FUZZY CONVENTIONS

MARCIN PĘSKI

ABSTRACT. We propose an equilibrium selection theory for Granovetter (1978)'s threshold adoption model on networks. In the model, each agent adopts a new behavior only if the fraction of her neighbors doing the same is larger than her i.i.d. threshold. A fuzzy convention x is a profile where, for (almost) all agents, approximately x fraction of their neighbors adopts a new behavior. A *random-utility (RU) dominant outcome* x^* is a maximizer of an integral of the distribution of thresholds. The definition generalizes Harsanyi and Selten (1988)'s risk dominance to coordination games with random utility. We show that each network, if the number of agents is large and each agent has sufficiently many neighbors, has a fuzzy convention x^*. On some networks, including a city network, all equilibria are fuzzy conventions x^*. Fuzzy convention x^* is the only profile with such properties, and the only profile robust to incomplete information about the network structure.

1. INTRODUCTION

In many social situations, people's behavior is chosen due to a combination of individual and social factors. An important recent example is the post-Covid-era mask-wearing: some people wear masks to protect themselves or others, others don't wear them because of inconvenience or personal beliefs, and many, including the author of this study, are affected by how many people around them wear masks. The latter reason is social and it turns mask-wearing into a game of coordination, with, possibly, multiple equilibria. This paper proposes a theory of equilibrium choice that is based on the distribution of individual tastes and the details of the network of interactions among agents.

A well-known model of such situations has been introduced in Granovetter (1978). Each agent i has a threshold τ_i and the agent will adopt a new behavior (say, wear a mask) if and only if more than fraction τ_i of the population adopts it as well. A threshold below 0 (or, above 1) means the agent always (never) wears a mask. The distribution of thresholds, or tastes, in the population is given by function P, where $P(x)$ is the fraction of the population with a threshold equal to or smaller than x. The top row of Figure 1 contains two examples of threshold distributions. In both cases, 30% of the population always wear masks, and another 30% never wear masks. Under P_1, the remaining 40% wear masks only if at least 0.55 of their neighbors wear masks. Under P_2, the remaining 40% wear masks only if at least 0.4 of their neighbors do the same. Granovetter (1978) characterized equilibria of this model as intersections with the 45°-line. In both cases, there are two stable equilibria: A with 0.3 and B with 0.7 fractions wearing masks. (In each case, there is also an unstable equilibrium in-between.) In order to explain why one or the other equilibrium may be realized, one must refer to outside-of-the-model factors, like a local history, expectations, sunspots, etc.

Following the seminal work of Granovetter (1978), a large literature generalizes Granovetter's model to networks. In the model, each agent is a single node on a network and wears a mask (i.e., adopts the new behavior) only if the fraction of her neighbors doing the same is larger than her threshold. Most papers, like Watts (2002), Jackson and Yariv (2007), or López-Pintado (2008) (among many others) study Granovetter's model on a random graph with heterogeneous degree distribution. One of the key observations is that the degree distribution affects the equilibrium behavior. At the same time, random graph-based models have limitations: they do not capture many important aspects of real-world networks, like clustering, or overlapping neighborhoods. The latter are well-known to play an important role in coordination and contagion-type phenomena (Ellison (1993), Blume (1995), orMorris (2000)).

This paper combines Granovetter's model with insights from the contagion literature. Consider two examples of networks. In a "city" network, people located on a two-dimensional grid interact with their nearest neighbors. The neighborhood sets overlap among neighbors. In the random graph (Erdős and Rényi (1959)), neighbors are randomly selected from the population. For each network and each realization of thresholds, we compute the average behavior in the lowest (least mask-wearing) and
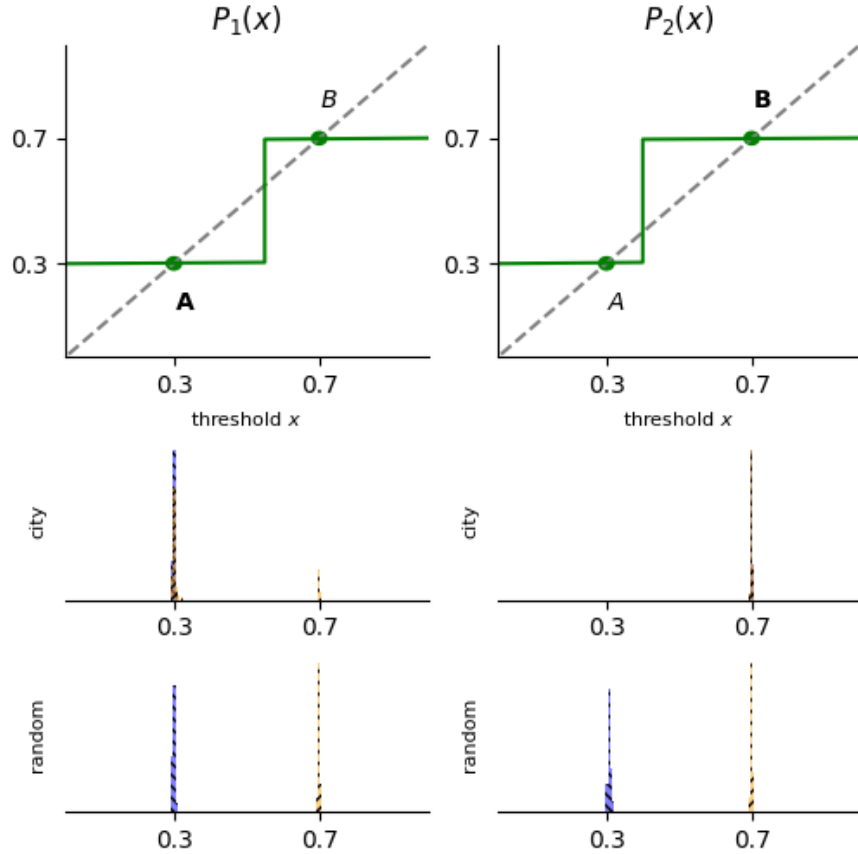
FIGURE 1. Monte-Carlo derived frequencies of the average behavior in the lowest (blue, "\" hatch) and highest (yellow, "/" hatch) equilibria for two different networks, city and random graph, and two threshold distributions. Both graphs have 60,000 agents and each agent has 120 neighbors.

the highest equilibria. By taking many realizations of individual thresholds, we compute the distribution of the sets of equilibrium average behavior for each network, and then compare the distributions across networks.

The two bottom rows of Figure 1 show the distribution of the lowest and highest average equilibrium behaviors for the two threshold distributions from the top row. A few patterns emerge. Not surprisingly, the lowest and the highest equilibria in the random graph correspond to the lowest (A) and highest (B) equilibria from the population model of Granovetter (1978), regardless of the threshold distribution. On the city network, the range of equilibrium behaviors is smaller and it depends on a
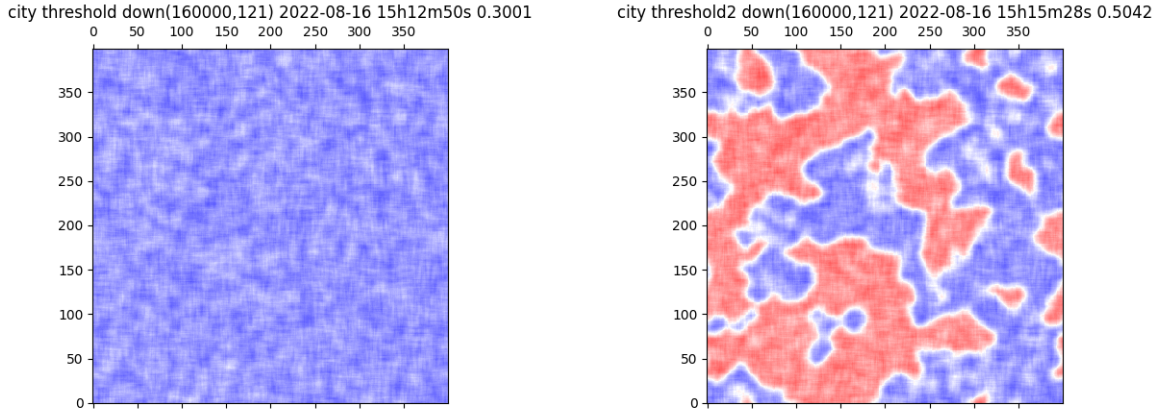
FIGURE 2. The average neighborhood behavior in a fuzzy convention (left panel) and in a not-fuzzy-convention-type (right panel) equilibrium on a city network with 160,000 agents and each agent has 120 neighbors. The shade of each pixel represents the fraction of neighbors wearing masks, with blue (red) shade corresponding to fractions below (above) 0.5.

threshold distribution: Under P_1, the lowest and the majority of realizations of the highest equilibria are concentrated around the continuum equilibrium A. Under P_2, the average behavior in the highest and lowest equilibria is equal to the behavior in the continuum equilibrium B.

Additional insight comes from a closer examination of the equilibria. For each equilibrium, we compute the average neighborhood behavior, i.e., the share of neighbors who wear masks in the neighborhood. The left panel shows a typical average neighborhood behavior in the highest equilibrium under P_1 on the city network. Almost all agents face roughly the same fraction of neighbors who wear masks. We refer to such an equilibrium as a *fuzzy convention*. In a fuzzy convention, there is considerable behavior heterogeneity on the micro, but on not the macro scale. The right panel shows an equilibrium that is not a fuzzy convention: there is visible macro-level heterogeneity, with mask-wearing more prevalent in some areas than others.

We have two results to explain the observed patterns. Define a *random utility-dominant*, or *RU-dominant*, outcome x^* as a solution to the maximization problem[1]

$$x^* \in \arg\max_x \int\limits_0^x \left(y - P^{-1}\left(y\right) \right) dy. \tag{1}$$

The RU-dominant outcome depends on the threshold distribution. On Figure 1, the RU-dominant outcomes are denoted in bold font: it is **A** for distribution P_1 and **B** for P_2. If the threshold distribution is concentrated on a single outcome (i.e., all agents' preferences are identical), then the RU-dominant outcome is equivalent to the risk-dominant outcome (Harsanyi and Selten (1988)) of a 2×2 coordination game.

The first result applies to all networks in which each agent has sufficiently many neighbors or, in short, to all sufficiently fine networks. We show that, for almost all realizations of taste thresholds, any such network has an equilibrium that is a fuzzy convention x^*, i.e., where the average neighborhood behavior is (roughly) equal to x^*. The proof relies on the characterization of Granovetter's model as a potential game. Such games are introduced in Monderer and Shapley (1996), where it is shown that any profile that is a local maximizer of the potential function is an equilibrium of the underlying game. In the proof, we show that, regardless of the structure of the network, the global maximizer of the potential function is a fuzzy convention x^* with a probability close to 1 (i.e., for almost all realizations of thresholds).

The above result does not preclude an existence of non-fuzzy-convention equilibria for general networks. A proper understanding of whether there are other non-fuzzy equilibria on fine networks, and what happens when neighborhood size is not large enough goes beyond this paper and it is left for future research.

At the same time, the second result shows that, on some networks, including the sufficiently large city network, all equilibria are fuzzy conventions x^*. Not only is non-fuzzy-convention behavior not an equilibrium, but all equilibria also look very similar. In the proof we show that, for each profile with an average behavior that is not a fuzzy convention, contagion-like best response dynamics would bring the neighborhood average behavior close to x^*. The proof uses an idea from Blume (1995) to show that a contagion wave spreads across lattice networks. The form of the relevant contagion

---

[1]Although our derivation is independent, (1) is equivalent to a formula from Morris and Shin (2006), where it is derived as a potential function for the continuum population version of the model.

wave adapted for RU-dominance is more complicated. We do not have an explicit construction and rely on an existence argument instead. Another complication is due to the random preferences. We first work with an approximate toy model, where each agent is replaced by a continuum population. This is additionally supplemented with explicit calculations of (a) the likelihood that a favorable configuration of payoff shocks may initiate a contagion wave, and (b) the likelihood that such a wave would not be stopped by an unfavorable configuration of payoff shocks (similar calculations are studied in percolation theory). The latter is a reason why the 1-dimensional "line" network of Ellison (1993) is not a good example for the result and a 2-dimensional "city" network is needed.

The two results together show that the single-element set $\{x^*\}$ is a tight lower bound on all sets of equilibrium average behaviors across all sufficiently fine networks. This leads to an *equilibrium selection* theory: x^* is the only average behavior that is robust to changes in the underlying network or incomplete information about the network.

The two results of this paper apply to networks where each agent has a large neighborhood. We choose this assumption for both methodological and practical reasons. First, we want to move one step away from Granovetter (1978), which studies equilibria on a continuum population. The characterization of average equilibrium on such networks is straightforward, and it relies on a continuum version of the law of large numbers. Our assumptions allow us to keep the power of the law of large numbers as much as possible, while testing how the characterization of Granovetter (1978) is affected by the network structure. As our second result shows, the network structure may eliminate some of the continuum population equilibria. Contrary to Granovetter (1978), our results allow for sparse networks, where the number of connections is significantly smaller than the size of the population. Second, many important networks have a large number of connections. Some are reasonably well approximated by the city network that is the object of the first result. Although our results are asymptotic (the exact bounds on the "fineness" of the networks can be derived from the proofs, but the proofs are not optimized for this goal), the simulations reported in Figure 1 suggest that natural parameters generate patterns that are consistent with our results.

1.1. **Literature review.** Coordination games form one of three main approaches in the literature that studies games on networks (Jackson and Zenou (2015)). The results of this paper are closely related to the literature on contagion in networks. Ellison

(1993) (see also Ellison (2000)) is the first to show that a risk-dominant action can spread from a small initial set of deviators to an entire 1-dimensional lattice network by a simple best response process. Blume (1995) and Lee and Valentinyi (2000) show that a risk-dominant outcome will spread to the entire 2-dimensional lattice if it is large and there is sufficient randomness in the initial configuration. Notice that the second requirement is satisfied in our case if the threshold distribution admits a non-zero probability for players with strictly dominant actions. Morris (2000) describes general properties of networks for which Ellison's contagion wave exists. Morris (2000) also shows that risk-dominated actions cannot spread through a best response process regardless of the geometry of the network.

Jackson and Yariv (2007) analyzes a Bayesian equilibrium, where the agents choose their action without knowing the thresholds of their neighbors (a similar approach to observability is taken, for instance in Galeotti *et al.* (2010)). This assumption improves model's tractability as agent's behavior does not depend on individual thresholds of her neighbors. At the same time, this assumption is not satisfactory if the equilibrium is to be interpreted as a long-term process as each agent may change her behavior when they observe the actions of their neighbors. This is different from our model, where an equilibrium is a steady state behavior *after* the thresholds are realized and actions are chosen. Because the neighbors in the Bayesian equilibrium of Jackson and Yariv (2007) are selected at random, the neighborhood structure looks like a random graph. For similar reasons to those discussed above (see Fig. 1), there are typically multiple equilibria.

Another literature studies evolutionary equilibrium selection in games with heterogeneous populations. For instance, Friedman (1991) describes a general framework with multiple continuum populations choosing actions and receiving payoffs and studies evolutionary steady states of continuous time adjustment dynamics. More closely related to this paper is Neary (2012), which studies a similar model to us but with two payoff shocks (more precisely, two subpopulations of deterministic size) and agents located on a complete graph. The paper presents conditions under which the evolutionary dynamics of Kandori *et al.* (1993) selects a fuzzy convention, i.e., an equilibrium where members of different subpopulations play different actions. Neary and Newton (2017) studies general payoff shocks and presents a sufficient condition under which the logit dynamics of Blume (1993) selects a fuzzy convention.

Evolutionary game theory (Kandori *et al.* (1993), Young (1993), Blume (1993), Newton (2021), and many others) studies the long-run behavior of perturbed best response processes, where players commit mistakes with a small probability, and instead of choosing a best response, take some other action. One of the key results of this literature is that risk-dominant coordination is (uniquely) stochastically stable regardless of the underlying network (Peski (2010)). Our current results (specifically, Theorems 2 and 1) are closely related, but with some key differences. On the one hand, there is a relation between "noise" in the behavioral rules of the evolutionary literature and "noise" in the payoffs of the current paper. On the other hand, there are two important differences: Here, we are interested in static equilibria instead of a dynamic adjustment process, and our payoff shocks are permanent instead of temporary mistakes. (The best response dynamic plays an important role in the proofs as a tool to identify equilibria.) Finally, the evolutionary literature is subject to the criticism that one may need to wait for a very long time before reaching a stochastically stable outcome (Ellison (1993)). That criticism does not apply to our static model.

2. Model

2.1. **Model.** There are N agents $i = 1, ..., N$. The network is defined as an undirected weighted graph with weights $g_{ij} = g_{ji} \geq 0$ for $i, j \leq N$. The weights represent a frequency of interactions between two agents. We assume that $g_{ii} = 0$ and that $g_i = \sum_j g > 0$ for each player i. We also assume that none of the players has significantly more connections than others, $\max_{i,j} g_i/g_i \leq w^* < \infty$. Each agent i has a threshold τ_i drawn i.i.d. from probability distribution P. Each network g, and each realization of thresholds τ leads to a many-player complete information static game $G(g, \tau)$.

Each agent chooses a binary action $a_i \in \{0, 1\}$. Examples of such choices are mask-wearing, planting a yard sign, mowing one's lawn, etc. For each action profile a, define a profile $\beta^a = (\beta_i^a)$ of average neighborhood fractions of agents who play action 1, i.e., $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$. An action profile is a Nash equilibrium if agents play action 1 (0) only if the average action in their neighborhood is larger (smaller) than their threshold, i.e., $a_i = 1 \, (0) \Rightarrow \beta_i^a \geq (\leq) \tau_i$.

The model is equivalent to a wide class of random utility binary coordination games on networks. The notion of equilibrium is a standard, static equilibrium of a complete information model. Although it is convenient to assume that players know the

thresholds and the network structure of the entire society, this assumption is neither realistic nor necessary. For the interpretation of the equilibrium, it is sufficient that agents observe the actions of their neighbors. Because ours is a coordination game, we can safely think about an equilibrium as a steady state of myopic best response adjustment process.

2.2. **Fuzzy convention.** For $\varepsilon > 0$ and $x \in [0, 1]$, a profile a is ε-*fuzzy convention* x if the fraction of agents who observe an average neighborhood behavior ε-away from x is not larger than ε:

$$\frac{1}{N} \left\{ i : |\beta_i^a - x| \geq \varepsilon \right\} \leq \varepsilon.$$

In a fuzzy convention, almost all agents experience approximately the same behavior of their neighbors, regardless of possibly complicated topology of the network, or their behavior thresholds.

2.3. *RU*-**dominant outcome.** An outcome $x^* \in [0, 1]$ is *random utility (RU) dominant* if

$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y) \right) dy. \tag{2}$$

(When P is not invertible, we define $P^{-1}(y) = \inf\{(x : P(x) \geq y)\}$.) It is *strictly RU-dominant*, if it is the unique maximizer. The integral (2) is equal to the difference in the areas below the cdf $P(.)$ and the 45° line. To compute it, we add areas between the two lines, such that the area below the 45° line and above $P(.)$ is added with a "−" sign and the area above the 45° line and below $P(.)$ is added with the "+" sign. The left panel of Fig. 3 illustrates such a calcuation for $x = 1$.

Generically, any maximizer of (2) is a stable fixed point of $P(x) = x$, and hence, it is equal to the average behavior in the continuum version of Granovetter (1978). However, even if the continuum model has multiple stable fixed points, generically, there exists only a unique *RU*-dominant outcome.

The definition generalizes the risk-dominance of (Harsanyi and Selten (1988)) from deterministic binary coordination games to games with heterogeneous payoffs. To see that, suppose that $P(.)$ is degenerate and concentrated on a single threshold τ (i.e., there is no uncertainty about thresholds). In such a case, our model is strategically equivalent to a deterministic binary coordination game, where each agent's best response is 1 if and only if fraction τ of his neighbors choose 1 as well. Figure 3 shows
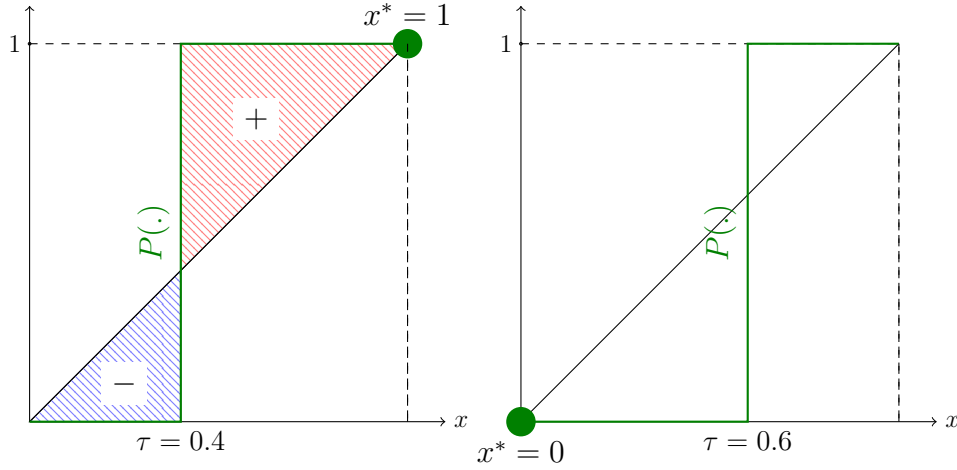
FIGURE 3. RU- and risk-dominance when P is degenerate for $\tau = 0.4$ (left panel) and $\tau = 0.6$ (right panel).

the distribution P for two values of τ. In both cases, the integral from expression (2)

is equal to $\begin{cases} -\frac{1}{2}x^2 & x \leq \tau \\ x - \frac{1}{2}x^2 - \tau & x \geq \tau \end{cases}$ and

- when $\tau = 0.4$, it is maximized at $x^* = 1$,
- when $\tau = 0.6$, it is maximized at $x^* = 0$.

In both cases, the RU-dominant outcome is identical to the risk-dominant one.

For future reference, note that any strictly RU-dominant outcome is also a unique maximizer of

$$\nu(x) = \frac{1}{2}(P(x))^2 - \int_0^x y\,dP(y). \tag{3}$$

Indeed, the maximizer of (3) must satisfy $P(x) = x$, and a change of variables shows that the two expressions are equal for such x.

3. RU-DOMINANT FUZZY CONVENTION

Define a bound on the importance of a single player in another player's neighborhood as

$$d(g) = \max_{i,j} \frac{g_{ij}}{g_i} \in [0,1].$$

For $d(g)$ to be small, each player must have many neighbors.

Theorem 1. *Suppose that x^* is the strictly RU-dominant outcome. For each $\eta > 0$, there is $d > 0$ such that, for each network g st. $d(g) \leq d$, with probability $1 - \eta$, there is an equilibrium that is η-fuzzy convention x^*.*

If the network is sufficiently fine, i.e. when $d(g)$ is small, then, for almost all realizations of thresholds, there is an equilibrium where almost all agents observe that approximately fraction x^* of their neighbors playing action 1.

The proof relies on the fact that Granovetter's model is a potential game (Monderer and Shapley (1996)). For each action profile a and threshold profile τ, define

$$V(a; \tau) = \frac{1}{2} \sum_{i,j} g_{ij} a_i a_j - \sum g_i a_i \tau_i.$$

Then, $V(a_i, a_{-i}; \tau) - V(a_i', a_{-i}; \tau) = g_i (\beta_i^a - \tau_i)(a_i - a_i')$, which implies that $V(1, a_{-i}; \tau) - V(0, a_{-i}; \tau) \geq 0$ if and only if 1 is a best response for player i. In other words, V is a potential function.

Monderer and Shapley (1996) shows that a profile is an equilibrium profile of a potential game if and only if it is a *local* maximizer of a potential function. Additionally, *global* maximizers of the potential function are equilibria selected by two different equilibrium selection arguments: robustness to incomplete information (Ui (2001)) and stochastic stability under logistic dynamics (Blume (1993), Blume (2018)).

In the proof, we show that, if the network is sufficiently large and fine, for almost all realizations of τ, any *global* maximizer of the potential function is a fuzzy convention x^*. Apart from demonstrating Theorem 1, our argument also shows that fuzzy convention x^* survives the two aforementioned equilibrium selection criteria.

3.1. **Concentration inequality.** We sketch the main steps of the proof. We start with a concentration inequality. Let $\mathcal{F}$ be the set of measurable functions $f : [0,1]^2 \to [0,1]$. For each $f \in \mathcal{F}$, each b, let $\mathbb{E} f(., b) = \int f(x, b) \, dP(x)$ denote the expectation of $f(., b)$ with respect to the distribution of thresholds P. The Hoeffding inequality implies that there exists constants $B < \infty$ and $c_\varepsilon > 0$ such that for each profile a and measurable function $f(\tau, \beta) \in [0, 1]$,

$$\text{Prob}\left(\left|\sum_i g_i f(\tau_i, \beta_i^a) - \sum_i g_i \, \mathbb{E} f(., \beta_i^a)\right| \geq \varepsilon \sum g_i \right) \leq B \exp(-c_\varepsilon N). \qquad (4)$$

Similarly, the Hanson-Wright inequality says that, for possibly different constants B and c_ε,

$$\text{Prob}\left(\left|\sum_{i,j} g_{ij}\left(\prod_{k=i,j} f\left(\tau_k, \beta_k^a\right)\right) - \sum_{i,j} g_{ij}\left(\prod_{k=i,j} \mathbb{E} f\left(., \beta_k^a\right)\right)\right| \geq \varepsilon \sum g_i\right) \leq B\exp\left(-c_\varepsilon N\right).$$
(5)

The above inequalities hold for each profile a separately. The next Lemma show that they can be strengthened to hold *uniformly* across all profiles.

Lemma 1. *There exist constants $B < \infty$ and $c\left(\varepsilon, K, d\right)$ for each $\varepsilon > 0$, $K < \infty$, and $d > 0$ such that $\liminf_{d \to 0} c_{\varepsilon, K, d} > 0$ and such that if $f \in \mathcal{F}$ is a K-Lipschitz function, then*

$$Prob\left(\sup_a \left|\sum_i g_i f\left(\tau_i, \beta_i^a\right) - \sum_i g_i \mathbb{E} f\left(., \beta\right)\right| \geq \varepsilon \sum g_i\right)$$
$$\leq B\,exp\left(-c_{\varepsilon, K, d(g)} N\right),$$

$$Prob\left(\sup_a \left|\sum_{i,j} g_{ij}\left(\prod_{k=i,j} f\left(\tau_k, \beta_k^a\right)\right) - \sum_{i,j} g_{ij}\left(\prod_{k=i,j} \mathbb{E} f\left(., \beta_k^a\right)\right)\right| \geq \varepsilon \sum g_i\right)$$
$$\leq B\,exp\left(-c_{\varepsilon, K, d(g)} N\right).$$

In the proof of the Lemma, the probability bounds are obtained as a product between bounds (4) and (5) and an appropriate measure of the size of the set of neighborhood profiles $\mathcal{B} = \{\beta^a : a \text{ is a profile}\}$. The basic idea is contained in the following inequality that holds for any function $F\left(\beta^a\right)$ of the profile of neighborhood average behavior β^a and the counting measure $|.|$ of the size of set $\mathcal{B}$:

$$\text{Prob}\left(\sup_a F\left(\beta^a\right)\right) = \text{Prob}\left(\sup_{\beta \in \mathcal{B}} F\left(\beta\right)\right)$$
$$\leq |\mathcal{B}| \sup_{\beta \in \mathcal{B}} \text{Prob}\left(F\left(\beta\right)\right) = |\mathcal{B}| \sup_a \text{Prob}\left(F\left(\beta^a\right)\right).$$

Because the counting measure is too large ($|\mathcal{B}| \sim \exp\left(2N\right)$), the proof instead relies on metric entropy in order to evaluate the "size" of $\mathcal{B}$. We show the metric entropy of $\mathcal{B}$ is of order $\exp\left(d\left(g\right) N\right)$. The use of metric entropy requires some modifications to the above argument, including the restriction to Lipschitz functions f.

3.2. **Estimates of the potential function.** We use Lemma 1 in two calculations below. First, we find the potential of profile a^*:

$$a_i^* = \mathbf{1}\left\{\tau_i < x^*\right\}. \tag{6}$$

In profile a^*, each player chooses an action optimally assuming that they face fraction x^* of their opponents playing 1. It can be shown that, with a large probability, a^* is a fuzzy convention x, i.e., $\beta_i^a \approx x^*$. Note that $\mathbb{E}\,\mathbf{1}\left\{. < x^*\right\} = P\left(x^*\right)$. Lemma 1 implies the following estimate:

$$
\begin{aligned}
V\left(a^*; \tau\right) &= \frac{1}{2}\sum_{i,j} g_{ij} a_i a_j - \sum g_i a_i \tau_i \\
&= \frac{1}{2}\sum_{i,j} g_{ij} \mathbf{1}\left\{\tau_i < x^*\right\}\mathbf{1}\left\{\tau_j < x^*\right\} - \sum g_i \mathbf{1}\left\{\tau_i < x^*\right\}\tau_i \\
&\approx \frac{1}{2}\sum_{i,j} g_i \left(P\left(x^*\right)\right)^2 - \sum g_i \int_0^{x^*} y\,dP\left(y\right) = \sum g_i \nu\left(x^*\right).
\end{aligned}
$$

(Because $\mathbf{1}\left\{. < x^*\right\}$ is not Lipschitz, the Lemma is applied to a Lipschitz approximation - the details are left for the Appendix).

Second, we estimate the potential for an arbitrary equilibrium profile. We show that, unless the equilibrium is a fuzzy convention x^*, its potential is strictly smaller than the one computed above. Indeed, take profile a such that $a_i = \mathbf{1}\left(\tau_i \leq \beta_i^a\right)$ for each i. Applying Lemma 1, we obtain

$$
\begin{aligned}
V\left(a; \tau\right) &= \frac{1}{2}\sum_{i,j} g_{ij} a_i a_j - \sum g_i a_i \tau_i \\
&= \frac{1}{2}\sum_{i,j} g_i \mathbf{1}\left(\tau_i \leq \beta_i^a\right)\mathbf{1}\left\{\tau_j < \beta_j^a\right\} - \sum g_i \mathbf{1}\left(\tau_i \leq \beta_i^a\right)\tau_i \\
&\approx \frac{1}{2}\sum_{i,j} g_{ij} P\left(\beta_i^a\right) P\left(\beta_j^a\right) - \sum g_i \int_0^{\beta_i^a} y\,dP\left(y\right).
\end{aligned}
$$

Because $2P\left(\beta_i^a\right)P\left(\beta_j^a\right) \leq P\left(\beta_i^a\right)^2 + P\left(\beta_j^a\right)^2$, the potential of a is not larger than

$$
\leq \frac{1}{2}\sum_{i,j} g_{ij}\left(P\left(\beta_i^a\right)\right)^2 - \sum g_i \int_0^{\beta_i^a} y\,dP\left(y\right) = \sum_i g_i \nu\left(\beta_i^a\right).
$$

By the last remark in Section 2.3, unless $\beta_i^a = x^*$, the above is strictly smaller than the potential of a^*. In other words, unless a is fuzzy convention x^*, it cannot maximize the potential.

4. RU-DOMINANT SELECTION

In the previous section, we showed that all sufficiently fine networks have equilibria that are fuzzy conventions x^*. Here, we show that there are networks where, with a large probability, all equilibria are fuzzy conventions x^*:

Theorem 2. *Suppose that x^* is the strictly RU-dominant outcome and that either $x^* > 0$ and $P(0) > 0$, or $x^* < 1$ and $P(1) < 1$. For each $\eta > 0$, there is a network g such that, with probability $1 - \eta$, each equilibrium is η-fuzzy convention x^*.*

The network constructed in the proof is a version of the city network described in the introduction. It is parameterized with M and m. There are M^2 agents living on square $\left[0, \frac{M}{m}\right]^2 \subseteq \mathbb{R}^2$ at fractional points $\left(\frac{k}{m}, \frac{l}{m}\right)$ for $k, l = 1, ..., M$. Any two agents i and j are connected, $g_{ij} = 1$, if the (Euclidean) distance between them is no larger than 1. To avoid separately dealing with the border cases, we assume that all distance calculations are done $\mod \frac{M}{m}$, which transforms the square $\left[0, \frac{M}{m}\right]^2$ into a torus. We show that, if m and $\frac{M}{m}$ are sufficiently large, then, for a large probability set of realizations, there is no equilibrium such that the average neighborhood actions are significantly higher than x^* for a significant group of agents. Together with an identical argument for the other side, this suffices to establish the theorem. Our argument extends to K-dimensional lattices for any $K \geq 2$, but not to $K = 1$.

If $P(0) > 0$, then, with positive probability, there are agents for whom action 1 is strictly dominant and it is played in any equilibrium. Similarly, if $P(1) < 1$, then, with a positive probability, there are agents for whom action 0 is strictly dominant. The assumption on the distribution ensures that, on a large network, there will be a group of agents who play one or the other action regardless of the behavior of their neighbors. Such agents play the same role as the initial infectors in Morris (2000) or the agents who make mistakes in evolutionary models like the one studied in Ellison (1993).

4.1. **Unique behavior on line.** First, we explain how the maximization problem (2) is connected to the (approximate) uniqueness of the equilibrium. For the intuition, we

work with a toy version of the line network from Ellison (1993) that we describe now. Suppose that agents are distributed uniformly along a line at discrete and equally spaced locations. Each location contains a continuum population of mass 1. The populations in locations i and j are connected with each other, with weights that depend only on the distance $g_{ij} = g_{i-j} =: g_{j-i}$. We assume there are no connections between agents in the same location, i.e., $g_0 = 0$, and the weights are normalized so that $\sum g_d = 1$. In this toy version of our model, the continuum assumption allows us to use the law of large numbers to compute the average best response action of agents in location i as $P\left(\sum_d g_d a_{i+d}\right)$, where a_j is the average current action played by agents in location j.

Suppose that, initially, the average action in locations $i \leq 0$ is x^* or below. We are going to show that, if agents in other locations best respond, their average behavior cannot be larger than x^*. For this, consider the largest profile of average actions such that $a_j \leq x^*$ for $j \leq 0$ and such that, in all locations $i > 0$, the behavior is not larger than the best response:

$$a_i \leq P\left(\sum_d g_d a_{i+d}\right).$$

Such largest profile exists due to the payoff complementarities. Clearly, $a_i \geq x^*$ for each i, and $a_i = x^*$ for $i \leq 0$. Due to the payoff complementarities again, a_i must be increasing in i. Let $a = \lim_{i \to \infty} a_i$ and, by contradiction, suppose that $a > x^*$. Taking the inverse, we obtain

$$P^{-1}(a_i) \leq \sum_d g_d a_{i+d} = x^* + \sum_j \left(\sum_{d \geq j-i} g_d\right)(a_{j+1} - a_j),$$

where the equality is due to a discrete version of the "integration by parts" formula and the fact that $a_i \geq x^*$ for each i. After multiplying by $a_{i+1} - a_i \geq 0$, and summing up across all locations i, we get

$$\sum_i \left(P^{-1}(a_i) - x^*\right)(a_{i+1} - a_i) \leq \sum_{i,j} \left(\sum_{d \geq j-i} g_d\right)(a_{i+1} - a_i)(a_{j+1} - a_j). \qquad (7)$$

The left-hand side of the inequality is approximately equal to $\int_{x^*}^a \left(P^{-1}(y) - x^*\right) dy$. To compute the right-hand side, notice that we can switch the roles of i and j in the summation without affecting its value. Together with the fact that $\sum_{d \geq j-i} g_d +$

$\sum_{d \geq i-j} g_d = \sum g_d = 1$, we get

$$\sum_{i,j} \left(\sum_{d \geq j-i} g_d \right) (a_{i+1} - a_i)(a_{j+1} - a_j)$$

$$= \frac{1}{2} \left(\sum_{i,j} \left(\sum_{d \geq j-i} g_d + \sum_{d \geq i-j} g_d \right) (a_{i+1} - a_i)(a_{j+1} - a_j) \right)$$

$$= \frac{1}{2} \left(\sum_{i,j} (a_{i+1} - a_i)(a_{j+1} - a_j) \right) = \frac{1}{2}(a - x^*)^2$$

$$= \frac{1}{2}(a - x^*)^2 = \int_{x^*}^{a} (y - x^*)\, dy.$$

Putting the two sides together, inequality (7) implies that

$$\int_{x^*}^{a} \left(y - P^{-1}(y) \right) dy \geq 0.$$

If $a > x^*$, this contradicts the fact that x^* is the unique maximizer of the integral on the right-hand side. The contradiction shows that the only equilibrium behavior that is consistent with the group of agents $j \leq 0$ playing $a_j \leq x^*$ (perhaps for non-equilibrium reasons) is that all locations play no more than x^*.

One can think about the largest profile as an outcome of a best-response revision process, where locations $i > 0$ start with average behavior $a_i^0 = 1$, and then, in each period, revise it to the current best responses (Morris (2000)). In such a case, the above observation shows that contagion spreads x^* across the entire line.

The same observation extends from the line to higher-dimensional lattices due to an elegant argument from Blume (1995) (see also Lee and Valentinyi (2000) and Morris (2000)). The idea is that if the initial group is sufficiently large, we can approximate it using a set with a smooth (i.e., low curvature) boundary. Then, we can analyze the spread of the contagion wave behavior in the direction that is normal to the boundary. This trick turns the problem into a one-dimensional one, and the above argument applies.

4.2. **Obstacles.** The continuum assumption used in the above argument ensures that the average best response action of agents in a location is a deterministic (rather than random) function of the average behavior in neighboring locations. At the same time,
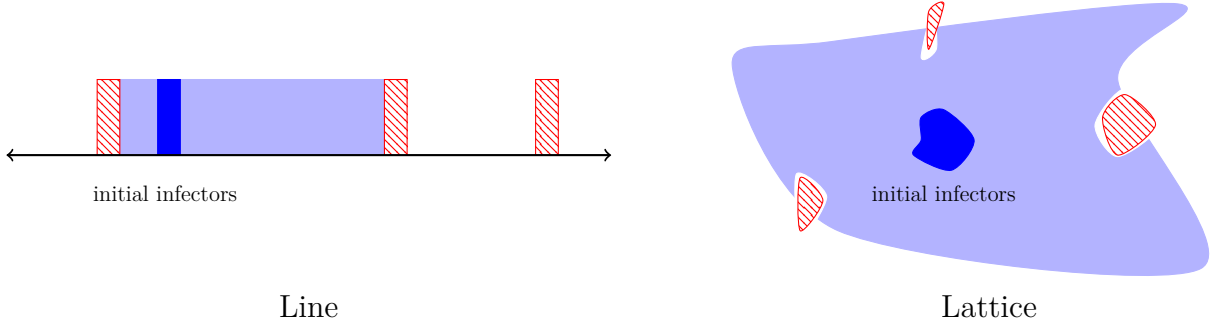
FIGURE 4. Obstacles to the contagion wave

the assumption ignores a positive probability of a contiguous group of "bad" agents for whom 1 is the strictly dominant action. If sufficiently large, such a group of "bad" agents will stop the best response revisions towards action 0 and block the contagion wave (see the left panel of Figure 4).

One could try to compare the relative frequency of initial infectors necessary to start the wave versus the sets of "bad" agents who may block it. Unfortunately, for some Ps, the latter are more frequent. As a result, the line network is not a good candidate example for Theorem 2.

At the same time, the "bad" sets are intuitively less likely to block the contagion wave on higher-dimensional lattices (see the right panel of Figure 4). The reason is that to block the wave, the "bad" sets would have to be arranged to surround it. Even if the number of "bad" sets is much larger than the number of initial infectors, the probability of a bad arrangement can be quite small. This intuition is clarified in the proof of Theorem 2.

5. Robustness of fuzzy conventions

The discussion after Theorem 1 shows that, for a large probability of threshold realizations, and for each sufficiently fine network, Granovetter's game has an equilibrium that is robust to incomplete information (Kajii and Morris (1997)): any perturbation of the original coordination game obtained by adding a small probability that players' payoffs are different, has a nearby equilibrium that is a fuzzy convention x^*.

We are going to show that, additionally, fuzzy convention x^* equilibria are typically robust to (possibly, large) uncertainty about the underlying network. In the next result, a^* is the profile defined in (6).

Theorem 3. *Suppose that P does not have an atom at x^*. For each $\delta > 0$, there is $\eta > 0$, and $N_0 < \infty$ such that if $N \geq N_0$, there is a set $T_\delta \subseteq [0,1]^N$ of threshold realizations such that $Prob\,(T_\delta) \geq 1 - \delta$ and, for each τ, if a is an η-fuzzy convention x^* equilibrium in game $G\,(g, \tau)$ for some network g, then $\frac{1}{N}\,|\{i : a_i \neq a_i^*\}| \leq \delta$.*

To interpret the Theorem, notice that profile a^* has a network-independent definition: each player is best responding *as if* fraction x^* of its neighbors play x, regardless of the structure of their neighborhoods and what their neighbors on the network are actually doing. According to the Theorem, the behavior in such a profile is close to any fuzzy convention x^* equilibrium, regardless of the underlying network. Together with Theorem 1, the above result shows that playing a^* is close to an equilibrium for a great majority of players, whatever is the true network of interactions, whether the agents know the network or not.

Because Theorem 2 shows that fuzzy conventions x^* are the only equilibria on some networks, the results of this paper lead to an *equilibrium selection* theory: the only equilibria that are robust to changes in the underlying networks are fuzzy conventions of the RU-dominant outcome x^*.

Appendix A. Proof of Theorem 1

A.1. **Proof of Lemma 1.** Define a distance on the space of (mixed) profiles: For any $a, b \in [0,1]^N$, let

$$d\,(a, b) = \sqrt{\frac{1}{\sum g_i^2} \sum g_i^2\,(a_i - b_i)^2}.$$

Recall that $\mathcal{B} = \{\beta^a : a \text{ is action profile}\}$ is the space of neighborhood fractions. For each $\delta > 0$, let $\mathcal{N}\,(\delta, \mathcal{B})$ be the covering number of $\mathcal{B}$, i.e., the smallest cardinality n of a list of profiles $b^1, ..., b^n \in \mathcal{B}$ such that, for each $b \in \mathcal{B}$, there is $l \leq n$ so that $d\,(b, b^l) \leq \delta$.

Lemma 2. *There exists a universal constant $c < \infty$ such that, for each $\delta > 0$, and each network g,*

$$\mathcal{N}\,(\delta, \mathcal{B}) \leq \exp\left(\frac{1}{\delta^2} c w^{*2} d\,(g)\,N\right).$$

Proof. We will use Sudakov's Minoration Inequality (Theorem 7.4.1 from Vershynin (2018)), which provides an upper bound on the covering number via the expectation of a certain Gaussian process. For this, let Z_i for each agent i be an i.i.d. standard

normal random variable. For each (possibly mixed) profile $a \in \mathcal{A}$, define

$$X_a = \frac{1}{\sqrt{\sum_i g_i^2}} \sum_i g_i a_i Z_i.$$

For any two profiles $a, b \in \mathcal{A}$,

$$\sqrt{\mathbb{E}(X_a - X_b)^2} = \sqrt{\frac{1}{\sum g_i^2} \mathbb{E}\left(\sum_i g_i (a_i - b_i) Z_i\right)^2}$$

$$= \sqrt{\frac{1}{\sum g_i^2} \sum_i g_i (a_i - b_i)^2} = d(a, b).$$

Given the definition and the above property, Sudakov's Minoration Inequality implies that, for some universal constant $c_1 > 0$ (i.e., a constant that is independent of parameters and the current problem),

$$\log \mathcal{N}(\delta, \mathcal{B}) \leq c_1 \frac{(\mathbb{E} \sup_{b \in \mathcal{B}} X_b)^2}{\delta^2}.$$

We compute

$$\mathbb{E} \sup_{b \in \mathcal{B}} X_b = \mathbb{E} \sup_{a \in \mathcal{A}} X_{\beta^a} = \mathbb{E}\left(\sup_{a \in \mathcal{A}} \frac{1}{\sqrt{\sum_i g_i^2}} \sum_i g_i Z_i \left(\frac{1}{g_i} \sum g_{ij} a_j\right)\right)$$

$$= \frac{1}{\sqrt{\sum_i g_i^2}} \mathbb{E}\left(\sup_{a \in \mathcal{A}} \sum_i a_i \left(\sum_j g_{ij} Z_j\right)\right) \leq \frac{1}{\sqrt{\sum_i g_i^2}} \mathbb{E} \sum_i \left|\sum_j g_{ij} Z_j\right|$$

$$\leq \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\sum_i g_i^2}} \sum_i \sqrt{\sum_j g_{ij}^2},$$

where the last inequality is due to a bound on the expectation of the absolute value of the normal variable $\sum g_{ij} Z_j$ via its standard deviation $\sigma_i = \sqrt{\sum_j g_{ij}^2}$. Because $\sum_j g_{ij}^2 \leq d(g) g_i^2$ and $(\sum_i g_i)^2 \leq N^2 w^{*2} g_{\min}^2 \leq N w^{*2} \sum g_i^2$, we have

$$\log \mathcal{N}(\delta, \mathcal{B}) \leq \sqrt{\frac{2}{\pi}} c_1 \frac{1}{\delta^2} \frac{1}{\sum_i g_i^2} \left(\sum_i \sqrt{d(g)} g_i\right)^2 d(g) \leq \frac{1}{\delta^2} \sqrt{\frac{2}{\pi}} c_1 w^{*2} d(g) N.$$

$$\square$$

We proceed with the proof of Lemma 1. For the first inequality, suppose f is K-Lipschitz. Fix $\varepsilon > 0$ and $\delta > 0$ so that $\delta = \frac{1}{12K\sqrt{w^*}} \varepsilon$. Find δ-cover $b^1, ..., b^n$ of $\mathcal{B}$.

Because $n \leq \mathcal{N}(\delta, \mathcal{B})$, Lemma 2 implies that

$$\text{Prob}\left(\sup_{l \leq n}\left|\sum_i g_i f\left(\tau_i, b_i^l\right) - \sum_i g_i \, \mathbb{E}\, f\left(., b_i^l\right)\right| \geq \frac{1}{2}\varepsilon \sum g_i\right)$$

$$\leq nB\exp\left(-c\left(\frac{1}{2}\varepsilon\right)N\right) \leq B\exp\left(-\left(c\left(\frac{1}{2}\varepsilon\right) - \frac{1}{\delta^2}cw^{*2}d\left(g\right)\right)N\right).$$

Assume that the complement of the event in the parantheses of the first line of the above inequality holds. For each action profile a, find l so that $d\left(b^l, \beta^a\right) \leq \delta$. Then, by the Jensen's inequality, and because $\frac{g_i}{\sum g_i} \leq w^* \frac{g_i^2}{\sum g_i^2}$,

$$\sum_i \frac{g_i}{\sum_i g_i}\left|\beta_i^a - b_i^l\right| \leq \sqrt{\sum \frac{g_i}{\sum g_i}\left(\beta_i^a - b_i^l\right)^2} \leq \sqrt{\sum w^* \frac{g_i^2}{\sum g_i^2}\left(\beta_i^a - b_i^l\right)^2} \leq \sqrt{w^*}\delta.$$

Hence,

$$\left|\sum_i g_i f\left(\tau_i, \beta_i^a\right) - \sum_i g_i \, \mathbb{E}\, f\left(., \beta_i^a\right)\right|$$

$$\leq \left|\sum_i g_i f\left(\tau_i, b_i^l\right) - \sum_i g_i \, \mathbb{E}\, f\left(., b_i^l\right)\right| + 2K\left|\sum_i g_i\left|\beta_i^a - b_i^l\right|\right|$$

$$\leq \left|\sum_i g_i f\left(\tau_i, b_i^l\right) - \sum_i g_i \, \mathbb{E}\, f\left(., b_i^l\right)\right| + 2K\sqrt{w^*}\delta\left(\sum_i g_i\right) \leq \varepsilon \sum_i g_i.$$

Take $c\left(\varepsilon, K, d\right) = c\left(\frac{1}{2}\varepsilon\right) - \frac{1}{\varepsilon^2}c\left(6Kw^*\right)^2 d$. The claim follows.

For the second inequality, we first derive a version of (4): we show that there exists constants $B < \infty$ and $c\left(\varepsilon\right) > 0$ such that, for each profile a and measurable function $f\left(\tau, \beta\right) \in [0, 1]$,

$$\text{Prob}\left(\left|\sum_i g_{ij}\left(\prod_{k=i,j} f\left(\tau_k, \beta_k^a\right)\right) - \sum_i g_{ij}\left(\prod_{k=i,j} \mathbb{E}\, f\left(., \beta_k^a\right)\right)\right| \geq \varepsilon \sum g_i\right) \leq B\exp\left(-c\left(\varepsilon\right)N\right).$$

(8)

Indeed, suppose that $X_i \in [-1, 1]$ is a collection of independent mean zero random variables. The Hanson-Wright inequality (Theorem 6.2.1 Vershynin (2018)) implies that there exists a universal constant $c > 0$ such that, for each $t > 0$,

$$\mathbb{P}\left(\left|\sum g_{ij}X_iX_j - \mathbb{E}\sum g_{ij}X_iX_j\right| \geq t\right) \leq 2\exp\left(-c\min\left(\frac{t^2}{\|G\|_F^2}, \frac{t}{\|G\|}\right)\right). \qquad (9)$$

where $G = [g_{ij}]$ is the adjacency matrix, $\|G\|_F$ is the Frobenius norm and $\|G\|$ is the operator L^2-norm. Let $X_i = f(\tau_i, \beta_i^a) - \mathbb{E} f(., \beta_i^a)$ and $t = \varepsilon_i \sum g_i$. Recall that X_i are independent and that $g_{ij} = g_{ji}$ and $g_{ii} = 0$ to obtain

$$\sum g_{ij} X_i X_j - \mathbb{E} \sum g_{ij} X_i X_j = \sum g_{ij} X_i X_j$$

$$= \sum g_{ij} \left(\prod_{k=i,j} f(\tau_k, \beta_k^a) \right) - \sum_i g_{ij} \left(\prod_{k=i,j} \mathbb{E} f(., \beta_k^a) \right)$$

$$- 2 \sum g_i \left(\mathbb{E} f(., \beta_j^a) \right) \left(f(\tau_i, \beta_i^a) - \mathbb{E} f(., \beta_i^a) \right).$$

Hence,

$$\text{Prob} \left(\left| \sum_i g_{ij} \left(\prod_{k=i,j} f(\tau_k, \beta_k^a) \right) - \sum_i g_{ij} \left(\prod_{k=i,j} \mathbb{E} f(., \beta_k^a) \right) \right| \geq \varepsilon \sum g_i \right)$$

$$\leq \text{Prob} \left(\left| \sum g_{ij} X_i X_j - \mathbb{E} \sum g_{ij} X_i X_j \right| \geq \frac{1}{2} \varepsilon \sum g_i \right)$$

$$+ \text{Prob} \left(\left| \sum g_i \left(\mathbb{E} f(., \beta_j^a) \right) \left(f(\tau_i, \beta_i^a) - \mathbb{E} f(., \beta_i^a) \right) \right| \geq \frac{1}{4} \varepsilon \sum g_i \right).$$

We apply (9) to the first bound (notice that $\|G\| \leq \|G\|_F \leq \sqrt{N} \|G\|$ and $g_{\min} \leq \|G\| \leq w^* g_{\min}$, where $g_{\min} = \min_i g_i$) and Hoeffding's inequality (4) to the second bound to obtain

$$\leq 2 \exp \left(-c \min \left(\varepsilon^2 \frac{(\sum g_i)^2}{N w^{*2} g_{\min}^2}, \varepsilon \frac{\sum g_i}{w^* g_{\min}} \right) \right) + B \exp \left(-c \left(\frac{1}{4} \varepsilon \right) N \right)$$

$$\leq 2 \exp \left(-c \frac{1}{w^{*2}} \varepsilon^2 N \right) + B \exp \left(-c \left(\frac{1}{4} \varepsilon \right) N \right).$$

This concludes the proof of (8).

Given (8), we conclude the proof of the second inequality of Lemma 1 in the same manner as in the case of the first inequality. In particular, if $d\left(b^l, \beta^a\right) \leq \delta$,

$$
\left|\sum g_{ij}\left(\prod_{k=i,j} f\left(\tau_k, \beta_k^a\right) - \prod_{k=i,j} f\left(a_k, b_k^l\right)\right)\right|
$$

$$
\leq \sum g_{ij} f\left(\tau_i, \beta_i^a\right)\left|f\left(\tau_j, \beta_j^a\right) - f\left(\tau_k, b_j^l\right)\right| + \sum g_{ij} f\left(\tau_j, b_j^l\right)\left|f\left(\tau_i, \beta_i^a\right) - f\left(\tau_i, b_i^l\right)\right|
$$

$$
\leq K\left(\sum_j\left(\sum_i g_{ij} f\left(\tau_i, \beta_i^a\right)\right)\left|\beta_j^a - b_j^l\right| + \sum_i\left(\sum_j g_{ij} f\left(\tau_j, b_j^l\right)\right)\left|\beta_i^a - b_i^l\right|\right)
$$

$$
\leq 2K\sum_i g_i\left|\beta_i^a - b_i^l\right| \leq 2K\sqrt{w^*}\delta\sum_i g_i \leq \frac{1}{2}\varepsilon\sum_i g_i.
$$

Similar calculations apply to $\sum_i g_{ij}\left(\prod_{k=i,j} \mathbb{E} f\left(., \beta_k^a\right)\right)$. Hence, if

$$
\sup_{l \leq n}\left|\sum_i g_{ij}\prod_{k=i,j} f\left(\tau_k, b_k^l\right) - \sum_i g_{ij}\mathbb{E}\prod_{k=i,j} f\left(., b_k^l\right)\right| \geq \frac{1}{2}\varepsilon\sum g_i, \text{ for each } l
$$

then

$$
\left|\sum_i g_{ij}\left(\prod_{k=i,j} f\left(\tau_k, \beta_k^a\right)\right) - \sum_i g_{ij}\left(\prod_{k=i,j} \mathbb{E} f\left(., \beta_k^a\right)\right)\right| \leq \varepsilon\sum_i g_i.
$$

The rest of the argument follows.

A.2. **Proof of Theorem 1.** Fix $\eta > 0$. For each $\delta > 0$, let $\nu_\delta^0 = \max_{x:|x-x^*|\leq\delta}\left(\nu\left(x^*\right) - \nu\left(x\right)\right)$ and let $\nu_\delta^1 = \min_{x:|x-x^*|\geq\delta}\left(\nu\left(x^*\right) - \nu\left(x\right)\right)$. Because x^* is the unique maximizer of $\nu\left(.\right)$, $\nu_\delta^1 > 0$ for each δ. Moreover, $\lim_{\delta\to 0}\nu_\delta^0 > 0$.

Let $\kappa > 0$ and define $\frac{1}{\kappa}$-Lipshitz functions:

$$
\mathbf{1}^-\left(\tau, \beta\right) = \max\left(0, \min\left(1, \frac{1}{\kappa}\left(\beta - \tau\right)\right)\right),
$$

$$
\mathbf{1}^+\left(\tau, \beta\right) = \max\left(0, \min\left(1, 1 + \frac{1}{\kappa}\left(\beta - \tau\right)\right)\right).
$$

Then, $\mathbf{1}\left(\tau \leq \beta - \kappa\right) \leq \mathbf{1}^-\left(\tau, \beta\right) \leq \mathbf{1}\left(\tau \leq \beta\right) \leq \mathbf{1}^+\left(\tau, \beta\right) \leq \mathbf{1}\left(\tau \leq \beta + \kappa\right)$.

For any any equilibirum profile $a_i = \mathbf{1}\left(\tau \leq \beta_i^a\right)$, we have

$$
V\left(a; \tau\right) \leq \frac{1}{2}\sum g_{ij}\mathbf{1}^+\left(\tau_i, \beta_i^a\right)\mathbf{1}^+\left(\tau_j, \beta_j^a\right) - \sum g_i\mathbf{1}^-\left(\tau_i, \beta_i^a\right)\tau_i.
$$

An application of probabilistic bounds (4) and (8) shows that, if N is sufficiently large, then, with a probability of at least $1 - \varepsilon$,

$$
\begin{aligned}
V\left(a^{*}; \tau\right) \geq & \frac{1}{2} \sum g_{ij} P\left(x^{*} - \kappa\right) P\left(x^{*} - \kappa\right) - \sum g_i \int_0^{x^{*}+\kappa} y dP\left(y\right) - \varepsilon \sum g_i \\
= & \sum g_i \left(\nu\left(x^{*} - \kappa\right) - \varepsilon - 2\kappa\right) \\
\geq & \sum g_i \left(\nu\left(x^{*}\right)\right) - \kappa N w^{*} g_{\min} \nu_\kappa^0 - \left(\varepsilon + 2\kappa\right) N w^{*} g_{\min},
\end{aligned}
$$

where in the last inequality, we use constants $\nu_{\cdot}^0$.

Because $\mathbb{E}\,\mathbf{1}^-\left(., b\right) \geq P\left(b - \kappa\right)$ and $\mathbb{E}\,\mathbf{1}^+\left(., b\right) \leq P\left(b + \kappa\right)$, an application of Lemma 1 shows that, for each $\varepsilon > 0$, there is $d > 0$ small enough such that if $d\left(g\right) < d$ (hence N is sufficiently large), then with probability of at least $1 - \varepsilon$, we have for each equilibrium profile a,

$$
\begin{aligned}
V\left(a; \tau\right) \leq & \frac{1}{2} \sum g_{ij} P\left(\beta_i^a + \kappa\right) P\left(\beta_j^a + \kappa\right) - \sum g_i \int_0^{b_i - \kappa} y dP\left(y\right) + \varepsilon \sum g_i \\
\leq & \frac{1}{2} \sum g_i \left(P\left(\beta_i^a + \kappa\right)\right)^2 - \sum g_i \int_0^{b_i - \kappa} y dP\left(y\right) + \varepsilon \sum g_i \\
= & \sum g_i \left(\nu\left(\beta_i^a + \kappa\right) + \varepsilon + 2\kappa\right).
\end{aligned}
$$

If an equilibrium profile $a = \mathbf{1}\left(\tau_i \leq \beta_i^a\right)$ is not an η-fuzzy convention, then we get

$$
V\left(a; \tau\right) \leq \sum g_i \left(\nu\left(x^{*}\right)\right) + N w^{*} g_{\min}\left(\varepsilon + 2\kappa\right) - \eta N g_{\min} \nu_{\frac{1}{2}\eta}^1
$$

,where we used the definition of constants $\nu_{\cdot}^1$. If κ and $\varepsilon \leq \frac{1}{2}\eta$ (and $d\left(g\right)$) are sufficiently small, $V\left(a; \tau\right) < V\left(a^{*}; \tau\right)$ with probability of at least $1 - 2\varepsilon \geq 1 - \eta$. In such a case, the potential maximizer must be an η-fuzzy convention x^{*}.

Appendix B. Proof of Theorem 2

In part B.1 of this Appendix, we formally define the city network (M, m) and also develop some of its properties. Part B.2 contains the probabilistic part of the proof: We establish the existence of a large connected component of the network that is also obstacle-free, i.e., without "bad" groups of agents. The last part elaborates on the contagion argument from the main body of the paper to conclude the proof of the Theorem.

B.1. **Lattice.** We start by formally defining the city network. For each $M \geq m$, the (M, m)-lattice is a network with

- $N = M^2$ nodes from the set $I_M = \{1, ..., M\}^2$. We define a distance on I_M by

$$d(i, j) = \frac{1}{m} \sqrt{\sum_l ((i_l - j_l) \mod M)^2},$$

  and a ball in this metric as $B(i, r) = \{y : d(x, y) \leq r\}$. The subtraction "mod$M$" turns the lattice into a subset of "discrete Euclidean torus" $\left[0, \frac{M}{m}\right]^2$,

- connections $g_{i,j} = 1 \iff j \in B(i, 1)$.

For each $i \in I_M$, and two sets $U, W \subseteq I_M$, let

$$d(i, W) = \min_{j \in W} d(i, j) \text{ and } d(U, W) = \min_{i \in U} \min_{j \in W} d(i, j). \tag{10}$$

For each set W, and each r, define the r-neighborhood of W:

$$B(W, r) = \{i : d(i, W) \leq r\} = \bigcup_{i \in W} B(i, r).$$

B.1.1. *Large m approximations.* For large m, the neighborhoods of each agent have similar properties as open balls on a Euclidean plane. This is formalized as follows. Let $B_{\mathbb{R}^2}(x, r)$ be the ball on the plane with center $x \in \mathbb{R}^2$ and radius r. Let $|A|$ be a Lebesgue measure of a measurable set $A \subseteq \mathbb{R}^2$. Let

$$f_0(d, r_1, r_2) = \frac{1}{\pi} |B_{\mathbb{R}^2}((0, 0), r_1) \cap B_{\mathbb{R}^2}((d, 0), r_2)|$$

be the mass of the intersection of two balls, with radii r_1 and r_2 respectively, separated by distance d, and normalized by the mass of the unit ball $B((0, 0), 1)$.

Lemma 3.   *(1) For each $\rho > 0$, there exists $C_\rho < \infty$ such that if $m \geq C_\rho$, then for any two agents i, j, for any $r_1 \leq 1 \leq r_2$, we have*

$$\left| \frac{|B(i, r_1) \cap B(j, r_2)|}{|B(i, 1)|} - f_0(d(i, j), r_1, r_2) \right| \leq \rho.$$

  (2) Function f_0 has the following properties:
  - *f_0 is Lipschitz over d and $r_1 \leq 1 \leq r_2$,*
  - *f_0 is decreasing in d, and*
  - *$f_0(d, r_1, r_2) = 0$ if $r_1 + r_2 \leq d$, and $f_0(d, r_1, r_2) = 1$ if $r_1 = 1$ and $d \leq r_2 - r_1$.*

  (3) Functions $f_1(x, r_1; r_2) = f_0(r_2 - x, r_1, r_2)$ for $r_1 \leq 1$ and $x \in \mathbb{R}$ converge uniformly to function $\lim_{r_2 \to \infty} f_1(x, r_1; r_2) = f_2(x, r_1)$. In particular, for each
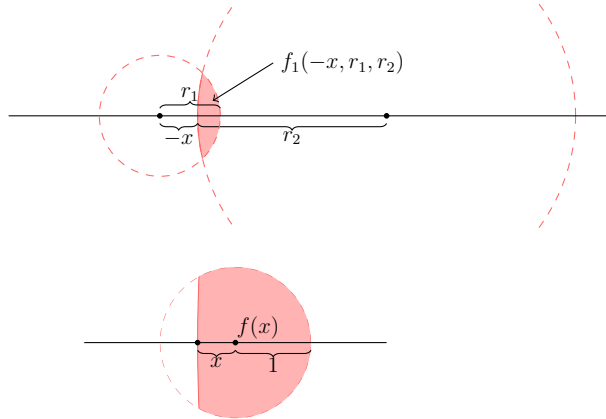
FIGURE 5. Illustrations of functions f_1 and f.

$\rho > 0$, there exists R_ρ such that, if $r_1 \leq 1$ and $r_2 \geq R_\rho$, then,

$$\sup_{r_1 \leq 1, x} |f_2(x, r_1) - f_1(x, r_1; r_2)| \leq \rho.$$

Functions f_1 and f_2 are Lipschitz over d and $r_1 \leq 1$ and increasing in x.

(4) Let $f(x) = f_2(x, 1)$. Then, $f(x) + f(-x) = 1$.

Proof. The properties of f_0, f_1, f_2, and f follow from their geometric interpretations and from the fact that the counting measure on I_M converges weakly to the Lebesgue measure on the torus. For example, $f_2(x, r_1)$ is a segment of radius r_1 ball with height equal to $r_1 + x$ for $x \in (-r_1, r_1)$. See Figure 5. $\qquad\square$

B.1.2. *Cubes.* let G be a (M, m)-lattice. We divide the lattice into disjoint areas that we refer to as *cubes*. We will assume there exists values b such that $0 \ll b \ll m$, and M is divisible by b. (This divisibility assumption simplifies the proof. The theorem remains valid without it, but the proof requires small modifications due to the existence of non-zero reminders from the division by b. We omit the details.) Each cube has b^2 elements and, because $b \ll m$, it is much smaller than the diameter of the neighborhood of each node so that the neighborhoods of nodes in the same cube are largely overlapping. At the same time, each cube contains a sufficiently large number of nodes so that the distribution of thresholds within the cube can be probabilistically approximated by its expected distribution.

Formally, for each real number x, let $\lfloor x \rfloor$ be the largest integer no larger than x. For each node i, the set of nodes

$$c^b(i) = \left\{ j \in \{1, ..., M\}^2 : \forall_l \lfloor i_l/b \rfloor = \lfloor j_l/b \rfloor \right\}$$

is referred to as a cube that contains i. Any two cubes are either disjoint or identical. Each cube c is uniquely identified by a pair of numbers $c_l = \lfloor i_l/b \rfloor$ for each $l = 1, 2$ and any $i \in c$. Due to the divisibility assumption, there are $\left(\frac{M}{b} \right)^2$ cubes on the (M, m)-lattice.

Let $\mathcal{G}^b = \left\{ c^b(i) : i \in G \right\}$ be the set of all cubes. We refer to the elements of $\mathcal{G}^b$ as *cubes*. The network of cubes $\mathcal{G}^b$ consists of cubes as vertices and edges between any two cubes that share one of their sides: for any $c, c' \in \mathcal{G}^b$, $g_{c,c'}^b = 1$ iff for some $l = 1, 2$, $c_l = c_l'$ and $\left| (c_{-l} - c_{-l}') \mod \frac{M}{b} \right| = 1$. Thus, each cube shares an edge with four other cubes.

Say that set $S \subseteq \mathcal{G}^b$ is *r-connected* if for any subset $A \subseteq S, A \neq S$, there is $c \in A$, $c' \in S \backslash A$, and at most an r-element path between c and c'. (A path is a tuple of cubes connected by the edges of the cube network.) S is connected if it is 1-connected.

For any two cubes, define a distance $d^b(c, c') = \max_l \left| (c_l - c_l') \mod \frac{M}{b} \right|$. For any $S, S' \subseteq \mathcal{G}^b$, let $d^b(S, S') = \min_{c \in S, c' \in S'} d^b(c, c')$ be the distance between two sets of cubes. Let $U(c, r) = \{ c' : d(c, c') \leq r \}$ be the r-neighborhood of c. Thus, each cube has 8 other cubes in its 1-neighborhood.

B.2. **Probabilistic part.** We will show that if the lattice is sufficiently large then, with arbitrarily high probability, we can find a set W of cubes that (a) contains almost all cubes and (b) is connected in the cube network, where (c) each cube in the set is far away from bad cubes, and (d) contains a large set of agents for whom action 0 is dominant. Properties (b)-(c) will allow the contagion wave to spread across the entire set W, property (a) will ensure that spreading to set W means spreading almost everywhere, and property (d) will ensure that the set contains sufficiently many "initial infectors" to start the contagion wave.

For each realization of threshold profile τ, define the empirical cdf of best response thresholds in cube $c \in \mathcal{G}^b$:

$$P_c(x|\tau) = \frac{1}{|c|} \sum_{i \in c} \mathbf{1} \{ \tau_i < x \}.$$

For $\gamma > 0$, say that a cube c is γ-bad if there exists x such that $P_c(x|\tau) > P(x) + \gamma$; otherwise, the cube is γ-good.

Agent x is *extraordinary* if action 0 is strictly dominant for such an agent. A cube $c \in \mathcal{G}^b$ is *extraordinary* if it only consists of extraordinary agents. In any equilibrium, $a(c) = 0$ for extraordinary cube c. Clearly, an extraordinary cube is γ-good for each $\gamma \geq 0$.

Say that set $W \subseteq \mathcal{G}^b$ of cubes is (γ, R)-*good* if

(a) W contains at least a fraction $(1 - \gamma)$ of cubes, $|W| \geq (1 - \gamma)\left|\mathcal{G}^b\right|$,

(b) W is connected as a subset of the cube network,

(c) if $c \in \mathcal{G}^b$ is γ-bad, then $d^b(c, c') > 3R$ for each $c' \in W$ (in particular, each cube in W is γ-good), and

(d) W contains a cube c_0 such that each cube c s.t. $d(c, c_0) \leq R$ is extraordinary.

We show that large good sets of cubes exist with high probability:

Lemma 4. *For each $\gamma, \rho > 0$, and $R < \infty$, there exist $m_{\gamma,\rho,R} > 0$, and for each $m > m_{\gamma,\rho,R}$, there exists $M_{\gamma,\rho,R}(m)$ such that, if $m \geq m_{\gamma,\rho,R}$ and $M \geq M_{\gamma,\rho,R}(m)$, then, if G is an (M, m)-lattice, $b = \lfloor \rho m \rfloor$, and $\mathcal{G}^b$ is the associated cube network, then*

$$\mathbb{P}\left(\text{there exists } (\gamma, R)\text{-good set } W \subseteq \mathcal{G}^b\right) \geq 1 - \gamma.$$

B.2.1. *Intermediate results.* We need two intermediate results. The first result provides a bound on the size of the largest connected component of the graph obtained from the network of cubes after removing a group of smaller and connected sets of cubes.

Lemma 5. *Suppose that $\{S_1, ..., S_J\}$ is a collection of connected subsets of $\mathcal{G}^b$ such that $S_i \cup S_j$ are not 2-connected for any $i \neq j$. Then, there is a connected subset $V \subseteq \mathcal{G}^b \backslash \bigcup S_j$ such that $\left|\mathcal{G}^b \backslash V\right| \leq \sum_j |S_j|^2$.*

Proof. First, observe that for each connected set S such that $|S|^2 < \left|\mathcal{G}^b\right|$, there is a set S' and a loop (i.e., a path with the same beginning and ending) $c_0^S, ..., c_n^S = c_0$ of cubes $c_l^S \notin S'$ such that

- $S' \supseteq S$ and $|S'| \leq |S|^2$, and
- loop $c_0^S, ..., c_n^S$ tightly surrounds set S' and separates it from the rest of the graph: $|\{c : d(c, S') = 1\}| \subseteq \left\{c_l^S\right\} \subseteq |\{c : d(c, S') \leq 2\}|$.

This observation follows from the Jordan Curve Theorem and from the fact that each connected set S such that $|S|^2 < |\mathcal{G}^b|$ can be contained in a $|S'^2|$-element "square" of cubes such that the set outside the square is connected.

For each set S_i from the hypothesis of the Lemma, find loop c^i and set S_i' as in the observation above. We will show that set $\mathcal{G}^b \backslash \bigcup S_j'$ is connected, which will conclude the proof of the Lemma. Take any two cubes $c, c' \in \mathcal{G}^b \backslash \bigcup S_j'$ and an arbitrary path $c = c_0, ..., c_n = c'$ between them. We will modify this path so that it avoids each set S_i. For each i, either the existing path avoids set S_i', or it intersects it. Find $l_0^i = \min\{l : d(c_l, S_i) = 1\}$ and $l_1^i = \max\{l : d(c_l, S_i) = 1\}$. Then, replace the interval $c_{l_0^i}, ..., c_{l_1^i}$ of the path with the path from $c_{l_0^i}$ to $c_{l_0^i}$ along path c^i. The new path avoids set S_i'. Because the modified part of the path stays within 2-distance of set S_i', the modification does not create new intersections with other sets S_j'. After possibly modifying the path for any i, we obtain a path between c and c' that avoids each set S_i'. Thus, set $\mathcal{G}^b \backslash \bigcup S_j'$ is connected. $\qquad\square$

The second result provides an upper bound on the number of different r-connected sets of cubes.

Lemma 6. *The number of r-connected sets in $\mathcal{G}_b$ of cardinality n is not larger than* $2^{2n}(2r+1)^n |\mathcal{G}_b|$.

Proof. We first find an encoding for each r-connected tuple. Let m_r be the size of the r-neighborhood of an element of $\mathcal{G}_b$. Then, $m_r \leq (2r+1)^2$. Consider tuples $(s_1, (l_2, ..., l_n), (k_2, ..., k_n))$ such that $s_1 \in \mathcal{G}_b$, $k_i \in \{1, .., m_r\}$, and $l_i \leq i$ and $l_i \leq l_j$ for each $2 \leq i \leq j$.

We show that each r-connected set can be encoded as one of the above tuples in such a way that any two different r-connected sets must have a different encoding. Let $e : \mathcal{G}^b \to \{1, ..., |\mathcal{G}^b|\}$ be an enumeration of set $\mathcal{G}^b$. For each $s \in \mathcal{G}^\llcorner$, let $e_s : \{s' : d(s, s') = 1\} \to \{1, ..., 4\}$ be the enumeration of the immediate neighborhood of s that has the same ranking in the neighborhood as enumeration e. Choose $s_1 = \arg\min_{s \in S} e(s)$. Suppose that $s_1, ...s_{i-1}$ are chosen for $1 < i < n$. For each $x \in S \backslash \{s_1, ..., s_{i-1}\}$, let $l(x) = \min_{d(x, s_l) = 1} l$ and let it equal ∞ if the set is empty. Then, $l(x) < i$ for at least one x. Let $k(x) = e_{s_{l(x)}}(x)$. Choose $s_i = \arg\min_{\text{lexicograpically}, x \in S}(l(x), k(x))$, so as to minimize lexicographically $(l(x), k(x))$ among all $x \in S \backslash \{s_1, ..., s_{i-1}\}$. Let $l_i = l(s_i)$ and $k_i = k(s_i)$.

We derive an upper bound on the number of encoding tuples. Say that a sequence $l_i, ..., l_n$ is (i, m)-sequence if it is increasing, $l_j < j$ for each j, and $l_i = i - m - 1$. Let $S(i, m)$ denote the number of different (i, m)-sequences. It is easy to see that

$$S(i, m) = \sum_{p=0}^{m+1} S(i + 1, p),$$

where $S(n, m) = 1$. We check by induction on i that $S(i, n) \leq 2^{2(n-i)+m}$.

The number of choices for s_1 is not larger than $|\mathcal{G}_b|$. By the above, the number of $(2, 0)$-sequences is not larger than $2^{2(n-2)}$. The number of choices of $k_2, ..., k_n$ is not larger than $(2r + 1)^{n-1}$. It follows that the total number of encodings, and hence the number of connected sets, is not larger than $2^{2n} (2r + 1)^n |\mathcal{G}_b|$. $\qquad\square$

B.2.2. *Proof of Lemma 4.* Lemma 4 follows from the following two results. The first result establishes the existence of a large connected component that is far from bad cubes. Let $B_\gamma = \left\{ c \in \mathcal{G}^b : c \text{ is } \gamma\text{-bad} \right\}$ be the (random) set of γ-bad cubes.

Lemma 7. *For each $\gamma > 0$ and $R < \infty$, there exists $b_{\gamma,R} > 0$ such that if $b > b_{\gamma,R}$, then*

$$\mathbb{P}\left(\exists W^0 \subseteq \mathcal{G}^b, \text{ st. } W^0 \text{ is connected, } \left|W^0\right| \geq (1 - \gamma)\left|\mathcal{G}^b\right|, \ d^b\left(W^0, B_\gamma\right) \geq 5R \right) \geq 1 - \frac{1}{4}\gamma.$$

Proof. Let $p_\gamma > 0$ be the probability that a cube is γ-bad. Due to the Dvoretzky–Kiefer–Wolfowitz–Massart inequality, the probability that a cube c is γ-bad is bounded by

$$p_\gamma \leq C\mathrm{e}^{-2b^2\gamma^2}$$

for some universal constant C.

Let $S_1^0, ..., S_n^0$ be the smallest division of the set of bad cubes $B_\gamma = \bigcup S_i^0$ into sets that are $11R$-connected and such that $S_i^0 \cup S_j^0$ are not $11R$-connected for $i \neq j$. Let $X = \sum |S_i^0|^2$. We compute the expected value of X. Let $m_n = \left(2^{2n} (11R + 1)^n |\mathcal{G}_b|\right)$ be an upper bound on the cardinality of all $11R$-connected sets (obtained from Lemma 6). Then,

$$\mathbb{E}\, X \leq \sum_{n \geq 1} n^2 m_n p_\gamma^n \leq |\mathcal{G}_b| \sum_{n \geq 1} 2^n 2^{2n} (6R + 1)^n p_\gamma^n$$

$$= |\mathcal{G}_b| \frac{8(11R + 1)\, p_\gamma}{1 - 8(11R + 1)\, p_\gamma}.$$

Let $S_i^1 \supseteq S_i^0$ be the smallest connected set such that sets $S_i^1 \cup S_j^1$ are not $11R$-connected for $i \neq j$ and such that $|S_i^1| \leq 11R |S_i^0|$. Such sets can be constructed by connecting elements of S_i^0 by a path inside the intersection of the $11R$-neighborhood of the two sets.

Let S_i be the $5R$-neighborhood of set S_i^1. Clearly, sets S_i are disjoint (and separated by R). Because each $5R$-neighborhood of an element of a set S_i^1 has no more than $(11R+1)^2 |S_i^1|$ cubes, the cardinality of S_i is at most $(11R+1)^2 |S_i^1| \leq (11R+1)^3 |S_i^0|$.

Let W^0 be the largest connected component of $\mathcal{G}^b$ that does not contain elements of sets S_i. By construction, each set S_i is connected, but sets $S_i \cup S_j$ are not 2-connected. By Lemma 5, the cardinality of W^0 is at least $|\mathcal{G}^b| - 4(11R+1)^6 X$. By the Markov's inequality,

$$\mathbb{P}\left(|W^0| \geq (1-\gamma)|\mathcal{G}^b|\right) \leq \mathbb{P}\left(4(11R+1)^6 X \leq \gamma |\mathcal{G}^b|\right)$$
$$\leq \frac{4(11R+1)^6 \mathbb{E} X}{\gamma |\mathcal{G}^b|} \leq \frac{1}{\gamma} \frac{32(11R+1)^7 p_\gamma}{1 - 8(11R+1)p_\gamma}.$$

Assume that $b_{\gamma,R} > 0$ is large enough so that for each $b > b_{\gamma,R}$, $\frac{1}{\gamma} \frac{32(11R+1)^7 C\mathrm{e}^{-2b^2\gamma^2}}{1 - 8(11R+1)C\mathrm{e}^{-2b^2\gamma^2}} \leq \frac{1}{4}\gamma$. $\qquad\square$

Say that cube $c \in G_R$ is an extraordinary center if all cubes in $U(c, R)$ are extraordinary.

Lemma 8. *There exists $K_{\gamma,R} < 0$ large enough so that if $\frac{M}{b} > K_{\gamma,R}$, then*

$$\mathbb{P}\left(\exists W \subseteq \mathcal{G}^b,\ st.\ \begin{array}{c} W \supseteq W^0,\ W\ \text{is connected},\ d^b(W, B_\gamma) \geq 3R \\ \text{and } W\ \text{contains an extraordinary center} \end{array}\right) \geq 1 - \gamma,$$

where W_0 inside the probability satisfied the conditions from Lemma 7.

Proof. Recall that $K = \frac{M}{b}$ is the number of cubes. If K is divisible by $(2R+1)$, we can find a grid of cubes $G_R \subseteq \mathcal{G}^b$ such that any two $c, c' \in G$, $d(c, c') = 2R$ and $\mathcal{G}^b = \bigcup_{c \in G_R} U(c, R)$. Because the $U(c, R)$ neighborhoods are disjoint, $|\mathcal{G}^b| = |G_R|(2R+1)^2$, where $(2R+1)^2$ is the size of each neighborhood. For simplicity, the rest of the arguments rely on the divisibility assumption. The argument is easily modified for the case when the divisibility does not hold (and b and $\frac{M}{b}$ are sufficiently large).

Let W^0 be the (random) set from Lemma 7. Let $W^1 = \bigcup_c U(c, R+1)$ and $W = \bigcup_c U(c, 2R+1)$. Then, $d(W, B_\gamma) > 2R$. Because for each $c' \in U(c, r)$ there is a path between c and c' that is inside set $U(c, r)$, W is connected.

We show that $|G_R \cap W^1| \geq (1-\gamma)|G_R|$. On the contrary, suppose that $|G_R \backslash W^1| > \gamma \, |G_R|$. Then, $A = \bigcup_{c \in G_R \backslash W} U(c, R) \subseteq \mathcal{G}^b \backslash W^0$. Moreover, $|A| > \gamma|G_R|(2R+1)^2 = \gamma \left|\mathcal{G}^b\right|$. However, this contradicts $\left|\mathcal{G}^b \backslash W^0\right| \leq \gamma \left|\mathcal{G}^b\right|$.

Let $q > 0$ be the probability that a cube c is an extraordinary center. Then, $q \geq P(0)^{(2R+1)^2 b^2}$. Let q^* be the probability that cube c is an extraordinary center, conditional on $c \in W^1$. Because being in $c \in W^1$ provides no other information about the distribution of taste shocks apart from c is not γ-bad and γ-bad cubes are not extraordinary, it must be that $q^* \geq q$. Similarly, conditional on $c, c' \in W^1$, if c and c' are separated by $2R+1$, the events that the two are extraordinary centers are independent. Hence, the probability that none of the cubes in $c \in G_R \cap W_1$ is an extraordinary center is at most

$$(1 - q^*)^{|G_R \cap W^1|} \leq \left(1 - P(0)^{(2R+1)^2 b^2}\right)^{(1-\gamma)K^2(2R+1)^{-2}}$$

$$\leq e^{-(1-\gamma)K_{\gamma,R}(2R+1)^{-2}P(0)^{(2R+1)^2 b^2}}.$$

If K is sufficiently large, the above is smaller than $\frac{1}{4}\gamma$. $\qquad\square$

To conclude the proof of the Lemma, we set $m_{\gamma,\rho,R} > \frac{1}{\rho} b_{\gamma,R}$ and then $M_{\gamma,\rho,R}(m) \geq \rho m K_{\gamma,R}$.

B.3. **Proof of Theorem 2.** Below, we will show the following Lemma.

Lemma 9. *For each $\varepsilon > 0$, there exists sufficiently small $\gamma, \rho > 0$ and sufficiently large $R > 0$ so that if $b = \lfloor \rho m \rfloor$, W is a (γ, R)-good set in the network of cubes $\mathcal{G}^b$, and a is an equilibrium profile, then for each $i \in c \in W$, $\beta_i^a \leq x^* + \varepsilon$.*

Together with Lemma 4, Lemma 9 shows that for each $\varepsilon > 0$, if m and $\frac{M}{m}$ are sufficiently large, with probability of at least $1 - \varepsilon$, if a is an equilibrium profile, then all but a ε-fraction of the population (i.e., all members of the "good" set W), $\beta_i^a \leq x^* + \varepsilon$.

A similar argument shows that $\beta_i^a \geq x^* - \varepsilon$ for elements of an analogously defined "good" set (with the appropriate modification of what good and extraordinary cubes

are). Together, the two arguments show that, with probability of at least $1 - 2\varepsilon$, a is a 2ε-fuzzy convention of x^*. Take $\varepsilon = \frac{1}{2}\eta$.

Proof. We divide the proof of the Lemma into two steps.

Preparation. Find $\varepsilon_0 > 0$, such that

$$\sigma^* = \max_{a \geq x^* + \frac{\varepsilon}{2}} \int_{x^* + \varepsilon_0}^{a} \left(P^{-1}(y) - y\right) dy > 0.$$

The existence of such $\varepsilon_0 \in \left(0, \frac{\varepsilon}{2}\right)$ comes from the definition of x^* as the unique maximizer of $\int_{x^*}^{a} (y - P^{-1}(y)) \, dy$. Let δ_ρ be a fraction of neighbors of i who are not members of a cube that is fully contained in the neighborhood of i. It is easy to see that $\delta_\rho \to 0$ as $\rho \to 0$.

Let a be an equilibrium profile. For each cube c, define

$$a_c = \frac{1}{|c|} \sum_{j \in c} a_j \text{ and } \beta_c = \frac{1}{|c|} \sum_{j \in c} \beta_j^a.$$

Then, $|\beta_c - \beta_i^a| \leq \delta_\rho$, and

$$\beta_c \leq \delta_\rho + \frac{|c|}{|B(i,1)|} \sum_{c \subseteq B(i,1)} a_c. \tag{11}$$

If cube c is γ-good, then

$$a_c = \frac{1}{|c|} \sum_{i \in c} \mathbf{1}\{\tau_i < \beta_i^a\} \leq \frac{1}{|c|} \sum_{i \in c} \mathbf{1}\{\tau_i < \beta_c + \delta_\rho\} \leq P(\beta_c + \delta_\rho) + \gamma. \tag{12}$$

From now on, assume that $W \subseteq \mathcal{G}^b$ is (γ, R)-good. If $d^b(c, W) \leq 3R$, then cube c is γ-good.

Define

$$C_0 = \{c : \forall c' \ d(c, c') \leq R \implies a_c \leq x^* + \varepsilon_0\}.$$

For each $i \in C_0$, the average behavior in all the cubes fully contained in the neighborhood of i is $\leq x^* + \varepsilon_0$, which, together with (11), implies that

$$\beta_i^a \leq (x^* + \varepsilon_0)(1 - \delta_\rho) + \delta_\rho \leq x^* + \varepsilon.$$

The last inequality holds when ρ is sufficiently small so that $\delta_\rho \leq \frac{\varepsilon}{2}$. Hence, to establish our claim, it is enough to show that $W \subseteq C_0$.

Notice that C_0 cannot be empty as it contains at least one extraordinary cube. For each $a > x^* + \frac{\varepsilon}{2}$, define

$$d(a) = \min_{c \in W : a_c \geq a} d^b(c, C_0) \geq R,$$

where the value is ∞ if the set over which the distance is minimized is empty.

On the contrary to our claim, suppose that there is a cube $c \in W_0$ such that $a_c > a > x^* + \frac{\varepsilon}{2}$. Then, there exists $a > x^* + \frac{\varepsilon}{2}$ such that $d(a) < \infty$. Find $a^* \geq x^* + \varepsilon_0$ such that $d(a^*) \leq 2R$ and $d\left(a^* + \frac{1}{R}\right) \geq d(a^*) + 1$. Such a^* exists: otherwise, if for each a such that $d(a) \leq 2R$, $d\left(a + \frac{1}{R}\right) \leq d(a) + 1$, then $d(a+1) \leq 2R$, which is impossible (as there is no cube with the action average strictly larger than 1).

Contagion wave. Notice that a_c takes discrete values $a \in A = \left\{0, \frac{1}{|c|}, ..., 1\right\}$, where $|c|$ is the size of a cube. Let $a_k = \frac{k}{|c|}$ be the enumeration of set $A \cap \left\{a : a \geq x^* + \frac{\varepsilon}{2}\right\}$. For each such cube c, and each $i \in c$, (11) implies

$$
\begin{aligned}
\beta_c \leq & \delta_\rho + \frac{|c|}{|B(i,1)|} \sum_{c \subseteq B(i,1)} a_c \\
\leq & \delta_\rho + \sum_{a \in A} a \frac{|\{c \subseteq B(i,1) : a_c = a\}|}{|B(i,1)|/c} \\
\leq & \delta_\rho + x^* + \varepsilon_0 + \sum_k (a_{k+1} - a_k) \frac{|\{c \subseteq B(i,1) : a_c \geq a\}|}{|B(i,1)|/c} \\
\leq & \delta_\rho + \delta_{R,\rho} + x^* + \varepsilon_0 + \sum_k (a_{k+1} - a_k) \left(1 - f\left(d(a_k) - d^b(c, C_0)\right)\right),
\end{aligned}
$$

where the third inequality is a consequence of a discrete version of the integration by parts (i.e., $\sum x_i(y_i - y_{i+1}) = \sum (x_{i+1} - x_i) y_{i+1}$), and the fourth one is due to Lemma 3, where $\delta_{R,\rho} \to 0$ as R is sufficiently large and ρ is sufficiently small. Let $\delta^1_{R,\rho} = \delta_\rho + \delta_{R,\rho}$.

Additionally, for each $a_l \in A$, $a_l \leq a^*$, find a cube c such that $d^b(c, C_0) = d_R(a_l) < 2R$ and $a_c \geq a_l$. Using the above inequality and (12), we obtain

$$
\begin{aligned}
P^{-1}(a_l - \gamma) \leq P^{-1}(a_c - \gamma) \leq & \beta_c + \delta_\rho \\
\leq & \delta^1_{R,\rho} + x^* + \varepsilon_0 + \sum_k (a_{k+1} - a_k)(1 - f(d(a_k) - d(a_l))).
\end{aligned}
$$

Let $k^* = \max\{k : a_k \leq a^*\}$. Then, the right-hand side is not larger than

$$\leq \delta^1_{R,\rho} + x^* + \varepsilon_0 + \sum_{k \leq k^*} (a_{k+1} - a_k)(1 - f(d(a_k) - d(a_l)))$$

$$+ \sum_{k > k^* : a_k \leq a^* + \frac{\varepsilon}{10}} (a_{k+1} - a_k)(1 - f(d(a_k) - d(a_l)))$$

$$+ \sum_{k : a_k > a^* + \frac{\varepsilon}{10}} (a_{k+1} - a_k)(1 - f(d(a_k) - d(a_l)))$$

$$\leq \delta^1_{R,\rho} + x^* + \varepsilon_0 + \frac{1}{R} + \sum_{k \leq k^*} (a_{k+1} - a_k)(1 - f(d(a_k) - d(a_l))),$$

due to the second term in the first line being not larger than $\frac{\varepsilon}{10}$, and the third term being equal to 0 (as $f(d(a_k) - d(a_l)) \geq f(1) = 1$).

Let $\Delta = a^* - (x^* + \varepsilon_0)$. Multiplying by $(a_{l+1} - a_l)$ and summing across $l \leq k^*$, we obtain

$$\sum_{l \leq K^*} P^{-1}(a_l - \gamma)(a_{l+1} - a_l)$$

$$\leq \left(\delta^1_{R,\rho + \frac{1}{R}} + x^*\right)\Delta + \sum_{l \leq K^*}\sum_{k \leq K^*} (a_{k+1} - a_k)(a_{l+1} - a_l)(1 - f(d_R(a_l) - d_R(a_k)))$$

$$= \left(\delta^1_{R,\rho} + \frac{1}{R} + x^*\right)\Delta + \frac{1}{2}\sum_{l,k \leq K^*} (a_{k+1} - a_k)(a_{l+1} - a_l) = \left(\delta^1_{R,\rho} + \frac{1}{R} + x^* + \varepsilon_0\right)\Delta + \frac{1}{2}\Delta^2$$

$$\leq \delta^1_{R,\rho} + \frac{1}{R} + \int_{x^* + \varepsilon_0}^{a^*} y \, dy.$$

To obtain the equality, we use the fact that f is balanced.

Because $P^{-1}(.-\gamma) \in [0,1]$ and $a_{l+1} - a_l = \frac{1}{|c|}$, the left-hand side of the above inequality is smaller than

$$\int_{x^* + \varepsilon_0}^{a^*} P^{-1}\left(y - \gamma - \frac{1}{|c|}\right) dy \geq \int_{x^* + \varepsilon_0 - \gamma - \frac{1}{|c|}}^{a^* - \gamma - \frac{1}{|c|}} P^{-1}(y) \, dy$$

Assuming that b is large enough so that $\frac{1}{|c|} \leq \gamma$, the above is not smaller than $\int_{x^*+\varepsilon_0}^{a^*} \left(P^{-1}(y) - y\right) dy - 2\gamma$. Putting it back into the main inequality, we obtain

$$\int_{x^*+\varepsilon_0}^{a^*} \left(P^{-1}(y) - y\right) dy \leq \delta_{R,\rho}^1 + \frac{1}{R} + 2\gamma.$$

If $\gamma, \rho > 0$ are sufficiently small and R sufficiently large, $\delta_{R,\rho}^1 + \frac{1}{R} + 2\gamma < \sigma^*$. The contradiction shows that $W \subsetneq C_0$, which concludes the proof of the Lemma. □

Appendix C. Proof of Theorem 3

For each $\eta > 0$, define $P_\eta = P(x : |x - x^*| \leq \eta)$ as the probability that the threshold realization is within η of x^*. If P does not have an atom at x^*, then, we can choose η_δ such that $P_{\eta_\delta} \leq \frac{1}{30}\delta$. Assume w.l.o.g. that $\eta_\delta \leq \delta$. Let

$$T_\delta = \left\{ \tau : \frac{1}{N} |\{\tau_i : |\tau_i - x^*| \leq \eta_\delta\}| \leq \frac{1}{3}\delta \right\}.$$

The Law of Large Numbers implies that for sufficiently high N, $\text{Prob}(T_\delta) \geq 1 - \delta$.

Fix threshold profile $\tau \in T_\delta$. Let $I_0 = \{i : |\tau_i - x^*| \leq \eta_\delta\}$. Suppose that a is $\frac{1}{3}\eta_\delta$-fuzzy convention x^*. Let $I(g) = \left\{ i : |\beta_i^a - x^*| > \frac{1}{3}\eta_\delta \right\}$ be the set of agents that is an equilibrium in game $G(g, \tau)$. Let $I = I_0 \cap I(g)$. Then, $\frac{1}{N}|I| \leq \frac{2}{3}\delta$. For each $i \notin I$, either

- $\tau_i > x^* + \eta_\delta$ and $\beta_i^a \leq x^* + \frac{1}{3}\eta_\delta$, which implies $a_i = a_i^* = 0$, or
  - $\tau_i < x^* - \eta_\delta$ and $\beta_i^a \geq x^* - \frac{1}{3}\eta_\delta$, which implies $a_i = a_i^* = 1$.

Hence, for any $i \notin I$, $a_i = a_i^*$. This concludes the proof of the Theorem.

References

Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and economic behavior*, **5** (3), 387–424, publisher: Elsevier.

— (1995). The statistical mechanics of best-response strategy revision. *Games and economic behavior*, **11** (2), 111–145, publisher: Elsevier.

— (2018). Population games. In *The economy as an evolving complex system II*, CRC Press, pp. 425–460.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, pp. 1047–1071, publisher: JSTOR.

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

Erdős, P. and Rényi, A. (1959). On Random Graphs. I. *Publicationes Mathematicae*, **6**, 290–297.

Friedman, D. (1991). Evolutionary games in economics. *Econometrica: Journal of the Econometric Society*, pp. 637–666.

Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F. and Yariv, L. (2010). Network games. *The review of economic studies*, **77** (1), 218–244, publisher: Wiley-Blackwell.

Granovetter, M. (1978). Threshold models of collective behavior. *American journal of sociology*, **83** (6), 1420–1443.

Harsanyi, J. C. and Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, Mass. and London: MIT Press.

Jackson, M. O. and Yariv, L. (2007). Diffusion of behavior and equilibrium properties in network games. *American Economic Review*, **97** (2), 92–98.

— and Zenou, Y. (2015). Games on networks. In *Handbook of game theory with economic applications*, vol. 4, Elsevier, pp. 95–163.

Kajii, A. and Morris, S. (1997). The Robustness of Equilibria to Incomplete Information. *Econometrica*, **65** (6), 1283–1309.

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

Lee, I. H. and Valentinyi, A. (2000). Noisy contagion without mutation. *The Review of Economic Studies*, **67** (1), 47–56, publisher: Wiley-Blackwell.

López-Pintado, D. (2008). Diffusion in complex social networks. *Games and Economic Behavior*, **62** (2), 573–590, publisher: Elsevier.

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

Morris, S. (2000). Contagion. *Review of Economic Studies*, **67** (1), 57–78.

— and Shin, H. S. (2006). Heterogeneity and Uniqueness in Interaction. *The Economy As an Evolving Complex System, III: Current Perspectives and Future Directions*, **3**, 207, publisher: Oxford University Press.

NEARY, P. R. (2012). Competing conventions. *Games and Economic Behavior*, **76** (1), 301–328.

— and NEWTON, J. (2017). Heterogeneity in preferences and behavior in threshold models. *Available at SSRN 3035289*.

NEWTON, J. (2021). Conventions under heterogeneous behavioural rules. *The Review of Economic Studies*, **88** (4), 2094–2118.

PESKI, M. (2010). Generalized risk-dominance and asymmetric dynamics. *Journal of Economic Theory*, **145** (1), 216–248, publisher: Elsevier.

UI, T. (2001). Robust equilibria of potential games. *Econometrica*, **69** (5), 1373–1380.

VERSHYNIN, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge: Cambridge University Press.

WATTS, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, **99** (9), 5766–5771, publisher: National Acad Sciences.

YOUNG, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pp. 57–84, publisher: JSTOR.