University of Cyprus
Department of Economics

Working Paper 11-2015

Diffusion by Imitation: The Importance of Targeting Agents

Nikolas Tsakas

Diffusion by Imitation: The Importance of Targeting Agents[☆]

Nikolas Tsakas

Department of Economics, University of Cyprus

Abstract

We study the optimal targeting strategy of a planner who seeks to maximize the diffusion of an action in a society where agents imitate successful past behavior of others. The agents face individual decision problems under uncertainty and interact locally, so that each agent affects only his neighbors. We find that the optimal targeting strategy depends on two parameters: (i) the likelihood of the action being more successful than its alternative and (ii) the planner's patience. More specifically, for an infinitely patient planner the optimal strategy is to concentrate all the targeted agents in one connected group when her preferred action has higher probability of being more successful than its alternative; whereas it is optimal to spread them across the population when this probability is lower. Interestingly, for an impatient planner the optimal targeting strategy is exactly the opposite.

1. Introduction

1.1. Motivation

The importance of social interactions for the diffusion of innovations, ideas and behavior is a topic that has attracted a lot of research interest over the years (see Jackson, 2008). Recent technological advances have made the collection and analysis of data related to the structure of relationships inside societies, as well as the rules

---

---

guiding the behavior of their members possible. The appropriate use of this information can provide helpful tools for the effective diffusion of products, technologies and ideas in societies.

In this paper, we describe the optimal intervention of an interested party (from now on called *planner*) who seeks to maximize the diffusion of a given action in a society where agents imitate successful past behavior of their neighbors. Effective design of social influence campaigns appears to be a crucial problem in several real–life situations, usually when the planner has limited resources available.

The direct example that comes to ones mind is a firm that produces a new product and wants to establish it in a new market. In such a setting, it is rather common for people to seek advice from previous users before purchasing such. Moreover, firms tend to advertise the product initially to a limited number of people that will then spread the word to the rest of the society. The correct selection of initial targets might be crucial for the success or not of the product. Furthermore, the effectiveness of this selection is affected by several factors, such as the horizon in which the firm expects to observe the outcome of the targeting strategies, or how the initial targets are going to be distributed around the society.

There are several other examples that fit the general idea of our work. For instance, an NGO or government that wishes to promote a new highly productive agricultural technology and for doing so it wants to select a few farmers in one or several villages to adopt the technology initially. Or else, a government that wishes to reduce criminal activity and is willing to sponsor a number of ex–criminals to change their lifestyle. Or even, a political or religious organization that wishes to propagate its ideology and needs to locate a number of initial seeds in the society to spread the word to their neighborhood. As one can see, the problem of optimal social influence is directly applicable to a bunch of different environments and seemingly unrelated areas.

The particular focus of this paper is twofold. First, we highlight the importance of the distribution of targeted agents in the society and not on their individual characteristics. As it will become apparent, this turns out to be a very crucial feature that has been overlooked until now. Second, we compare the optimal strategy of a short–sighted versus that of a far–sighted planner, which turn out to have very distinct characteristics. To the best of our knowledge this is the first paper to discuss thoroughly these two aspects; thus opening a number of interesting avenues for further research.

More specifically, most of the existing literature on targeting has focused on the importance of central agents (see Ballester et al., 2006). Having a high or a low number of connections (see Galeotti and Goyal, 2009, Chatterjee and Dutta, 2011), or diffusing information to many others who are poorly connected (see Galeotti et al., 2011) are some usual characteristics of influential agents. The importance of these characteristics is obvious and beyond doubt. Nevertheless, we show that there is another important factor with significant effect on diffusion. This is whether the targeted agents are concentrated all together, in the sense that they are connected between them, or they are spread across the society.

Furthermore, throughout our analysis we highlight the differences between optimal targeting strategies of a *patient* versus an *impatient planner*, i.e. who care about diffusion in the long–run and short–run respectively. It turns out that the two cases

differ sharply and these differences persist independently of the parameters' values. This comparison is important in several realistic scenarios, since different targeting strategies may be appropriate depending on the time horizon.

Undoubtedly, in order to obtain tractable and intuitive results we need to make a set of simplifying assumptions, which might reduce the applicability of our analysis to certain problems. Nevertheless, we provide a framework that can help us understand better which are those parameters that affect problems of social influence crucially and we illustrate how beneficial the knowledge about society's structure may be for the efficient design of marketing and general social influence campaigns.

1.2. Setting and Results

Formally, we consider a finite population of behaviorally homogeneous agents located around a circle. In each period, all agents choose simultaneously between two alternative actions. The stage payoff each action yields is uncertain and depends on a random shock, which is common for all the agents who have chosen the same action in that period.[1] Shocks are independent across actions and across periods. There are no strategic interactions between agents. After making their decisions, all agents observe the chosen actions and the realized payoffs of their two immediate neighbors. Subsequently, they update their choice myopically, imitating the action that yielded the highest payoff within their neighborhood in the preceding period.

There are several reasons why economic agents adopt simple behavioral rules, such as the imitation of successful past behavior. For example, they often need to make decisions without knowing the potential gains or losses of their possible choices. Furthermore, such situations may arise with high frequency and the agents' computational capabilities are limited, then they tend to rely on information received from past experience of others, rather than experimenting themselves.[2] These arguments are also supported by a recent, but growing, empirical and experimental literature which provides strong evidence in favor of the fact that in several decision problems the agents tend to imitate those who have been particularly successful (see Apesteguia et al., 2007, Conley and Udry, 2010, Bigoni and Fort, 2013).

A simple problem that fits our model particularly well is that of the diffusion of agricultural technologies.[3] Farmers' harvests depend mostly on common factors such as the weather and the fertility of the land. Moreover, it is normal to assume that farmers are aware of the technologies and crops used by their neighbors, as well as the payoffs they receive. In particular, Conley and Udry (2010) show that farmers tend to imitate those who have been very successful in the past, whereas as it is pointed out by Ellison and Fudenberg (1993) the farmers' technology decisions are guided mainly by short–term considerations, especially when capital markets are poorly developed or malfunctioning.

The planner is interested in maximizing the diffusion of her preferred action in the population. She can be either infinitely patient, therefore interested in the

---

[1]This assumption is made only in order to facilitate the tractability of the results.

[2]These are some of the reasons why imitation has been subject to extensive theoretical study in different environments (see Ellison and Fudenberg, 1993, 1995, Vega-Redondo, 1997, Eshel et al., 1998, Schlag, 1998, Alós-Ferrer and Weidenholzer, 2008, Duersch et al., 2012).

[3]See also Ellison and Fudenberg (1993).

diffusion of the action in the long run; or impatient, therefore interested in the diffusion of the action after just one period. She is assumed to know the structure of the society, as well as how agents behave and can intervene by enforcing a change at the initial choice of a subset of the population. Ideally, she would like to target the whole population, but doing so in reality would be extremely costly. Hence, our goal is to identify the planner's optimal targeting strategy given the number of agents that can be targeted.

Observe that, all the agents are identical with respect to any measure of centrality. In fact, none of them has any positional advantage or disadvantage compared to the rest of the population. This is an important feature, given that we want to focus on the distribution of targeted agents in the society. Despite this fact, we find that expected diffusion changes substantially depending on the subset of the population that has been targeted by the planner.

We show that the optimal targeting strategy depends on two parameters: (i) the likelihood of the planner's preferred action being more successful than its alternative and (ii) the planner's patience. In fact, we observe a sharp contrast between the optimal strategies of a patient planner versus that of an impatient one. More specifically, when the planner's preferred action has higher probability of being more successful than its alternative, then the optimal targeting strategy for a patient planner is to concentrate all the targeted agents in one connected group;[4] whereas when this probability is lower it is optimal to spread them uniformly across the population. Interestingly, for an impatient planner, the optimal targeting strategy is exactly the opposite.

The intuition is relatively simple and depends on the fact that in the long run only one of the two actions survives. Therefore, when the action is likely to be successful, then an infinitely patient planner wants to prevent its disappearance due to a few consecutive negative shocks in the first periods. For this reason she prefers to concentrate all the targeted agents together. To the contrary, if the action is unlikely to be successful, then the optimal strategy for the planner is to try and take advantage of a possible sequence of successful shocks during the first periods. By concentrating all targeted agents together, she would only manage to make the action disappear more slowly, since for its diffusion a large number of consecutive successful shocks would be needed, which is rather unlikely to happen.

For an impatient planner the arguments are reversed. When the preferred action is likely to be successful, then the planner wants to make it visible to as many agents as possible, therefore she should spread the targeted agents around the society. On the other hand, if the action is more likely to be unsuccessful, then the planner wants to prevent as many of the targeted agents as possible from observing the alternative action, therefore she should concentrate them all together.

At this stage, one could question how important is the role of the particular behavioral rule for obtaining these results and how robust they would be under either Bayesian (see Gale and Kariv, 2003) or De Groot (see DeGroot, 1974, Golub and Jackson, 2010) learning rules. A crucial aspect for the current results is that the agents do not accumulate information through time, which has two negative

---

[4]In the circle a connected group is a segment of the circle.

4

and one partially positive effect. On the one hand, the society is vulnerable to misguidance by certain unexpected events even at later stages, which for example should not be the case if the agents perform Bayesian learning. Moreover, under any initial conditions there is no guarantee that the society would converge to the planner's desired action (even if this is the socially optimal). On the other hand, the process is less path dependent than De Groot learning, where initial opinions may drive a society towards an inefficient state, sometimes even with certainty. Once again, this would not be a problem under Bayesian updating, which however has been acknowledged by large part of the literature of learning in networks to require extremely complex calculations.

We extend our analysis in many different directions.[5] We discuss the optimal strategies of planners with intermediate levels of patience, thus intending to identify how the transition between the two extreme cases occurs. Moreover, we quantify the practical meaning of infinite patience by characterizing the expected waiting time before convergence occurs. We observe that, for those cases in which the planner's optimal strategy is to concentrate all the targeted agents in one group the process is slowed down substantially. In addition to this, we discuss what happens if we allow for inertia and show that the results remain unchanged. This extension captures many realistic features, such as the existence of switching costs and some forms of conformity. Finally, we repeat our analysis for the line and the star and provide numerical simulations for other network structures, as an attempt to identify the effect of centrality on our results.

1.3. Related Literature

The role of influential agents in environments with local interactions has been studied in different disciplines, such as computer science (see Kempe et al., 2003, Richardson and Domingos, 2002), marketing (see Kirby and Marsden, 2006) and physics (see Bagnoli et al., 2001), as well as in economics. Intuitively, a crucial feature is the centrality of an agent, which depicts either its number of immediate neighbors or its important importance for the connectivity of the society (see Ballester et al., 2006).

Our paper is closely related to Galeotti and Goyal (2009) and Goyal and Kearns (2012). Although the modeled processes are different, the research questions are similar. However, neither of the papers focuses on either of our two main points of interest, i.e. the distribution of targeted agents and the contrast between short–run and long–run strategies. Galeotti and Goyal (2009) study the problem of a firm that seeks to maximize the diffusion of its product in a society, under word–of–mouth communication and social conformism. The authors focus exclusively on short–run strategies and on the degree distribution of the agents, assuming that agents meet randomly. This feature is crucial, since it does not allow us to study effect of the distribution of targeted agents across the population. Similarly, Goyal and Kearns (2012) study competition between two firms who distribute their resources trying to maximize the long–run diffusion of a product in a social network. The paper focuses more on how efficiently the resources are allocated, as well as on the effect of budget asymmetries on the equilibrium allocations.

---

[5]A thorough discussion of the extensions can be found in the online appendix.

As it has already been mentioned, large part of the literature has focused on the identification of particular characteristics of influential agents, highlighting the importance of different measures of centrality.[6] In another recent paper, in an setting where agents are able to learn from their neighbors about the quality of a product, Tsakas (2014) finds decay centrality to be a crucial measure for the characterization of optimal strategies both in the short and in the long–run. Furthermore, in an environment with strategic information transmission, Galeotti et al. (2011) show that influential agents are those who diffuse information to many others, who themselves are poorly connected.

More broadly related are also the areas of cellular automata and voter models, used mainly in the physics literature but also in economics. For instance Bagnoli et al. (2001) study the long-run behavior and the phase transition of a system with characteristics similar to ours, without focusing on the initial conditions,[7] which is the main focus of our paper. In economics, Yildiz et al. (2011) generalize the standard voter model by introducing "stubborn" agents, i.e. who never change their choice, and they discuss the problem of optimal placement of such agents. Also, Ortuño (1993) considers a standard voter model setting, where agents are located in a two dimensional infinite lattice in which a planner seeks the diffusion of a technology. This is the only article where centrality does not play a role.

Regarding agents' behavior, we focus on imitation of successful past behavior. Similar rules have been studied extensively in several theoretical settings,[8] with the focus being mostly on the characterization of stochastically stable configurations. There is also a recent but growing empirical literature[9] that provides empirical evidence on the adoption of this behavior in real environments. In a more general framework, the current analysis builds upon the work on learning from neighbors,[10] where most of the papers focus mainly conditions under which efficient actions spread to the whole population and not on optimal influence strategies.

The rest of the paper is organized as follows. In Section 2 we define the model formally. Sections 3 and 4 contain the characterizations of optimal targeting strategies for an impatient and a patient planner respectively. In Section 5 we briefly discuss some extensions and Section 6 concludes. All proofs can be found in the Appendix. A thorough study of the extensions can be found in the online appendix.[11]

2. The Model

2.1. The Agents

There is a finite set of agents $N = \{1, ..., n\}$, referred to as *population* of the society. Agents are located around a circle. Each agent can observe his two immediate neighbors, i.e. one to his left and one to his right. At time $\tau = 1, 2, ...$, each agent

---

[6]Bonacich centrality in Ballester et al. (2006), decay centrality in Chatterjee and Dutta (2011).

[7]Given that they refer to initial positions of particles.

[8]See Vega-Redondo (1997), Alós-Ferrer and Weidenholzer (2008), Eshel et al. (1998), Fosco and Mengel (2011).

[9]See Apesteguia et al. (2007), Conley and Udry (2010), Bigoni and Fort (2013).

[10]See Banerjee (1992), Banerjee and Fudenberg (2004), Bala and Goyal (1998), Chatterjee and Xu (2004), Ellison and Fudenberg (1993, 1995), Gale and Kariv (2003).

[11]https://sites.google.com/site/nikolastsakas/research

$i \in N$ chooses between two alternative actions, $a_i^\tau \in \{A, B\}$. Each action yields random payoff. The payoff of agent i is independent of the other agents' choices. Therefore, interactions among agents are not strategic and their connections represent only an exchange of information.[12] Moreover, agents who choose the same action at a given period receive equal payoffs.[13] The payoffs of both actions change in each period, with the realizations being independent across periods. Action B yields strictly higher payoff than action A with probability $p \in (0,1)$, while action A yields strictly higher payoff than B with probability q. For the derivation of the main results we focus on the case where $q = 1 - p$. Relaxation of this assumption does not affect the results, as it is only about the case where the distributions of the payoffs are not atomless, so that both actions yield the same payoff with positive probability. In fact, the important factor for our analysis is the ratio between p and q, which we define as $r = \frac{p}{q}$. From now on, we will say that there is a *success* (*failure*) in period τ if action B (A) yielded higher payoff in this period. [14]

The *planner* is an agent, outside of the population, who seeks to maximize the diffusion of action B in the population.[15] She can do so by changing to her favor the choice of a subset of the population before the beginning of the first period.[16] Throughout the paper, these are mentioned as *targeted agents*. Optimally, she would like to affect the whole population, but in reality this would be extremely costly. This cost enters implicitly if we assume that the cardinality of the subset she can affect is fixed exogenously.[17] More specifically, given that at period $\tau = 0$ all the agents are choosing action A, the planner can target $t \leq n$ agents from the population and make them choose B in period $\tau = 1$. After that, the planner cannot affect the society anymore. The goal of this paper is to characterize the planner's optimal targeting strategy.

The planner can be either impatient or infinitely patient. A planner is called *impatient* if she cares about the diffusion of her preferred action after only one period. Similarly, a planner is called *infinitely patient* if she cares only about the

---

[12]Non–strategic nature of communication is a behavioral assumption often employed in the literature for two reasons. First, because it facilitates the tractability of the results and second because a possible strategic behavior of the agents would interact with their imperfect knowledge of the network structure in very complex ways.

[13]This assumption is imposed mainly in order to facilitate the tractability of the results. Doing so allows us to summarize all stochasticity in one parameter, as well as to reduce the dimensionality of the stochastic processes that describe it.

[14]Later on, we will also define success and failure in a random walk in a similar way.

[15]Throughout the paper, by diffusion of an action we refer to the expected number of adopters over a certain number of periods, either one or infinite. Implicitly, this means that the planner is assumed to be risk neutral.

[16]Assuming that the planner intervenes only once is a simplifying assumption. Nevertheless, if the planner could target agents in multiple periods, a similar analysis would be repeated several times. Intuitively, multiple periods of targeting could allow the planner to be more risky in the beginning, in the sense of targeting larger number of groups, thus intending to attract the whole society quickly and then condition her future strategy on the realized history. In a slightly different environment, where the planner does not know perfectly the likelihood of success of each action, multiple periods of targeting could help the planner improve the accuracy of her beliefs over this. Such an explanation would be more plausible for environments such as in Tsakas (2014).

[17]Later on, we endogenize the number of affected agents and we discuss the returns to investment for different values of it.

diffusion of her preferred action in the long–run. We find that the optimal behavior of an impatient planner is exactly the opposite to that of an infinitely patient planner. Later on, we also discuss some intermediate levels of planner's patience.

2.2. The Behavior

At the end of each period, the agents observe the actions and realized payoffs of themselves and their neighbors. Subsequently, they have the opportunity to revise their choices. Revision occurs happen simultaneously for all agents.[18] According to these observations, the agents revise their choices by imitating the most successful action within their neighborhood in the preceding period. Notice that, an agent never switches to an action that she did not observe, i.e. that neither her nor any of her neighbors chose in the previous period. Moreover, if an action disappears from the population it never reappears.

The important aspect of this myopic behavior is that the agents discard most of the available information. They ignore whatever has happened before the previous period, hence they are unable to form beliefs about the underlying payoff distributions of their alternative choices.

2.3. The Problem

After the planner has chosen its targets, the population consists of s agents choosing action A (from now on called *non–adopters*) and t agents choosing action B (from now on called *adopters*); obviously $s + t = n$. We call a *group* a sequence of neighboring agents all of whom choose the same action and are surrounded by agents choosing the opposite action. The population is formed of m groups of neighboring agents who choose action A, with sizes $\{s_1, s_2, \ldots, s_m\}$, where $\sum_{k=1}^{m} s_k = s$ and analogously m groups of neighboring agents who all choose action B, with sizes $\{t_1, t_2, \ldots, t_m\}$, where $\sum_{k=1}^{m} t_k = t$.[19] We refer to these groups as groups *of type A* and *type B* respectively. The numbering of the groups is based on their size in increasing order, $s_1 \leq s_2 \leq \cdots \leq s_m$ and $t_1 \leq t_2 \leq \cdots \leq t_m$. With some abuse of notation we also use $s_1, s_2, \ldots, s_m$ and $t_1, t_2, \ldots, t_m$ to name the groups.

Our aim is to find the optimal initial size of all s_k and t_k for $k = 1, \ldots, m$, their optimal initial position (if it matters), as well as the optimal initial number of groups, m.

In order to avoid unnecessary complications in the calculations (which arise without the gain of any additional intuition) we assume that every group must have an *even* number of agents. The qualitative intuition of the results would be the same even without imposing this assumption, and in most of the cases it would not affect them at all. It is only imposed for better exposition of the results.[20] Formally:

Assumption 1 (A1). s_i *and* t_i *are even numbers for all* $i \in \{1, \ldots, m\}$.

---

[18]Simultaneous updating is assumed to keep the problem tractable. In extensions, we provide a simple intuition of how this, as well as not perfectly correlated payoffs, would affect them quantitatively, but not qualitatively.

[19]Notice that, the fact that the network is a circle and there exist exactly two actions ensures that the number of groups is the same for both actions.

[20]See the analysis of the line in the online appendix for a more detailed motivation regarding this assumption.
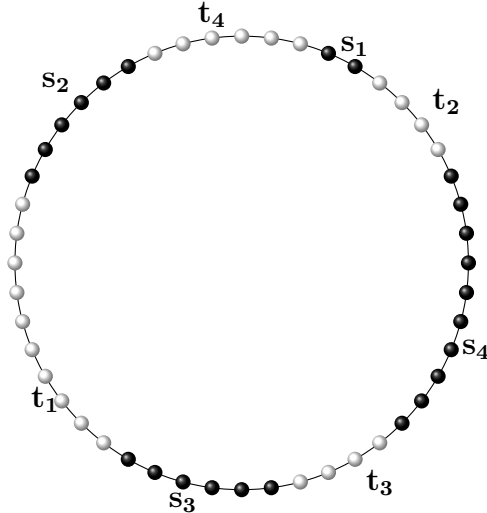
Figure 1: Example of an initial configuration: White nodes represent agents choosing action B and black nodes agents choosing action A.

This restriction is imposed only on the initial configuration, nevertheless it ensures that this will be the case in all the periods, given that the size of a group in each period either increases by two or decreases by two.

3. Results for an Impatient Planner

In this section, we study the optimal targeting strategy of a planner who cares about maximizing the expected number of agents choosing action B after exactly one period. Figure 2 shows the two possible configurations after one period. White dots represent the agents who choose initially action B and black dots those agents who choose initially action A. Observe that only those agents who are on the boundary of a group can change their choice. In fact, for m denoting the total number of groups, in case action B is more successful in the first period, then there will be $2m$ additional adopters in the next period; whereas, in case action A is more successful, the number of adopters will decrease by $2m$. The probabilities of ending in each of the two possible states is p and $q = 1 - p$ respectively.

Figure 2: Initial configuration and the two possible configurations after one period.

Hence, the objective function of the impatient planner is as follows:

$$EN_B(1) = t + 2mp - 2m(1-p) = t + 2m(2p-1)$$

It is easy to see that the optimal targeting strategy depends on p. Namely, for $p > 1/2$ the objective function is strictly increasing in m.[21] Therefore, it is optimal to have as many groups as possible. On the other hand, if $p < 1/2$ the objective function is strictly decreasing in m and therefore it is optimal to locate all targeted agents in one group.[22] This result is formally stated in the following proposition:

Proposition 1. *Under (A1), then for an impatient planner and for a given number of targeted agents, t*

- *If $r > 1$, the optimal targeting strategy is to spread the targeted agents to as many groups as possible, i.e.*

    - *If $t < s$, then $m = \frac{t}{2}$, with $t_1 = \cdots = t_m = 2$.*
    - *If $t > s$, then $m = \frac{s}{2}$, with $s_1 = \cdots = s_m = 2$.*

- *If $r < 1$, the optimal targeting strategy is to concentrate all the targeted agents in one group, i.e. $m = 1$, $t_1 = t$ and $s_1 = s$.*

Observe that, as long as the planner creates the maximum number of groups, the allocation of the agents inside these groups is not important. This feature will change slightly in the case of infinite patience.

Intuitively, this result suggests the following. If an action is likely to be successful, then the planner should try to make it directly visible to as many non–adopters as possible. Doing so, she will manage to attract the maximum number of additional adopters in case of success. To the contrary, if an action is unlikely to be successful, then the planner should prevent most of the targeted agents from observing the opposite action. As a result, even upon an unsuccessful realization, most of them will not change their choice in the second period. As we will see, this optimal strategy changes sharply if the planner is infinitely patient.

4. Results for an Infinitely Patient Planner

In this section, we study the optimal behavior of an infinitely patient planner, i.e. one who cares only about the diffusion of her preferred action in the long run. A crucial feature of this setting is that such a planner disregards completely the speed of the procedure. Before beginning our analysis, it is useful to state two prior results.

---

[21]If $q \neq 1 - p$, we would simply have to replace $2p - 1$ with $p - q$.

[22]Throughout the paper, we disregard the case of $p = \frac{1}{2}$. This is because for $p = \frac{1}{2}$ and given Assumption 1 every targeting strategy of the planner yields exactly the same result.

4.1.1. Diffusion when Agents Imitate-the-Best Neighbor

The present behavioral rule constitutes a special case of imitate–the–best neighbor, applied in a setting of individual decision–making under uncertainty without strategic interactions between agents. Agents observe the choices of their neighbors and the payoff they yield. Subsequently, they revise their choices repeatedly according to these observations. In particular, they do so by imitating the action that yielded the highest payoff within their neighborhood in the preceding period.

In such a setting, it turns out that the population converges with probability one to a steady state where all the agents choose the same action (see Tsakas, 2013). Moreover, if $p \in (0,1)$ then any of the actions can be the one to survive in the long run. This is based on the fact that all actions are vulnerable to a sequence of negative shocks, which can lead to their disappearance. Given that an action which disappears from the population never reappears, it turns out that only one of them survives at the end.

In our case, this result ensures that only one of the two actions will survive in the long run and that both of them have a positive probability to be the ones succeeding. Hence, the optimal strategy for an infinitely patient planner is the one that maximizes the probability that action B gets diffused to the whole population in the long run. We define this probability as follows:

Definition 1. $P_B(\cdot)$ *is the probability that action B will be diffused to the whole population in the long run.*

This probability will depend not only on the size of the population n, the number of targeted agents t and the probability of success p, but also on the choice of the planner about which agents to target. Notice as well that maximizing this probability is equivalent to maximizing the expected number of agents choosing action B in the long run; a remark that will clarify the analogy between our short run and long run analysis.

4.1.2. Results on Random Walks with Absorbing Barriers

A technical result which turns out to be particularly useful comes from Kemeny and Snell (1960). It refers to the properties of a finite one–dimensional random walk with absorbing barriers, which in the current context is defined as a Markov chain whose state space is given by the integers $j \in \{0, 1, \ldots, n\}$ and its initial state is i. For some numbers p and q satisfying $0 < p, q < 1$, the transition probabilities are given by $P_{j,j+1} = p$, $P_{j,j-1} = q$ and $P_{j,j} = 1 - p - q$ for all $j \neq 0, n$ and $P_{0,0} = P_{n,n} = 1$. The endpoints of a random walk are called absorbing barriers because upon reaching one of them the process eventually stays there forever. Those two states are the only absorbing ones. We denote the probability of absorption at state n (respectively at state 0), when the process initiates at state i, by $P_n(i)$ (respectively $P_0(i)$).Specifically, Kemeny and Snell (1960) compute the random walk's probabilities of absorption at each one of the two absorbing states as follows:

Lemma 1 (Kemeny and Snell (1960)). *Consider a random walk with state space* $\{0, 1, \ldots, n\}$, *where both barriers* 0 *and* n *are absorbing. If the probability of moving to the right (from* j *to* $j + 1$*) is* p*, the probability of moving to the left (from* j *to* $j - 1$*) is* q*, and* $r = \frac{p}{q}$*, then the probability of absorption at state* n*, when starting from state* i *is:*

$$P_n(i) = \begin{cases} \frac{r^n - r^{n-i}}{r^n - 1} & \text{if } p \neq q \ \text{(or equivalently } r \neq 1\text{)} \\ \frac{i}{n} & \text{if } p = q \ \text{(or equivalently } r = 1\text{)} \end{cases} \tag{1}$$

Analogously, the probability of absorption at the state 0 *is* $P_0(i) = 1 - P_n(i)$.

For the moment, $p \neq q$ is equivalent to $p \neq \frac{1}{2}$. The results are completely analogous in the general case.

To help us understand how this result can be used to express the current diffusion process we consider a line (see Figure 3), with agents named $\{1, 2, \ldots, n\}$, where each agent has two neighbors, except of agents 1 and n who have one neighbor each. At period $\tau = 1$, agents 1 to i choose action B and the rest choose action A. Hence, every period only two agents may revise their choice (for example in the first period those are the agents i and $i + 1$). The border fluctuates until either agent n chooses B, or agent 1 chooses A. The position of the right border of adopters follows a random walk with absorbing barriers 0 and n. Notice that, in order for agent 1 to choose A the left barrier must be located at the artificial node 0, which is going to be omitted in most of the graphs for expositional simplicity. Hence, we can use the result stated above to describe the probability of diffusion for each of the two actions. We call a random walk *successful* (*unsuccessful*) if it ends up in the absorbing state where all the nodes included in the walk choose action $B(A)$.

This result is particularly helpful for our analysis, because any initial targeting strategy induces a stochastic process that can be expressed as a sequence of conditionally independent random walks with absorbing barriers, similar to the one described above. Despite having multiple borders between groups, all of them fluctuate synchronously, because the payoffs for each action are perfectly correlated and therefore all agents on the boundaries make the same choice in each period. To grasp a better understanding of this argument recall how the process evolves at Figure 2, focusing the attention on the borders of each group. One can notice that a group is either enlarged or reduced by two agents, one at each side. Therefore, if one splits the group in the middle, then the group will evolve on the same way to the left and to the right, until either it disappears, or connects the smaller of its neighboring groups does so. The same holds for every other group of each type. Hence, one can consider that the colors of the nodes in each border follow a random and the fact that the payoffs are perfectly correlated, allows us to consider the whole process following a unidimensional random walk. This initial walk is able to describe the process until some group disappears, as for example it happens in Figure 2 with the bottom left white group after a failure. Once again dure to the payoff correlation we can be sure that the first group to disappear will be either the smallest group of adopters, with size t_1 or the smallest group of non–adopters, with size s_1 disappears. Therefore, the process until one of these two happens can be represented by the random walk that is shown in Figure 6, where if the random walk hits the right (left) corner then the group s_1 (t_1) has disappeared. When either
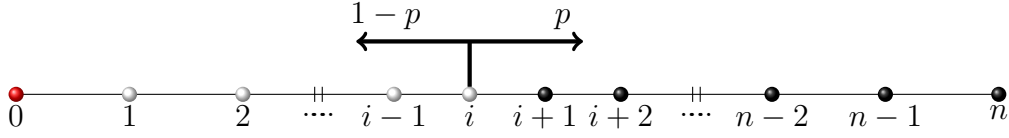
Figure 3: Initial configuration of a random walk with absorbing barriers 0 and n.
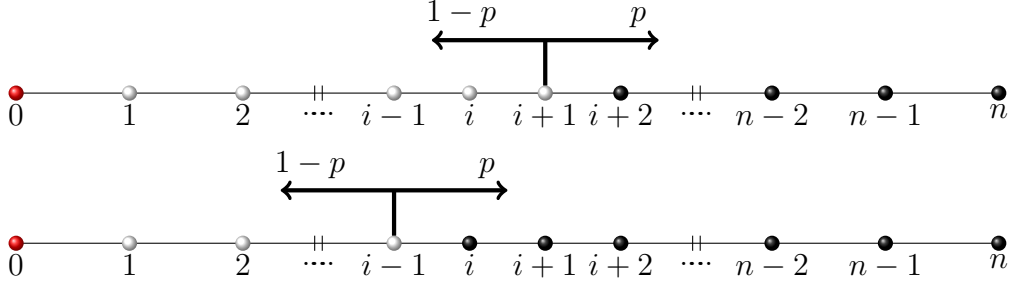

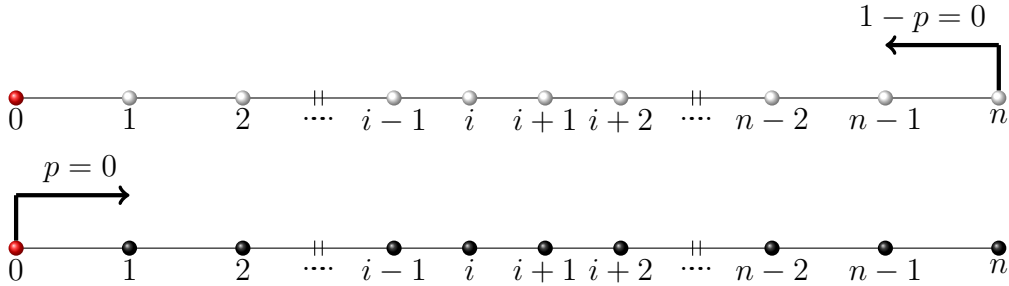Figure 4: The two possible configurations after one period.

Figure 5: The two possible configurations after absorption.

of this happens the group of each type is reduced by one, once and for all, thus changing the possible fluctuations of the process. This can be considered, as the described random walk having absorbing barriers. Upon reaching one of the two barriers, the process starts fluctuating according to a new random walk, with the steps to the next absorbing barriers being determined by the sizes of the new smallest groups. Notice that the probabilities of absorption for this new random walk are conditionally independent on what happened before. Overall, the process can be described as a sequence of conditionally independent unidimensional random walks with absorbing barrier, where in the final stage there is only one group of each type and absorption is associated with total diffusion of one of the two actions.
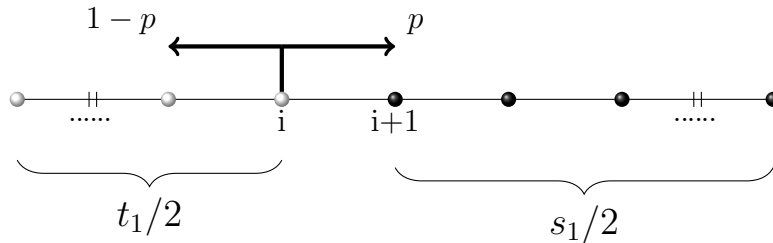

Figure 6: The first random walk

13

4.2. Main Results

Not surprisingly, the ratio $r = \frac{p}{q} = \frac{p}{1-p}$, which describes the likelihood of action B being more successful than action A, is a crucial parameter. However, surprisingly enough, this is the only parameter that affects the optimal targeting strategy and more specifically whether r is higher or lower than 1.[23] It is also interesting though that the optimal strategy of the infinitely patient planner is in complete contrast to that of the impatient planner. In particular, we observe that, for $r > 1$ the optimal targeting strategy of the planner is to concentrate all the targeted agents in a single group, whereas for $r < 1$, the optimal strategy is to spread them as much as possible across the population, splitting them into as many groups as possible and as symmetrically as possible. Observe that, the results are not just different, but are exactly opposite to those found for the impatient planner. In fact, the optimal strategy for an infinitely patient planner is the worst possible strategy for an impatient planner and vice versa.

For a better exposition of the general results, we split the problem into three sub-problems. First, we consider the symmetric case where the groups are restricted to be have equal sizes. Then, we consider the asymmetric case where the planner can target up to two groups with potentially unequal sizes and finally we consider the general asymmetric case.

4.2.1. Symmetric cases

First, we consider the symmetric case, where the groups of agents choosing the same action are restricted to have equal sizes, namely $s_1 = \cdots = s_m = \frac{s}{m}$ and $t_1 = \cdots = t_m = \frac{t}{m}$. Assuming no problems of divisibility we find the optimal number of groups, m. As we have mentioned already, the optimal targeting strategy depends only on the ratio r. In particular, when $r > 1$ it is optimal to concentrate all targeted agents in one group, whereas when $r < 1$, it is optimal to split them in as many groups as possible, i.e. $m = \min\{s/2, t/2\}$. Formally:

Proposition 2. *Under (A1) and given $s_1 = \cdots = s_m = \frac{s}{m}$ and $t_1 = \cdots = t_m = \frac{t}{m}$, then for an infinitely patient planner:*

- *If $r > 1$, the optimal targeting strategy is to concentrate all the targeted agents in one group.*

- *If $r < 1$, the optimal targeting strategy is to spread the targeted agents to as many groups as possible.*

All proofs can be found in the Appendix.

Intuitively, this proposition suggests that when the probability of success is high, it is beneficial to concentrate all targeted agents together. This prevents the disappearance of the preferred action upon the realization of a sequence of negative shocks during the first periods. The opposite strategy is optimal when the probability of success is low. Then the planner wants to take advantage of some potential good shocks during the first periods, which will spread the action to as many agents as possible.

---

[23]Even though, at the moment $r > 1$ is identical to $p > 1/2$, we keep this notation because it facilitates the extension to cases where $q \neq 1 - p$.

4.2.2. Asymmetric Cases

We now turn attention towards the more general asymmetric cases. At first, consider the case where the planner is restricted to target at most two groups of each type, with sizes s_1, s_2 and t_1, t_2 respectively. Then the configuration is as shown in Figure 7. Recall that $s_1 \leq s_2$ and $t_1 \leq t_2$.

It turns out that the optimal decision depends completely on the value of r. More specifically, if $r > 1$ it is optimal to concentrate all the agents in one group, while if $r < 1$ the optimal strategy is to have two groups of equal sizes for each action.
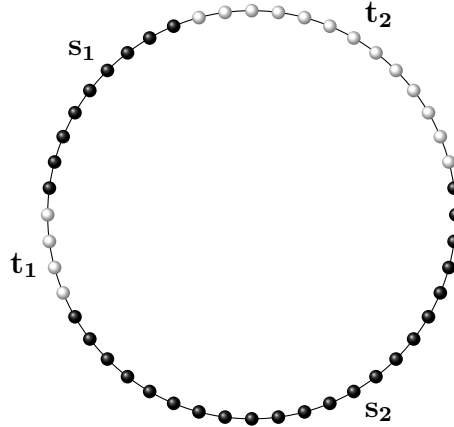

Figure 7: Example of an initial configuration with two groups of each type

Proposition 3. *Under (A1) and given $m \leq 2$, then for an infinitely patient planner:*

- *If $r > 1$, the optimal targeting strategy is to concentrate all targeted agents in one group.*

- *If $r < 1$, then the optimal targeting strategy is to split the targeted agents into two as equal as possible groups, and locate them in the population as symmetrically as possible, i.e. $s_2 - s_1 \leq 2$ and $t_2 - t_1 \leq 2$*

The intuition is similar to that of the symmetric case. However, an interesting finding is that this result does not hold for all restrictions on m. For example, if we restrict the number of groups to be not greater than three, then it is not optimal to split the agents into three equal groups of each type. Hence, it is not the case that we always prefer symmetric configurations compared to asymmetric ones. Notice that, this would be a sufficient condition for the proof of our main result, but it does not always hold. Nevertheless, this does not affect our general result which is stated below.

The two propositions help us construct the main theorem of the paper which describes the optimal targeting strategy in the general case of m initial groups of each type, allowing them to be of different sizes. The result is in line with the previous findings and suggests that the optimal choice is to concentrate all the targeted agents in one group when $r > 1$; and to spread them uniformly across the population in as many and as equal groups as possible when $r < 1$. Namely:

Theorem 1. *Under (A1), for an infinitely patient planner*

- *If $r > 1$, the optimal targeting strategy is to concentrate all the targeted agents in one group, i.e. $t_m = t$ and $t_1 = \cdots = t_{m-1} = 0$ for any m.*

- *If $r < 1$, the optimal targeting strategy is to spread the targeted agents in as many groups as possible and locate these groups as symmetrically as possible, i.e.*

  - *If $t < s$, then $m = \frac{t}{2}$, with $t_1 = \cdots = t_m = 2$ and $s_m - s_1 \leq 2$,*
  - *If $t > s$, then $m = \frac{s}{2}$, with $s_1 = \cdots = s_m = 2$ and $t_m - t_1 \leq 2$*

As it has been mentioned already, the importance of this result lies in the complete contrast between the optimal strategy of an infinitely patient planner in comparison to an impatient one. An infinitely patient planner prefers to protect an action which is more likely to be successful from some initial negative shocks, whereas she prefers to spread as much as possible an action which is more likely to be unsuccessful, trying to take advantage of a few positive shocks in the first periods. When the probability of success is low, she knows that by concentrating all the targeted agents together, a lot of positive shocks will be needed in order to capture the whole population, which is rather improbable for an action that is expected to be often unsuccessful.

5. Extensions

In this section, we briefly present the results of several extensions that address specific questions related to the problem of interest.[24]

A natural question that arises from the contrast between the optimal behavior of an impatient and an infinitely patient planner is what happens for intermediate levels of patience. First of all, we observe that the expected diffusion is very sensitive to small changes in the initial configuration and therefore it becomes particularly hard to construct a general strategy for all intermediate levels of patience. Nevertheless, we discuss the optimal targeting strategy of a planner who cares about the diffusion of action B after three periods and we obtain an enlightening result. Namely, if $r > 1$, then in some cases the planner prefers to spread the targeted agents in groups consisting of four, instead of two, agents. Whereas, if $r < 1$, she always needs to compare between the two extreme cases, i.e. concentrating all of them in one group or spread them to as many groups as possible. This result provides a useful starting point to understand how the optimal targeting strategy changes as the planner becomes more patient.

Furthermore, we discuss what happens if t becomes endogenous. To do this it is necessary to define explicitly the profit function of the planner. We focus on linear cost of targeting and we find that for sufficiently low (high) unit cost the planner prefers to target all the (none of the) agents. For intermediate unit costs (with the bounds depending on the different parameters, but mostly on p) the planner prefers to target an intermediate number of agents, $0 < t^* < n$. This result is quite

---

[24]A thorough discussion of these extensions can be found in the Online Appendix.

intuitive if one thinks that targeting additional agents increases the probability of total diffusion of action B, but never ensures it. Therefore, if a planner has targeted sufficiently many agents, then the additional expected benefit from targeting one more might not compensate the cost of targeting this agent.

Another crucial aspect is the practical meaning of infinite patience. In reality, no planner can wait literally infinitely many periods. We try to identify the expected time before total diffusion of one action occurs and how the planner's strategies affect this expected time. Not surprisingly, we find that a larger number of groups leads to faster total diffusion and therefore the optimal strategy for $r > 1$ has the drawback of maximizing also the expected time before total diffusion occurs. On the other hand, for $r < 1$ the optimal strategy is also the one that leads to the fastest expected time of total diffusion. An interesting feature, which is in line with standard results in the analysis of random walks, is that the expected time of total diffusion explodes as r gets very close to 1. In general we find that for different configurations the expected time to total diffusion may vary substantially and therefore one must be very careful when acting as an infinitely patient planner.

In addition to this, one might argue that fast total diffusion might not be always optimal for the planner, because this might lead to a fast disappearance of her desired action from the population. For this reason, we perform a simple exercise where the planner obtains a positive gain from total diffusion of her preferred action and a negative one from total diffusion of the alternative action. We find that for extreme values (either high or low) of p the optimal strategies of the planner might be determined by the speed at which diffusion occurs, rather than by the probability of successful diffusion. In fact, irrespectively of the level of patience, if p is very low the planner prefers to locate all the agents together, whereas if it is very high prefers to spread them as much as possible. The intuition is similar to that described for the impatient planner. This can be also seen as an additional attempt to characterize optimal strategies for intermediate levels of patience, but should be approached with caution since it is only a partial result.

Moreover, we discuss cases where inertia is possible. This generalization allows us to capture some realistic scenarios, such as the possibility of both actions having equally good realizations, existence of switching costs and some forms of conformity. Such settings can be captured by allowing $q \neq 1 - p$, where q is the probability of action A being more successful than action B. As we have already mentioned, our results are not affected by this feature.

A similar question to this would be how the results would be affected either if the payoffs were not perfectly correlated in each period, or if updating was not occurring simultaneously.Extending the analysis to independent shocks would require further structure on the realized payoffs, which would made necessary the use of additional parameters. This would not be a problem per se, but in addition to this we could not simplify the problem reducing it to one–dimensional random walks. Trying to work with these multidimensional walks would lead to horrendous expressions even for very simple cases.[25] Intuitively, either of these two scenarios would induce more

---

[25]An expression for the probabilities of diffusion for both actions regarding a very similar process and only for one group of initial adopters can be found in Donnelly and Welsh (1983), where the complexity of the problem already becomes more apparent.

realizations of random shocks, which in turn should favor quantitatively the action that is more likely to be successful. However, given that all the targeting strategies would be affected on the same way, the results should not be affected qualitatively. Nevertheless, one should be cautious when making such claims and a concrete answer would be possible only after a systematic analysis of these scenarios.

Additionally, we discuss the optimal targeting strategy of a planner for some slightly modified social structures, so as to get an idea of how the results would be affected by the presence of central agents. First, we study the line assuming that the planner can target a single group of agents. Once again, we find a sharp contrast between the optimal targeting strategies of an impatient and an infinitely patient planner. In particular, for an infinitely patient planner, if $r > 1$ it is optimal to target one of the two corners, whereas if $r < 1$ it is optimal to target the agents who are located around the center. To the contrary, for an impatient planner, if $r > 1$ it is optimal to target any segment of the line that does not include any of the corner agents, whereas if $r < 1$ it is optimal to target one of the two corners. In this part of our analysis, we also drop the assumption of groups having an even number of agents and we show why dropping this assumption complicates our analysis without providing additional insights. We also discuss briefly the star, in which the vast importance of very central agents becomes apparent.

Finally, we run a set of simulations to test the robustness of our results to the addition of a few links in the circle and we find that our conclusions remain valid. In particular, if $r > 1$ it is always optimal to target only one group of connected agents, with the optimal location of the group depending on the position of the additional links. Conversely, if $r < 1$ it is almost always optimal to target the subset of agents that minimizes the number of successful draws needed to capture the whole population. This is in line with our previous findings and provides intuitions which can be useful for the study of more general network structures.

6. Conclusion

We have analyzed the optimal intervention of a planner who seeks to maximize the diffusion of an action in a society where agents imitate successful past behavior of their neighbors. It turns out that there is room for strategic targeting even in environments where all agents are completely identical. We find that the optimal decision depends almost completely on two parameters. On the likelihood r of the preferred action being more successful and on how patient the planner is. Changes in these two parameters lead to completely opposite optimal behavior.

Assuming that the planner knows the exact social structure, as well as the location of each agent, might seem quite strong and tough to be satisfied in large populations. Nevertheless, it is exactly this assumption that allows us to focus on the importance of the targeted agents' relative positions, which in turn brings into consideration several new insights on the planner's problem. One can see this paper as a first step on understanding the effect of agents' exact positions in a society for certain diffusion processes and use it as a benchmark for problems where this assumption can be partially relaxed. A similar argument applies to the assumption regarding the particular behavioral rule, which turns attention from Bayesian (see Gale and Kariv, 2003) and DeGroot (see DeGroot, 1974, Golub and Jackson, 2010)

learning rules towards more naive ones. A natural step one would consider is to construct targeting mechanisms that are robust to uncertainty over the social structure and behavioral rule. In particular, an interesting avenue for future research would be to identify targeting strategies that are able to perform well (are close to optimal) for a broad family of different structures and under uncertainty over the rules that govern the behavior of the agents.

Moreover, throughout the paper we have disregarded completely the risk aversion of the planner. We have assumed the planner to be risk neutral, caring only about the expected number of adopters. For a risk averse planner, we would expect the optimal behavior to contain more dispersed targets than for the risk neutral one, but this remains an open question for future research.

The current paper constitutes a first attempt to explore targeting possibilities in settings where agents imitate successful behavior. A natural extension would be to explore which of the current features are still present and which of these are changing when passing to more general social structures. It is apparent that centrality features arising in more complex structures will play an important role. However the exact characteristics remain to be studied.

Appendix - Proofs

Proof of Proposition 2. Under (A1), $\frac{s}{m}$ and $\frac{t}{m}$ are even numbers. Then, the process is equivalent to having a line of $\frac{n}{2m}$ agents, consisting of one group of $\frac{t}{2m}$ adjacent agents choosing B and another group of $\frac{s}{2m}$ adjacent agents choosing A.
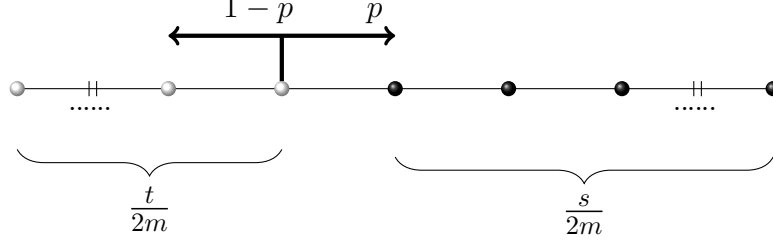

Figure 8: The random walk that describes the process in the symmetric case.

By Lemma 1, the probability of successful diffusion becomes:

$$P_B(m|s,t,n,r) = \frac{r^{\frac{n}{2m}} - r^{\frac{s}{2m}}}{r^{\frac{n}{2m}} - 1}$$

Despite the fact, that we are interested only in the integer values of m, t and n, the function $P_B(\cdot)$ is well-defined and smooth for all $r \neq 1$ and $m \geq 1$. Hence, we can check its monotonicity by differentiating with respect to m.

$$\frac{dP_B}{dm} = \frac{\left[r^{\frac{n}{2m}}\ln r\left(-\frac{n}{2m^2}\right) - r^{\frac{s}{2m}}\ln r\left(-\frac{s}{2m^2}\right)\right]\left(r^{\frac{n}{2m}} - 1\right) - \left(r^{\frac{n}{2m}} - r^{\frac{s}{2m}}\right)\left[r^{\frac{n}{2m}}\ln r\left(-\frac{n}{2m^2}\right)\right]}{\left(r^{\frac{n}{2m}} - 1\right)^2}$$

$$= \frac{\ln r}{\left(r^{\frac{n}{2m}} - 1\right)^2}\left[\left(-\frac{n}{2m^2}r^{\frac{n}{2m}} + \frac{s}{2m^2}r^{\frac{s}{2m}}\right)\left(r^{\frac{n}{2m}} - 1\right) - \left(r^{\frac{n}{2m}} - r^{\frac{s}{2m}}\right)\left(-\frac{n}{2m^2}r^{\frac{n}{2m}}\right)\right]$$

$$= \frac{\ln r}{\left(r^{\frac{n}{2m}} - 1\right)^2}\left(-\frac{n}{2m^2}r^{\frac{2n}{2m}} + \frac{s}{2m^2}r^{\frac{s+n}{2m}} + \frac{n}{2m^2}r^{\frac{n}{2m}} - \frac{s}{2m^2}r^{\frac{s}{2m}} + \frac{n}{2m^2}r^{\frac{2n}{2m}} - \frac{n}{2m^2}r^{\frac{s+n}{2m}}\right)$$

$$= \frac{\ln r}{2m^2\left(r^{\frac{n}{2m}} - 1\right)^2}\left(sr^{\frac{s+n}{2m}} + nr^{\frac{n}{2m}} - sr^{\frac{s}{2m}} - nr^{\frac{s+n}{2m}}\right)$$

$$= \frac{\ln r}{2m^2\left(r^{\frac{n}{2m}} - 1\right)^2}\left[sr^{\frac{s}{2m}}\left(r^{\frac{n}{2m}} - 1\right) - nr^{\frac{n}{2m}}\left(r^{\frac{s}{2m}} - 1\right)\right]$$

$$= \frac{\ln r}{2m^2\left(r^{\frac{n}{2m}} - 1\right)^2}\left(r^{\frac{n}{2m}} - 1\right)\left(r^{\frac{s}{2m}} - 1\right)\left(\frac{sr^{\frac{s}{2m}}}{r^{\frac{s}{2m}} - 1} - \frac{nr^{\frac{n}{2m}}}{r^{\frac{n}{2m}} - 1}\right)$$

If we call $\frac{s}{2m} = s'$ and $\frac{n}{2m} = n'$, then the following lemma helps us conclude the argument.

Lemma 2. $f(x) = \frac{2mxr^x}{r^x - 1}$ *is strictly increasing for $x \geq 1$, for all $r \neq 1$ and $m \geq 1$*

Proof. Let $r \neq 1$ and $m \geq 1$, then

$$\frac{df}{dx} = \frac{2m}{(r^x - 1)^2}\left[(r^x + xr^x\ln r)(r^x - 1) - xr^x(r^x)\ln r\right]$$

$$= \frac{2m}{(r^x - 1)^2}(r^{2x} - r^x - xr^x\ln r) = \frac{2mr^x}{(r^x - 1)^2}(r^x - 1 - x\ln r) > 0 \text{ for all } x \geq 1$$

To show this, we define $g(x) = r^x - 1 - x \ln r$, which is strictly increasing for $x \geq 1$ because $\frac{dg}{dx} = r^x \ln r - \ln r = \ln r (r^x - 1) > 0$. So it attains minimum for $x = 1$, which is $g(1) = r - 1 - \ln r$. Moreover, $g(1) > 0$ for all $r \neq 1$ because it holds that $h(r) = r - 1 - \ln r > 0$ for all $r \neq 1$. This holds because $\frac{dh}{dr} = 1 - \frac{1}{r}$ is strictly positive when $r > 1$ and strictly negative when $r < 1$. So, h attains global minimum for $r = 1$, with value $h(1) = 0$. Hence, $g(x) > 0$ for all $x \geq 1$, which means that also $\frac{df}{dx} > 0$ for all $x \geq 1$ and this concludes the argument. $\square$

By Lemma 2, given that $n > s$, we get that $\left(\frac{sr^{\frac{s}{2m}}}{r^{\frac{s}{2m}} - 1} - \frac{nr^{\frac{n}{2m}}}{r^{\frac{n}{2m}} - 1} \right) < 0$ always, so we can conclude that $\frac{dP_B}{dm} < 0$ if $r > 1$ and $\frac{dP_B}{dm} > 0$ if $r < 1$. Hence, for $r > 1$ the $P_B(m|s,t,n,r)$ is decreasing in m, so $\arg\max_m P_B(m|s,t,n,r) = 1$, i.e. the optimal choice is to target a single group of targeted agents. On the other hand, for $r < 1$, P is increasing in m, so we would like to split the targeted agents in as many groups as possible, i.e. $\arg\max_m P_B(m|s,t,n,r) = \min\{s/2, t/2\}$. $\square$

Proof of Proposition 3. First, we have to construct the probability of successful diffusion. For $r \neq 1$, the process again can be described as a sequence of random walks with absorbing barriers. At the beginning, we have a random walk of $(s_1 + t_1)/2$ nodes, starting from node $t_1/2$, until it disappears either t_1 or s_1. By Lemma 1, the probability of successful absorption of this walk is $\frac{r^{\frac{s_1+t_1}{2}} - r^{\frac{s_1}{2}}}{r^{\frac{s_1+t_1}{2}} - 1}$. In case of successful absorption we get a random walk of $n/2$ nodes starting from the node $(t + 2s_1)/2$. Otherwise, in case of unsuccessful absorption we get a random walk of $n/2$ nodes as well, but starting from node $(t - 2t_1)/2$. Again by Lemma 1, the probabilities of successful absorption in these two scenarios are $\frac{r^{\frac{n}{2}} - r^{\frac{n-t-2s_1}{2}}}{r^{\frac{n}{2}} - 1}$ and $\frac{r^{\frac{n}{2}} - r^{\frac{s+2t_1}{2}}}{r^{\frac{n}{2}} - 1}$ respectively. If the second walk, in either of the two scenarios, is unsuccessful then action B disappears. Figure 9 depicts the process we just described. Notice that (A1) solves all the problems of divisibility.

The histories that lead to full diffusion of action B are (i) success in both the first and the second walk and (ii) failure in the first and success in the second walk. Therefore the probability of successful diffusion for $r \neq 1$ can be written as follows:

$$P_B(s_1, t_1 | s, t, n, r) = \frac{r^{\frac{s_1+t_1}{2}} - r^{\frac{s_1}{2}}}{r^{\frac{s_1+t_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-t-2s_1}{2}}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{s_1}{2}} - 1}{r^{\frac{s_1+t_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{s+2t_1}{2}}}{r^{\frac{n}{2}} - 1}$$

Now, we compute the derivatives with respect to t_1 and s_1. As usually, we are only interested in integer points, but the function P_B is a well-behaved smooth function for $r \neq 1$, so we can study its monotonicity.
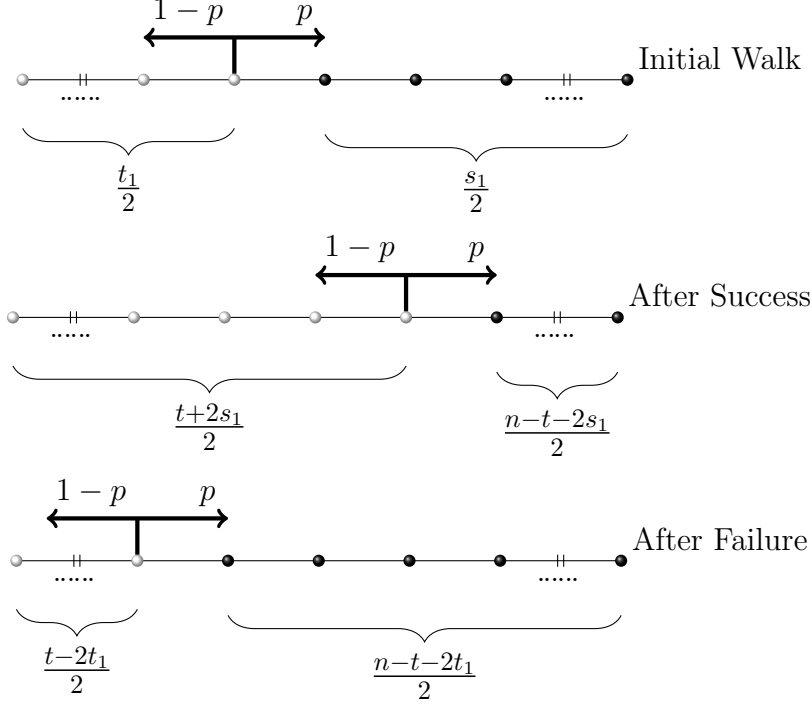
Figure 9: The random walks that describe the process in the asymmetric case with two groups.

$$\frac{\partial P_B}{\partial s_1} = \left[\frac{r^{\frac{n}{2}}\left(r^{\frac{t_1}{2}} - 1\right)}{r^{\frac{n}{2}} - 1} \right] \left[\frac{\left(\frac{\ln r}{2}r^{\frac{s_1}{2}} + \frac{\ln r}{2}r^{\frac{-s_1-t}{2}}\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right) - \left(r^{\frac{s_1}{2}} - r^{\frac{-s_1-t}{2}}\right)\frac{\ln r}{2}r^{\frac{s_1+t_1}{2}}}{\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} \right] +$$

$$+ \left(\frac{r^{\frac{n}{2}} - r^{\frac{s+2t_1}{2}}}{r^{\frac{n}{2}} - 1} \right) \left[\frac{\frac{\ln r}{2}r^{\frac{s_1}{2}}\left(r^{\frac{s_1+t_1}{2}} - 1\right) - \left(r^{\frac{s_1}{2}} - 1\right)\frac{\ln r}{2}r^{\frac{s_1+t_1}{2}}}{\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} \right] =$$

$$= \frac{r^{\frac{n}{2}}\left(r^{\frac{t_1}{2}} - 1\right)\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{2s_1+t_1}{2}} + r^{\frac{-t+t_1}{2}} - r^{\frac{s_1}{2}} - r^{\frac{-t-s_1}{2}} - r^{\frac{2s_1+t_1}{2}} + r^{\frac{-t+t_1}{2}} \right) +$$

$$+ \frac{\left(r^{\frac{n}{2}} - r^{\frac{s+2t_1}{2}}\right)\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{2s_1+t_1}{2}} - r^{\frac{s_1}{2}} - r^{\frac{2s_1+t_1}{2}} + r^{\frac{s_1+t_1}{2}} \right) =$$

$$= \frac{\ln r\, r^{\frac{n}{2}}\left(r^{\frac{t_1}{2}} - 1\right)\left[2r^{\frac{-t+t_1}{2}} - r^{\frac{s_1}{2}} - r^{\frac{-t-s_1}{2}} + r^{\frac{s_1}{2}} - r^{\frac{-t+2t_1+s_1}{2}}\right]}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} =$$

$$= \frac{\ln r\left(r^{\frac{t_1}{2}} - 1\right)\left(2r^{\frac{n-t+t_1}{2}} - r^{\frac{n-t-s_1}{2}} - r^{\frac{s+s_1+2t_1}{2}}\right)}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} = \frac{\ln r\left(r^{\frac{t_1}{2}} - 1\right)r^{\frac{s-s_1}{2}}\left(2r^{\frac{s_1+t_1}{2}} - 1 - r^{s_1+t_1}\right)}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} =$$

$$= \frac{\ln r\left(r^{\frac{t_1}{2}} - 1\right)r^{\frac{s-s_1}{2}}\left[-\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2\right]}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} = -\frac{\ln r\left(r^{\frac{t_1}{2}} - 1\right)r^{\frac{s-s_1}{2}}}{2\left(r^{\frac{n}{2}} - 1\right)} \begin{cases} < 0 & \text{if } r > 1 \\ > 0 & \text{if } r < 1 \end{cases}$$

22

Analogously for t_1 we have:

$$
\frac{\partial P_B}{\partial t_1} = \frac{\ln r \left[(r^{\frac{n+s_1}{2}} - r^{\frac{s-s_1}{2}}) r^{\frac{t_1}{2}} (r^{\frac{s_1+t_1}{2}} - 1) - (r^{\frac{n+s_1}{2}} - r^{\frac{s-s_1}{2}})(r^{\frac{t_1}{2}} - 1) r^{\frac{s_1+t_1}{2}} \right]}{2(r^{\frac{n}{2}} - 1)(r^{\frac{s_1+t_1}{2}} - 1)^2} +
$$

$$
+ \frac{\ln r \left[(r^{\frac{s_1}{2}} - 1)(-2r^{\frac{s+2t_1}{2}})(r^{\frac{s_1+t_1}{2}} - 1) - (r^{\frac{s_1}{2}} - 1)(r^{\frac{n}{2}} - r^{\frac{s+2t_1}{2}}) r^{\frac{s_1+t_1}{2}} \right]}{2(r^{\frac{n}{2}} - 1)(r^{\frac{s_1+t_1}{2}} - 1)^2} =
$$

$$
= \frac{\ln r}{2(r^{\frac{n}{2}} - 1)(r^{\frac{s_1+t_1}{2}} - 1)^2} \left(r^{\frac{n+s_1}{2}} - r^{\frac{s-s_1}{2}} \right) \left(-r^{\frac{t_1}{2}} + r^{\frac{s_1+t_1}{2}} \right) +
$$

$$
+ \frac{\ln r}{2(r^{\frac{n}{2}} - 1)(r^{\frac{s_1+t_1}{2}} - 1)^2} \left(r^{s_1/2} - 1 \right) \left(-r^{\frac{s+s_1+3t_1}{2}} + 2r^{\frac{s+2t_1}{2}} - r^{\frac{n+s_1+t_1}{2}} \right) =
$$

$$
= \frac{\ln r \left(r^{\frac{s_1}{2}} - 1 \right) \left(-r^{\frac{s+t_1-s_1}{2}} - r^{\frac{s+s_1+3t_1}{2}} + 2r^{\frac{s+2t_1}{2}} \right)}{2(r^{\frac{n}{2}} - 1) \left(r^{\frac{s_1+t_1}{2}} - 1 \right)^2} =
$$

$$
= \frac{\ln r \left(r^{\frac{s_1}{2}} - 1 \right) r^{\frac{s+t_1}{2}}}{2r^{\frac{s_1}{2}} \left(r^{\frac{n}{2}} - 1 \right) \left(r^{\frac{s_1+t_1}{2}} - 1 \right)^2} \left(2r^{\frac{s_1+t_1}{2}} - r^{\frac{2s_1+2t_1}{2}} - 1 \right) =
$$

$$
= \frac{\ln r \left(r^{\frac{s_1}{2}} - 1 \right) r^{\frac{s+t_1}{2}} \left[-\left(r^{\frac{s_1+t_1}{2}} - 1 \right)^2 \right]}{2r^{\frac{s_1}{2}} \left(r^{\frac{n}{2}} - 1 \right) \left(r^{\frac{s_1+t_1}{2}} - 1 \right)^2} = -\frac{\ln r \left(r^{\frac{s_1}{2}} - 1 \right) r^{\frac{s+t_1}{2}}}{2r^{\frac{s_1}{2}} \left(r^{\frac{n}{2}} - 1 \right)} \begin{cases} < 0 & \text{if } r > 1 \\ > 0 & \text{if } r < 1 \end{cases}
$$

Hence, given that $0 \leq s_1 \leq s_2$ and $0 \leq t_1 \leq t_2$ we conclude that for $r > 1$ the optimal targeting strategy is $(s_1, t_1) = (0, 0)$, whereas for $r < 1$ it is $s_2 - s_1 \leq 2$ and $t_2 - t_1 \leq 2$.

$\square$

Proof of Theorem 1. For the case of $r > 1$ we proceed by induction. First, we recall the result by Proposition 2, which states that if we can target up to two groups, then the optimal choice is to concentrate all the targeted agents in one group. Remember also that $s_1 \leq s_2 \leq s_3$ and $t_1 \leq t_2 \leq t_3$. Now suppose that we can target up to three groups ($m \leq 3$). Then again at first we are interested in the two smallest groups of each type and we have the following random walk:
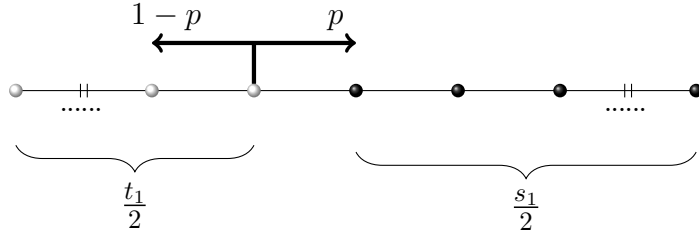
The system fluctuates in this direction until either s_1 or t_1 disappears. Depending on the successful or unsuccessful absorption of this walk we get one of the configurations depicted in Figure 10, with only two groups of each type left.

By Proposition 2, in both of these cases we know that the optimal choice would be to eliminate one of the two groups of adopters. Hence, we would like to choose s_2 and t_2 (as functions of s_1 and t_1 respectively), in such a way that the probability of diffusion is maximized in both of these cases. Recalling that $s_1 \le s_2 \le s_3$ and $t_1 \le t_2 \le t_3$, we see that this can be achieved if $s_2 = s_1$ and $t_2 = t_1$, where the optimal s_1 and t_1 remained to be determined. Notice that, by construction, $s_3 = s - s_1 - s_2$ and $t_3 = t - t_1 - t_2$.

So now, we can rewrite the probability of diffusion as a function of s_1 and t_1 only.

$$P_B(s_1, t_1 | s, t, n, r, m = 3) = \frac{r^{\frac{t_1+s_1}{2}} - r^{\frac{s_1}{2}}}{r^{\frac{t_1+s_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-t-3s_1}{2}}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{s_1}{2}} - 1}{r^{\frac{t_1+s_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{s+3t_1}{2}}}{r^{\frac{n}{2}} - 1}$$

Like before we study the monotonicity of the function with respect to s_1 and t_1

$$\frac{\partial P_B}{\partial s_1} = \frac{\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} \left[\left(r^{\frac{t_1}{2}} - 1\right)\left(r^{\frac{n+s_1}{2}} + 2r^{\frac{s-2s_1}{2}}\right) + \left(r^{\frac{n}{2}} - r^{\frac{s+3t_1}{2}}\right)r^{\frac{s_1}{2}}\right]\left(r^{\frac{s_1+t_1}{2}} - 1\right) -$$

$$- \frac{\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2} \left[\left(r^{\frac{t_1}{2}} - 1\right)\left(r^{\frac{n+s_1}{2}} - r^{\frac{s-2s_1}{2}}\right) + \left(r^{\frac{n}{2}} - r^{\frac{s+3t_1}{2}}\right)\left(r^{\frac{s_1}{2}} - 1\right)\right]r^{\frac{s_1+t_1}{2}} =$$

$$= \frac{\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{t_1}{2}} - 1\right)\left[\left(r^{\frac{n+s_1}{2}} + 2r^{\frac{s-2s_1}{2}}\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right) - \left(r^{\frac{n+s_1}{2}} - r^{\frac{s-2s_1}{2}}\right)r^{\frac{s_1+t_1}{2}}\right] +$$

$$+ \frac{\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{n}{2}} - r^{\frac{s+3t_1}{2}}\right)\left[r^{\frac{s_1}{2}}\left(r^{\frac{s_1+t_1}{2}} - 1\right) - \left(r^{\frac{s_1}{2}} - 1\right)r^{\frac{s_1+t_1}{2}}\right] =$$

$$= \frac{\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{t_1}{2}} - 1\right)\left(3r^{\frac{s+t_1-s_1}{2}} - 2r^{\frac{s-2s_1}{2}} - r^{\frac{s+s_1+3t_1}{2}}\right) =$$

$$= \frac{r^{\frac{s}{2}}\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{t_1}{2}} - 1\right)\left(3r^{\frac{t_1-s_1}{2}} - 2r^{\frac{-2s_1}{2}} - r^{\frac{s_1+3t_1}{2}}\right) =$$

$$= \frac{r^{\frac{s}{2}}\ln r}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left(r^{\frac{t_1}{2}} - 1\right)\left[-\frac{\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2\left(r^{\frac{s_1+t_1}{2}} + 2\right)}{r^{s_1}}\right] =$$

$$= -\frac{r^{\frac{s}{2}}\ln r\left(r^{\frac{t_1}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} + 2\right)}{2\left(r^{\frac{n}{2}} - 1\right)r^{s_1}} < 0 \text{ for } r > 1$$

Hence the optimal choice is $s_1 = s_2 = 0$ and $s_3 = s$.

24
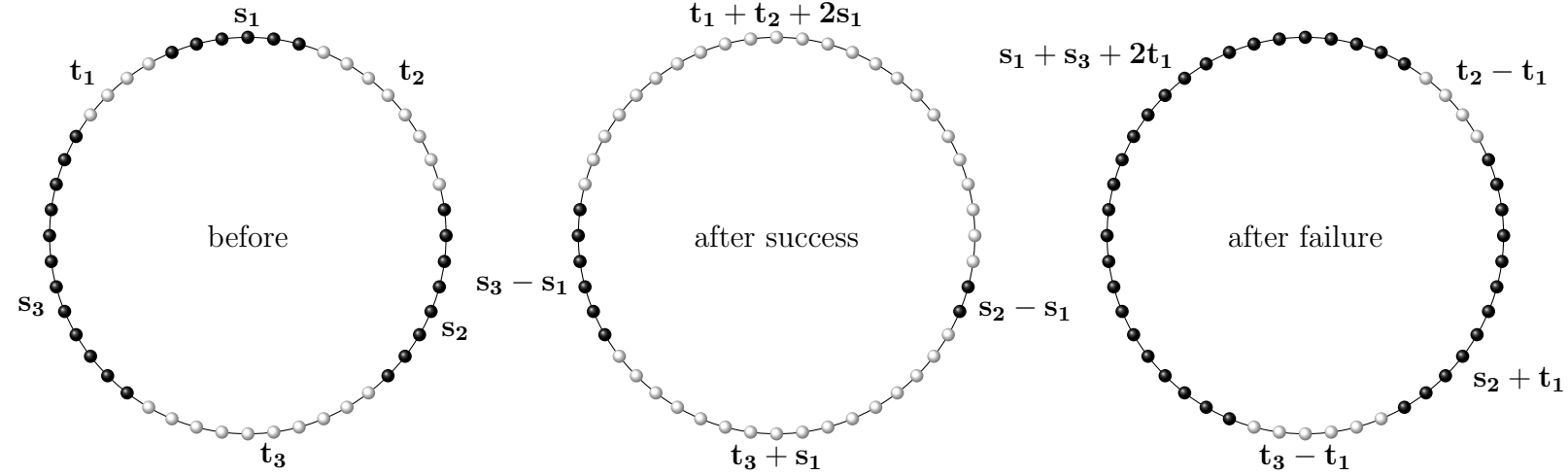
Figure 10: Configurations with 3 groups of targeted agents.

Analogously for t_1 we get the following:

$$\frac{\partial P_B}{\partial t_1} = \frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-2s_1}{2}}\right)\left[r^{\frac{t_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)-r^{\frac{s_1+t_1}{2}}\left(r^{\frac{t_1}{2}}-1\right)\right]+$$

$$+\frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{s_1}{2}}-1\right)\left[-3r^{\frac{s+3t_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)-\left(r^{\frac{n}{2}}-r^{\frac{s+3t_1}{2}}\right)r^{\frac{s_1+t_1}{2}}\right]=$$

$$=\ln r\frac{\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-2s_1}{2}}\right)\left(r^{\frac{s_1+t_1}{2}}-r^{\frac{t_1}{2}}\right)+\left(r^{\frac{s_1}{2}}-1\right)\left(3r^{\frac{s+3t_1}{2}}-2r^{\frac{s+s_1+4t_1}{2}}-r^{\frac{n+s_1+t_1}{2}}\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(3r^{\frac{s+3t_1}{2}}-2r^{\frac{s+s_1+4t_1}{2}}-r^{\frac{s+t_1-2s_1}{2}}\right)=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s}{2}}r^{\frac{t_1}{2}}}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(3r^{\frac{2t_1}{2}}-2r^{\frac{s_1+3t_1}{2}}-r^{\frac{-2s_1}{2}}\right)=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s}{2}}r^{\frac{t_1}{2}}}{2r^{s_1}\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(3r^{\frac{2s_1+2t_1}{2}}-2r^{\frac{3s_1+3t_1}{2}}-1\right)=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s_2-s_1}{2}}r^{\frac{t_1}{2}}\left[-\left(r^{\frac{s_1+t_1}{2}}-1\right)^2\left(2r^{\frac{s_1+t_1}{2}}+1\right)\right]}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}<0 \text{ for } r>1.$$

The last step comes from the observation that $-2x^3+3x^2-1=-(x-1)^2(2x+1)$, where in this case $x=r^{\frac{s_1+t_1}{2}}$. Hence, P_B is always decreasing in t_1, and given that $t_1=t_2$ the optimal choice is $t_1=t_2=0$ and $t_3=t$. This concludes the argument

for the case where $m = 3$. We will generalize this argument by induction.

Formally, given that the argument holds for $m = 3$, it suffices to show that if it holds for $m = k - 1 \geq 3$ then it holds as well for $m = k$.

At the beginning of the process we care only about the two smallest groups of each type s_1 and t_1 and the system fluctuates, as in the previous cases, until one of the two disappears. Figure 11 shows the possible configurations after the disappearance of either s_1 or t_1. The location of the groups around the circle comes without loss of generality.
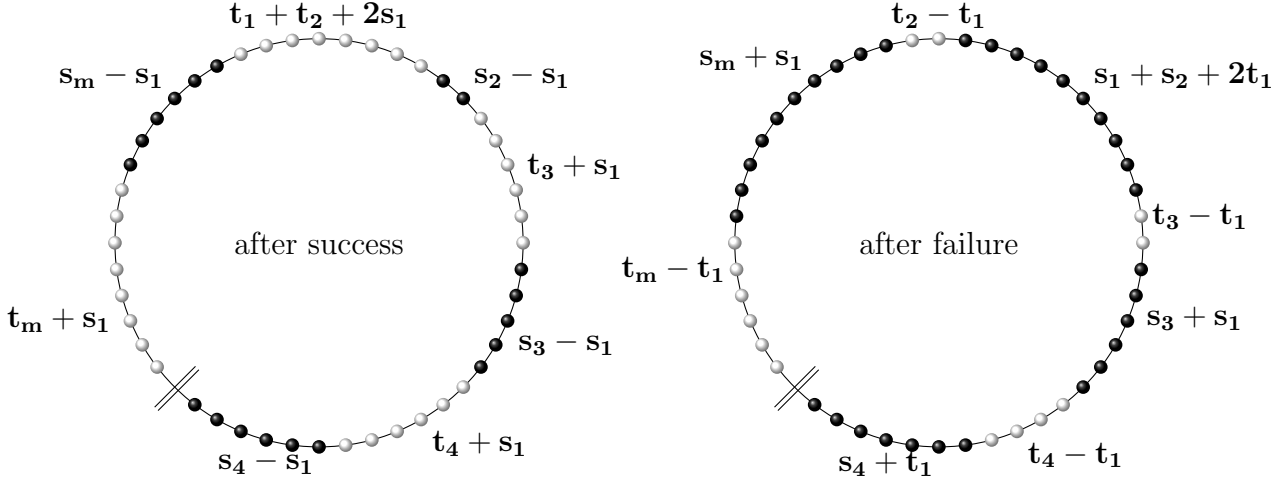

Figure 11: Configurations after the disappearance of s_1 or t_1 with m groups.

Given that the argument holds for $k-1$ groups then we know that $s_1 = \cdots = s_{k-1}$ and $t_1 = \cdots = t_{k-1}$. Therefore, we only need to find the optimal size for s_1 and t_1. The probability of diffusion becomes:

$$P_B(s_1, t_1 | s, t, n, r, m = k) = \frac{r^{\frac{t_1+s_1}{2}} - r^{\frac{s_1}{2}}}{r^{\frac{t_1+s_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-t-ks_1}{2}}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{s_1}{2}} - 1}{r^{\frac{t_1+s_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{s+kt_1}{2}}}{r^{\frac{n}{2}} - 1}$$

By calculations which are omitted because they are identical to the case where $m = 3$, we get:

$$\frac{\partial P_B}{\partial s_1} = \frac{r^{\frac{s}{2}} \left(r^{\frac{t_1}{2}} - 1 \right) \ln r}{2 r^{(k-1)\frac{s_1}{2}} \left(r^{\frac{n}{2}} - 1 \right) \left(r^{\frac{s_1+t_1}{2}} - 1 \right)^2} \left[kr^{\frac{s_1+t_1}{2}} - r^{k\frac{s_1+t_1}{2}} - (k-1) \right] \leq 0 \text{ for } r > 1$$

and equality holds only if $s_1 = t_1 = 0$. For the argument to hold we need $kr^{\frac{s_1+t_1}{2}} - r^{k\frac{s_1+t_1}{2}} - (k-1)$ to be negative. So, let $x = r^{\frac{s_1+t_1}{2}}$ and take the function $f(x) = kx - x^k - (k-1)$ for some $k \geq 3$ and $x \geq 0$. Now, $\frac{df}{dx} = k - kx^{k-1}$ is positive if $x < 1$ and negative if $x > 1$, hence f attains global max at $x = 1$ equal to

$f(1) = k - 1^k - (k-1) = 0$, hence $f(x) < 0$ for all $x \neq 1$. Now given that $x = r^{\frac{s_1+t_1}{2}}$, with $r > 1$ and $s_1, t_1 \geq 0$ the function is always strictly negative and becomes equal to zero only when $s_1 = t_1 = 0$. So, the optimal choice is $s_1 = \cdots = s_{k-1} = 0$ and $s_k = s$.

Analogously for t_1 we get that:

$$\frac{\partial P_B}{\partial t_1} = \frac{r^{\frac{s}{2}} r^{\frac{t_1}{2}} \left(r^{\frac{s_1}{2}} - 1 \right) \ln r}{2 r^{(k-1)\frac{s_1}{2}} \left(r^{\frac{n}{2}} - 1 \right) \left(r^{\frac{s_1+t_1}{2}} - 1 \right)^2} \left[k r^{(k-1)\frac{s_1+t_1}{2}} - (k-1) r^{k\frac{s_1+t_1}{2}} - 1 \right] \leq 0 \text{ for } r > 1$$

again equality holds only when $s_1 = t_1 = 0$ and to ensure the result we need that $k r^{(k-1)\frac{s_1+t_1}{2}} - (k-1) r^{k\frac{s_1+t_1}{2}} - 1 \leq 0$ for all s_1 and t_1 with equality holding only in case they are both equal to zero. As before, let $x = r^{\frac{s_1+t_1}{2}}$ and define the function $g(x) = k x^{k-1} - (k-1)x^k - 1$. Then $\frac{df}{dx} = k(k-1)x^{k-2} - k(k-1)x^{k-1}$ which is strictly negative for $x > 1$ and strictly positive for $x < 1$, then g attains unique maximum at $x = 1$ equal to $g(1) = 0$. So $g(x) < 0$ for all $x \neq 1$. Given again that $x = r^{\frac{s_1+t_1}{2}}$ then $x = 1$ only if $s_1 = t_1 = 0$. So again the optimal choices are $t_1 = \cdots = t_{k-1} = 0$ and $t_k = t$, which completes the inductive argument. Hence, when $r > 1$, for any possible number of groups m, the optimal choice is to concentrate all the targeted agents in one group, i.e. $s_1 = \cdots = s_{m-1} = 0$ and $s_m = s$, as well as $t_1 = \cdots = t_{m-1} = 0$ and $t_m = t$.

Now, we turn our attention towards the case where $r < 1$. We tackle this case in a different way. Namely, we construct an upper bound for the probability of successful diffusion and we show that the actual probability is equal to this upper bound for the same configurations that this upper bound is maximized. Hence this has to be the maximum value of the probability as well.

In order to proceed, we need to construct the upper bound for the value of the probability of successful diffusion of action B. We solve it first for $t < s$ and then for $s < t$.

Let $t < s$, then allowing for the existence of $m = \frac{t}{2}$ groups, the circle will have the form of Figure 12. Notice that, the fact that s_i can have size equal to zero, allows us to construct any possible configuration. For example, if $s_1 = 0$ then the two groups next to s_1 merge to one group with four agents. According to this structure, the process will initially follow a random walk with $\frac{s_1}{2}$ black steps and one white. By Lemma 1, the probability of success in this first walk is equal to $\frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1}$. In Figure 12 we also see how the society will look like if the first walk is successful. Unsuccessful absorption in the first walk leads to the disappearance of action B from the population, because all t's have the same size. After success, the process will move according to the random walk of Figure 13, with the probability of success in this walk is equal to $\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1}$.

We depict as well the two possible configurations that arise after success or failure in the second walk (see Figure 14). It is important to notice that the probability of successful diffusion after two successes is obviously weakly lower than 1 and it is
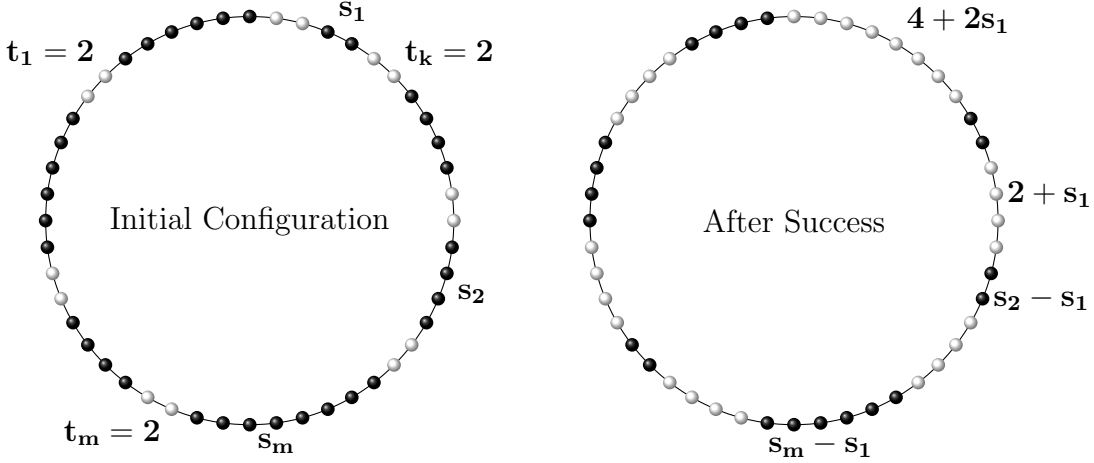
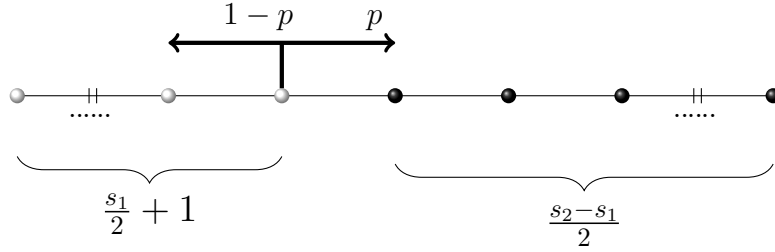Figure 12: General Initial Configuration and Result After Success, for $t < s$.

Figure 13: Random Walk After Success.

strictly lower as long as $s_m - s_1 > 2$, where s_m is the size of the largest group and s_1 is the size of the smallest one.

Hence, we can construct the probability of successful diffusion, which is equal to:

$$P_B(\cdot) = \frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1} \left[\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1} P_B(\cdot|s,s) + \frac{r^{\frac{s_2-s_1}{2}} - 1}{r^{\frac{s_2}{2}+1} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}}{r^{\frac{n}{2}} - 1} \right]$$

where $P_B(\cdot|s,s)$ stands for the probability of diffusion of B after two successes in the first two random walks. Given that $P_B(\cdot|s,s) \leq 1$ we get the following upper bound of P_B, denoted by $\widetilde{P_B(\cdot)}$, which is equal to:

$$\widetilde{P_B(\cdot)} = \frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1} \left[\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1} + \frac{r^{\frac{s_2-s_1}{2}} - 1}{r^{\frac{s_2}{2}+1} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}}{r^{\frac{n}{2}} - 1} \right]$$

Before performing any calculations it is important to simplify the expression of $\widetilde{P_B(\cdot)}$. Specifically,
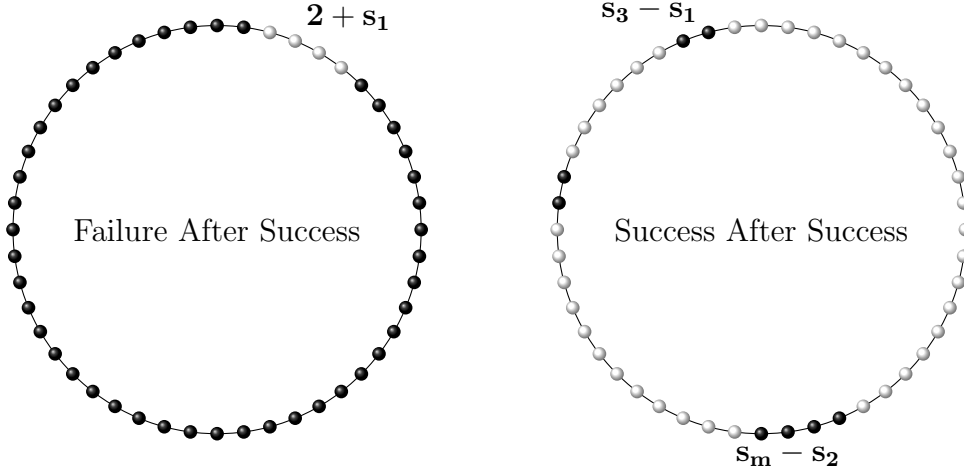
Figure 14: Resulting Configurations given Failure or Success in the second random walk, given successful first walk, for $t < s$

$$
\widetilde{P_B(\cdot)} = \frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1} \left[\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1} + \frac{r^{\frac{s_2-s_1}{2}} - 1}{r^{\frac{s_2}{2}+1} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}}{r^{\frac{n}{2}} - 1} \right] =
$$

$$
= \frac{r-1}{r^{\frac{n}{2}}-1} \left[\frac{\left(r^{\frac{s_1+s_2}{2}+1} - r^{\frac{s_2}{2}} \right)\left(r^{\frac{n}{2}} - 1 \right) + \left(r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1} \right)\left(r^{\frac{s_2}{2}} - r^{\frac{s_1}{2}} \right)}{\left(r^{\frac{s_1}{2}+1} - 1 \right)\left(r^{\frac{s_2}{2}+1} - 1 \right)} \right] =
$$

$$
= \frac{r-1}{r^{\frac{n}{2}}-1} \left[\frac{\left(r^{\frac{n+s_2}{2}} + r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}} - r^{\frac{n}{2}-1} \right)\left(r^{\frac{s_1}{2}+1} - 1 \right)}{\left(r^{\frac{s_1}{2}+1} - 1 \right)\left(r^{\frac{s_2}{2}+1} - 1 \right)} \right] =
$$

$$
= \frac{r-1}{r^{\frac{n}{2}}-1} \left[\frac{r^{\frac{n}{2}-1}\left(r^{\frac{s_2}{2}+1} - 1 \right)}{r^{\frac{s_2}{2}+1} - 1} + \frac{r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}}}{r^{\frac{s_2}{2}+1} - 1} \right] =
$$

$$
= \frac{r-1}{r^{\frac{n}{2}}-1} \left[r^{\frac{n}{2}-1} + \frac{r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}}}{r^{\frac{s_2}{2}+1} - 1} \right]
$$

Notice that $s_2 = s - s_1 - s_3 - \cdots - s_m$, hence $\frac{\partial s_2}{\partial s_1} = -1$. And now we can differentiate $\widetilde{P_B(\cdot)}$ with respect to s_1.

$$
\frac{\partial \widetilde{P_B(\cdot)}}{\partial s_1} = \frac{r-1}{r^{\frac{n}{2}}-1} \left[\frac{\left(-r^{\frac{n+s_2-s_1}{2}-1}\ln r + \frac{\ln r}{2}r^{\frac{s_2}{2}} \right)\left(r^{\frac{s_2}{2}+1} - 1 \right) - \left(r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}} \right)\left(-\frac{\ln r}{2}r^{\frac{s_2}{2}+1} \right)}{\left(r^{\frac{s_2}{2}+1} - 1 \right)^2} \right] =
$$

$$
= \frac{\ln r \,(r-1)\, r^{\frac{s_2}{2}}}{2\left(r^{\frac{n}{2}}-1 \right)\left(r^{\frac{s_2}{2}+1} - 1 \right)^2} \left(2r^{\frac{n-s_1}{2}-1} - r^{\frac{n+s_2-s_1}{2}} - 1 \right) > 0, \text{ for } r < 1.
$$

The fact that the term $2r^{\frac{n-s_1}{2}-1} - r^{\frac{n+s_2-s_1}{2}} - 1$ is always negative is not obvious and is proven here. Substituting s_2, we can rewrite it as:

29

$$r^{-s_1-1}\left[2r^{\frac{n+s_1}{2}} - r^{\frac{n+s-s_3-\cdots-s_m}{2}+1} - r^{2\frac{s_1}{2}+1}\right]$$

If we denote $x = r^{\frac{s_1}{2}}$ then we get a polynomial of degree two with respect to x. The discriminant of this polynomial is equal to:

$$\Delta = 4r^{2\frac{n}{2}} - 4r^{\frac{n}{2}+\frac{s-s_3-\cdots-s_m}{2}+2} = 4r^{\frac{n}{2}}\left(r^{\frac{n}{2}} - r^{\frac{s-s_3-\cdots-s_m}{2}+2}\right) = 4r^{\frac{n}{2}}\left(r^{\frac{n}{2}} - r^{\frac{s_1+s_2}{2}+2}\right) < 0$$

Because $r < 1$ and $\frac{n}{2} > \frac{s_1+s_2}{2} + 2$ for $m \geq 3$. For $m = 2$ this holds with equality, but we have already analyzed this case. So, this polynomial has no roots and given that the factor of the quadratic term is negative $(-r)$, we can conclude that for $r < 1$ the polynomial is always negative. Therefore, $\widetilde{P_B}(\cdot)$ takes its maximum value when s_1 is maximized. For this value of s_1, the real probability of successful diffusion is equal to this upper bound as long as $s_m - s_1 \leq 2$. Therefore, remembering that $\widetilde{P_B}(\cdot) \geq P_B(\cdot)$ always, it has to be that $P_B(\cdot)$ is also maximized for when both s_1 is maximized and $s_m - s_1 \leq 2$.

In case m divides s exactly, then the maximum of s_1 is equal to $\frac{s}{m}$ and the optimal choice is $s_1 = \cdots = s_m = \frac{s}{m}$. If m does not divide s exactly, then we have $s = mq + d$, where q is the quotient of the division and d is the remainder. In this case, P_B is maximized if we have $m - \frac{d}{2}$ groups of size $q = \frac{s-d}{m}$ and $\frac{d}{2}$ groups with size $q + 2 = \frac{s-d}{m} + 2$, so again the difference in the size of any two groups is no larger than four. We still remain to describe what is the optimal position of the groups that have the two additional agents. The result will become apparent after we analyze the case for $t > s$.

Now, we prove the result for $t > s$ in a completely analogous way. In this case, the initial configuration is as in the left part of Figure 15. A success in the first random walk leads to the total diffusion of action B, while a failure leads to a configuration as in the right part of the same figure. The probability of success in the first walk is $\frac{r^{\frac{t_1}{2}+1}-r}{r^{\frac{t_1}{2}+1}-1}$. Figure 16 shows the possible configurations after success or failure in the second random walk, given a failure in the first one. The probability of success in the second walk is $\frac{r^{\frac{t_2}{2}+1}-r^{\frac{t_1}{2}+1}}{r^{\frac{t_2}{2}+1}-1}$. Therefore, we can construct again an upper bound for the probability of successful diffusion, equal to:

$$\widetilde{P_B}(\cdot) = \frac{r^{\frac{t_1}{2}+1}-r}{r^{\frac{t_1}{2}+1}-1} + \frac{r-1}{r^{\frac{t_1}{2}+1}-1}\left[\frac{r^{\frac{t_2}{2}+1}-r^{\frac{t_1}{2}}}{r^{\frac{t_2}{2}+1}-1}\frac{r^{\frac{n}{2}}-r^{\frac{t_1}{2}-1}}{r^{\frac{n}{2}}-1} + \frac{r^{\frac{t_1}{2}}-1}{r^{\frac{t_2}{2}+1}-1}\right]$$

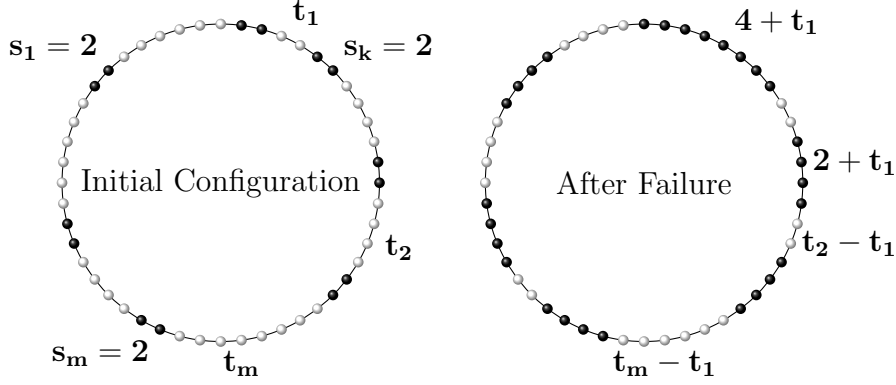This expression can be transformed in a similar manner as before:

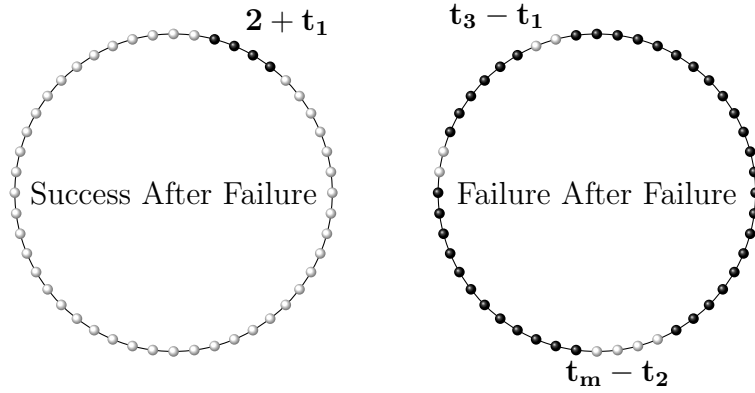Figure 15: General Initial Configuration and Result After Success, for $t > s$.

Figure 16: Resulting configurations after success or failure in the second random walk, given failure in the first random walk, for $t > s$

$$\widetilde{P_B(\cdot)} = \frac{r^{\frac{t_1}{2}+1} - r}{r^{\frac{t_1}{2}+1} - 1} + \frac{r - 1}{r^{\frac{t_1}{2}+1} - 1}\left[\frac{r^{\frac{t_2}{2}+1} - r^{\frac{t_1}{2}}}{r^{\frac{t_2}{2}+1} - 1}\frac{r^{\frac{n}{2}} - r^{\frac{t_1}{2}-1}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{t_1}{2}} - 1}{r^{\frac{t_2}{2}+1} - 1}\right] =$$

$$= \frac{1}{r^{n/2} - 1}\left[\frac{(r^{\frac{t_2}{2}+1} - 1)(r^{n/2} - 1) - (r^{\frac{t_2}{2}+1} - r^{\frac{t_1}{2}+1})(r - 1)}{r^{\frac{t_2}{2}+1} - 1}\right] =$$

$$= \frac{1}{r^{n/2} - 1}\left[r^{n/2} - r + (r - 1)\frac{r^{\frac{t_1}{2}+1} - 1}{r^{\frac{t_2}{2}+1} - 1}\right]$$

Notice again, that the upper bound becomes equal to the actual probability if $t_m - t_1 \leq 2$, where t_m is the size of the largest group of type B and t_1 the smallest one.

Now, we can differentiate the expression with respect to t_1, remembering that $t_2 = t - t_1 - t_3 - \cdots - t_m$:

$$\frac{\partial \widetilde{P_B(\cdot)}}{\partial t_1} = \frac{(r - 1)\ln r}{2(r^{n/2} - 1)}\left[\frac{r^{\frac{t_1}{2}+1}(r^{\frac{t_2}{2}+1} - 1) + (r^{\frac{t_1}{2}+1} - 1)r^{\frac{t_2}{2}+1}}{(r^{\frac{t_2}{2}+1} - 1)^2}\right] > 0, \text{ for all } r < 1.$$

31

The upper bound is increasing in t_1. For this value of t_1, the real probability of successful diffusion is equal to this upper bound as long as $t_m - t_1 \leq 2$. Therefore, remembering that $\widetilde{P_B(\cdot)} \geq P_B(\cdot)$ always, it has to be that $P_B(\cdot)$ is also maximized for when both t_1 is maximized and $t_m - t_1 \leq 2$.

In case m divides s exactly, then the maximum of t_1 is equal to $\frac{t}{m}$ and the optimal choice is $t_1 = \cdots = t_m = \frac{t}{m}$. If m does not divide t exactly, then we have $t = mq + d$, where q is the quotient of the division and d is the remainder. In this case, P_B is maximized if we have $m - \frac{d}{2}$ groups of size $q = \frac{t-d}{m}$ and $\frac{d}{2}$ groups with size $q + 2 = \frac{t-d}{m} + 2$, so again the difference in the size of any two groups is no larger than four.

To complete the proof we need to explain the optimal location of the groups which have the two additional agents. For the case where $t < s$ we need to notice that after successful absorption in the first random walk, now the population consists of $\frac{d}{2}$ groups of each type, where all the groups of type A have exactly two agents. Hence, we fall into the analysis of the case where $t > s$, where we would like the groups of type B to be as equal as possible. In order to succeed this we should have located the groups of type A with more agents as symmetrically as possible around the circle.

An example can be illustrated in Figure 17. We have targeted 14 out of 48 agents, having seven groups of two agents of type B, three groups of six agents and four groups of four agents of type A. After successful absorption in the first walk, there will be left only three groups of two agents of type A, which we want to be located as symmetrically as possible. For this reason we do not put two groups of six agents one next to the other in the initial configuration. However, notice that we cannot make the configuration that arises after success totally symmetric, due to the restriction on the sizes of the groups. But again we want it to be as symmetric as possible, by maximizing the smallest group and minimizing its difference with the largest one. The argument for the case where $t > s$ is completely analogous.
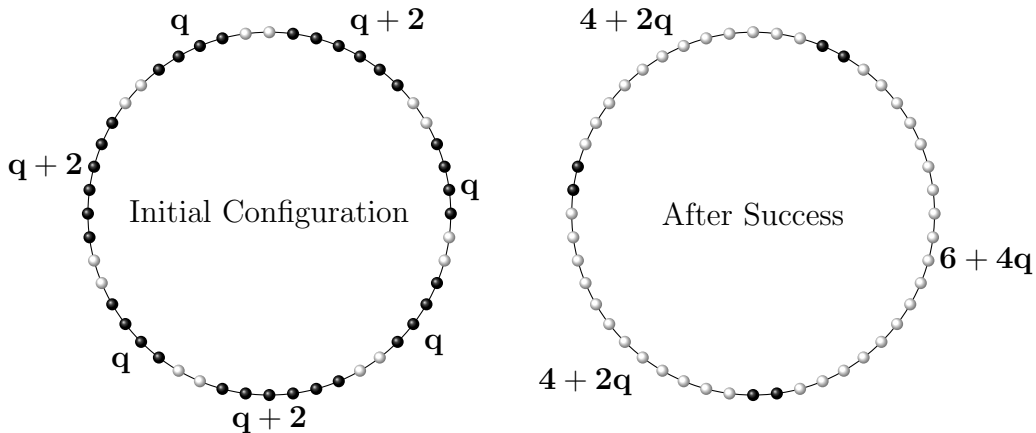

Figure 17: Optimal Initial Configuration with $s = mq + d$, for $s > t$.

□

References

ALÓS-FERRER, C. & WEIDENHOLZER, S. (2008). Contagion and efficiency. *Journal of Economic Theory* 143, 251–274.

APESTEGUIA, J., HUCK, S. & OECHSSLER, J. (2007). Imitation – Theory and experimental evidence. *Journal of Economic Theory* 136, 217–235.

BAGNOLI, F., BOCCARA, N. & RECHTMAN, R. (2001). Nature of phase transition in a probabilistic cellular automaton with two absorbing states. *Physical Review E* 63(4), 046116.

BALA, V. & GOYAL, S. (1998). Learning from Neighbours. *Review of Economic Studies* 65, 595–621.

BALLESTER,C., CALVO-ARMENGOL, A. & ZENOU, Y. (2006). Who's who in networks. Wanted: The key player *Econometrica* 74(5), 1403–1417.

BANERJEE, A. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107, 797–817.

BANERJEE, A. & FUDENBERG, D. (2004). Word-of-mouth learning. *Games and Economic Behavior* 46, 1–22.

BIGONI, M. & FORT, M. (2013). Information and Learning in Oligopoly. *IZA Discussion Paper No. 7125*.

CHATTERJEE, K. & DUTTA, B. (2011). Credibility and Strategic Learning in Networks. *Warwick Economic Research Papers* no.972.

CHATTERJEE, K. & XU, S.H. (2004). Technology diffusion by learning from neighbors. *Advances in Applied Probability* 36(2), 355–376.

CONLEY, T. & UDRY, C. (2010). Learning About a New Technology: Pineapple in Ghana. *American Economic Review* 100, 35–69.

DEGROOT, M.H. (1974). Reaching a Consensus. *Journal of the American Statistical Association* 69, 118–121.

DOMINGOS, P. AND RICHARDSON, M. (2001). Mining the network value of customers. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 57–66.

DONNELLY, P. & WELSH, D. (1983). Finite Particle Systems and Infection Models, *Mathematical Proceedings of the Cambridge Philosophical Society* 94(1), 167–182.

DUERSCH, P., OECHSSLER, J. AND SCHIPPER, B.C. (2012). Unbeatable Imitation, *Games and Economic Behavior* 76(1), 88–96.

ELLISON, G. & FUDENBERG, D. (1993). Rules of Thumb for Social Learning. *Journal of Political Economy* 101, 612–644.

—— (1995). Word-of-mouth communication and social learning. *Quarterly Journal of Economics* 109, 93–125.

ESHEL, I., SAMUELSON, L. & SHAKED A. (1998). Altruists, egoists and hooligans in a local interaction model. *American Economic Review* 88, 157–179.

FOSCO, C. & MENGEL, F. (2011). Cooperation through imitation and exclusion in networks. *Journal of Economic Dynamics and Control* 35(5), 641–658.

GALE, D. & KARIV, S. (2003) Bayesian Learning in Social Networks. *Games and Economic Behavior* 45, 329–346.

GALEOTTI, A. & GHIGLINO, C. & SQUINTANI, F. (2011). Strategic Information Transmission in Networks. *Working Paper*.

GALEOTTI, A. & GOYAL, S. (2009). Influencing the influencers: a theory of strategic diffusion. *RAND Journal of Economics* 40(3), 509–532.

GIGERENZER, G. & SELTEN, R. (2002). *Bounded Rationality: The Adaptive Toolbox*. The MIT Press.

GOLUB, B. & JACKSON, M.O. (2010). Naive Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics* 2, 112–149.

GOYAL, S. & KEARNS, M. (2012). Competitive contagion in networks. In *Proceedings of the 44th symposium on Theory of Computing* 759–774, New York, NY.

KEMENY, J.G. & SNELL, J.L. (1960). *Finite Markov Chains*. Springer.

KEMPE, D., KLEINBERG, J. AND TARDOS, E. (2003). Maximizing the Spread of Influence in a Social Network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 137–146, New York, NY.

KEMPE, D., KLEINBERG, J. AND TARDOS, E. (2005). Influential nodes in a diffusion model for social networks. In *ICALP 2005: Proceedings of the 32nd International Colloquium on Automata, Languages and Programming*, 1127–1138.

KIRBY, J. & MARSDEN, P. (2006). *Connected marketing*. Elsevier.

JACKSON, M.O. (2008). *Social and Economic Networks*. Princeton University Press.

ORTUÑO, I. (1993), Multiple equilibria, local externalities and complexity, mimeo.

RICHARDSON, M. AND DOMINGOS, P. (2002). Mining Knowledge–Sharing Sites for Viral Martketing. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 61–70. Edmonton, Canada 2002.

SÁKOVICS, J. & STEINER, J. (2012). Who matters in coordination problems?. *American Economic Review*, forthcoming.

SCHLAG, K. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory* 78, 130–156.

TSAKAS, N. (2013). Imitating the most successful neighbor in social networks. *Review of Network Economics* 12(4), 403–435.

TSAKAS, N. (2014). Optimal Influence under Observational Learning, mimeo.

VEGA-REDONDO, F. (1997). The evolution of Walrasian Behavior. *Econometrica* 65, 375–384.

WEIBULL, J. (1995). *Evolutionary Game Theory*. MIT Press, Cambridge, MA.

YILDIZ, E., ACEMOGLU, D., OZDAGLAR, A. E., SABERI, A. AND SCAGLIONE, A. (2011). Discrete Opinion Dynamics with Stubborn Agents.http://dx.doi.org/10.2139/ssrn.1744113