

*Fondazione Eni Enrico Mattei*  
*Working Papers*  
Fondazione Eni Enrico Mattei

---

*Year 2009*

*Paper 294*

---

Optimal Equilibria of the Best Shot  
Game

Paolo Pin\*      Luca Dall'Asta†  
Abolfazl Ramezani‡

\*Università degli Studi di Siena

†The Abdus Salam International Centre for Theoretical Physics

‡Politecnico di Torino

This working paper site is hosted by The Berkeley Electronic Press (bepress).

<http://www.bepress.com/feem/paper294>

Copyright ©2009 by the authors.

# Optimal Equilibria of the Best Shot Game

## **Abstract**

We consider any network environment in which the “best shot game” is played. This is the case where the possible actions are only two for every node (0 and 1), and the best response for a node is 1 if and only if all her neighbors play 0. A natural application of the model is one in which the action 1 is the purchase of a good, which is locally a public good, in the sense that it will be available also to neighbors. This game will typically exhibit a great multiplicity of equilibria. Imagine a social planner whose scope is to find an optimal equilibrium, i.e. one in which the number of nodes playing 1 is minimal. To find such an equilibrium is a very hard task for any non-trivial network architecture. We propose an implementable mechanism that, in the limit of infinite time, reaches an optimal equilibrium, even if this equilibrium and even the network structure is unknown to the social planner.

# Optimal equilibria of the best shot game

Luca Dall'Asta\*    Paolo Pin<sup>†</sup>    Abolfazl Ramezanzpour<sup>‡</sup>

February 17, 2009

## Abstract

We consider any network environment in which the “best shot game” is played. This is the case where the possible actions are only two for every node (0 and 1), and the best response for a node is 1 if and only if all her neighbors play 0. A natural application of the model is one in which the action 1 is the purchase of a good, which is locally a public good, in the sense that it will be available also to neighbors. This game will typically exhibit a great multiplicity of equilibria. Imagine a social planner whose scope is to find an optimal equilibrium, i.e. one in which the number of nodes playing 1 is minimal. To find such an equilibrium is a very hard task for any non-trivial network architecture. We propose an implementable mechanism that, in the limit of infinite time, reaches an optimal equilibrium, even if this equilibrium and even the network structure is unknown to the social planner.

**JEL classification code:** C61, C63, D85, H41.

**Keywords:** networks, best shot game, simulated annealing.

---

\*The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34014 Trieste, Italy

<sup>†</sup>Dipartimento di Economia Politica, Università degli Studi di Siena, Piazza San Francesco 7, 53100 Siena, Italy – and Max Weber Programme, European University Institute, Via Delle Fontanelle 10, 50014 San Domenico di Fiesole (FI), Italy

<sup>‡</sup>Dipartimento di Fisica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

# 1 Introduction.

Take an exogenous network in which otherwise homogeneous players (nodes) play a public good game, which is the one defined *Best shot game* in Galeotti et alii (2008).<sup>1</sup> The best shot game is a particular case, with restricted strategy profiles, of the model in Bramoullé and Kranton (2007) and of the second stage of the game in Galeotti and Goyal (2008), where the action of each node  $i$  is an effort  $x_i$  and her payoff will depend on the aggregate effort of herself and that of her neighbors, minus some cost for her own effort.

In particular, we will restrict strategy profiles to the two specialized actions:  $x_i \in \{0, 1\}$ .<sup>2</sup> In this way  $\vec{x}$ , a vector of specialized actions whose length is given by the number of nodes, will characterize any possible configuration of the system. We will consider the class of incentives such that, in Nash equilibrium (NE), agent  $i$  will play action  $x_i$  according to the following rule:

$$\begin{cases} x_i = 1 & \text{if } x_j = 0 \text{ for any neighbor } j \text{ of node } i; \\ x_i = 0 & \text{otherwise.} \end{cases} \quad (1)$$

We will study all the NE of the game; that is all those action profiles in which, for any link, not both nodes of the link put in effort 1; but at the same time for any node, if we consider the set including itself and its neighborhood, at least one node in this set puts in effort 1. Mathematically, the subset of nodes playing 1 in a NE will then be a *maximal independent set* of the network, as it is called in graph theory.

The next example will give some insight on the maximal independent sets, our NE, for simple networks.

**Example 1** *A network of 9 nodes.*

Figure 1 shows four possible NE for the same network of 9 nodes. Red nodes are those playing 1, while all the others are playing 0. The bottom-right NE is the only one in which only three nodes play action 1. If we assume action 1 to be a costly action, interpreting it as the purchase of a local public good, then the bottom-right NE is socially optimal, at least regarding costs.  $\square$

---

<sup>1</sup>Example 2, page 13, of the July 2008 version of the working paper. The name *Best shot game* comes from Hirschleifer (1983), where it is however described as a non-network game.

<sup>2</sup>One result in Bramoullé and Kranton (2007) is actually that, even when the possible actions of nodes are continuous, in equilibrium every agents would play either 0 or a fixed value  $e^* > 0$  which can be normalized to 1.

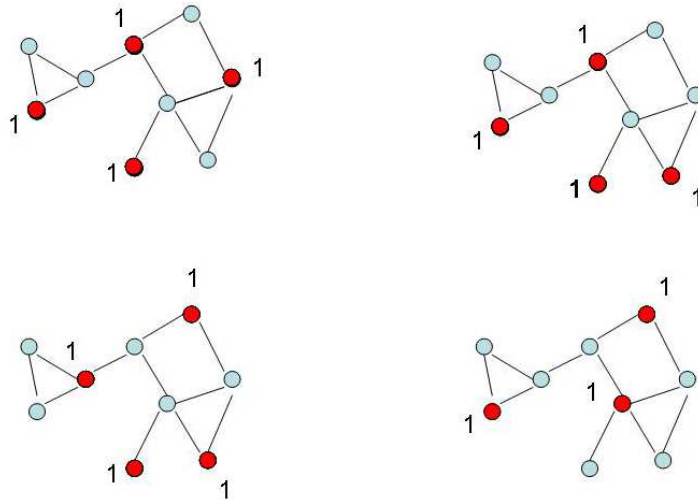


Figure 1: Four NE for a 9–nodes network.

By considering this last example, a first intuition is that when more connected nodes play 1, then the number of 1–players in equilibrium is reduced. The extremal case of this will happen on a star–shaped network, as shown in the next example.

**Example 2** *The star.*

It is easy to see that the star has only two maximal independent sets (see Figure 2): one in which the center alone plays 1, and another one in which the spokes do so. If we are looking for efficiency (defined as fewer 1s, which are supposed to be costly) it is very easy to find that the first case is the best one. Suppose that we are in the *bad* NE (spokes exerting the costly effort), then a social planner could shift to the *good* equilibrium by incentivating a contribution from the center. When the center is contributing, then, by best response, all the spokes will stop doing so. This mechanism will be formalized in the next section, but the idea is that of incentivating a contribution from agents that were not doing so in a NE, thus the system will move to a new NE, which may reduce the social cost of being in equilibrium.  $\square$

The problem of finding all the maximal independent set of a general network is however not an easy one. This problem is actually NP–complete,<sup>3</sup>

<sup>3</sup>Consider a general problem whose object (input) is characterized by a certain size  $N$

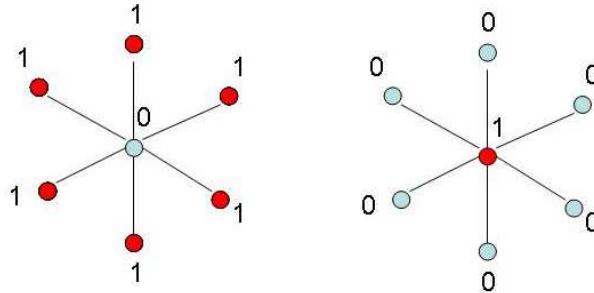


Figure 2: The two NE of a star network.

as is the problem of finding those maximal independent sets with more or less nodes playing 1 (even in Example 1 it is not immediately straightforward at first sight.). In a companion paper, Dall’Asta, Pin and Ramezanpour (2009), we discuss this in more detail for a particular class of random networks. The next example may give a hint of this, for a case which is large but apparently simple.

**Example 3** *Simulations on large regular graphs.*

Consider a regular graph of degree  $K$ , i.e. a network in which every node has exactly  $K$  links, even if the network has not the regularity of a lattice and the connections are randomly drawn. In Dall’Asta, Pin and Ramezanpour (2009) we consider such networks consisting of  $N = 10.000$  nodes. These networks have many possible maximal independent sets, of different density (that is: the percentage of nodes playing 1). Call this density  $\rho$  and define  $s(\rho)$  as the quantity for which the number of equilibria of a given  $\rho$  is equal to  $e^{Ns(\rho)}$  ( $s$  is called “entropy”; it is just a probability distribution re-scaled logarithmically, which gives a more intuitive graphical representation). Figure 3 shows the cases for degree  $K = 3, 4,$  and  $5$ .

In the other paper we use an analytic approximation in order to compute probability distributions.<sup>4</sup> Then we run some heuristic algorithms to

---

(as could be the number of nodes in our case). Here is given a non-rigorous definition: The problem is called NP-complete if there is no algorithm that can find a solution to the problem, for any possible input of size  $N$ , in a time that grows at most polynomially in  $N$ . An NP-complete problem is one in which the time required to find a solution typically grows exponentially in  $N$ . In practice this means that, even if a good computer can solve the problem in a reasonable time for  $N = 1.000$ , the case  $N = 10.000$  may take years to be solved.

<sup>4</sup>López-Pintado (2008) identifies instead the mode of this distribution by adopting a mean field analysis.

actually find all the real distributions of maximal independent sets, in order to confirm the analytical predictions. The predictions are very accurate on the internal part of the curves, but the algorithms are unable to find upper and lower extrema in finite time (it can be proven that this problem is NP-complete). The colored dots represents the upper and lower limits that, given the timing constraint we imposed, the algorithms were able to reach. This is the starting point for the idea of implementing a mechanism inspired by heuristic optimization algorithms.  $\square$

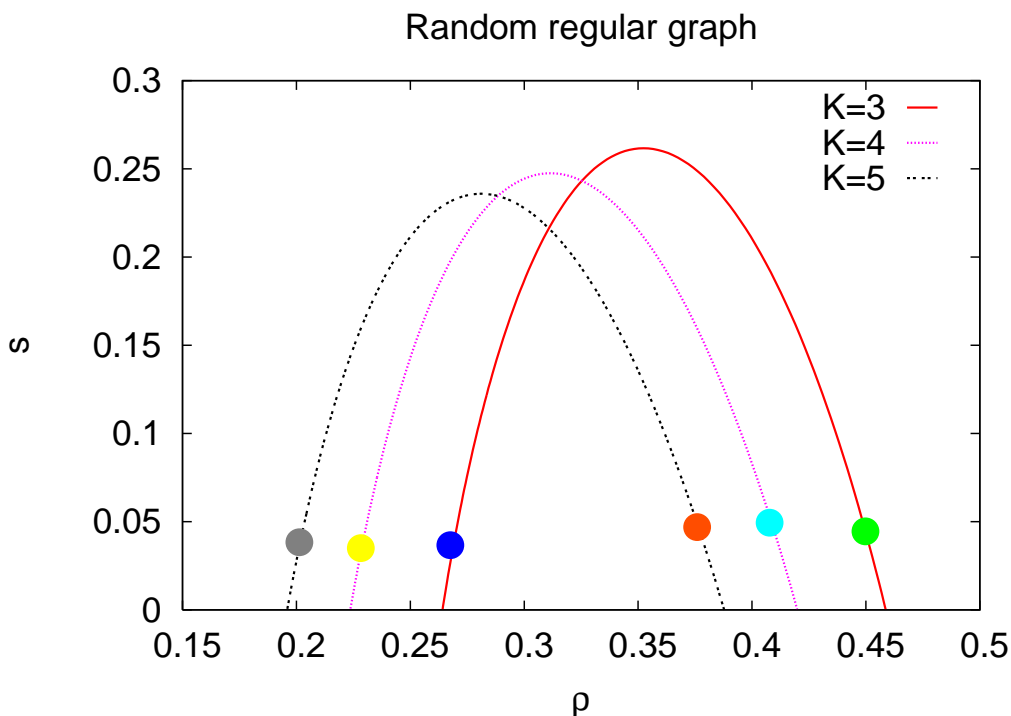


Figure 3: Density versus entropy (probability distribution re-scaled logarithmically) for all the maximal independent sets on regular graphs of 10.000 nodes.  $K$  is the degree. Colored dots are the limits of our algorithm with respect to the analytical predictions.

From the point of view of economics, the rule specified in (1) is not *behavioral* and could be justified by many modelling choices with rational agents. Up to now we have defined (pure) Nash equilibria without explicitly defining actions and payoffs; this however could easily be done. One possibility is the following. Any agent attributes utility  $v$  to a homogeneous good, if she

has access to it (independently of whether it is provided by herself or by any of her neighbors), and her utility is satiated by one unit of it. Finally, the cost of providing the good is a positive value  $c < v$ . Since utilities are satiated, and in equilibrium every agent has local access to the good, then considering efficiency from the point of view of minimal aggregated costs is enough to achieve global efficiency (Bramoullé and Kranton (2007) consider non-satiated utility functions and finds the typical public-good discrepancy between efficient strategy profiles and equilibria). It may seem that we exclude full rationality when we assume that agents respond to changes with a best response rule that considers only the present configuration but is myopic and not strategic on possible future new changes. Consider, however, that another explanation for agents not being interested in future expected payoffs is a high rate  $\delta$  of temporal discount.

The kind of situation we have in mind is that of every agent deciding whether or not to exert a fixed costly effort that is beneficial to herself and also to her neighbors, so that a typical situation of free riding incentives arises. This could be the case with farmers or firms adopting new technologies, with an information network and a cost for possible failures.<sup>5</sup> Another application could be that of several municipalities in a given region; the public good could be a library or a fire brigade, and two municipalities are linked if the public good in one of them makes the same public good undesirable in the other one because of geographical proximity. Finally, since the mechanism we will propose requires low costs of shifting between strategies and repeated interaction, a good application could be that of a big firm encouraging people to share cars in order to minimize parking places. Action 1 would mean ‘take the car’ and an employee would play 0 if a friend gives her a lift. Generally, in any of these applications there could be a planner whose objective could reasonably be that of minimizing costs.

Suppose that the planner considers all the possible NE of the game (all the maximal independent sets of the network) and wants to minimize among them the number of nodes exerting effort 1 (i.e. find a maximal independent set of minimal cardinality: MNE). She could impose the proper action on the agents, and the resulting configuration, being a NE, would be stable without imposing more incentives. Suppose, however, that the planner does not know such an optimal distribution (remember that the theoretical problem is typically a complex one) or that moreover she may not even know anything about the network. Assuming that we also have a time dimension, our question is: would it still be possible for the planner to build a *mechanism*

---

<sup>5</sup>This is the application proposed in Bramoullé and Kranton (2007).

that would incentivate the agents to move towards an optimal MNE?<sup>6</sup> Our answer is only theoretical but positive: at the limit of infinite time such a mechanism exists, and it will lead to a MNE with probability 1.

What we assume is that the social planner's goal is to minimize the costs of a NE, when she has the possibility of incentivating players' actions out of equilibrium, but she is not able to modify the structure of the network. It is clear that if the planner had the possibility of changing the network structure, directly or by incentives, at a reasonable cost (as is the case considered on a different network game by Haag and Lagunoff (2006)) then the problem would look very different. It would be enough to approximate a star-like configuration such as the one analyzed in Example 2, and the solution would easily be found.

In the next section we show how we obtain our result. We show that our setup is included in the hypothesis of a theorem first proved in Geman and Geman (1984) and presented here in Appendix A. The proof of this equivalence is based on three lemmas, whose proofs are in Appendix B. Section 3 concludes the paper.

## 2 Main result

The mechanism we study is the following in discrete time,  $t = 1, 2, 3, \dots$  Every time step is characterized by a configuration  $\vec{x}_t$  of nodes' actions satisfying (1) for every node, and hence NE. Suppose then that at time 1 the system is in a NE, so that  $x_{i,1} \in \{0, 1\}$  is a best response for every agent  $i$ , as specified in (1). The planner does not know anything about the network, the only thing she observes at any step  $t$  in time is the action of each player and hence the aggregate number of agents playing 1; call it  $M_t = \sum_i x_{i,t}$ . What she will do is, at every time step, pick an agent  $i_t$  playing 0, at random with uniform probabilities, and force her to flip her strategy to 1.<sup>7</sup> In consequence of this flip, all the other nodes in the network will change their strategy according to a simple *best response* rule. When the system is stable again, i.e. again in a new NE, the planner will observe a new configuration  $\vec{x}_t^{new}$  and the new aggregate quantity of 1's, call it  $M_t^{new}$ . The planner will accept the new

---

<sup>6</sup>We will use the term *mechanism* to differentiate it from *algorithm*. While the latter is intended as a computational technique, the former is a plausible implementation of any single step of such a technique into a real system, also allowing the interaction of self-interested agents.

<sup>7</sup>This can easily be done through incentives. The reason why the planner is looking for a minimum could be that she is financing all the agents exerting effort; in this case she could raise her contribution to the agent up to the desired threshold level.

configuration with probability

$$\begin{cases} 1 & \text{if } M_t^{new} < M_t \text{ ;} \\ t^{-\epsilon(M_t^{new}-M_t)} & \text{otherwise,} \end{cases} \quad (2)$$

where  $\epsilon > 0$  is a constant. The second probability in (2) identifies the level of rejection of non-improving changes.

We start by proving that  $\vec{x}_t^{new}$  is always a NE for any  $t$  (see Lemma 1 below). If the planner accepts the new configuration, then  $\vec{x}_{t+1} = \vec{x}_t^{new}$  and  $M_{t+1} = M_t^{new}$ , otherwise she will impose reverse incentives so that we return to the original configuration,<sup>8</sup> i.e.  $\vec{x}_{t+1} = \vec{x}_t$  and  $M_{t+1} = M_t$ .

In the limit  $t \rightarrow \infty$ , the second probability in (2) goes to 0 and the mechanism will converge to any member of a precise subset of NE. Call the subset of such possible NE *local minima*.<sup>9</sup> Every MNE is also a local minimum. The question is whether the local minimum in which the process ends is also a MNE. The aim of this paper is to show under which conditions the answer is positive.

The structure of the proof is the following. We show that we meet the conditions required for the application of a known theorem.

**Lemma 1** *if we start from a NE and invert the action of one node from 0 to 1, then the best response rule of all the other nodes in the network will imply a new NE.*

**Lemma 2** *if we start from a NE and invert the action of one node from 0 to 1, then the best response rule of all the other nodes in the network will be limited to the neighborhood of order 2 of the original node (i.e. the change is only local).*

**Lemma 3** *it is possible to reach any NE from any other NE with a finite number of the following procedures: flip the action of a single node from 0 to 1 and obtain, by best response of the nodes, a new NE.*

**Proposition 4** *the probability  $\pi(\epsilon)$  that the mechanism ends in a MNE, at the limit  $t \rightarrow \infty$ , is strictly positive for any  $\epsilon > 0$ ; it is decreasing in  $\epsilon$ ; and finally, there exists an  $\bar{\epsilon} > 0$  such that, for any  $\epsilon < \bar{\epsilon}$ , we have that  $\pi(\epsilon) = 1$  independently on the initial conditions.*

---

<sup>8</sup>This can be done by reverting all incentives to the nodes who changed; they are, by following Lemma 2, restricted to a local neighborhood.

<sup>9</sup>It is also possible that the mechanism, at the limit  $t \rightarrow \infty$ , alternates between more than one single NE, if all of them have the same number of 1's. Without loss of generality, such subsets of NE can simply be included among *local minima*.

The lemmas are proven in Appendix B, by applying the discrete mathematics of network theory. Lemmas 1 and 2 also guarantee that the proposed mechanism is well defined.

The main proposition is obtained by including our setup in the general hypothesis of the theory of *simulated annealing*, first proposed and formalized in Kirkpatrick, Gelatt and Vecchi (1983). Simulated annealing is a heuristic algorithm based essentially on the increasing rejection probability in a Monte Carlo step, as the probability  $t^{-\epsilon(M_t^{new} - M_t)}$  in (2), for our case. Simulated annealing works exactly as described above, finding a global minimum of a certain function, avoiding local minima. Theory tells us that, if the number of possible configurations is finite, and it is possible to reach any configuration from any other with basic steps, then a generalization of the above proposition holds. A rigorous proof that applies to our model can be found in Theorem B of Geman and Geman (1984), which we discuss in Appendix A. The original proof takes various pages, its intuition is that we are analyzing a Markov chain of finite possible configurations (all the NE of the game) which is ergodic for any finite  $t$ .

In our case, we consider all the NE as the possible states of the system; they are finite because the network is finite. Lemmas 1 and 2 define a stochastic process between the states of the system, and this process is ergodic by Lemma 3. We thus meet the conditions that apply in Appendix A.

### 3 Short considerations

The problem of finding a MNE among all the NE is in general not a trivial one, and the difference between the aggregate number of nodes playing 1 in NE could vary dramatically even in homogeneous networks, as examined in the companion paper, Dall'Asta, Pin and Ramezanpour (2009). The star structure (Example 2) is a trivial but dramatic example: there are two NE, one in which the center alone plays 1, and another in which all the spokes do so and the center free rides.

The main practical problem in the implementation of the mechanism we propose is clearly the necessity of infinite time. This paper is only theoretical. However, simulated annealing is used in practice in many optimization problems.<sup>10</sup> Consider that for any  $\epsilon > 0$  the system will reach a local minimum, which can be easily identified even in finite time (the higher the  $\epsilon$  the faster the convergence). Noting that the values  $\epsilon < \bar{\epsilon}$  are typically unrealistically low, and that the algorithm therefor converges very slowly, the choice of a proper heuristic  $\epsilon > \bar{\epsilon}$  could be appropriate. This choice would depend on a

---

<sup>10</sup>Crama and Schyns (2003) is a good example related to finance.

profit/costs comparison but also, in the case of finite time, on the structure of the network (e.g. the star needs a single flip to move from the bad NE to the MNE).

Finally, even if the planner does initially not know the real structure of the network, she could infer it link by link as the steps of the mechanism are played. In this way she could mix the mechanism with a theoretical investigation, and could target nodes non-randomly in order to maximize the likelihood of finding the desired MNE. The analysis of such a sophisticated approach would be much more complicated. What we give here is an upper bound that, we prove, works (even if at the limit). Any improvement on this naïve mechanism will work as well, faster, but not in finite *short* time for any possible network, because the original problem is NP-complete.

## Appendices

### A Theorem B in Geman and Geman (1984)

Geman and Geman (1984) is a pioneering theoretical paper on computer graphics, studying the best achievable quality of images. Sections X to XII are devoted to the general case of optimization among a finite number of states. We find there a general theorem (Theorem B at page 731) proving a conjecture on the *Simulated Annealing* heuristic algorithm proposed by Kirkpatrick, Gelatt and Vecchi (1983). The arising popularity of Simulated Annealing has attested the success of Geman and Geman (1984), which is now cited (according to *scholar.google.com* in January 2009) by almost 8000 papers from all disciplines.

In this appendix we summarize what is necessary for us from this result, with some of the original notation but avoiding most of the thermodynamics *jargon*. Suppose that there is a finite set  $\Omega$  of states, and a function  $U : \Omega \rightarrow R_+$ , so that, for any  $\omega \in \Omega$ ,  $U(\omega)$  is a positive number. Call  $U^* \equiv \max_{\omega \in \Omega} U(\omega)$  the maximal value of  $U$ ,  $U_* \equiv \min_{\omega \in \Omega} U(\omega)$  its minimal value, and  $\Omega_0 \equiv \arg \min_{\omega \in \Omega} U(\omega)$  those states whose value is  $U_*$ . Suppose moreover that we have a fixed transition matrix  $X$  between all the elements of  $\Omega$  and that this stochastic matrix  $X$  is ergodic, i.e. there is a positive probability of reaching any state  $\omega' \in \Omega$  from any other state  $\omega'' \in \Omega$ . Given any  $\omega \in \Omega$ , call  $X(\omega)$  all those states that can be reached from  $\omega$  with positive probability, through  $X$ , with a single step.

Consider now a discrete time flow with  $t = 1, 2, \dots$  and the following new stochastic process.  $\omega_1$  is any member of  $\Omega$ . Imagine that, at time  $t$ , the process is in the state  $\omega_t$ , then apply  $X$  from  $\omega_t$ , obtaining a state that we

call  $\omega_t^{new}$ . We now define  $\omega_{t+1}$  as

$$\omega_{t+1} \equiv \begin{cases} \omega_t^{new} & \text{with probability } \begin{cases} 1 & \text{if } U(\omega_t^{new}) < U(\omega_t), \\ t^{-\epsilon(U(\omega_t^{new})-U(\omega_t))} & \text{otherwise;} \end{cases} \\ \omega_t & \text{otherwise.} \end{cases} \quad (3)$$

The probability  $t^{-\epsilon(U(\omega_t^{new})-U(\omega_t))}$  in (3) identifies the level of acceptance of non-improving changes, which is declining in time at a rate that depends on the constant  $\epsilon > 0$ . Any such stochastic process will be identified by  $\omega_0$  and  $\epsilon$ : call it  $P_{\omega_0, \epsilon}$ .

It is easy to prove that at the limit  $t \rightarrow \infty$  any realization of  $P_{\omega_0, \epsilon}$  will end up in a set of local minima  $\Omega_\epsilon \subseteq \Omega$ .  $\Omega_\epsilon$  is such that, for any  $\omega', \omega'' \in \Omega_\epsilon$  and  $\omega_X \in X(\omega'')$ ,  $U(\omega') = U(\omega'')$  and  $U(\omega') \leq U(\omega_X)$ .

The theorem imposes a single condition on  $\epsilon$  so that the local minima obtained through  $P_{\omega_0, \epsilon}$  are also global minima.

**Theorem B:** call  $N_\Omega$  the cardinality of  $\Omega$  and  $\Delta \equiv U^* - U_*$ . If  $\epsilon < \bar{\epsilon} \equiv \frac{1}{N_\Omega \Delta}$ , then  $\Omega_\epsilon \subseteq \Omega_0$  for any realization of  $P_{\omega_0, \epsilon}$ , independently of  $\omega_0$ .

The proof is by no means trivial, it takes various pages and it is heavily based on the ergodicity of the system. In Geman and Geman's notation, what they call *temperature* is  $\frac{1}{\epsilon \log t}$ . They prove, moreover, that, in the presence of more global minima, the probabilities of ending in any one of them are uniform.

## B Proof of Lemmas

Consider a finite network and call  $x_i \in \{0, 1\}$  the action of node  $i$ , so that  $\vec{x}$  is the vector of the actions of all the nodes. Call  $N_i^1$  the set of nodes which are first neighbors of node  $i$ , and  $N_i^2$  those which are second neighbors of node  $i$ .

We also need the following definitions. A set of nodes in a network is an *independent set* if, for every link of the network, not both its nodes are in the set. A set  $C$  of nodes in a network is a *covering* if, for every node  $i$ ,  $C \cap (\{i\} \cup N_i^1) \neq \emptyset$  (i.e. if for any node  $i$  we consider the set made of  $i$  itself and its first neighbors, then at least one of them is also a member of  $C$ ). A set of nodes in a network is a *maximal independent set* if it is both an independent set and a covering. In our notation a maximal independent set is characterized by those nodes playing 1 in a NE  $\vec{x}$ .

**Proof of Lemmas 1 and 2:** suppose that  $x_i = 1$ , and we flip this action so that  $x_i^{new} = 0$ . Consider now any node  $j$  in  $N_i^1$ , it is clear that  $x_j = 0$  since  $x_i = 1$ . For all those  $j \in N_i^1$  such that  $x_k = 0$  for any  $k \in N_j^1 \setminus \{i\}$ , we will have  $x_j^{new} = 1$ . In the case that two such  $j$ 's that flipped from 0 to 1 will be linked together, by best response only some of them will flip to 1 (this is the only random part in the best response rule). If  $j$  is such that  $x_j = 0$  and  $x_j^{new} = 1$ , it is surely the case that any  $k \in N_j^1 \setminus \{i\}$  was playing  $x_k = 0$  and remains at  $x_k^{new} = 0$ . The propagation of the best response is then limited to  $N_i^1$ .

**Note:** a best response from 0 to 1 applies only to nodes that are playing 0, are linked to a node which is shifting from 1 to 0, and that node is the only neighbor they have who is originally playing 1.

Suppose now that  $x_i = 0$  and we flip this action so that  $x_i^{new} = 1$ . The nodes  $j$  in  $N_i^1$  who were playing  $x_j = 0$  will continue to do so. Any node  $j$  in  $N_i^1$  (at least one) who was playing  $x_j = 1$  will move to  $x_j^{new} = 0$ . By the previous point this will create a propagation to some  $k \in N_j^1$ , but not  $i$ . This proves that the propagation of the best response is limited to  $N_i^2$  (and ends in a new NE).  $\square$

**Proof of Lemma 3:** we proceed by defining intermediate NE  $\vec{x}^1, \vec{x}^2 \dots$  between any two NE  $\vec{x}$  and  $\vec{x}'$ .  $\vec{x}^{n+1}$  will be obtained from  $\vec{x}^n$  by flipping one node from 0 to 1 and waiting for the best response.

If two NE  $\vec{x}$  and  $\vec{x}'$  are different, it must be that there is at least one  $i_1$  such that  $x_{i_1} = 0$  and  $x'_{i_1} = 1$  (it is easy to check that any strict subset of a maximal independent set is not a covering any more). Change the action of that node so that  $x_{i_1}^1 = x'_{i_1} = 1$ . By previous proof this will propagate deterministically to  $N_{i_1}^1$  and, for all  $j \in N_{i_1}^1$ , we will have  $x_j^1 = x'_j = 0$ . Propagation may also affect  $N_{i_1}^2$  but this is of no importance for our purposes.

If still  $\vec{x}^1 \neq \vec{x}'$ , then take another node  $i_2$  such that  $x_{i_2}^1 = 0$  and  $x'_{i_2} = 1$  ( $i_2$  is clearly not a member of  $N_{i_1}^1 \cup \{i_1\}$ ). Pose  $x_{i_2}^2 = x'_{i_2} = 1$ , this will change some other nodes by best response, but not  $j \in N_{i_1}^1 \cup \{i_1\}$ , because any  $j \in N_{i_1}^1$  can rely on  $x_{i_1}^1 = 1$ , and then also  $x_{i_1}^2 = x_{i_1}^1 = 1$  is fixed.

We can go on as long as  $\vec{x}^n \neq \vec{x}'$ , taking any node  $i_{n+1}$  for which  $x_{i_{n+1}}^n = 0$  and  $x'_{i_{n+1}} = 1$ . This process will converge to  $\vec{x}^n \rightarrow \vec{x}'$  in a finite number of steps because:

- when  $i_{n+1}$  shifts from 0 to 1, the nodes  $j \in \bigcup_{h=1}^n (N_{i_h}^1 \cup \{i_h\})$  will not change, since they are either 0-players with a 1-player beside already (the 1-player is some  $i_h$ , with  $h \leq n$ ), or a 1 (some  $i_h$ ) surrounded by frozen 0's;

- by construction it is never the case that  $i_{n+1} \in \bigcup_{h=1}^n (N_{i_h}^1 \cup \{i_h\})$ , because for all  $j \in \bigcup_{h=1}^n (N_{i_h}^1 \cup \{i_h\})$  we have that  $x_j^n = x_j'$ ;
- the network is finite.  $\square$

In the above proof, the shift from  $\vec{x}$  to  $\vec{x}'$  is done by construction re-defining the covering of any  $\vec{x}^n$  from the covering of  $\vec{x}'$ . It is always certain that, by best response, any  $\vec{x}^n$  is also an independent set.

## References

- [1] Bramoullé, Y., and R. Kranton, 2007. Public goods in networks. *Journal of Economic Theory* 135, 478–494.
- [2] Crama, Y., and M. Schyns, 2003. Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research* 150, 546-571.
- [3] Galeotti, A., and S. Goyal, 2008. The Law of the Few, Mimeo.
- [4] Galeotti, A., S. Goyal, M. Jackson, F. Vega-Redondo and L. Yariv, 2008. Network Games, Mimeo.
- [5] Geman, S., and D. Geman, 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 721–741.  
[http://www.cis.jhu.edu/publications/papers\\_in\\_database/GEMAN/GemanPAMI84.pdf](http://www.cis.jhu.edu/publications/papers_in_database/GEMAN/GemanPAMI84.pdf)
- [6] Dall’Asta, L., P. Pin and A. Ramezanzpour, 2009. Statistical Mechanics of maximal independent sets, Mimeo.
- [7] Haag, M., and R. Lagunoff, 2006. Social norms, local interaction and neighborhood planning. *International Economic Review*, 47 (1), 265–296.
- [8] Hirshleifer, J., 1983. From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods. *Public Choice*, 41 (3), 371–386.
- [9] Kirkpatrick, S., C. D. Gelatt Jr. and M. Vecchi, 1983. Optimization by Simulated Annealing. *Science*, 220, 671-680.
- [10] López-Pintado, D., 2008. The Spread of Free-Riding Behavior in a Social Network. *Eastern Economic Journal* 34 (4), 464-479.