Submitted to the Annals of Applied Probability

CONNECTIVITY AND EQUILIBRIUM IN RANDOM GAMES

BY CONSTANTINOS DASKALAKIS*,
ALEXANDROS G. DIMAKIS† AND ELCHANAN MOSSEL‡

We study *how the structure of the interaction graph* of a game affects the existence of pure Nash equilibria. In particular, for a fixed interaction graph, we are interested in whether there are pure Nash equilibria arising when random utility tables are assigned to the players. We provide conditions for the structure of the graph under which equilibria are likely to exist and complementary conditions which make the existence of equilibria highly unlikely. Our results have immediate implications for many deterministic graphs and generalize known results for random games on the complete graph. In particular, our results imply that the probability that bounded degree graphs have pure Nash equilibria is exponentially small in the size of the graph and yield a simple algorithm that finds small non-existence certificates for a large family of graphs. Then we show that in any strongly connected graph of n vertices with expansion $(1+\Omega(1))\log_2(n)$ the distribution of the number of equilibria approaches the Poisson distribution with parameter 1, asymptotically as $n \to +\infty$.

In order to obtain a refined characterization of the degree of connectivity associated with the existence of equilibria, we also study the model in the random graph setting. In particular, we look at the case where the interaction graph is drawn from the Erdős-Rényi, $G(n,p)$, model where each edge is present independently with probability p. For this model we establish a *double phase transition* for the existence of pure Nash equilibria as a function of the average degree pn, consistent with the non-monotone behavior of the model. We show that when the average degree satisfies $np > (2+\Omega(1))\log_e(n)$, the number of pure Nash equilibria follows a Poisson distribution with parameter 1, asymptotically as $n \to \infty$. When $1/n << np < (0.5-\Omega(1))\log_e(n)$, pure Nash equilibria fail to exist with high probability. Finally, when $np = O(1/n)$ a pure Nash equilibrium exists with constant probability.

*EECS, MIT. Supported by a Sloan Fellowship in Computer Science, and NSF award CCF-0953960 (CAREER). Part of this work was done while the author was at EECS, UC Berkeley, with the support of a Microsoft Research Fellowship, and NSF awards CCF-0635319, DMS 0528488 and DMS 0548249 (CAREER).

†EECS, USC. Part of this work was done while the author was at EECS, UC Berkeley. Supported by NSF award DMS 0528488 and a Microsoft Research Fellowship.

‡UC Berkeley and the Weizmann Institute of Science. Supported by a Sloan fellowship in Mathematics, NSF awards DMS 0528488 and DMS 0548249 (CAREER), and ONR grant N0014-07-1-05-06.

AMS 2000 subject classifications: Primary 91A; secondary 60, 68Q

Keywords and phrases: Game Theory, Graphical Games, Connectivity, Phase Transitions, Random Constraint Satisfaction Problems

1. Introduction. In recent years there has been a convergence of ideas from computer science and the social sciences aiming to model and analyze large complex networks such as the web graph, social networks and recommendation systems. From the computational perspective, it has been recognized that the successful design of algorithms performed on such networks, including routing, ranking and recommendation algorithms, must take into account the social dynamics and economic incentives as well as the technical properties that govern these networks [20, 24, 27].

Game theory has been very successful in modeling strategic behavior in large systems of economically incentivized entities. In the context of routing, for instance, it has been employed to study the effect of selfishness on the efficiency of a network, whereby the performance of the network at equilibrium is compared to the performance when a central authority can simply dictate a solution [7, 30–32]. The effect of selfishness has been studied in several other settings, e.g. load balancing [8, 9, 21, 29], facility location [34], and network design [3].

A simple way to model interactions between agents in a large network is with a *graphical game* [19]: a graph $G = (V, E)$ is defined whose vertices represent the players of the game and an edge $(v, w) \in E$ corresponds to the strategic interaction between players v and w; each player $v \in V$ has a finite set of strategies S_v, which throughout this paper will be assumed to be binary so that there are two possible strategies for each player. A *utility, or payoff, table* u_v for player v assigns a real number $u_v(\sigma_v, \sigma_{\mathcal{N}(v)})$ to every selection of strategies by player v and the players in v's neighborhood, i.e. the set of nodes v' such that $(v, v') \in E$, denoted by $\mathcal{N}(v)$. A *pure Nash equilibrium* (PNE) of the game is some state, or *strategy profile*, σ of the game, assigning to every player v a single strategy $\sigma_v \in S_v$, such that no player has a unilateral incentive to deviate. Equivalently, for every player $v \in V$,

$$u_v(\sigma_v, \sigma_{\mathcal{N}(v)}) \geq u_v(\sigma'_v, \sigma_{\mathcal{N}(v)}), \text{ for every strategy } \sigma'_v \in S_v. \tag{1}$$

When condition (1) is satisfied, we say that the strategy σ_v is a *best response to the strategies* $\sigma_{\mathcal{N}(v)}$.

The concept of the pure strategy Nash equilibrium is more compelling, decision theoretically, than the concept of the mixed strategy Nash equilibrium—its counterpart that allows players to choose distributions over their strategy sets. This is because it is not always meaningful in applications to assume that the players of a game may adopt randomized strategies. Unfortunately, unlike mixed Nash equilibria, PNE do not always exist. It is then an important problem to study how the existence of PNE depends on the properties of the game.

The focus of this paper is to understand how the connectivity of the underlying graph affects the existence of a PNE. We obtain two kinds of results. The first concerns the existence of a PNE in an ensemble of random graphical games

defined on a random—$G(n,p)$—graph. We obtain a characterization of the probability that a PNE exists as a function of the density of the graph. The second set of results concerns random graphical games on deterministic graphs. Here, we obtain conditions on the structure of the graph under which a PNE does not exist with high probability, suggesting also an efficient algorithm for finding witnesses of the non-existence of a PNE. We also give complementary conditions on the structure of the graph under which a PNE exists with constant probability. Our results are described in detail in Section 1.3.

Comparison to Typical Constraint Satisfaction Problems. Graphical games provide a more compact way for representing large networks of interacting agents, than normal form games, in which the game is described as if it were played on the complete graph. Besides the compact description, one of the motivations for the introduction of graphical games is their intuitive affinity to graphical statistical models; indeed, several algorithms for graphical games do have the flavor of algorithms for solving Bayes nets or constraint satisfaction problems [10, 13, 16, 22, 23].

In the other direction, the notion of a PNE provides a *new genre of constraint satisfaction problems*; notably one in which, for any assignment of strategies (values) to the neighborhood of a player (variable), there is always a strategy (value) for that player which makes the constraint (1) corresponding to that player satisfied (i.e. being in best response). The reason why it might be hard to satisfy simultaneously the constraints corresponding to all players is the long range correlations that may arise between players. Indeed, deciding whether a PNE exists is NP-hard even for very sparse graphical games [16].

Viewed as a constraint satisfaction problem, the problem of the existence of PNE poses interesting challenges. First, for natural random ensembles over payoff tables such as the one adopted in this paper (see Definition 1.2), the *expected number* of PNE is 1 for *any graph* (this is shown for our model in the main body of the paper; see Eq. (10)). On the contrary, for typical constraint satisfaction problems, the expected number of solutions is exponential in the size of the graph with different exponents corresponding to different density parameters. Second, unlike typical constraint satisfaction problems studied before, the existence of PNE is a-priori not a monotone property of the connectivity. It is surprising that given these novel features of the problem it is possible to obtain a result establishing a double phase transition on the existence of PNE as described below.

1.1. *Our Model.* We define the notion of a graphical game and proceed to describe the ensemble of random graphical games studied in this paper.

DEFINITION 1.1 (Graphical Game). *Given a graph $G = (V, E)$, we define the* neighborhood of node $v \in V$ *to be the set $\mathcal{N}(v) = \{v' \mid (v, v') \in E\}$. If S_v is*

a set associated with vertex v, for all $v \in V$, we denote by $S_{\mathcal{N}(v)} := \times_{v' \in \mathcal{N}(v)} S_{v'}$ the Cartesian product of the sets associated with the nodes in v's neighborhood.

A graphical game *on G is a collection $(S_v, u_v)_{v \in V}$, where S_v is the* strategy set of node v *and $u_v : S_v \times S_{\mathcal{N}(v)} \to \mathbb{R}$ the* utility (or payoff) function (or table) of player v. *We also define* the best response function (or table) of player v *to be the function* $\mathrm{BR}_v : S_v \times S_{\mathcal{N}(v)} \to \{0,1\}$ *such that*

$$\mathrm{BR}_v(\sigma_v, \sigma_{\mathcal{N}(v)}) = 1 \Leftrightarrow \sigma_v \in \arg_x \max\{u_v(x, \sigma_{\mathcal{N}(v)})\},$$

for all $\sigma_v \in S_v$ and $\sigma_{\mathcal{N}(v)} \in S_{\mathcal{N}(v)}$.

DEFINITION 1.2 (Random Graphical Games on a Fixed Graph). *Given a graph $G = (V, E)$ and an atomless distribution $\mathcal{F}$ over $\mathbb{R}$, the probability distribution $\mathcal{D}_{G,\mathcal{F}}$ over graphical games $(S_v, u_v)_{v \in V}$ on G is defined as follows:*

- *$S_v = \{0,1\}$, for all $v \in V$;*
- *the payoff values $\{u_v(\sigma_v, \sigma_{\mathcal{N}(v)})\}_{v \in V, \sigma_v \in S_v, \sigma_{\mathcal{N}(v)} \in S_{\mathcal{N}(v)}}$ are mutually independent and identically distributed according to $\mathcal{F}$.*

REMARK 1.3 (Invariance under Payoff Distributions). *It is easy to see that the existence of a PNE is only determined by the* best response tables *of the game; see Condition (1). In particular, given that the distributions considered in this paper are atomless, we can study PNE under $\mathcal{D}_{G,\mathcal{F}}$, for any atomless $\mathcal{F}$, by restricting our attention (up to probability 0 events) to the measure $\mathcal{D}_G$ over best response tables, defined as follows*

- *$\{\mathrm{BR}_v(0, \sigma_{\mathcal{N}(v)})\}_{v \in V, \sigma_{\mathcal{N}(v)} \in S_{\mathcal{N}(v)}}$ are mutually independent and uniform in $\{0,1\}$;*
- *$\mathrm{BR}_v(1, \sigma_{\mathcal{N}(v)}) = 1 - \mathrm{BR}_v(0, \sigma_{\mathcal{N}(v)})$, for all $\sigma_{\mathcal{N}(v)} \in S_{\mathcal{N}(v)}$.*

We will sometimes refer to a graphical game defined in terms of its best response tables as an underspecified graphical game. *Other times, we will overload our terminology and just call it a graphical game. We use $\mathbb{P}_G[.]$ and $\mathbb{E}_G[.]$ to denote probabilities of events and expectations respectively under the measure $\mathcal{D}_G$.*

REMARK 1.4 (Invariance under Payoff Distributions II). *Given our observation in Remark 1.3, it follows that it is not important to use a common distribution $\mathcal{F}$ for sampling the payoffs of all the players of the game. All our results in this work are true if different players have different distributions as long as these distributions are atomless and all payoffs values are sampled independently.*

Extending the Model to Random Graphs. One of the goals of this paper is to investigate what average degree is required in a graph for a graphical game played

on this graph to have a PNE. To study this question, it is natural to consider families of graphs with different densities and relate the probability of PNE existence with the density of the graph. We consider graphical games on graphs drawn from the Erdős-Rényi, $G(n,p)$, model, with varying values of the edge probability p. The ensemble of graphical games we consider is formally the following.

DEFINITION 1.5. *Given $n \in \mathbb{N}$, $p \in [0,1]$ and an atomless distribution $\mathcal{F}$ over $\mathbb{R}$, we define a measure $\mathcal{D}_{(n,p,\mathcal{F})}$ over graphical games. A graphical game is drawn from $\mathcal{D}_{(n,p,\mathcal{F})}$ as follows:*

- *a graph G is drawn from $G(n,p)$;*
- *a random graphical game is drawn from $\mathcal{D}_{G,\mathcal{F}}$.*

REMARK 1.6 (Invariance under Payoff Distributions III). *Given our discussion in Remark 1.3, it follows that in order to study PNE in the random ensemble of Definition 1.5 it is sufficient to consider a measure that fixes only the best response tables of the players in the sampled games.*

For a given $n \in \mathbb{N}$ and $p \in [0,1]$, we define the measure $\mathcal{D}_{(n,p)}$ over underspecified graphical games. An underspecified graphical game is drawn from $\mathcal{D}_{(n,p)}$ as follows:

- *a graph G is drawn from $G(n,p)$;*
- *a random underspecified graphical game is drawn from $\mathcal{D}_G$.*

We use $\mathbb{P}_{(n,p)}[.]$ to denote probabilities of events under the measure $\mathcal{D}_{(n,p)}$ and $\mathbb{P}_{\mathcal{G}}[.]$ for probabilities of events measurable under $G(n,p)$.

In the model defined in Definition 1.5 and Remark 1.6, there are two sources of randomness: the selection of the graph, determining what players interact with each other, and the selection of the payoff tables. Note that in the two-stage process that samples a graphical game from our distribution, the payoff tables can only be realized once the graph is fixed. This justifies the subscript G in the measure $\mathcal{D}_G$ defined above.

1.2. *Discussion.*

Non-Monotonicity. Observe that the existence of a PNE is a *non-monotone property* of p: any graphical game on the empty graph has a PNE for trivial reasons; on the complete graph a random graphical game has a PNE with asymptotic probability $1-\frac{1}{e}$ (see [12, 28]); but our results indicate that, when p is in some intermediate regime, a PNE does not exist with probability approaching 1 as $n \to +\infty$ (see Theorem 1.10).

The non-monotonicity in the average degree of the existence of a PNE makes the relation between PNE and connectivity non-obvious. Surprisingly, we show (Theorem 1.9) that the convergence to a Poisson distribution of the distribution of the number of PNE in complete graphs [26, 33] extends to much sparser graphs, as long as the average degree is at least *logarithmic* in the number of players. If the sparsity increases further, we show (Theorem 1.10) that a PNE does not exist with high probability, while if the graph is essentially empty, PNE exist with probability 1 (Theorem 1.11). Our results establish a *double phase transition* consistent with the non-monotonicity of the model.

Methodological Challenges. Our study here is an instance of studying the satisfiability of constraint satisfaction problems (CSPs). The generic question is to investigate the effect of the structure of the constraint graph on the satisfiability of the problems defined on that graph, as well as their computational complexity. In the context of CNF formulas (corresponding to the SATisfiability problem) the graph property most commonly studied in the literature is the density of the hypergraph that contains an edge for each clause of the formula, see e.g. [14]. In other settings, different structural properties of the constraint graph are relevant, e.g. measures of cyclicity of the graph [6, 17]. In our case, studying the average degree reveals an interesting, non-monotonic behavior of the model, as described above.

In a typical CSP, to show that a solution does not exist one either uses the first moment method to exhibit that the expected number of solutions is tiny [2], or finds a witness of unsatisfiability that exists with high probability. To show that a satisfying assignment *does* exist it is quite common to use the second moment method or its refinements, which have provided some of the best bounds for satisfiability to date [1]. In our model the expected number of satisfying assignments turns out to be 1 for any graph (see Eq. (10) below). This suggests that the analysis of the problem should be harder, since in particular we cannot use the first moment method to establish the non-existence of a PNE. Our proof of the non-existence of PNE (Theorems 1.10 and 1.16) uses succinct non-existence witnesses that appear with high probability in sufficiently sparse graphs. These witnesses are specific subgame structures that do not possess a PNE with high probability. To establish the existence of a PNE for sufficiently large densities (Theorems 1.9 and 1.13) we use the second moment method. Further, we use Stein's [4] method to establish that the distribution of the number of PNE converges asymptotically to a $\mathrm{Poisson}(1)$ distribution in this case.

1.3. *Outline of Main Results.* We describe first our results for random graphs (for the measure $\mathcal{D}_{(n,p)}$ defined in Remark 1.6), and proceed with our results for deterministic graphs (for the measure $\mathcal{D}_G$ defined in Remark 1.3).

PNE on Random Graphs. We study how the connectivity probability p influences the existence of PNE for games sampled from $\mathcal{D}_{(n,p)}$. The transition is described by the following theorems applying to different levels of graph connectivity. Before stating the theorems, we introduce some notation.

REMARK 1.7 (Order Notation). *Let $f(x)$ and $g(x)$ be two functions defined on some subset of the real numbers. One writes $f(x) = O(g(x))$ if and only if, for sufficiently large values of x, $f(x)$ is at most a constant times $g(x)$ in absolute value. That is, $f(x) = O(g(x))$ if and only if there exists a positive real number M and a real number x_0 such that*

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

Similarly, we write $f(x) = \Omega(g(x))$ if and only if there exists a positive real number M and a real number x_0 such that

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

We casually use the order notation *$O(\cdot)$ and $\Omega(\cdot)$ throughout the paper. Whenever we use $O(f(n))$ or $\Omega(f(n))$ in some bound, there exists a constant $c > 0$ such that the bound holds true for sufficiently large n if we replace the $O(f(n))$ or $\Omega(f(n))$ in the bound by $c \cdot f(n)$.*

REMARK 1.8 (Order Notation Continued). *If $g(n)$ is a function of $n \in \mathbb{N}$, then we denote by $\omega(g(n))$ any function $f(n)$ such that $f(n)/g(n) \to +\infty$, as $n \to +\infty$; similarly, we denote by $o(g(n))$ any function $f(n)$ such that $f(n)/g(n) \to 0$, as $n \to +\infty$. Finally, for two functions $f(n)$ and $g(n)$, we write $f(n) >> g(n)$ whenever $f(n) = \omega(g(n))$.*

THEOREM 1.9 (High Connectivity). *Let Z denote the number of PNE in a graphical game sampled from $\mathcal{D}_{(n,p)}$, where $p = \frac{(2+\epsilon)\log_e(n)}{n}$, $\epsilon = \epsilon(n) > 0$. For an arbitrary constant $c > 0$ we assume that $\epsilon(n) > c$ and (in order for $p \leq 1$) $\epsilon(n) \leq \frac{n}{\log_e(n)} - 2$.*

Under the above assumptions, for all finite n, with probability at least $1 - 2n^{-\epsilon/8}$ over the random graph sampled from $G(n,p)$, it holds that the total variation distance between Z and a $\mathrm{Poisson}(1)$ r.v. W is bounded by:

$$||Z - W|| \leq O(n^{-\epsilon/8}) + \exp(-\Omega(n)). \tag{2}$$

In other words,

$$\mathbb{P}_{\mathcal{G}}\left[||Z - W|| \leq O(n^{-\epsilon/8}) + \exp(-\Omega(n))\right] \geq 1 - 2n^{-\epsilon/8}. \tag{3}$$

In particular, the distribution of Z converges in total variation distance to a $\mathrm{Poisson}(1)$ *distribution, as* $n \to +\infty$.

(Note that the two terms on the right hand side of (2) *can be of the same order when* ϵ *is of the order of* $n/\log_e(n)$*.)*

THEOREM 1.10 (Medium Connectivity). *For all* $p = p(n) \leq 1/n$*, if a graphical game is sampled from* $\mathcal{D}_{(n,p)}$*, the probability that a PNE exists is bounded by:*

$$\exp(-\Omega(n^2 p)).$$

For $p(n) = g(n)/n$*, where* $\log_e(n)/2 > g(n) > 1$*, the probability that a PNE exists is bounded by:*

$$\exp(-\Omega(e^{\log_e(n)-2g(n)})).$$

In particular, the probability that a PNE exists goes to 0 *as* $n \to +\infty$ *for all* $p = p(n)$ *satisfying*

$$\frac{1}{n^2} << p < (0.5 - \epsilon'(n))\frac{\log_e(n)}{n},$$

where $\epsilon'(n) = \omega\left(\frac{1}{\log_e(n)}\right)$.

THEOREM 1.11 (Low Connectivity). *For every constant* $c > 0$*, if a graphical game is sampled from* $\mathcal{D}_{(n,p)}$ *with* $p \leq \frac{c}{n^2}$*, the probability that a PNE exists is at least*

$$\left(1 - \frac{c}{n^2}\right)^{\frac{n(n-1)}{2}} \longrightarrow e^{-\frac{c}{2}}.$$

Note that our upper and lower bounds for $G(n,p)$ leave a small gap, between $p \approx \frac{0.5\log_e(n)}{n}$ and $p \approx \frac{2\log_e(n)}{n}$. The behavior of the number of PNE in this range of p remains open. We establish the non-existence of PNE for medium connectivity graphs via a simple structure that prevents PNE from arising, called the 'indifferent matching pennies game' (see Definition 1.18 below). It is natural to ask whether our 'indifferent matching pennies' witnesses are (similarly to isolated vertices in connectivity) the smallest structures that prevent the existence of PNE and the last ones to disappear.

General Graphs. We give conditions on the structure of a graph implying the (likely) existence or non-existence of a PNE in a random game played on that graph. The existence of a PNE is guaranteed by sufficient connectivity of the underlying graph. The connectivity that we require is captured by the notion of (α, δ)-*expansion* given next.

DEFINITION 1.12 ((α,δ)-Expansion). *A graph $G=(V,E)$ has (α,δ)-*expansion *iff every set V' such that $|V'| \leq \lceil \delta |V| \rceil$ has $|\mathcal{N}(V')| \geq \min(|V|, \alpha|V'|)$ neighbors. Here we let*

$$\mathcal{N}(V') = \{w \in V : \exists u \in V' \text{ with } (u,w) \in E\}.$$

(Note in particular that $\mathcal{N}(V')$ may intersect V').

We show the following result.

THEOREM 1.13 (Strongly Connected Graphs). *Let Z denote the number of PNE in a graphical game sampled from $\mathcal{D}_G$, where G is a graph on n vertices that has (α,δ)-expansion with $\alpha = (1+\epsilon)\log_2(n)$, $\delta = \frac{1}{\alpha}$ and $\epsilon > 0$. Then the total variation distance between the distribution of Z and the distribution of a* $\mathrm{Poisson}(1)$ *r.v. W is bounded by:*

$$||Z - W|| \leq O(n^{-\epsilon}) + O(2^{-n/2}). \tag{4}$$

Next we provide a complementary condition for the non-existence of PNE. The condition will be given in terms of the following structure.

DEFINITION 1.14 (d-Bounded Edge). *An edge $e=(u,v) \in E$ of a graph $G(V,E)$ is called* d-bounded *if both u and v have degrees smaller or equal to d.*

We bound the probability that a PNE exists in a game sampled from $\mathcal{D}_G$ as a function of the number of d-bounded edges in G. For the stronger version of our theorem, we also need the notion of a *maximal weighted independent edge-set* defined next.

DEFINITION 1.15 (Maximal Weighted Independent Edge-Set). *Given a graph $G(V,E)$, a subset $\mathcal{E} \subseteq E$ of the edges is called* independent *if no pair of edges in $\mathcal{E}$ are adjacent.*

If $w : E \to \mathbb{R}$ is a function assigning weights to the edges of G, we extend w to subsets of edges by assigning to each $\mathcal{E} \subseteq E$ the weight $w_{\mathcal{E}} = \sum_{e\in\mathcal{E}} w(e)$. Then we call a subset $\mathcal{E} \subseteq E$ of edges a maximal weighted independent edge-set *if $\mathcal{E}$ is an independent edge-set with maximal weight among independent edge-sets.*

THEOREM 1.16. *A random game sampled from $\mathcal{D}_G$, where G is a graph with at least m vertex disjoint d-bounded edges, has no PNE with probability at least*

$$1 - \exp\left(-m\left(\frac{1}{8}\right)^{2^{2d-2}}\right). \tag{5}$$

In particular, if G has at least m edges that are d-bounded, then a game sampled from $\mathcal{D}_G$ has no PNE with probability at least

$$1 - \exp\left(-\frac{m}{2d}\left(\frac{1}{8}\right)^{2^{2d-2}}\right). \tag{6}$$

Moreover, there exists an algorithm of complexity $O(n^2+m2^{d+2})$ for proving that a PNE does not exist, which has success probability given by (5) and (6) respectively.

More generally, let us assign to every edge $(u,v) \in E$ the weight

$$w_{(u,v)} := -\log_e\left(1 - p_{(u,v)}\right),$$

for $p_{(u,v)} = 8^{-2^{d_u+d_v-2}}$, where d_u and d_v are respectively the degrees of u and v. Given these weights, suppose that $\mathcal{E}$ is a maximal weighted independent edge-set with value $w_{\mathcal{E}}$. Then the probability that there exists no PNE is at least

$$1 - \exp\left(-w_{\mathcal{E}}\right).$$

An easy consequence of this result is that many sparse graphs, such as the line and the grid, do not have a PNE with probability tending to 1 as the number of players increases.

The proof of Theorem 1.16 is based on a small witness for the non-existence of PNE, called the *indifferent matching pennies game*. As the name implies this game is inspired by the simple *matching pennies game*. Both games are described next.

DEFINITION 1.17 (The Matching Pennies Game). *We say that two players a and b play the* matching pennies game *if their payoff matrices are the following, up to permuting the players' names.*

Payoff table of player a :

	b plays 0	b plays 1
a plays 0	1	0
a plays 1	0	1

Payoff table of player b :

	b plays 0	b plays 1
a plays 0	0	1
a plays 1	1	0

DEFINITION 1.18 (The Indifferent Matching Pennies Game). *We say that two players a and b that are adjacent to each other in a graphical game play the* indifferent matching pennies game *if, for all strategy profiles $\sigma_{\mathcal{N}(a)\cup\mathcal{N}(b)\setminus\{a,b\}}$ in the*

neighborhood of a and b, the players a and b play a matching pennies game against each other.

In other words, for all fixed $\sigma := \sigma_{\mathcal{N}(a)\cup\mathcal{N}(b)\setminus\{a,b\}}$, the payoff tables of a and b projected on $\sigma_{\mathcal{N}(a)\setminus\{b\}}$ and $\sigma_{\mathcal{N}(b)\setminus\{a\}}$ respectively are the following, up to permuting the players' names.

Payoffs to player a :

	b *plays* 0, *other neighbors play* $\sigma_{\mathcal{N}(a)\setminus\{b\}}$	b *plays* 1, *other neighbors play* $\sigma_{\mathcal{N}(a)\setminus\{b\}}$
a plays 0	1	0
a plays 1	0	1

Payoffs to player b :

	a *plays* 0, *other neighbors play* $\sigma_{\mathcal{N}(b)\setminus\{a\}}$	a *plays* 1, *other neighbors play* $\sigma_{\mathcal{N}(b)\setminus\{a\}}$
b plays 0	0	1
b plays 1	1	0

Observe that if a graphical game contains an edge (u, v) so that players u and v play the indifferent matching pennies game then the game has no PNE. In particular, the indifferent matching pennies game provides a *small witness* for the non-existence of a PNE, which is a coNP-complete problem for bounded degree graphical games [16]. Our analysis implies that, with high probability over bounded degree graphical games, there are short proofs for the non-existence of PNE which can be found efficiently. A related analysis and randomized algorithm was introduced for mixed Nash equilibria in 2-player games by Bárány et al. [5].

1.4. *Related Work.* The number of PNE in random games with i.i.d. payoffs has been extensively studied in the literature prior to our work: Goldberg et al. [15] characterize the probability that a two-player random game with i.i.d. payoff tables has a PNE, as the number of strategies tends to infinity. Dresher [12] and Papavassilopoulos [25] generalize this result to n-player random games on the complete graph. Powers [26] and Stanford [33] generalize the result further, showing that the distribution of the number of PNE approaches a $\mathrm{Poisson}(1)$ distribution as the number of strategies increases. Finally, Rinott et al. [28] investigate the asymptotic distribution of PNE for a more general ensemble of random games on the complete graph where there are positive or negative dependencies among the players' payoffs.

Our work generalizes the above results for i.i.d. payoffs beyond the complete graph to random graphical games on random graphs and several families of deterministic graphs. Parallel to our work, Bistra et al. [11] studied the existence of PNE in certain families of deterministic graphs, and Hart et al. [18] obtained results for evolutionarily stable strategies in random games. These results are related but not directly comparable to our results.

1.5. *Acknowledgement.* We thank Martin Dyer for pointing out an error in a previous formulation of Theorem 1.16. We also thank the anonymous referee for comments that helped improve the presentation of this work.

2. Random Graphs.

2.1. *High Connectivity.* In this section we study the number of PNE in graphical games sampled from $\mathcal{D}_{(n,p)}$. We show that, when the average degree is $pn = (2 + \epsilon(n)) \log_e(n)$, where $\epsilon(n) > c$ and $c > 0$ is any fixed constant, the distribution of the number of PNE converges to a $\mathrm{Poisson}(1)$ random variable, as n goes to infinity. This implies in particular that a PNE exists with probability converging to $1 - \frac{1}{e}$ as the size of the network increases.

As in the study of the complete graph in [28], we use the following result of Arratia et al. [4], established using Stein's method. For two random variables Z, Z' supported on $0, 1, \ldots$ we define their *total variation distance* $||Z - Z'||$ as

$$||Z - Z'|| := \sum_{i=0}^{\infty} |Z(i) - Z'(i)|.$$

LEMMA 2.1 ([4]). *Consider arbitrary Bernoulli random variables X_i, $i = 0, \ldots, N$. For each i, define some* neighborhood of dependence *B_i of X_i such that B_i satisfies that $(X_j : j \in B_i^c)$ are* independent *of X_i. Let*

$$Z = \sum_{i=0}^{N} X_i, \quad \lambda = \mathbb{E}[Z], \tag{7}$$

and assume that $\lambda > 0$. Also, let

$$b_1 = \sum_{i=0}^{N} \sum_{j \in B_i} \mathbb{P}[X_i = 1]\mathbb{P}[X_j = 1]$$

$$\text{and } b_2 = \sum_{i=0}^{N} \sum_{j \in B_i \setminus \{i\}} \mathbb{P}[X_i = 1, X_j = 1].$$

Then the total variation distance between the distribution of Z and a Poisson *random variable W_λ with mean λ is bounded by*

$$||Z - W_\lambda|| \leq 2(b_1 + b_2). \tag{8}$$

Proof of Theorem 1.9: For ease of notation, we identify the players of the graphical game with the indices $1, 2, \ldots, n$. We also identify pure strategy profiles with the integers in $\{0, \ldots, 2^n - 1\}$, mapping each integer to a strategy profile. The mapping is defined so that, if the binary expansion of i is $i(1) \ldots i(n)$, player k plays $i(k)$.

Next, to each strategy profile $i \in \{0, \ldots, N\}$, where $N = 2^n - 1$, we assign an indicator random variable X_i which is 1 if the strategy profile i is a PNE. Then the counting random variable

$$Z = \sum_{i=0}^{N} X_i \tag{9}$$

corresponds to the number of PNE. Hence the existence of a PNE is equivalent to the random variable Z being positive.

Let us condition on a realization of the graph G of the graphical game, but not its best response tables. For a given strategy profile i, each player is in best response with probability $1/2$ over the selection of her best response table; * therefore $\mathbb{E}_G[X_i] = 2^{-n}$, for all i, where recall that $\mathbb{E}_G$ denotes expectation under the measure $\mathcal{D}_G$. Hence, conditioning on G the expected number of PNE is

$$\mathbb{E}_G[Z] = 1. \tag{10}$$

Since this holds for any realization of the graph G it follows that $\mathbb{E}[Z] = 1$.

In Lemma 2.2 that follows, we characterize the neighborhood of dependence B_i of the variable X_i in order to be able to apply Lemma 2.1 on the collection of variables $X_0, \ldots, X_N$. Note that this neighborhood depends on the graph realization, but is independent of the realization of the payoff tables.

LEMMA 2.2. *For a fixed graph G, we can choose the neighborhoods of dependence for the random variables $X_0, \ldots, X_N$ as follows:*

$$B_0 = \{j : \exists k \text{ such that } \forall k' \text{ with } (k, k') \in E(G) \text{ it holds that } j(k') = 0\}$$

and

$$B_i = i \oplus B_0 = \{i \oplus j : j \in B_0\},$$

where $i \oplus j = (i(1) \oplus j(1), \ldots, i(n) \oplus j(n))$ and $\oplus$ is the exclusive or operation.

*This follows directly from our model (Remark 1.6), following our assumption of atomless payoff distributions (Definition 1.5).

REMARK 2.3. *Intuitively, when the graph G is realized, the neighborhood of dependence of the strategy profile 0 (variable X_0) contains all strategy profiles j (variables X_j) assigning 0 to all the neighbors of at least one player k. If such a player k exists, then whether 0 or $j(k)$ is a best response to the all-0 neighborhood are dependent random variables (over the selection of the best response table of player k). The definition of B_i in terms of B_0 is justified by the symmetry of our model.*

PROOF OF LEMMA 2.2. By symmetry, it is enough to show that X_0 is independent of $\{X_i\}_{i\notin B_0}$. Fix some $i \notin B_0$. Observe that in i, each player k of the game has at least one neighbor k' playing strategy 1. By the definition of measure $\mathcal{D}_G$, it follows that whether strategy 0 is a best response for player k in strategy profile 0 is independent of whether strategy $i(k)$ is a best response for player k in strategy profile i, since these events depend on different strategy profiles of the neighbors of k. ■

Now, for a fixed graph G, the functions $b_1(G)$ and $b_2(G)$ (corresponding to b_1 and b_2 in Lemma 2.1) are well-defined. We proceed to bound the expectation of these functions over the sampling of the graph G.

$$\begin{aligned}
\mathbb{E}_{\mathcal{G}}[b_1(G)] &= \mathbb{E}_{\mathcal{G}}\left[\sum_{i=0}^{N}\sum_{j\in B_i}\mathbb{P}_G[X_i=1]\mathbb{P}_G[X_j=1]\right] \\
&= \mathbb{E}_{\mathcal{G}}\left[\frac{1}{(N+1)^2}\sum_{i=0}^{N}|B_i|\right] \\
&= \frac{\mathbb{E}_{\mathcal{G}}[|B_0|]}{N+1}; \qquad (11)\\
\mathbb{E}_{\mathcal{G}}[b_2(G)] &= \mathbb{E}_{\mathcal{G}}\left[\sum_{i=0}^{N}\sum_{j\in B_i\setminus\{i\}}\mathbb{P}_G[X_i=1,X_j=1]\right] \\
&= (N+1)\sum_{j\neq 0}\mathbb{E}_{\mathcal{G}}\left[\mathbb{P}_G[X_0=1,X_j=1]\mathbb{I}[j\in B_0]\right]. \qquad (12)
\end{aligned}$$

In the last line of both derivations we made use of the symmetry of the model. Invoking symmetry again, we observe that the expectation

$$\mathbb{E}_{\mathcal{G}}\left[\mathbb{P}_G[X_0=1,X_j=1]\mathbb{I}[j\in B_0]\right]$$

in (12) depends only on the number of 1's in the strategy profile j, denoted s below. Let us write Y_s for the indicator that the strategy profile j_s, where the first s players play 1 and all the other players play 0, is a PNE. Also, write I_s for the indicator

that this strategy is in B_0 (note that I_s is a function of the graph only). Using this notation, we obtain:

$$\mathbb{E}_{\mathcal{G}}[b_2(G)] = 2^n \sum_{s=1}^{n} \binom{n}{s} \mathbb{E}_{\mathcal{G}}[I_s \mathbb{P}_G[Y_0 = 1, Y_s = 1]]; \tag{13}$$

$$\text{and} \quad \mathbb{E}_{\mathcal{G}}[b_1(G)] = 2^{-n} \sum_{s=0}^{n} \binom{n}{s} \mathbb{E}_{\mathcal{G}}[I_s]. \tag{14}$$

LEMMA 2.4. *$\mathbb{E}_{\mathcal{G}}[b_1(G)]$ and $\mathbb{E}_{\mathcal{G}}[b_2(G)]$ are bounded as follows.*

$$\mathbb{E}_{\mathcal{G}}[b_1(G)] \le R(n,p) := \sum_{s=0}^{n} \binom{n}{s} 2^{-n} \min(1, n(1-p)^{s-1});$$

$$\mathbb{E}_{\mathcal{G}}[b_2(G)] \le S(n,p) := \sum_{s=1}^{n} \binom{n}{s} 2^{-n} \left[(1+(1-p)^s)^{n-s} - (1-(1-p)^s)^{n-s}\right].$$

PROOF. We begin with the study of $\mathbb{E}_{\mathcal{G}}[b_1(G)]$. Clearly, it suffices to bound $\mathbb{E}[I_s]$ by $n(1-p)^{s-1}$, for $s > 0$. For the strategy profile j_s to belong in B_0 it must be that there is at least one player who is not connected to any player in the set $S := \{1, 2, \ldots, s\}$. The probability that a specific player k is not connected to any player in S is either $(1-p)^s$ or $(1-p)^{s-1}$, depending on whether $k \in S$; so it is always at most $(1-p)^{s-1}$. By a union bound it follows that the probability there is at least one player not connected to S is at most $n(1-p)^{s-1}$.

We now analyze $\mathbb{E}_{\mathcal{G}}[I_s \mathbb{P}_G[Y_0 = 1, Y_s = 1]]$. Recall from the previous paragraph that $I_s = 1$ only when there exists a player k who is not connected to any player in S. If there exists such a player k with the extra property that $k \in S$, then $\mathbb{P}_G[Y_0 = 1, Y_s = 1] = 0$, since it cannot be that both 0 and 1 are best responses for player k when all her neighbors play 0.

Therefore the only contribution to $\mathbb{E}_{\mathcal{G}}[I_s \mathbb{P}_G[Y_0 = 1, Y_s = 1]]$ is from the event every player in S is connected to at least one other player in S. Conditioning on this event, in order for $I_s = 1$ it must be that at least one of the players in $S^c := V \setminus S$ is not adjacent to any player in S.

Let us define $p_s := \mathbb{P}_{\mathcal{G}}[\nexists$ isolated node in the subgraph induced by $S]$ and let t denote the number of players in S^c, which are not connected to any player in S. Since every player outside S is non-adjacent to any player in S with probability $(1-p)^s$, the probability that exactly t players are not adjacent to S is

$$\binom{n-s}{t} [(1-p)^s]^t (1-(1-p)^s)^{n-s-t}.$$

Moreover, conditioning on the event that exactly t players in S^c are not adjacent to any player in S, we have that the probability that $Y_0 = 1$ and $Y_s = 1$ is:

$$\frac{1}{2^t}\frac{1}{2^{n-t}}\frac{1}{2^{n-t}}.$$

Putting these together we obtain:

$$\begin{aligned}
&\mathbb{E}_{\mathcal{G}}[I_s \mathbb{P}_G[Y_0 = 1, Y_s = 1]] \\
&\qquad = p_s \sum_{t=1}^{n-s} \binom{n-s}{t} [(1-p)^s]^t (1-(1-p)^s)^{n-s-t} \frac{1}{2^t}\frac{1}{4^{n-t}}, \\
&\qquad = \frac{p_s}{4^n}\left((2(1-p)^s + (1-(1-p)^s))^{n-s} - (1-(1-p)^s)^{n-s}\right) \\
&\qquad = \frac{p_s}{4^n}\left((1+(1-p)^s)^{n-s} - (1-(1-p)^s)^{n-s}\right);
\end{aligned}$$

therefore

$$\mathbb{E}_{\mathcal{G}}[b_2(G)] = \sum_{s=1}^{n} 2^{-n}\binom{n}{s} p_s \left[(1+(1-p)^s)^{n-s} - (1-(1-p)^s)^{n-s}\right] \leq S(n,p).$$

■

In the appendix we show that

LEMMA 2.5.

$$S(n,p) \leq O(n^{-\epsilon/4}) + \exp(-\Omega(n)),$$

and

$$R(n,p) \leq O(n^{-\epsilon/4}) + \exp(-\Omega(n)).$$

Given the above bounds on $\mathbb{E}_{\mathcal{G}}[b_1(G)]$ and $\mathbb{E}_{\mathcal{G}}[b_2(G)]$, Markov's inequality implies that with probability at least $1 - n^{-\epsilon/8} - 2^{-n}$ over the selection of the graph G from $G(n,p)$ we have

$$\max(b_1(G), b_2(G)) \leq O(n^{-\epsilon/8}) + \exp(-\Omega(n)). \tag{15}$$

Let us condition on the event that Condition (15) holds. Under this event, Lemma 2.1 implies that:

$$||Z - W|| \leq 2(b_1(G) + b_2(G)) \leq O(n^{-\epsilon/8}) + \exp(-\Omega(n))$$

as needed. Noting that $1 - n^{-\epsilon/8} - 2^{-n} \geq 1 - 2n^{-\epsilon/8}$, we obtain

$$\mathbb{P}_{\mathcal{G}}\left[||Z - W|| \leq O(n^{-\epsilon/8}) + \exp(-\Omega(n))\right] \geq 1 - 2n^{-\epsilon/8}. \tag{16}$$

Using the pessimistic upper bound of 2 on the total variation distance when Condition (15) fails, we obtain

$$||Z - W|| \leq O(n^{-\epsilon/8}) + \exp(-\Omega(n)).$$

Taking the limit of the above bound as $n \to +\infty$ we obtain our asymptotic result. This concludes the proof of Theorem 1.9. ■

2.2. *Medium Connectivity.*

PROOF OF THEOREM 1.10. Recall the *matching pennies game* from Definition 1.17. It is not hard to see that this game does not have a PNE. Hence, if a graphical game contains two players who are connected to each other, are isolated from all the other players, and play matching pennies against each other, then the graphical game will have no PNE. The existence of such a witness for the non-existence of PNE is precisely what we use to establish our result. In particular, we show that with high probability a random game sampled from $\mathcal{D}_{(n,p)}$ will contain an isolated edge between two players playing a matching pennies game.

We use the following exposure argument. Label the vertices of the graph with the integers in $[n] := \{1, \ldots, n\}$. Set $\Gamma_1 = [n]$ and perform the following operations, which iteratively define the sets of vertices Γ_i, $i \geq 2$. If $|\Gamma_i| \leq n/2$, for some $i \geq 2$, stop the process and do not proceed to iteration i: [†]

- Let j be the minimal value such that $j \in \Gamma_i$.
- If j is adjacent to more than one vertex or to none, let $\Gamma_{i+1} = \Gamma_i \setminus (\{j\} \cup \mathcal{N}(j))$. Go to the next iteration.
- Otherwise, let j' be the unique neighbor of j. If j' has a neighbor $\neq j$, let $\Gamma_{i+1} = \Gamma_i \setminus (\{j, j'\} \cup \mathcal{N}(j'))$. Go to the next iteration.
- Otherwise check if the players j and j' play a matching pennies game. [‡] If this is the case, declare NO NASH. Let $\Gamma_{i+1} = \Gamma_i \setminus \{j, j'\}$. Go to the next iteration.

Observe that the number of vertices removed at some iteration of the process can be upper bounded (formally, it is stochastically dominated) by

$$2 + \mathrm{Bin}(n, p),$$

[†]Throughout the process Γ_i represents the set of vertices that could be adjacent to an isolated edge, given the information available to the process at the beginning of iteration i.

[‡]More precisely, check if the best response tables of the players j and j' are the same with the best response tables of the players a and b of the matching pennies game from Definition 1.17 (up to permutations of the players' names).

where $\mathrm{Bin}(n,p)$ is a random variable distributed according to the Binomial distribution with n trials and success probability p. This follows from the fact that the vertices removed at some iteration of the process are either the examined vertex j and j's neighbors (the number of those is stochastically dominated by a $\mathrm{Bin}(n,p)$ random variable), or—if j has a single neighbor j'—the removed vertices are j, j' and the neighbors of j' (the number of those is also stochastically dominated by a $\mathrm{Bin}(n,p)$ random variable). Letting $m := \lceil 0.02n/(np+1) \rceil$, the probability that the process runs for at most m iterations is bounded by

$$\Pr\left[2m + \mathrm{Bin}(mn,p) \geq n/2\right] \leq \exp(-\Omega(n)).$$

Condition on the information known to the exposure process up until the beginning of iteration i, and assume that $|\Gamma_i| > n/2$. Let j be the vertex with the smallest value in Γ_i. Now reveal all the neighbors of j, and if j has only one neighbor j' reveal also the neighbors of j'. The probability that j is adjacent to a node j' who has no other neighbors is at least $\frac{n}{2}p(1-p)^{2n} =: p_{\mathrm{iso}}$; note that we made use of the condition $|\Gamma_i| > n/2$ in this calculation. Conditioning on this event, the probability (over the selection of the payoff tables) that j and j' play a matching pennies game is $\frac{1}{8} =: p_{\mathrm{mp}}$. Hence, the probability of outputting NO NASH in iteration i is at least $\frac{1}{8}\frac{1}{2}np(1-p)^{2n} =: p_{\mathrm{imp}}$.

The probability that the game has a PNE is upper bounded by the probability that the process described above does not return NO NASH, at any point through its completion. To upper bound the latter probability, let us imagine the following alternative process:

1. *Stage 1:* Toss n coins independently at random with head probability p_{iso}. Let $\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_n \in \{0,1\}$, where 1 represents 'heads' and 0 represents 'tails', be the outcomes of these coin tosses.
2. *Stage 2:* Toss n coins independently at random with head probability p_{mp}. Let $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n \in \{0,1\}$, be the outcomes of these coin tosses.
3. *Stage 3:* Run through the exposure process in the following way. At each iteration i:
   - conditioning on the information available to the exposure process at the beginning of the iteration, compute the probability p_j that the vertex j corresponding to the smallest number in Γ_i is adjacent to an isolated edge; given the discussion above it must be that $p_j \geq p_{\mathrm{iso}}$;
   - if $\mathcal{I}_i = 1$, then create an isolated edge connecting the player j to a random vertex $j' \in \Gamma_i \setminus \{j\}$, forbidding all other edges from j or j' to any other player, and make the players j and j' play a matching pennies game if $\mathcal{M}_i = 1$; if they do output NO NASH;

- if $\mathcal{I}_i = 0$, then sample the neighborhood of j from the following modified model:
  - with probability $\frac{p_j - p_{\rm iso}}{1-p_{\rm iso}}$, create an isolated edge connecting the player j to a random vertex $j' \in \Gamma_i \setminus \{j\}$, forbidding all other edges from j or j' to any other player, and make the players j and j' play a matching pennies game with probability $p_{\rm mp}$; if both of these happen, output NO NASH;
  - with the remaining probability, sample the neighborhood of j and the neighborhood of the potential unique neighbor j' from $G(n,p)$, conditioning on j not being adjacent to an isolated edge.
- Define Γ_{i+1} from Γ_i appropriately and exit the process if $|\Gamma_{i+1}| \le n/2$.

It is clear that the process given above can be coupled with the process defined earlier to exhibit the same behavior. But it is easier to analyze. In particular, letting $\mathcal{S} := \sum_{i=1}^{m} \mathcal{I}_i \mathcal{M}_i$, the probability that a Nash equilibrium does not exists can be lower bounded as follows:

$$\begin{aligned}
\mathbb{P}_{\mathcal{G}}[\nexists \text{ a PNE}] &\ge \Pr\left[\mathcal{S} \ge 1 \ \wedge \ \begin{matrix}\text{process runs for} \\ \text{at least } m \text{ steps}\end{matrix}\right] \\
&\ge \Pr\left[\mathcal{S} \ge 1\right] - \Pr\left[\begin{matrix}\text{process runs for} \\ \text{less than } m \text{ steps}\end{matrix}\right] \\
&\ge 1 - (1-p_{\rm imp})^m - \exp(-\Omega(n)).
\end{aligned}$$

Hence, the probability that a PNE exists can be upper bounded by

$$\begin{aligned}
&\exp(-\Omega(n)) + \left(1 - \frac{1}{16} np(1-p)^{2n}\right)^m \\
&\qquad \le \exp(-\Omega(n)) + \exp(-\Omega(mnp(1-p)^{2n})) \\
&\qquad \le \exp(-\Omega(mnp(1-p)^{2n})).
\end{aligned}$$

For $p \le 1/n$ the last expression is

$$\exp(-\Omega(n^2 p)),$$

while for $p = g(n)/n$ where $g(n) \ge 1$ the expression is

$$\exp(-\Omega(n(1-p)^{2n})) = \exp(-\Omega(ne^{-2g(n)})) = \exp(-\Omega(e^{\log_e(n) - 2g(n)})).$$

This concludes the proof of Theorem 1.10. ■

2.3. *Low Connectivity.*

PROOF OF THEOREM 1.11. Note that if the graphical game is comprised of isolated edges that are not matching pennies games then a PNE exists. (This can be checked easily by enumerating all best response tables for a 2×2 game.) We wish to lower bound the probability of this event. To do this, it is convenient to sample the graphical game in two stages as follows: At the first stage we decide for each of the possible $\binom{n}{2}$ edges whether the edge is *present* (with probability p) and whether it is *predisposed* to be a matching pennies game (independently with probability $1/8$); by 'predisposed' we mean that the edge will be set to be a matching pennies game if the edge turns out to be isolated. At the second stage, we do the following: for an edge that is both isolated and predisposed, we assign random payoff tables to its endpoints conditioning on the resulting game being a matching pennies game; for an isolated edge that is not predisposed, we assign random payoff tables to its endpoints conditioning on the resulting game *not* being a matching pennies game; finally, for any node that is part of a connected component with 0 or at least 2 edges we assign random payoff tables to the node. The probability that there is no edge in the first stage that is both present and predisposed is

$$(1 - p/8)^{\binom{n}{2}}.$$

Conditioning on this event, all present edges are not predisposed. Note also that, when c is fixed, the probability that there exists a pair of adjacent edges is $o(1)$. It follows that the probability that all present edges are not predisposed and no pair of edges intersect can be lower bounded as

$$(1 - p/8)^{\binom{n}{2}} - o(1) = \left(1 - \frac{c}{8n^2}\right)^{\frac{n(n-1)}{2}} - o(1).$$

But, as explained above if all edges are isolated and none of them is a matching pennies game a PNE exists. Hence, the probability that a PNE exists is at least

$$\left(1 - \frac{c}{8n^2}\right)^{\frac{n(n-1)}{2}} - o(1) \quad \longrightarrow \quad e^{-\frac{c}{16}}.$$

■

3. Deterministic Graphs.

3.1. *A Sufficient Condition for Existence of Equilibria: Strong Connectivity.*

PROOF OF THEOREM 1.13. We use the same notation as in the proof of Theorem 1.9, except that we make the slight modification of setting $N := 2^n - 1$. Recall that X_i, $i = 0, 1, \ldots, N-1$, is the indicator random variable of the event that the strategy profile encoded by the number i is a PNE. It is rather straightforward (see the proof of Theorem 1.9) to show that

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{i=0}^{N-1} X_i\right] = 1.$$

As in the proof of Theorem 1.9, to establish our result, it suffices to bound the following quantities.

$$\begin{aligned} b_1(G) &= \sum_{i=0}^{N-1} \sum_{j \in B_i} \mathbb{P}[X_i = 1]\mathbb{P}[X_j = 1], \\ b_2(G) &= \sum_{i=0}^{N-1} \sum_{j \in B_i \setminus \{i\}} \mathbb{P}[X_i = 1, X_j = 1], \end{aligned}$$

where the neighborhoods of dependence B_i are defined as in Lemma 2.2. For $S \subseteq \{1, \ldots, n\}$, denote by $i(S)$ the strategy profile in which the players of the set S play 1 and the players not in S play 0. Then writing $1(j \in B)$ for the indicator of the event that $j \in B$ we have:

$$\begin{aligned} b_2(G) &= \sum_{i=0}^{N-1} \sum_{j \in B_i \setminus \{i\}} \mathbb{P}[X_i = 1, X_j = 1] \\ &= \sum_{i=0}^{N-1} \sum_{j \neq i} \mathbb{P}[X_i = 1, X_j = 1] 1(j \in B_i) \\ &= N \sum_{j \neq 0} \mathbb{P}[X_0 = 1, X_j = 1] 1(j \in B_0) \qquad \text{(by symmetry)} \\ &= N \sum_{k=1}^{n} \sum_{S, |S|=k} \mathbb{P}[X_0 = 1, X_{i(S)} = 1] 1(i(S) \in B_0). \end{aligned}$$

We will bound the sum above by bounding

$$N \sum_{k=1}^{\lfloor \delta n \rfloor} \sum_{S, |S|=k} \mathbb{P}[X_0 = 1, X_{i(S)} = 1] 1(i(S) \in B_0), \tag{17}$$

and

$$N \sum_{k=\lfloor \delta n \rfloor + 1}^{n} \sum_{S, |S|=k} \mathbb{P}[X_0 = 1, X_{i(S)} = 1] 1(i(S) \in B_0) \tag{18}$$

separately.

Note that if some set S satisfies $|S| \leq \lfloor \delta n \rfloor$ then $|\mathcal{N}(S)| \geq \alpha |S|$ since the graph has (α, δ)-expansion. Moreover, each vertex (player) of the set $\mathcal{N}(S)$ is playing its best response to the strategies of its neighbors in both profiles 0 and $i(S)$ with probability $\frac{1}{4}$, since its environment is different in the two profiles. On the other hand, each player not in that set is in best response in both profiles 0 and $i(S)$ with probability at most $\frac{1}{2}$. Hence, we can bound (17) by

$$
\begin{aligned}
N \sum_{k=1}^{\lfloor \delta n \rfloor} \sum_{S, |S|=k} & \mathbb{P}[X_0 = 1, X_{i(S)} = 1] \\
& \leq N \sum_{k=1}^{\lfloor \delta n \rfloor} \sum_{S, |S|=k} \left(\frac{1}{2}\right)^{n-\alpha k} \left(\frac{1}{4}\right)^{\alpha k} = \sum_{k=1}^{\lfloor \delta n \rfloor} \binom{n}{k} \left(\frac{1}{2}\right)^{\alpha k} \\
& < \left(1 + \left(\frac{1}{2}\right)^{\alpha}\right)^{n} - 1 \leq e n^{-\epsilon}.
\end{aligned}
$$

To bound the second term, notice that, if some set S satisfies $|S| \geq \lfloor \delta n \rfloor + 1$, then since the graph has (α, δ)-expansion $\mathcal{N}(S) \equiv V$ and, therefore, the environment of every player is different in the two profiles 0 and $i(S)$. Hence, $1(i(S) \in B_0) = 0$. By combining the above we get that

$$
b_2(G) \leq e n^{-\epsilon}.
$$

It remains to bound the expression $b_1(G)$. We have

$$
\begin{aligned}
b_1(G) - 2^{-n} &= \sum_{i=0}^{N-1} \sum_{j \in B_i \setminus \{i\}} \mathbb{P}[X_i = 1] \mathbb{P}[X_j = 1] \\
&= \sum_{i=0}^{N-1} \sum_{j \neq i} \mathbb{P}[X_i = 1] \mathbb{P}[X_j = 1] 1(j \in B_i) \\
&= 2^{-n} \sum_{j \neq 0} 1(j \in B_0) \\
&= 2^{-n} \sum_{k=1}^{\lfloor \delta n \rfloor} \sum_{S, |S|=k} 1(i(s) \in B_0) + 2^{-n} \sum_{k=\lfloor \delta n \rfloor + 1}^{n} \sum_{S, |S|=k} 1(i(s) \in B_0).
\end{aligned}
$$

The second term is zero as before. For all large enough n the first summation contains at most $2^{n/2}$ terms and is therefore bounded by $2^{-n/2}$. It follows that

$$
b_1(G) + b_2(G) \leq e n^{-\epsilon} + 2^{-n/2}.
$$

An application of the result by Arratia et al. [4] concludes the proof of Theorem 1.13. ■

3.2. *A Sufficient Condition for the Non-Existence of Equilibria: Indifferent Matching Pennies.* In this section we provide a proof of Theorem 1.16. Recall that an edge of a graph is called *d-bounded* if both adjacent vertices have degrees smaller or equal to d. Theorem 1.16 specifies that any graph with many such edges is unlikely to have PNE. We proceed to the proof of the claim.

PROOF OF THEOREM 1.16. Consider a d-bounded edge in a game connecting two players a and b; suppose that each of these players interacts with $d-1$ (or fewer) other players denoted by $a_1, a_2 \dots a_{d-1}$ and $b_1, b_2 \dots b_{d-1}$. [§] Recall that if a and b play an indifferent matching pennies game against each other then the game has no PNE. The key observation is that a d-bounded edge is an indifferent matching pennies game with probability at least $(\frac{1}{8})^{2^{2d-2}} =: p_{imp}$—since a random two-player game is a matching pennies game with probability $\frac{1}{8}$ and there are at most 2^{2d-2} possible pure strategy profiles for the players $a_1, a_2 \dots a_{d-1}$, $b_1, b_2 \dots b_{d-1}$; for each of these pure strategy profiles the game between a and b must be a matching pennies game.

For a collection of m vertex disjoint edges, observe that the events that each of them is an indifferent matching pennies game are independent. Hence, the probability that the game has a PNE is upper bounded by the probability that none of these edges is an indifferent matching pennies game, which is upper bounded by

$$(1-p_{imp})^m \leq \exp(-mp_{imp}) = \exp\left(-m\left(\frac{1}{8}\right)^{2^{2d-2}}\right).$$

For the second claim of the theorem note that, if there are m d-bounded edges, then there must be at least $m/(2d)$ vertex disjoint d-bounded edges.

The algorithmic statement follows from the fact that we may find all nodes with degree $\leq d$ in time $O(n^2)$, and then find all edges joining two such nodes in another $O(n^2)$ time, with the use of the appropriate data structures; these edges are the d-bounded edges of the graph. Then in time $O(m2^{d+2})$ we can check if the endpoints of any such edge play an indifferent matching pennies game.

The final claim of the theorem has a similar proof where now the potential witnesses for the non-existence of a PNE are the edges in $\mathcal{E}$. ■

Many random graphical games on deterministic graphs such as players arranged on a line, grid, or any other bounded degree graph (with $\omega(1)$ edges) are special cases of the above theorem and hence are unlikely to have PNE asymptotically.

[§] We allow these lists to share players.

References.

[1] D. Achlioptas and Y. Peres. The threshold for random k-sat is 2k (ln 2 - o(k)). In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 223–231, New York, NY, USA, 2003. ACM Press.

[2] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[3] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. In *FOCS '04: Proceedings of the 45th Symposium on Foundations of Computer Science*, 2004.

[4] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for poisson approximations: The chen-stein method. *The Annals of Probability*, 17:9–25, 1989.

[5] I. Bárány, S. Vempala, and A. Vetta. Nash equilibria in random games. In *FOCS '05: Proceedings of the of the 46th Symposium on Foundations of Computer Science*, pages 134–145, 2005.

[6] C. Beeri, R. Fagin, D. Maier, and M. Yannakakis. On the desirability of acyclic database schemes. *J. ACM*, 30(3):479–513, 1983.

[7] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4):961–976, Nov. 2004.

[8] A. Czumaj, P. Krysta, and B. Vöcking. Selfish traffic allocation for server farms. In *STOC '02: Proceedings of the 34th ACM Symposium on the Theory of Computing*, pages 287–296, 2002.

[9] A. Czumaj and B. Vöcking. Tight bounds for worst-case equilibria. In *SODA '02: Proceedings of the 13th ACM-SIAM Symposium on Discrete algorithms*, pages 413–420, 2002.

[10] C. Daskalakis and C. H. Papadimitriou. Computing pure nash equilibria in graphical games via markov random fields. In *EC '06: Proceedings of the 7th ACM conference on Electronic commerce*, pages 91–99, New York, NY, USA, 2006. ACM Press.

[11] B. Dilkina, C. P. Gomes, and A. Sabharwal. The impact of network topology on pure nash equilibria in graphical games. In *AAAI-07: Proceedings of the 22nd Conference on Artificial Intelligence*, pages 42–49, Vancouver, BC, 2007.

[12] M. Dresher. Probability of a pure equilibrium point in n-person games. *Journal of Combinatorial Theory*, 8:134–145, 1970.

[13] E. Elkind, L. A. Goldberg, and P. Goldberg. Nash equilibria in graphical games on trees revisited. In *EC '06: Proceedings of the 7th ACM conference on Electronic commerce*, pages 100–109, New York, NY, USA, 2006. ACM Press.

[14] E. Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *J. Amer. Math. Soc.*, 12(4):1017–1054, 1999. With an appendix by Jean Bourgain.

[15] K. Goldberg, A. Goldman, and M. Newman. The probability of an equilibrium point. *Journal of Research of the National Bureau of Standards*, 72B:93–101, 1968.

[16] G. Gottlob, G. Greco, and F. Scarcello. Pure nash equilibria: hard and easy games. In *TARK '03: Proceedings of the 9th conference on Theoretical aspects of rationality and knowledge*, pages 215–230, New York, NY, USA, 2003. ACM Press.

[17] G. Gottlob, N. Leone, and F. Scarcello. Hypertree decompositions and tractable queries. *J. Comput. Syst. Sci.*, 64(3):579–627, 2002.

[18] S. Hart, Y. Rinott, and B. Weiss. Evolutionarily stable strategies of random games, and the vertices of random polygons. *Annals of Applied Probability*, 18:259–287, 2008.

[19] M. J. Kearns, M. L. Littman, and S. P. Singh. Graphical models for game theory. In *UAI '01: Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*, pages 253–260, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[20] J. Kleinberg. The emerging intersection of social and technological networks: Open questions

and algorithmic challenges. In *FOCS '06: Proceedings of the of the 47th Symposium on Foundations of Computer Science*, 2006.

[21] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *STACS '99: Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, volume 1563, pages 404–413, 1999.

[22] M. L. Littman, M. J. Kearns, and S. P. Singh. An efficient, exact algorithm for solving tree-structured graphical games. In *NIPS*, pages 817–823, 2001.

[23] L. E. Ortiz and M. J. Kearns. Nash propagation for loopy graphical games. In *NIPS*, pages 793–800, 2002.

[24] C. H. Papadimitriou. Algorithms, games, and the internet. In *STOC '01: Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 749–753, 2001.

[25] G. P. Papavassilopoulos. On the probability of existence of pure equilibria in matrix games. *Journal of Optimization Theory and Applications*, 87:419–439, 1995.

[26] I. Y. Powers. Limiting distributions of the number of pure strategy nash equilibria in n-person games. *International Journal of Game Theory*, 19:277–286, 1990.

[27] P. Raghavan. The changing face of web search: algorithms, auctions and advertising. In *STOC '06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, page 129, 2006.

[28] Y. Rinott and M. Scarsini. On the number of pure strategy nash equilibria in random games. *Games and Economic Behavior*, 33:274–293, 2000.

[29] T. Roughgarden. Stackelberg scheduling strategies. In *STOC '01: Proceedings of the 33th ACM Symposium on Theory of Computing*, pages 104–113, 2001.

[30] T. Roughgarden. The price of anarchy is independent of the network topology. In *STOC '02: Proceedings of the 34th ACM Symposium on the Theory of Computing*, pages 428–437, 2002.

[31] T. Roughgarden and E. Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[32] T. Roughgarden and E. Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47(2):389–403, 2004.

[33] W. Stanford. A note on the probability of k pure nash equilibria in matrix games. *Games and Economic Behavior*, 9:238–246, 1995.

[34] A. Vetta. Nash equilibria in competitive societies with applications to facility location. In *FOCS '02: Proceedings of the 43rd Symposium on Foundations of Computer Science*, pages 416–425, 2002.

APPENDIX A: OMITTED PROOFS

PROOF OF LEMMA 2.5. We need to bound the functions $S(n,p)$ and $R(n,p)$. We begin with S.

Bounding S. Recall that

$$S(n,p) := \sum_{s=1}^{n} \binom{n}{s} 2^{-n} \left[(1+(1-p)^s)^{n-s} - (1-(1-p)^s)^{n-s}\right].$$

We split the range of the summation into four regions and bound the sum over each region separately. We begin by choosing $\alpha = \alpha(\epsilon)$ as follows

(i) if $\epsilon \leq \frac{1790}{105}$, we choose $\alpha = \left(\frac{\epsilon}{2+\epsilon}\right)^{20}$;
(ii) if $\epsilon > \frac{1790}{105}$, we choose $\alpha = \frac{\epsilon}{2+\epsilon}$.

Given our choice of $\alpha = \alpha(\epsilon)$ we define the following regions in the range of s (where—depending on ϵ—Regions I and/or III may be empty and Region IV may have overlap with Region II):

I. $\{s \in \mathbb{N} \mid 1 \leq s < \frac{\epsilon}{(2+\epsilon)p}\}$;
II. $\{s \in \mathbb{N} \mid \frac{\epsilon}{(2+\epsilon)p} \leq s < \alpha n\}$;
III. $\{s \in \mathbb{N} \mid \alpha n \leq s < \frac{1}{2+\epsilon} n\}$;
IV. $\{s \in \mathbb{N} \mid \frac{1}{2+\epsilon} n \leq s < n\}$.

We then write

$$S(n,p) \leq S_{\mathrm{I}}(n,p) + S_{\mathrm{II}}(n,p) + S_{\mathrm{III}}(n,p) + S_{\mathrm{IV}}(n,p),$$

where $S_{\mathrm{I}}(n,p)$ denotes the sum over region I etc., and bound each term separately.

Region I. The following lemma will be useful.

LEMMA A.1. *For all $\epsilon > 0$, $p \in (0,1)$ and s such that $1 \leq s < \frac{\epsilon}{(2+\epsilon)p}$,*

$$(1-p)^s \leq 1 - \frac{(2+0.5\epsilon)sp}{2+\epsilon}.$$

PROOF. First note that, for all $k \geq 1$,

$$\binom{s}{2k+2} p^{2k+2} \leq \binom{s}{2k+1} p^{2k+1}. \tag{19}$$

To verify the latter note that it is equivalent to

$$s \leq 2k + 1 + \frac{2k+2}{p},$$

which is true since $s \leq \frac{\epsilon}{(2+\epsilon)p} = \frac{1}{(\frac{2}{\epsilon}+1)p} \leq \frac{1}{p}$.

Using (19), it follows that

$$(1-p)^s \leq 1 - \binom{s}{1}p + \binom{s}{2}p^2. \tag{20}$$

Note finally that

$$\frac{0.5\epsilon}{2+\epsilon}sp > \frac{s(s-1)}{2}p^2,$$

which applied to (20) gives

$$(1-p)^s \leq 1 - \frac{(2+0.5\epsilon)sp}{2+\epsilon}.$$

■

Assuming that Region I is non-empty and applying Lemma A.1 we get:

$$
\begin{aligned}
S_{\mathrm{I}}(n,p) &\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \binom{n}{s} 2^{-n}(1+(1-p)^s)^{n-s} \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \binom{n}{s} 2^{-n}\left(1+1-\frac{(2+0.5\epsilon)sp}{2+\epsilon}\right)^{n-s} \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \binom{n}{s} 2^{-s}\left(1-\frac{(1+0.25\epsilon)sp}{2+\epsilon}\right)^{n-s} \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \binom{n}{s} 2^{-s}\exp\left(-\frac{(1+0.25\epsilon)sp}{2+\epsilon}(n-s)\right) \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \binom{n}{s} 2^{-s}\exp\left(-\frac{(1+0.25\epsilon)sp}{2+\epsilon}n\right)\exp\left(\frac{(1+0.25\epsilon)sp}{2+\epsilon}s\right) \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \binom{n}{s} 2^{-s}\exp\left(-(1+0.25\epsilon)\log_e(n)\ s\right)\exp\left(\frac{(1+0.25\epsilon)\epsilon}{(2+\epsilon)^2}s\right) \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} n^s 2^{-s} n^{-(1+0.25\epsilon)s}\exp\left(\frac{1}{2}s\right) \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \left(\frac{\sqrt{e}}{2}\right)^s n^{-0.25\epsilon s} \\
&\leq \sum_{s<\frac{\epsilon}{(2+\epsilon)p}} \left(\frac{\sqrt{e}}{2}\right)^s n^{-0.25\epsilon} \\
&\leq n^{-0.25\epsilon} \sum_{s<\frac{2\epsilon}{(2+\epsilon)p}} \left(\frac{\sqrt{e}}{2}\right)^s \\
&= O(n^{-0.25\epsilon}) \qquad\qquad \left(\text{since } \frac{\sqrt{e}}{2}<1\right).
\end{aligned}
$$

Region II. We have

$$\begin{aligned} S_{\text{II}}(n,p) &\leq \sum_{\frac{\epsilon}{(2+\epsilon)p}\leq s<\alpha n} \binom{n}{s} 2^{-n}(1+(1-p)^s)^n \\ &\leq \sum_{\frac{\epsilon}{(2+\epsilon)p}\leq s<\alpha n} \binom{n}{s} 2^{-n}\left(1+e^{-ps}\right)^n \\ &\leq \sum_{\frac{\epsilon}{(2+\epsilon)p}\leq s<\alpha n} \binom{n}{\alpha n} 2^{-n}\left(1+e^{-p\frac{\epsilon}{(2+\epsilon)p}}\right)^n \\ &\leq \alpha n \binom{n}{\alpha n}\left(\frac{1+e^{-\frac{\epsilon}{2+\epsilon}}}{2}\right)^n \\ &\leq \alpha n 2^{nH(\alpha)}(n+1)\left(\frac{1+e^{-\frac{\epsilon}{2+\epsilon}}}{2}\right)^n \\ &\leq \alpha n(n+1)\left(2^{H(\alpha)}\cdot\frac{1+e^{-\frac{\epsilon}{2+\epsilon}}}{2}\right)^n. \end{aligned} \tag{21}$$

In the above derivation $H(\cdot)$ represents the entropy function, and for the second to last derivation we used the fact that:

$$\binom{n}{k} \leq (n+1)2^{nH(\frac{k}{n})}. \tag{22}$$

Our definition of the function $\alpha = \alpha(\epsilon)$ guarantees that when $\epsilon \leq \frac{1790}{105}$:

$$\left(2^{H(\alpha)}\cdot\frac{1+e^{-\frac{\epsilon}{2+\epsilon}}}{2}\right) \leq 0.999,$$

while when $\epsilon > \frac{1790}{105}$:

$$\left(2^{H(\alpha)}\cdot\frac{1+e^{-\frac{\epsilon}{2+\epsilon}}}{2}\right) \leq 0.99.$$

Using the above and (21) we obtain

$$S_{\text{II}}(n,p) = \exp(-\Omega(n)). \tag{23}$$

Region III. Let us assume that the region is non-empty. We show that each positive term in the summation $S_{\text{III}}(n,p)$ is exponentially small. Since there are

$O(n)$ terms in the summation it follows then that $S_{\text{III}}(n,p)$ is exponentially small.

$$\begin{aligned}
\binom{n}{s}2^{-n}(1+(1-p)^s)^n &\le \binom{n}{s}2^{-n}(1+e^{-ps})^n \\
&\le \binom{n}{s}2^{-n}(1+e^{-p\alpha n})^n \\
&\le \binom{n}{s}2^{-n}\left(1+e^{-(2+\epsilon)\alpha\log_e(n)}\right)^n \\
&= \binom{n}{s}2^{-n}\left(1+\frac{1}{n^{(2+\epsilon)\alpha}}\right)^n \\
&= \binom{n}{s}2^{-n}\left(1+\frac{1}{n^{(2+\epsilon)\alpha}}\right)^{n^{(2+\epsilon)\alpha}n^{1-(2+\epsilon)\alpha}} \\
&\le \binom{n}{s}2^{-n}e^{n^{1-(2+\epsilon)\alpha}} \\
&\le (n+1)2^{nH\left(\frac{s}{n}\right)}2^{-n}e^{n^{1-(2+\epsilon)\alpha}} \\
&\le (n+1)2^{nH\left(\frac{1}{2+\epsilon}\right)}2^{-n}e^{n^{1-(2+\epsilon)\alpha}} \\
&= (n+1)2^{n\left(H\left(\frac{1}{2+\epsilon}\right)-1\right)}e^{n^{1-(2+\epsilon)\alpha}}, \qquad (24)
\end{aligned}$$

where in the third-to-last line of the derivation we employed the bound of Equation (22). Notice that the RHS of (24), seen as a function of $\epsilon > 0$ and $\alpha > 0$, is decreasing in both. Since $\epsilon > c$, our choice of $\alpha = \alpha(\epsilon)$ implies that $\alpha > \left(\frac{c}{c+2}\right)^{20}$. Hence, we can bound the RHS of (24) as follows:

$$(n+1)2^{-n\left(1-H\left(\frac{1}{2+c}\right)\right)}e^{n^{1-(2+c)\left(\frac{c}{c+2}\right)^{20}}} = \exp(-\Omega(n)),$$

where we used the fact that c is a constant, and therefore the factor $e^{n^{1-(2+c)\left(\frac{c}{c+2}\right)^{20}}}$ is sub-exponential in n, while the factor $2^{-n\left(1-H\left(\frac{1}{2+c}\right)\right)}$ is exponentially small in n.

Region IV. Note that, if $xk \le 1$, then by the mean value theorem

$$(1+x)^k-(1-x)^k \le 2x \max_{1-1/k\le y\le 1+1/k} ky^{k-1} = 2kx(1+1/k)^{k-1} \le 2ekx.$$

We can apply this for $k = n-s$ and $x = (1-p)^s$ since

$$(n-s)(1-p)^s \le (n-s)e^{-ps} \le (n-s)e^{-\frac{(2+\epsilon)\log_e(n)}{n}\frac{n}{2+\epsilon}} \le \frac{n-s}{n} \le 1.$$

Hence, $S_{\mathrm{IV}}(n,p)$ is bounded as follows.

$$\begin{aligned}
S_{\mathrm{IV}}(n,p) &\leq \sum_{\frac{n}{2+\epsilon}\leq s\leq n} \binom{n}{s} 2^{-n} 2e(n-s)(1-p)^s \\
&\leq 2e\cdot 2^{-n}\cdot n \sum_{\frac{n}{2+\epsilon}\leq s\leq n} \binom{n}{s}(1-p)^s \\
&\leq 2e\cdot 2^{-n}\cdot n(1+(1-p))^n \\
&\leq 2en\left(1-\frac{p}{2}\right)^n \\
&\leq 2ene^{-\frac{p}{2}n} \\
&\leq 2ene^{-\frac{(2+\epsilon)\log_e(n)}{2n}n} \\
&\leq 2enn^{-\frac{2+\epsilon}{2}} \\
&\leq 2en^{-\frac{\epsilon}{2}}.
\end{aligned}$$

Putting everything together. Combining the above we get that

$$S(n,p)\leq O(n^{-\epsilon/4})+\exp(-\Omega(n)).$$

Bounding R. Observe that

$$R(n,p)=2^{-n}+\sum_{s=1}^{n}\binom{n}{s}2^{-n}\min(1,n(1-p)^{s-1}).$$

We bound R as follows.

$$
\begin{aligned}
R(n,p) - 2^{-n} &\le \sum_{s=1}^{n} \binom{n}{s} 2^{-n} \min(1, n\exp(-p(s-1))) \\
&\le 2^{-n} \sum_{1 \le s \le \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}} \binom{n}{s} + 2^{-n} \sum_{s > \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}} \binom{n}{s} n\exp(-p(s-1)) \\
&\le 2^{-n} \sum_{1 \le s \le \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}} (n+1) 2^{nH(s/n)} + 2^{-n} \sum_{s > \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}} \binom{n}{s} n\exp(-p(s-1)) \\
&\le n(n+1) 2^{-n} 2^{nH\left(\frac{3+\epsilon}{6+3\epsilon}\right)} + 2^{-n} \sum_{s > \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}} \binom{n}{s} n\exp(-p(s-1)) \\
&\le \exp(-\Omega(n)) + 2^{-n} \sum_{s > \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}} \binom{n}{s} n\exp(-p(s-1)),
\end{aligned}
$$

where in the last line of the derivation we used that $\epsilon > c > 0$ for some absolute constant c. To bound the last sum we observe that when $s > \frac{n}{\frac{6+3\epsilon}{3+\epsilon}}$ we have

$$
\begin{aligned}
n\exp(-p(s-1)) &\le n\exp\left(-\frac{(2+\epsilon)\log_e(n)}{n}\left(\frac{n}{\frac{6+3\epsilon}{3+\epsilon}} - 1\right)\right) \\
&\le n \cdot n^{-\frac{2+\epsilon}{\frac{6+3\epsilon}{3+\epsilon}}} \cdot \exp\left(\frac{(2+\epsilon)\log_e(n)}{n}\right) \\
&\le n^{-\epsilon/3} \cdot n^{2/n} \cdot n^{\epsilon/n} = O(n^{-\epsilon/4}).
\end{aligned}
$$

Using this bound and the fact $\sum_{s=0}^{n} \binom{n}{s} = 2^n$ concludes the proof. ■

ADDRESS OF THE FIRST AUTHOR
COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY
DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
MIT, CAMBRIDGE, MA 02139
E-MAIL: costis@csail.mit.edu

ADDRESS OF THE SECOND AUTHOR
DEPARTMENT OF ELECTRICAL ENGINEERING - SYSTEMS
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CA 90089
E-MAIL: dimakis@usc.edu

ADDRESS OF THE THIRD AUTHOR
DEPARTMENT OF STATISTICS,
UC BERKELEY
BERKELEY, CA 94720
E-MAIL: mossel@stat.berkeley.edu