\mathbf{Y}_{-i}^{t-1}) = 0$: set $Y_i^t = X_i^t$ and $R_i(\mathbf{Y}_{-i}^{t-1}) = Y_i^t$
 If $R_i(\mathbf{Y}_{-i}^{t-1}) \neq 0$: set $Y_i^t = R_i(\mathbf{Y}_{-i}^{t-1})$
-

Algorithm 2 is a “clockwork random walk” on the set of action profiles \mathcal{M} . The walk starts at some randomly drawn initial profile \mathbf{X}^0 and, at each time t , moves in direction $s_c(t)$ to a profile chosen uniformly at random from among the $m_{s_c(t)}$ profiles in that direction. A path generated according to this process does not have memory.

Algorithm 3 describes the coupled dynamic. The process starts at the same initial profile as the clockwork random walk. For each player i and environment \mathbf{a}_{-i} , we set the initial “response” value $R_i(\mathbf{a}_{-i})$ to zero and update it at step (3c) of the algorithm in the following manner: if the response value to the current environment \mathbf{Y}_{-i}^{t-1} is zero, then the environment was never encountered before and, in that case, player i ’s response value is set to X_i^t , the action drawn by the clockwork random walk at time t . If, on the other hand, the response value to the current environment \mathbf{Y}_{-i}^{t-1} is non-zero (i.e. the environment was encountered before), then this value is the action that i takes at time t . In other words, $\langle \vec{\mathbf{Y}}, s_c \rangle$ has the same memory property that is characteristic of paths generated according to Algorithm 1.

Algorithm 1 essentially draws a best-response digraph “up-front”, then selects an initial profile and traces a path by traveling along the edges of the digraph starting at the initial profile and moving in direction $s_c(t)$ at step t . In contrast, Algorithm 3 starts with an empty digraph and then generates its edges in an “online” manner. Nevertheless, both algorithms induce the same distribution over paths, as summarized in the following remark.

Remark 2 Let $\langle \vec{\mathbf{A}}, s_c \rangle$ and $\langle \vec{\mathbf{Y}}, s_c \rangle$ be generated according to Algorithms 1 and 3 respectively. Then $\langle \vec{\mathbf{A}}, s_c \rangle$ and $\langle \vec{\mathbf{Y}}, s_c \rangle$ have the same distribution.

By construction, the sequences $\vec{\mathbf{X}}$ and $\vec{\mathbf{Y}}$ must agree at least up to (but not including) the time at which some player encounters an environment for the second time. At such a time, under Algorithm 3, the player must play the action determined by their response function evaluated at that environment but, under Algorithm 2, the next action may be any of the available actions for that player. Remark 3 summarizes the key relationship between the clockwork random walk and the coupled dynamic.

Remark 3 $F_{\langle \vec{\mathbf{X}}, s_c \rangle} = F_{\langle \vec{\mathbf{Y}}, s_c \rangle}$.

The lemma below, which concerns paths $\langle \vec{\mathbf{X}}, s_c \rangle$ that are generated by the clockwork random walk, is useful for proving Lemmas 1 and 2. Under the clockwork sequence, player $i \in [n]$ plays at time $h_i(k) := i + (k - 1)n$ for $k \in \mathbb{N}$. For any $i \in [n]$ and any time $t \geq i$, define

$$k_i^*(t) := 1 + \left\lfloor \frac{t - i}{n} \right\rfloor.$$

So $k_i^*(t)$ is the largest $k \in \mathbb{N}$ such that $h_i(k) \leq t$. Between times 1 and t (inclusive), player $i \in [n]$ plays at times $h_i(1), h_i(2), \dots, h_i(k_i^*(t))$ and encounters environments $\mathbf{X}_{-i}^{h_i(1)-1}, \mathbf{X}_{-i}^{h_i(2)-1}, \dots, \mathbf{X}_{-i}^{h_i(k_i^*(t))-1}$. Lemma 3 establishes bounds on the probability that these environments are all distinct.

Define $\mu := \prod_{i \in [n]} m_i$ to be the cardinality of \mathcal{M} .

Lemma 3 For any $i \in [n]$ and $t \in \mathbb{N}$,

$$1 - \frac{m_i}{\mu} \frac{\left(\left\lfloor \frac{t}{n} \right\rfloor\right)^2}{2} \leq \Pr \left[\mathbf{X}_{-i}^{h_i(k)-1} \text{ for } k \in \{1, \dots, k_i^*(t)\} \text{ are all distinct} \right] \leq \exp \left\{ -\frac{m_i \left(\left\lfloor \frac{t}{n} \right\rfloor - 1\right)^2}{2\mu} \right\}.$$

Proof For any $i \in [n]$, the environments $\mathbf{X}_{-i}^{h_i(1)-1}, \mathbf{X}_{-i}^{h_i(2)-1}, \dots, \mathbf{X}_{-i}^{h_i(k_i^*(t))-1}$ are independent because they are disjoint subsets of the draws of the clockwork random walk. Each environment is distributed uniformly on \mathcal{M}_{-i} , and since \mathcal{M}_{-i} has cardinality $\frac{\mu}{m_i}$,

$$\Pr \left[\mathbf{X}_{-i}^{h_i(k)-1} \text{ for } k \in \{1, \dots, k_i^*(t)\} \text{ are all distinct} \right] = \prod_{k=1}^{k_i^*(t)-1} \left(1 - \frac{m_i k}{\mu} \right). \tag{9}$$

If $k_i^*(t) > 1 + \frac{\mu}{m_i}$ then the probability in (9) must be zero, and the lemma holds trivially ($k_i^*(t) > 1 + \frac{\mu}{m_i}$ implies $\left\lfloor \frac{t-i}{n} \right\rfloor > \frac{\mu}{m_i}$ which, in turn, implies $\left\lfloor \frac{t}{n} \right\rfloor > \frac{\mu}{m_i}$, so the lower bound in the statement of the lemma is negative and the upper bound is positive). We will therefore consider the case in which $k_i^*(t) \leq 1 + \frac{\mu}{m_i}$.

We obtain the following upper bound:

$$\begin{aligned} \prod_{k=1}^{k_i^*(t)-1} \left(1 - \frac{m_i k}{\mu} \right) &\leq \prod_{k=1}^{k_i^*(t)-1} \exp \left\{ -\frac{m_i k}{\mu} \right\} \\ &\leq \exp \left\{ -\frac{m_i (k_i^*(t) - 1)^2}{2\mu} \right\} \leq \exp \left\{ -\frac{m_i \left(\left\lfloor \frac{t}{n} \right\rfloor - 1\right)^2}{2\mu} \right\}. \end{aligned}$$

The first step follows from $\exp\{x\} \geq 1 + x$ for all x . The final inequality follows from $k_i^*(t) - 1 = \left\lfloor \frac{t-i}{n} \right\rfloor \geq \left\lfloor \frac{t-n}{n} \right\rfloor = \left\lfloor \frac{t}{n} \right\rfloor - 1$.

We now turn to the lower bound:

$$\prod_{k=1}^{k_i^*(t)-1} \left(1 - \frac{m_i k}{\mu} \right) \geq 1 - \sum_{k=1}^{k_i^*(t)-1} \frac{m_i k}{\mu} \geq 1 - \frac{m_i k_i^*(t)^2}{2} \geq 1 - \frac{m_i \left(\left\lfloor \frac{t}{n} \right\rfloor\right)^2}{2}.$$

The first step is an application of the Weierstrass product inequality. The final inequality follows from the fact that $k_i^*(t) = 1 + \left\lfloor \frac{t-i}{n} \right\rfloor \leq 1 + \left\lfloor \frac{t-1}{n} \right\rfloor = \left\lfloor \frac{t}{n} \right\rfloor$. □

Define $m^* := \max_{i \in [n]} m_i$, so that $q_{n,m} = \frac{\mu}{m^*}$.

Proof of Lemma 1 Recall that $T_{\langle \vec{A}, s_c \rangle}$ is the first time at which the path $\langle \vec{A}, s_c \rangle$ hits a PNE or a best-response cycle. So $T_{\langle \vec{A}, s_c \rangle} > t$ is the event that $\langle \vec{A}, s_c \rangle$ hits PNE($G_{n,m}$) or a best-response cycle only after time t . It follows that

$$\Pr \left[\langle \bar{\mathbf{A}}, s_c \rangle \text{ hits PNE}(G_{n,m}) \text{ after } t \right] \leq \Pr \left[T_{\langle \bar{\mathbf{A}}, s_c \rangle} > t \right].$$

By Remarks 1, 2, and 3,

$$\Pr \left[T_{\langle \bar{\mathbf{A}}, s_c \rangle} > t \right] \leq \Pr \left[F_{\langle \bar{\mathbf{A}}, s_c \rangle} > t \right] = \Pr \left[F_{\langle \bar{\mathbf{Y}}, s_c \rangle} > t \right] = \Pr \left[F_{\langle \bar{\mathbf{X}}, s_c \rangle} > t \right].$$

Now, let us focus on the path $\langle \bar{\mathbf{X}}, s_c \rangle$ and consider a player i satisfying $m_i = m^*$. The environments that player i faces between times 1 and t are given in the sequence $\mathbf{X}_{-i}^{h_i(1)-1}, \mathbf{X}_{-i}^{h_i(2)-1}, \dots, \mathbf{X}_{-i}^{h_i(k_i^*(t))-1}$. The event $F_{\langle \bar{\mathbf{X}}, s_c \rangle} > t$ implies that the environments in this sequence are all distinct. Hence

$$\Pr \left[F_{\langle \bar{\mathbf{X}}, s_c \rangle} > t \right] \leq \Pr \left[\mathbf{X}_{-i}^{h_i(k)-1} \text{ for } k \in \{1, \dots, k_i^*(t)\} \text{ are all distinct} \right] \leq \exp \left\{ - \frac{\left(\left\lfloor \frac{t}{n} - 1 \right\rfloor \right)^2}{2q_{n,m}} \right\},$$

where the final step follows from Lemma 3. □

The proof of Lemma 1 illustrates why we study a coupled system. Finding an upper bound on the probability that $\langle \mathbf{A}, s_c \rangle$ hits $\text{PNE}(G_{n,m})$ after t is central to our proof of Theorem 1. Our key step consists in showing that this probability is bounded above by the probability that the environments $\mathbf{X}_{-i}^{h_i(1)-1}, \mathbf{X}_{-i}^{h_i(2)-1}, \dots, \mathbf{X}_{-i}^{h_i(k_i^*(t))-1}$, which are generated by the clockwork random walk, are all distinct. This latter probability is easy to work out because the environments are independent uniform random draws. To avoid coupling, one might be tempted to argue that since the probability that $\langle \bar{\mathbf{A}}, s_c \rangle$ hits $\text{PNE}(G_{n,m})$ after t is bounded above by the probability that the environments $\mathbf{A}_{-i}^{h_i(1)-1}, \mathbf{A}_{-i}^{h_i(2)-1}, \dots, \mathbf{A}_{-i}^{h_i(k_i^*(t))-1}$ generated by Algorithm 1 are all distinct, one only needs to work out this latter probability. But this probability is not straightforward to work out: *these* environments are not independent uniform random draws since they are generated by a path-dependent process.

To prove Lemma 2, we introduce a slight modification of Algorithm 3. Algorithm 4, which describes a dynamic that is also coupled with the clockwork random walk, is identical to Algorithm 3 except that for some particular profile \mathbf{x} the algorithm is initialized with $R_i(\mathbf{x}_{-i}) = x_i$ for all $i \in [n]$. This effectively “plants” a sink in the digraph (at \mathbf{x}).

Algorithm 4 Coupled dynamic with sink \mathbf{x}

- (1) Set $R_i(\mathbf{a}_{-i}) = 0$ for all $i \in [n]$ and $\mathbf{a}_{-i} \in \mathcal{M}_{-i}$
 - (2) Set $R_i(\mathbf{x}_{-i}) = x_i$ for all $i \in [n]$
 - (3) Set the initial action profile to $\mathbf{Z}^0 = \mathbf{X}^0$
 - (4) For $t \in \mathbb{N}$:
 - (a) Set $i = s_c(t)$
 - (b) Set $\mathbf{Z}_{-i}^t = \mathbf{Z}_{-i}^{t-1}$
 - (c) If $R_i(\mathbf{Z}_{-i}^{t-1}) = 0$: set $Z_i^t = X_i^t$ and $R_i(\mathbf{Z}_{-i}^{t-1}) = Z_i^t$
 If $R_i(\mathbf{Z}_{-i}^{t-1}) \neq 0$: set $Z_i^t = R_i(\mathbf{Z}_{-i}^{t-1})$
-

In the remaining steps, Algorithm 4 selects a random initial profile and starts tracing a path by traveling along edges that (other than those edges already pointing to \mathbf{x} in the initialization) are generated in an online manner. The paths traced by the clockwork random walk and this coupled dynamic with a sink at \mathbf{x} must agree at least up to (but not including) the time at which *either* an environment is encountered by a player for the second time *or* the environment is \mathbf{x}_{-i} for some player i .

$\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ denotes a path generated according to Algorithm 4.

Remark 4 Consider an arbitrary action profile $\mathbf{x} \in \mathcal{M}$ and let $\langle \vec{\mathbf{A}}, s_c \rangle$ and $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ be generated according to Algorithms 1 and 4 respectively. Then the distribution of $\langle \vec{\mathbf{A}}, s_c \rangle$ conditional on \mathbf{x} being a pure Nash equilibrium, i.e. conditional on $\mathbf{x} \in \text{PNE}(G_{n,m})$, is the same as the distribution of $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$.

Proof of Lemma 2 For any $t \in \mathbb{N}$,

$$\begin{aligned}
 & \Pr \left[\langle \vec{\mathbf{A}}, s_c \rangle \text{ hits } \text{PNE}(G_{n,m}) \text{ by } t \right] \\
 &= \sum_{\mathbf{x} \in \mathcal{M}} \Pr \left[\langle \vec{\mathbf{A}}, s_c \rangle \text{ hits } \{\mathbf{x}\} \text{ by } t \text{ and } \mathbf{x} \in \text{PNE}(G_{n,m}) \right] \\
 &= \sum_{\mathbf{x} \in \mathcal{M}} \Pr \left[\langle \vec{\mathbf{A}}, s_c \rangle \text{ hits } \{\mathbf{x}\} \text{ by } t \mid \mathbf{x} \in \text{PNE}(G_{n,m}) \right] \Pr \left[\mathbf{x} \in \text{PNE}(G_{n,m}) \right] \quad (10) \\
 &= \sum_{\mathbf{x} \in \mathcal{M}} \underbrace{\Pr \left[\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle \text{ hits } \{\mathbf{x}\} \text{ by } t \right]}_{(10.1)} \underbrace{\Pr \left[\mathbf{x} \in \text{PNE}(G_{n,m}) \right]}_{(10.2)}
 \end{aligned}$$

The final step follows from Remark 4; namely, the probability that $\langle \vec{\mathbf{A}}, s_c \rangle$ hits $\{\mathbf{x}\}$ by time t conditional on $\mathbf{x} \in \text{PNE}(G_{n,m})$ is equal to the probability that $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ hits $\{\mathbf{x}\}$ by time t . We now analyze the expressions (10.1) and (10.2).

For (10.2), since payoffs are drawn identically and independently according to the atomless distribution \mathbb{P} , we have that

$$\Pr [\mathbf{x} \in \text{PNE}(G_{n,m})] = \prod_{i=1}^n \Pr \left[U_i(\mathbf{x}) \geq \max_{x'_i \in [m_i]} U_i(x'_i, \mathbf{x}_{-i}) \right] = \frac{1}{\mu}. \tag{11}$$

We now find upper and lower bounds on (10.1) by relating $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ to the clockwork random walk path $\langle \vec{\mathbf{X}}, s_c \rangle$. We start with the upper bound. Notice that $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ cannot hit $\{\mathbf{x}\}$ by time t unless $\mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)}$ for some $\tau \leq t$. Therefore

$$\begin{aligned} \Pr \left[\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle \text{ hits } \{\mathbf{x}\} \text{ by } t \right] &\leq \Pr \left[\bigcup_{\tau=1}^t \{ \mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)} \} \right] \\ &\leq \sum_{\tau=1}^t \Pr \left[\mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)} \right] \\ &= \sum_{\tau=1}^t \frac{m_{s_c(\tau)}}{\mu} \\ &\leq \frac{t}{q_{n,m}}. \end{aligned} \tag{12}$$

The penultimate step follows from the fact that $\mathbf{X}_{-s_c(\tau)}^{\tau-1}$ consists of $n - 1$ independent uniform random variables (one action for each player other than $s_c(\tau)$), so $\mathbf{X}_{-s_c(\tau)}^{\tau-1}$ is itself uniformly drawn from $\mathcal{M}_{-s_c(\tau)}$, and $\mathcal{M}_{-s_c(\tau)}$ has cardinality $\frac{\mu}{m_{s_c(\tau)}}$.

We now turn to the lower bound. If $F_{\langle \vec{\mathbf{X}}, s_c \rangle} > t$ and $\mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)}$ for some $\tau \leq t$ then $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ must hit $\{\mathbf{x}\}$ by time t . In other words, if no environments are repeated for any player and the environment is \mathbf{x}_{-i} for some player i by time t , then $\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle$ must hit $\{\mathbf{x}\}$ by time t . Therefore,

$$\begin{aligned} \Pr \left[\langle \vec{\mathbf{Z}}, \mathbf{x}, s_c \rangle \text{ hits } \{\mathbf{x}\} \text{ by } t \right] &\geq \Pr \left[\bigcup_{\tau=1}^t \{ \mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)} \} \text{ and } F_{\langle \vec{\mathbf{X}}, s_c \rangle} > t \right] \\ &= \Pr \left[\bigcup_{\tau=1}^t \{ \mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)} \} \mid F_{\langle \vec{\mathbf{X}}, s_c \rangle} > t \right] \Pr \left[F_{\langle \vec{\mathbf{X}}, s_c \rangle} > t \right]. \end{aligned} \tag{13}$$

To bound the first term in (13), select a player i satisfying $m_i = m^*$ and notice that $\mathbf{x}_{-i}^{h_i(k)-1} = \mathbf{x}_{-i}$ for some $k \in \{1, \dots, k_i^*(t)\}$ implies that $\mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)}$ for some $\tau \leq t$. Therefore

$$\begin{aligned}
 \Pr \left[\bigcup_{\tau=1}^t \{ \mathbf{X}_{-s_c(\tau)}^{\tau-1} = \mathbf{x}_{-s_c(\tau)} \} \mid F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t \right] &\geq \Pr \left[\bigcup_{k=1}^{k_i^*(t)} \{ \mathbf{X}_{-i}^{h_i(k)-1} = \mathbf{x}_{-i} \} \mid F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t \right] \\
 &= \sum_{k=1}^{k_i^*(t)} \Pr \left[\mathbf{X}_{-i}^{h_i(k)-1} = \mathbf{x}_{-i} \mid F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t \right] \\
 &= \sum_{k=1}^{k_i^*(t)} \frac{m^*}{\mu} \\
 &\geq \left\lfloor \frac{t}{n} \right\rfloor \frac{1}{q_{n,m}}.
 \end{aligned}
 \tag{14}$$

The second line follows from the fact that since all the environments for our chosen player i are distinct, the events in the union are mutually exclusive. The next step follows from the fact that our process is invariant under symmetry. So for any $k \in \{1, \dots, k_i^*(t)\}$ and for all \mathbf{x}_{-i} and \mathbf{y}_{-i} , $\Pr[\mathbf{X}_{-i}^{h_i(k)-1} = \mathbf{x}_{-i} \mid F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t] = \Pr[\mathbf{X}_{-i}^{h_i(k)-1} = \mathbf{y}_{-i} \mid F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t]$ which implies that $\Pr[\mathbf{X}_{-i}^{h_i(k)-1} = \mathbf{x}_{-i} \mid F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t] = \frac{m^*}{\mu} = \frac{1}{q_{n,m}}$. The last step follows from $k_i^*(t) = 1 + \left\lfloor \frac{t-i}{n} \right\rfloor \geq 1 + \left\lfloor \frac{t}{n} - 1 \right\rfloor = \left\lfloor \frac{t}{n} \right\rfloor$.

To bound the second term in (13), notice that if for each $i \in [n]$ the environments $\mathbf{X}_{-i}^{h_i(1)-1}, \mathbf{X}_{-i}^{h_i(2)-1}, \dots, \mathbf{X}_{-i}^{h_i(k_i^*(t))-1}$ are all distinct then $F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t$. Therefore

$$\begin{aligned}
 \Pr[F_{\langle \bar{\mathbf{x}}, s_c \rangle} > t] &\geq \Pr \left[\bigcap_{i \in [n]} \{ \mathbf{X}_{-i}^{h_i(k)-1} \text{ for } k \in \{1, \dots, k_i^*(t)\} \text{ are all distinct} \} \right] \\
 &= 1 - \Pr \left[\bigcup_{i \in [n]} \{ \mathbf{X}_{-i}^{h_i(k)-1} \text{ for } k \in \{1, \dots, k_i^*(t)\} \text{ are not all distinct} \} \right] \\
 &\geq 1 - \sum_{i \in [n]} \Pr \left[\mathbf{X}_{-i}^{h_i(k)-1} \text{ for } k \in \{1, \dots, k_i^*(t)\} \text{ are not all distinct} \right] \\
 &\geq 1 - \frac{\left(\left\lfloor \frac{t}{n} \right\rfloor \right)^2 \sum_{i=1}^n m_i}{2 \mu} \\
 &\geq 1 - \frac{\left(\left\lfloor \frac{t}{n} \right\rfloor \right)^2 n}{2 q_{n,m}}.
 \end{aligned}
 \tag{15}$$

The penultimate step follows from Lemma 3.

Gathering the results (10), (11), (12), (14), and (15) together yields the desired conclusion. \square

Results for 2-player games

In games with $n = 2$ players, the action taken by player $s_c(t)$ at t corresponds exactly to the environment that player $s_c(t + 1)$ faces at $t + 1$. We take advantage of this property in our proof of Theorem 2 below.

Proof of Theorem 2 Let η_t denote the probability under the 2-player clockwork random walk that, by time t , no player plays an action that corresponds to an environment that was ever encountered by the other player. For $t \geq 1$ we have

$$\eta_{t+1} = \eta_t \times \left(1 - \frac{\lfloor \frac{t}{2} \rfloor}{m_{s_c(t+1)}} \right).$$

The term in parentheses is the probability that player $s_c(t + 1)$ does not repeat any of the $\lfloor \frac{t}{2} \rfloor$ environments encountered by player $s_c(t)$ by time t . Solving with $\eta_1 = 1$ yields

$$\eta_t = \prod_{i=1}^t \left(1 - \frac{1}{m_{s_c(i)}} \lfloor \frac{i}{2} \rfloor \right),$$

and, evidently, η_t is non-negative provided $t \leq 2m_*$.

For the path to hit a $2k$ -cycle at time t , it must be that (i) by time $t + 2k - 2$, no player plays an action that corresponds to an environment that was ever encountered by the other player, and (ii) the action taken by player $s_c(t + 2k - 1)$ at time $t + 2k - 1$ is equal to the environment encountered by player $s_c(t)$ at time t . So, the probability that the clockwork sequence best-response dynamic converges to a $2k$ -cycle at time t is $\frac{\eta_{t+2k-2}}{m_{s_c(t+2k-1)}}$, which completes the proof. \square

For the remaining proofs, we employ the following standard notation for asymptotics: we write $f(n) = o(g(n))$ if $f(n)/g(n) \rightarrow 0$ as $n \rightarrow \infty$, $f(n) \sim g(n)$ if $f(n)/g(n) \rightarrow 1$ as $n \rightarrow \infty$, and $f(n) = O(g(n))$ if there is $M > 0$ and N such that $|f(n)| \leq Mg(n)$ for all $n \geq N$.

Proof of Theorem 3 η_t is precisely the probability that the clockwork best-response dynamic does not hit a $2k$ -cycle (for any k) until at least time t . With $m_1 = m_2 = m$ we can write η_t as

$$\prod_{i=1}^t \left(1 - \frac{1}{m} \lfloor \frac{i}{2} \rfloor \right) = \begin{cases} \frac{m!^2}{(m - \frac{t+1}{2})!^2 m^{t+1}} & \text{if } t \text{ is odd} \\ \left(\frac{m - \frac{t}{2}}{m} \right) \frac{m!^2}{(m - \frac{t}{2})!^2 m^t} & \text{if } t \text{ is even} \end{cases}.$$

Using Stirling’s formula which states that

$$n! \sim \sqrt{2\pi n} \cdot n^n \exp\{-n\},$$

as $n \rightarrow \infty$, we obtain

$$\frac{m!^2}{(m - \frac{t+1}{2})!^2 m^{t+1}} \sim \left(\frac{m - \frac{t+1}{2}}{m}\right)^{t-2m} \exp\{-(t+1)\}, \tag{16}$$

and

$$\left(\frac{m - \frac{t}{2}}{m}\right) \frac{m!^2}{(m - \frac{t}{2})!^2 m^t} \sim \left(\frac{m - \frac{t}{2}}{m}\right)^{t-2m} \exp\{-t\}, \tag{17}$$

whenever $m - t \rightarrow \infty$. Taking a logarithm of the last expression we get

$$\begin{aligned} -t + (t - 2m) \ln\left(1 - \frac{1}{m} \frac{t}{2}\right) &= -t + (t - 2m) \left(-\frac{1}{2} \frac{t}{m} - \frac{1}{8} \frac{t^2}{m^2} + O\left(\frac{t^3}{m^3}\right)\right) \\ &= -\frac{1}{4} \frac{t^2}{m} + O\left(\frac{t^3}{m^2}\right). \end{aligned}$$

Provided that $t = o(m^{2/3})$, the second term goes to zero and therefore Eq. (17) behaves asymptotically like $\exp\{-t^2/(4m)\}$. An identical argument shows that, under the same conditions, (16) is also asymptotically $\exp\{-t^2/(4m)\}$. Hence,

$$\prod_{i=1}^t \left(1 - \frac{1}{m} \left\lfloor \frac{i}{2} \right\rfloor\right) \sim \exp\left\{-\frac{t^2}{4m}\right\}. \tag{18}$$

This completes the proof of Theorem 3. Note that approximation (18) holds uniformly in the range $[1, o(m^{2/3})]$. □

Proof of Proposition 3 Let $T = T(m)$ satisfy $T = o(m^{2/3})$ and $k = o(T)$. We assume that $T \geq \frac{m^{2/3}}{\ln(m)}$ so that T is not too small, and we split the summation in (2) into two ranges: $t \leq T$ and $t > T$. Since (18) holds uniformly in our first range, we have

$$\frac{1}{m} \sum_{t=1}^T \prod_{i=1}^{t+2(k-1)} \left(1 - \frac{1}{m} \left\lfloor \frac{i}{2} \right\rfloor\right) \sim \frac{1}{m} \sum_{t=1}^T \exp\left\{-\frac{(t+2(k-1))^2}{4m}\right\}.$$

We now approximate the summation on the right-hand side with an integral. Firstly, note that

$$\begin{aligned}
 \frac{1}{m} \int_1^{T+1} \exp \left\{ -\frac{(t + 2(k - 1))^2}{4m} \right\} dt &= \sqrt{\frac{2}{m}} \int_{\frac{2k-1}{\sqrt{2m}}}^{\frac{T+1+2(k-1)}{\sqrt{2m}}} \exp \left\{ -\frac{x^2}{2} \right\} dx \\
 &\sim \sqrt{\frac{2}{m}} \int_{\frac{2k-1}{\sqrt{2m}}}^{\infty} \exp \left\{ -\frac{x^2}{2} \right\} dx \quad (19) \\
 &= 2\sqrt{\frac{\pi}{m}} \left(1 - \Phi \left(\frac{2k - 1}{\sqrt{2m}} \right) \right),
 \end{aligned}$$

where the first step uses the transformation $x = (t + 2(k - 1))/\sqrt{2m}$. Furthermore,

$$\frac{1}{m} \int_0^1 \exp \left\{ -\frac{(t + 2(k - 1))^2}{4m} \right\} dt \leq \frac{1}{m},$$

which goes to zero faster than (19). Since

$$\int_1^{T+1} f(t) dt \leq \sum_{i=1}^T f(i) \leq \int_0^T f(t) dt \leq \int_1^{T+1} f(t) dt + \int_0^1 f(t) dt,$$

for any positive and decreasing function $f(\cdot)$, it follows that

$$\frac{1}{m} \sum_{i=1}^T \exp \left\{ -\frac{(t + 2(k - 1))^2}{4m} \right\} \sim 2\sqrt{\frac{\pi}{m}} \left(1 - \Phi \left(\frac{2k - 1}{\sqrt{2m}} \right) \right).$$

It remains for us to show that the summation (2) over the second range is negligible. Since $\exp\{x\} \geq 1 + x$ and $\lfloor x \rfloor > x - 1$ for all x , we obtain the following upper bound:

$$\begin{aligned}
 \frac{1}{m} \sum_{i=T+1}^{2(m-k+1)} \prod_{i=1}^{i+2(k-1)} \left(1 - \frac{1}{m} \lfloor \frac{i}{2} \rfloor \right) &\leq \frac{1}{m} \sum_{i=T+1}^{2(m-k+1)} \prod_{i=1}^{T+1+2(k-1)} \left(1 - \frac{1}{m} \lfloor \frac{i}{2} \rfloor \right) \\
 &\leq \frac{1}{m} \sum_{i=T+1}^{2(m-k+1)} \exp \left\{ -\frac{1}{m} \sum_{i=1}^{T+1+2(k-1)} \left(\frac{i}{2} - 1 \right) \right\} \\
 &\leq \frac{2m - 2k - T + 1}{m} \exp \left\{ -\frac{1}{4m} (T + 2(k - 1) - 2)^2 \right\}.
 \end{aligned}$$

This expression is small compared to the other half of the sum. □

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.


References

- Alon N, Rudov K, Yariv L (2021) Dominance solvability in random games. [arXiv:2105.10743](https://arxiv.org/abs/2105.10743) (arXiv preprint)
- Amiet B, Collevocchio A, Hamza K (2021) When better is better than best. *Oper Res Lett* 49(2):260–264
- Amiet B, Collevocchio A, Scarsini M, Zhong Z (2021) Pure nash equilibria and best-response dynamics in random games. *Math Oper Res* 20:20
- Apt KR, Simon S (2015) A classification of weakly acyclic games. *Theory Decis* 78(4):501–524
- Arieli I, Babichenko Y (2016) Random extensive form games. *J Econ Theory* 166:517–535
- Arratia R, Goldstein L, Gordon L et al (1989) Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann Probab* 17(1):9–25
- Babichenko Y (2013) Best-reply dynamics in large binary-choice anonymous games. *Games Econ Behav* 81:130–144
- Berg J, Weigt M (1999) Entropy and typical properties of nash equilibria in two-player games. *EPL (Europhys Lett)* 48(2):129–135
- Berger N, Feldman M, Neiman O, Rosenthal M (2011) Dynamic inefficiency: Anarchy without stability. In: *International symposium on algorithmic game theory*. Springer, pp 57–68
- Blume LE et al (1993) The statistical mechanics of strategic interaction. *Games Econ Behav* 5(3):387–424
- Boucher V (2017) Selecting equilibria using best-response dynamics. *Econ Bull* 37(4):2728–2734
- Candogan O, Ozdaglar A, Parrilo PA (2013) Dynamics in near-potential games. *Games Econ Behav* 82:66–90
- Chauhan A, Lenzner P, Melnichenko A, Molitor L (2017) Selfish network creation with non-uniform edge cost. In: *International symposium on algorithmic game theory*. Springer, pp 160–172
- Christodoulou G, Mirrokni VS, Sidiropoulos A (2012) Convergence and approximation in potential games. *Theoret Comput Sci* 438:13–27
- Cohen JE (1998) Cooperation and self-interest: pareto-inefficiency of Nash equilibria in finite random games. *Proc Natl Acad Sci* 95(17):9724–9731
- Coucheny P, Durand S, Gaujal B, Touati C (2014) General revision protocols in best response algorithms for potential games. In: *2014 7th international conference on NETWORK Games, CONTROL and OPTimization (NetGCoop)*. IEEE, pp 239–246
- Daskalakis C, Dimakis AG, Mossel E et al (2011) Connectivity and equilibrium in random games. *Ann Appl Probab* 21(3):987–1016
- Dindoš M, Mezzetti C (2006) Better-reply dynamics and global convergence to Nash equilibrium in aggregative games. *Games Econ Behav* 54(2):261–292
- Dresher M (1970) Probability of a pure equilibrium point in n -person games. *J Combin Theory* 8(1):134–145
- Durand S, Garin F, Gaujal B (2019) Distributed best response dynamics with high playing rates in potential games. *Perform Eval* 129:40–59
- Durand S, Gaujal B (2016) Complexity and optimality of the best response algorithm in random potential games. In: *International symposium on algorithmic game theory*. Springer, pp 40–51
- Fabrikant A, Jaggard AD, Schapira M (2013) On the structure of weakly acyclic games. *Theory Comput Syst* 53(1):107–122
- Feldman M, Snappir Y, Tamir T (2017) The efficiency of best-response dynamics. In: *International symposium on algorithmic game theory*. Springer, pp 186–198
- Feldman M, Tamir T (2012) Convergence of best-response dynamics in games with conflicting congestion effects. In: *International workshop on internet and network economics*. Springer, pp 496–503
- Friedman JW, Mezzetti C (2001) Learning in games by random sampling. *J Econ Theory* 98(1):55–84
- Galla T, Farmer JD (2013) Complex dynamics in learning complicated games. *Proc Natl Acad Sci* 110(4):1232–1236
- Goemans M, Mirrokni V, Vetta A (2005) Sink equilibria and convergence. In: *46th annual IEEE symposium on foundations of computer science (FOCS'05)*. IEEE, pp 142–151
- Goldberg K, Goldman A, Newman M (1968) The probability of an equilibrium point. *J Res Natl Bureau Standards* 72(2):93–101
- Goldman A (1957) The probability of a saddlepoint. *Am Math Mon* 64(10):729–730

- Kash IA, Friedman EJ, Halpern JY (2011) Multiagent learning in large anonymous games. *J Artif Intell Res* 40:571–598
- Kulti K, Salonen H, Vartiainen H (2011) Distribution of pure Nash equilibria in n -person games with random best responses. Technical Report 71, Aboa Centre for Economics. Discussion Papers
- McLennan A (2005) The expected number of Nash equilibria of a normal form game. *Econometrica* 73(1):141–174
- McLennan A, Berg J (2005) Asymptotic expected number of Nash equilibria of two-player normal form games. *Games Econ Behav* 51(2):264–295
- Mirroknii VS, Skopalik A (2009) On the complexity of Nash dynamics and sink equilibria. In: Proceedings of the 10th ACM conference on Electronic commerce, pp 1–10
- Monderer D, Shapley LS (1996) Potential games. *Games Econ Behav* 14(1):124–143
- Nisan N, Schapira M, Valiant G, Zohar A (2011) Best-response auctions. In: Proceedings of the 12th ACM conference on Electronic Commerce, pp 351–360
- Pangallo M, Heinrich T, Farmer JD (2019) Best reply structure and equilibrium convergence in generic games. *Sci Adv* 5(2):eaat1328
- Pei T, Takahashi S (2019) Rationalizable strategies in random games. *Games Econ Behav* 118:110–125
- Powers IY (1990) Limiting distributions of the number of pure strategy Nash equilibria in n -person games. *Int J Game Theory* 19(3):277–286
- Quattropiani M, Scarsini M (2020) Efficiency of equilibria in games with random payoffs. [arXiv:2007.08518](https://arxiv.org/abs/2007.08518) (arXiv preprint)
- Quint T, Shubik M, Yan D (1997) Dumb bugs vs. bright noncooperative players: a comparison. In: Albers W, Güth W, Hammerstein P, Moldvanu B, van Damme E (eds) *Understanding strategic interaction*. Springer, Berlin, pp 185–197
- Rinott Y, Scarsini M (2000) On the number of pure strategy Nash equilibria in random games. *Games Econ Behav* 33(2):274–293
- Sanders JB, Farmer JD, Galla T (2018) The prevalence of chaotic dynamics in games with many players. *Sci Rep* 8(1):4902
- Sandholm WH (2010) *Population games and evolutionary dynamics*. MIT Press, New York
- Stanford W (1995) A note on the probability of k pure Nash equilibria in matrix games. *Games Econ Behav* 9(2):238–246
- Stanford W (1996) The limit distribution of pure strategy Nash equilibria in symmetric bimatrix games. *Math Oper Res* 21(3):726–733
- Stanford W (1997) On the distribution of pure strategy equilibria in finite games with vector payoffs. *Math Soc Sci* 33(2):115–127
- Stanford W (1999) On the number of pure strategy Nash equilibria in finite common payoffs games. *Econ Lett* 62(1):29–34
- Swenson B, Murray R, Kar S (2018) On best-response dynamics in potential games. *SIAM J Control Optim* 56(4):2734–2767
- Takahashi S (2008) The number of pure Nash equilibria in a random game with nondecreasing best responses. *Games Econ Behav* 63(1):328–340
- Takahashi S, Yamamori T (2002) The pure Nash equilibrium property and the quasi-acyclic condition. *Econ Bull* 3(22):1–6
- Wiese SC, Heinrich T (2022) The frequency of convergent games under best-response dynamics. *Dyn Games Appl* 12(2):689–700
- Young HP (1998) *Individual strategy and social structure*. Princeton University Press, Princeton

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Torsten Heinrich^{1,2,3} · **Yoojin Jang**^{2,4} · **Luca Mungo**^{2,5} · **Marco Pangallo**⁶  · **Alex Scott**⁵ · **Bassel Tarbush**⁷ · **Samuel Wiese**^{2,4}

Torsten Heinrich
torsten.heinrich@wirtschaft.tu-chemnitz.de

Yoojin Jang
yjluca98@gmail.com

Luca Mungo
luca.mungo@maths.ox.ac.uk

Alex Scott
scott@maths.ox.ac.uk

Samuel Wiese
samuel.wiese@cs.ox.ac.uk

- ¹ Faculty for Economics and Business Administration, Chemnitz University of Technology, Chemnitz, Germany
- ² Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, Oxford, UK
- ³ Oxford Martin Programme on Technological and Economic Change (OMPTEC), Oxford Martin School, University of Oxford, Oxford, UK
- ⁴ Department of Computer Science, University of Oxford, Oxford, UK
- ⁵ Mathematical Institute, University of Oxford, Oxford, UK
- ⁶ CENTAI Institute, Torino, Italy
- ⁷ Department of Economics, University of Oxford, Oxford, UK