



HAL
open science

Dynamic competition over social networks

Antoine Mandel, Xavier Venel

► **To cite this version:**

| Antoine Mandel, Xavier Venel. Dynamic competition over social networks. 2017. <halshs-01524453>

HAL Id: halshs-01524453

<https://shs.hal.science/halshs-01524453v1>

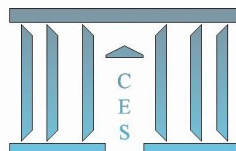
Submitted on 18 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Dynamic competition over social networks

Antoine MANDEL, Xavier VENEL

2017.21



Dynamic competition over social networks

Antoine Mandel*, Xavier Venel †

Abstract

We provide an analytical approach to the problem of influence maximization in a social network when two players compete by means of dynamic targeting strategies. We formulate the problem as a two-player zero-sum stochastic game. We prove the existence of the uniform value: if the players are sufficiently patient, both players can guarantee the same mean-average opinion without knowing the exact discount factor. Further, we put forward some elements for the characterization of equilibrium strategies. In general, players must implement a trade-off between a forward-looking perspective, according to which they shall aim at maximizing the future spread of their opinion in the network, and a backward-looking perspective, according to which they shall aim at counteracting their opponent's previous actions. When the influence potential of players is small, an equilibrium strategy is to systematically target the agent with the largest eigenvector centrality.

Keywords: Social Network, Dynamic games, Targeting, Stochastic games

JEL codes: SC71, D85

1 Introduction

"*Delivering the right message to the right person at the right time*" is a strong motto among influencers in politics, lobbying, marketing. The large increase in information about individual characteristics and social interactions brought about by the development of internet and of online social networks has generated a tremendous interest, among both practitioners and scientists, on the problem of identifying appropriate targets to maximize influence. In marketing, targeted advertising on social networks has become a cornerstone of the industry in less than a decade. In politics, the campaign of Barack Obama for the 2008 U.S presidential election has illustrated the tremendous influence potential embedded in social networks (Cogburn and Espinoza-Vasquez, 2011). In the academic literature, identifying the key target in a network in order to gain control, influence or market shares has become a central focus in computer science, economics, operations research. Most of the existing literature focuses on optimal/efficient strategies for a single agent (Kempe et al., 2003; Ballester et al., 2006). More recently, the competitive nature of the targeting problem has been emphasized and game-theoretic contributions have started to analyze the behavior of players competing for prominence over a network (Goyal et al., 2014; Bimpikis et al., 2016; Grabisch et al., 2017).

Yet, a common feature of existing approaches is their focus on the identification of the key target, i.e. "*the right person*". The temporal dimension, "*the right time*", hasn't yet been approached by the social network literature. Its role is nevertheless crucial in applications. In political campaigning, as emphasized by Granato and Wong (2004), "*the relation between voters and campaign strategists is dynamic and evolves until voters' views on a candidate crystallize.*" In marketing, an important element of a firm's strategy is the sequence according to which a product is offered to potential buyers (Hartline et al., 2008).

*Paris School of Economics, Centre d'économie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, 106-110 Boulevard de de l'hôpital, 75647 Paris Cedex 13, France E.mail: antoine.mandel@univ-paris1.fr

†Paris School of Economics, Centre d'économie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, 106-110 Boulevard de de l'hôpital, 75647 Paris Cedex 13, France E.mail: xavier.venel@univ-paris1.fr

The main objective of this paper is to include this temporal dimension in the analysis of the targeting problem on networks. Therefore, we place ourselves in a competitive setting where two players choose a sequence of targets in a social network in order to influence the average opinion that gets formed on a given topic, e.g a political candidate, a product or the truth value of an information. The network is formed by non-strategic agents whose opinion is represented by a real number and who update their opinions using a weighted average of their neighbors' opinions (as in Golub and Jackson, 2010). The strategic players choose a sequence of targets among non-strategic agents. Hence, they influence directly the opinion of the targeted agent and indirectly the whole network, through the dynamics of opinion formation.

In this setting, we are concerned with two main issues: the existence of an uniform value in the underlying (infinite-horizon) game and the characterization of optimal strategies, i.e the determination of an optimal sequence of targets. A priori, both issues shall be sensitive, on the one hand to the influence potential of the two strategic players, which measures the extent to which strategic players are able to change the opinion of the targeted agent and which might differ between players, and on the other hand to the informational structure of the game, that is what each player observes from his opponent's actions.

The existence of an uniform value is non-trivial in dynamic/stochastic games with compact state-spaces. In general, it may not exist as shown by (Ziliotto, 2016). The first section of the paper is devoted to the issue. We highlight the relation between the linear updating of the vector of opinions using the network's matrix of weights that we consider and the conventional stochastic game obtained when weights are interpreted as transition probabilities. We show that if the matrix is primitive, strategic influence "contracts" the vector of opinions in the network. This suffices to show that there exists an uniform value and that this value is independent of the initial distribution of beliefs. Furthermore, this result holds independently of the informational structure of the game.

As for the characterization of equilibrium strategies, i.e the determination of an optimal sequence of targets, a key observation is that players must implement a trade-off between a forward-looking perspective, according to which they shall aim at maximizing the future spread of their opinion in the network, and a backward-looking perspective, according to which they shall aim at counteracting their opponent's previous actions. It is likely impossible to provide a closed-form analytical characterization of dynamics strategies implementing this trade-off in the general case. We focus on specific cases where both strategic players have the same level influence and an equilibrium in pure strategies can be characterized.

In this setting, we first show that if the players do not have information about each other's actions, an equilibrium in pure strategies must be "purely" forward looking. If such an equilibrium exists, it is in fact also an equilibrium for any alternative informational structure. Yet, examples show that such "purely forward looking" equilibria do not necessarily exist and, more generally, that the existence of a pure equilibrium might depend on the informational structure. Nevertheless, we prove a general result in the case where the level of influence is small: it is then optimal for both players to systematically target the node with the largest eigenvector centrality in the network.

The results of the paper also highlight the importance of the temporal structure of strategies. For certain networks, there exist pure equilibria with dynamic targeting strategies whereas no pure equilibria exist when a fixed target must be chosen (as in Grabisch et al., 2017). This stems from the fact that dynamic strategies are less easily counteracted than static ones. More generally, our results suggest that optimal targeting sequences can be determined when the forward-looking perspective dominates whereas mixed strategies ought to be used when one's influence can be neutralized by the opponent.

The remaining of the paper is organized as follows. Section 2 briefly reviews the related literature. Section 3 presents a game-theoretic model of dynamic competitive targeting. Section 4 proves the existence of an uniform value for this game. Section 5 provides a partial characterization of equilibrium strategies. Section 6 concludes.

2 Related literature

Since the seminal work in quantitative sociology of the 1970s (notably Bonacich, 1972; Freeman, 1977), the identification of central agents has been a key research question in the analysis of socio-economic networks. With the raise of digital social networks and the correlative flow of data, numerous applications in communication and marketing have been identified, and the problem has been approached from a more goal-oriented perspective at the interface between economics and computer-science. The targeting problem, i.e. the choice of an optimal target in view of the diffusion of a product or a message, has been widely investigated from an algorithmic perspective in computer science following on the early contributions of Domingos and Richardson (2001) and Kempe et al. (2003, 2005). In economics, two important contributions are Ballester et al. (2006), which characterize a key target through an “intercentrality measure” that takes into account both a player’s centrality and his contribution to the centrality of others and Banerjee et al. (2013), which develop a model that discriminates between information passing and endorsement and derive a measure of “Diffusion centrality” that characterize efficient targets.

Now, the bulk of the literature has approached the targeting problem as an optimization problem for a single influencing agent. A competitive perspective has been introduced in a few recent contributions, which develop game-theoretic models in which agents choose, ex ante, a targeting strategy to maximize their influence on the outcome of the ensuing opinion dynamics. Bharathi et al. (2007) and Goyal et al. (2014) offer an algorithmic perspective on the issue while (Lever, 2010; De Meyer et al., 2014; Bimpikis et al., 2016; Grabisch et al., 2017) focus on the characterization of optimal targets. In particular, Grabisch et al. (2017) is a direct predecessor of this work and puts forward “intermediacy centrality” as a minmax characterization of the key target that accounts for the fact that, in a competitive context, agents should focus on relative, rather than absolute, centrality.

The importance of adopting a dynamic perspective when designing influence strategies is also clearly identified in the literature. In political science, Box-Steffensmeier and Lin (1996) emphasizes the fact that *“at different stages of the campaign process candidates have different goals, and their expenditures should have different effects on the final election outcome”*. Granato and Wong (2004) provides a simulation model of the dynamics of an electoral campaign from this perspective while Demange and Van der Straeten (2009); Aragonés et al. (2015) and Amorós and Puy (2013) investigate in a game-theoretic setting how candidates should allocate their campaign time on the different issues of concern for the electorate. In the computer science literature, Zhuang et al. (2013) considers the problem of network influence maximization in a dynamic setting but puts forward an adaptive approach where a single influencer periodically probes the network in order to adjust its influence strategy. Yadav et al. (2016) approaches the same problem but from the more formal perspective of Partially observable Markov decision process (POMDP). From a more economic perspective, Hartline et al. (2008) analyses the marketing strategy of a firm over a social network in a dynamic programming framework and emphasizes the complementarity between pricing and targeting: *“In general it is advantageous to get influential buyers to buy the item early in the sequence; it even makes sense to offer such buyers smaller prices to get them to buy the item.”* There also is a wide literature on competition in advertising. Chakrabarti and Haller (2011) provides a simple model of advertising wars, which have a long history documented in Beard (2010). Doraszelski and Markovitch (2007) model dynamic competition among firms through advertising and analyze its impact on industrial dynamics. Finally, Pastine and Pastine (2002) analyze the use of advertising by two competing firms as an expectation coordination device of purchasing decision of consumers in a setting with consumption externalities.

However, this paper is, to our knowledge, the first that brings together a dynamic and a competitive perspective on the problem of targeting in a social network. As such, it raises a number of theoretical issues. The first one is the definition of an appropriate concept of solution. We choose to focus on the mean-average payoff and on the uniform approach (Mertens and Neyman, 1981), which seems to us well fitted to extend the results of Grabisch et al. (2017) to a dynamical framework. However, the existence of the uniform value in our framework, where the set of states is compact, is a complicated question. It has been solved positively in several classes of stochastic games (Venel, 2014) or as intermediary result for stochastic games with incomplete information (Renault, 2012; Aumann et al., 1995). However none

of these results apply here. Moreover, it was recently shown that even under very regular assumption, the uniform value might not exist (Ziliotto, 2016).

3 Model

3.1 Notation

Given a finite set K , we denote by $e' = (1, \dots, 1)$ the vector of size K composed of 1. Given a subset S of K , we denote by δ_S the indicator function of S . If $S = \{i\}$ for some $i \in K$ then we denote it by δ_i .

3.2 Opinion game

We consider two strategic players who compete for influence on a social network through dynamic targeting strategies. The network is formed by K non-strategic agents whose social ties are represented by a row-stochastic matrix M of size $K \times K$. Each of the non-strategic agent is characterized by its opinion on an issue of concern for the two strategic players: e.g. the relative quality of two products or the proximity with respect to two opposite electoral platforms. This opinion is represented by a number in $[-1, 1]$ and, by default, evolves according to the social influence exerted on the agent, i.e. each agent updates his opinion by combining linearly the opinions of his neighbors proportionally to the corresponding row of the social network matrix M . This boundedly rational model of opinion formation has been introduced by DeGroot (1974) and has received a wide attention in the economic literature as a model of opinion formation in social networks (see e.g. Golub and Jackson, 2010, and references therein). It is particularly well-suited to contrast the myopic behavior of a "crowd" with the strategic behavior of external influencers such as lobbies, advertisers or political campaigners (see also Bimpikis et al., 2016; Grabisch et al., 2017).

In our framework, two strategic influencers have a fixed opinion, 1 and -1 respectively, and are accordingly referred to as agent 1 and agent -1 . They aim at bringing the average opinion in the network as close as possible to their own opinion, e.g. to trigger the purchase of their product or a vote for their political platform. Therefore, they perturb the process of opinion formation by targeting each period one of the non-strategic agents in order to influence his opinion. The non-strategic agent $k \in \{1, \dots, K\}$ then assigns a weight $D(k) \in \mathbb{R}_+$ to the opinion of his neighbors and, if targeted, a weight $\lambda \in \mathbb{R}_+$ to the opinion of agent 1 and a weight $\mu \in \mathbb{R}_+$ to the opinion of agent -1 (note that the influence potential of the two strategic players can differ). There are then four different cases to compute the updated opinion of an agent k in the perturbed process:

- If agent k is not the target of any influencer, then his opinion is the M -weighted average of the opinions of his neighbors.
- If agent k is the target of influencer 1 then his opinion is the convex combination of 1 with weight $\frac{\lambda}{\lambda + D(k)}$ and the M -weighted average of the opinions of his neighbors with weight $\frac{D(k)}{D(k) + \lambda}$.
- If agent k is the target of influencer -1 then his opinion is the convex combination of -1 with weight $\frac{\mu}{\mu + D(k)}$ and the M -weighted average of the opinions of his neighbors with weight $\frac{D(k)}{D(k) + \mu}$.
- If agent k is the target of both influencers then his opinion is the convex combination of 1 with weight $\frac{\lambda}{\lambda + \mu + D(k)}$, -1 with weight $\frac{\mu}{\lambda + \mu + D(k)}$ and the M -weighted average of the opinions of his neighbors with weight $\frac{D(k)}{D(k) + \mu + \lambda}$.

Remark 3.1 *An interesting particular case to consider is the one where each non strategic agent listens uniformly to his neighbors and the share of influence received from the network is proportional to the number of neighbors. In this case, $D(k)$ is the degree of agent k in the network and all non-zero entries of $M_{k,\cdot}$ have value $1/D(k)$.*

Hence, a strategic influencer faces a number of trade-off. At each stage, he must first choose whether to influence an easily influenceable agent, with low $D(k)$, or a central agent, which might however have a high $D(k)$ and be harder to influence (in particular in the case of remark 3.1 above). Second, he must choose whether to confront the other influencer by choosing the same target or to shield away by choosing another target or playing a mixed strategy (e.g. if he has a much weaker influence potential than his opponent). Both trade-offs are already present in the static targeting problem where the influenced agent is fixed through time (Bimpikis et al., 2016; Grabisch et al., 2017). The key issue when strategic influencers can use dynamic strategies is the trade-off between adopting a forward-looking or a backward-looking perspective. Influencers adopt a purely forward looking perspective if they only focus on the forward diffusion of their influence via the social network and neglect their opponent's previous actions. They adopt a purely backward looking perspective if their sole focus is to prevent the spread through the network of the influence precedently exerted by their opponent. In general, they ought to use a strategy, which implements a trade-off between both perspectives. We shall analyze these issues by formalizing the problem as a two-player stochastic game.

Definition 3.2 *The two player stochastic game $\Gamma = (X, I, J, q, g)$ is defined as follows.*

- *The set of state $X = [-1, 1]^K$ represents the vector of opinions among non-strategic agents.*
- *The set of actions of player 1 is $I = K$, i.e. the choice of a target among non-strategic agents.*
- *The set of actions of player -1 is $J = K$, i.e. the choice of a target among non-strategic agents.*
- *The transition $q : X \times I \times J \rightarrow X$ represents the deterministic evolution of the vector of opinions under the joint influence of the social network and of the strategic influencers. It is defined as follows.*
 - *For a fixed $i \in I$ and $j \in J$, we define $B(i, j)$ the $K \times K$ matrix such that for every $(k, l) \in K \times K$,*

$$B(i, j)_{kl} = \begin{cases} \frac{D(k)M_{kl}}{D(k)+\lambda\delta_i(k)+\mu\delta_j(k)} & \text{if } D(k) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

and $A_1(i, j)$ and $A_{-1}(i, j)$ the two vectors such that

$$A_1(i, j)_k = \begin{cases} \frac{\lambda}{D(k)+\lambda+\mu\delta_j(k)} & \text{if } k = i, \\ 0 & \text{otherwise.} \end{cases}$$

and

$$A_{-1}(i, j)_k = \begin{cases} \frac{\mu}{D(k)+\lambda\delta_i(k)+\mu} & \text{if } k = j, \\ 0 & \text{otherwise.} \end{cases}$$

They correspond respectively to the weights in the opinion updating process of the social network and of both influencers, given these target i and j respectively.

- *One then defines the transition q as:*

$$\begin{pmatrix} 1 \\ -1 \\ q(p, i, j) \end{pmatrix} = \begin{pmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ A_1(i, j) & A_{-1}(i, j) & B(i, j) \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ p \end{pmatrix} \quad (3.1)$$

- *the payoff function $g : X \rightarrow [-1, 1]$ is given by the mean average opinion in the network*

$$g(p) = \frac{1}{K} \sum_{k=1}^K p_k.$$

Remark 3.3 *The stochastic game Γ is similar (state space, action space, payoff) to the auxiliary game associated to POMDP in the dark as was introduced in Rosenberg et al. (2002) or stochastic game with "state-blind" players like in Venel (2014). A key difference is in the transition, which is non-expansive for the $\|\cdot\|_\infty$ -norm in the present case.*

3.3 Information structure and strategies

The sequence of actions of players in the game induces an history. For every $n \geq 1$, we denote by $\mathcal{H}_n = (I \times J)^n$ the set of finite n -stage histories. We then extend the definition of q to histories of finite length by induction:

$$\forall p \in X, \forall h_{n-1} \in \mathcal{H}_{n-1}, \forall i \in I, \forall j \in J, q(p, h_{n-1}, i, j) = q(q(p, h_{n-1}), i, j).$$

Hence, given an initial vector of opinions, a finite history of actions induces deterministically a finite history of opinions. Influencers might then condition their strategies on both the history of actions and of opinions. Yet, they might imperfectly observe the actions of their opponents or the vector of opinions in the network. More specifically, we define the information structure of the game and the strategies of players as follows.

Definition 3.4 *An information structure is a triple (C^1, C^{-1}, s) of two finite sets C^1 and C^{-1} and a function $s : I \times J \rightarrow \Delta(C^1 \times C^{-1})$.*

- *Agent 1 observes the signal in C^1 and his own actions. Hence, for every integer $n \geq 1$, the history observed by player 1 is $H_n^1 = (C^1 \times I)^{n-1} \times C^1$ and a strategy for player 1 is a function $\sigma : \cup_{n=1}^{+\infty} H_n^1 \rightarrow \Delta(I)$.*
- *Agent -1 observes the signal in C^{-1} and his own actions. Hence, for every integer $n \geq 1$, the history observed by player -1 is $H_n^{-1} = (C^{-1} \times J)^{n-1} \times C^{-1}$ and a strategy for player -1 is a function $\tau : \cup_{n=1}^{+\infty} H_n^{-1} \rightarrow \Delta(J)$.*
- *We denote respectively by Σ and \mathcal{T} the sets of strategies of players 1 and -1.*
- *We say that a strategy is pure if the image of every finite history is a Dirac.*

Remark 3.5 *Given the transition function is deterministic, information about the history of opinions in the network can be deduced from information about the history of actions in the game provided that the initial vector of opinions is known. Yet, as we shall see below, equilibrium can be determined independently of the initial vector of opinions. Therefore, one can restrict attention to the case where the influencers are informed only (and possibly incompletely) about the history of actions.*

Remark 3.6 *One could extend the class of information structures to consider cases where the type of signal received depends on the state of the network, e.g. in cases where an influencer receives information only about the states of players whose opinion is sufficiently close to his. This is outside the scope of this paper.*

Two structures of information will be of particular concern in the following.

- *The complete information structure where $C^1 = J$, $C^{-1} = I$ and for every $(i, j) \in I \times J$, $s(i, j) = (j, i)$. Then, both players are perfectly informed about the history of past actions in the game.*
- *The blind information structure where $C^1 = C^{-1} = \{1\}$ and s is the constant mapping. Then, the players have no information about their opponent's actions and only take into consideration their own history of play.*

3.4 Evaluations

In order to evaluate this model, we will adopt two different notions coming from the literature of stochastic games: asymptotic value and uniform value.

For every $n \geq 1$, we define the average n -stage payoff as:

$$\gamma_n(p_1, \sigma, \tau) = \mathbb{E}_{p_1, \sigma, \tau} \left(\frac{1}{n} \sum_{t=1}^n g(p_t) \right),$$

and for every $\rho \in (0, 1]$, the discounted payoff as:

$$\gamma_\rho(p_1, \sigma, \tau) = \mathbb{E}_{p_1, \sigma, \tau} \left(\rho \sum_{t=1}^{+\infty} (1 - \rho)^{t-1} g(p_t) \right),$$

It is clear that for every $n \geq 1$, (resp. $\rho \in (0, 1]$), the game with payoff γ_n (resp. γ_ρ) has a value denoted v_n (resp. v_ρ):

$$v_n(p_1) = \sup_{\sigma \in \Sigma} \inf_{\tau \in \mathcal{T}} \gamma_n(p_1, \sigma, \tau) = \inf_{\tau \in \mathcal{T}} \sup_{\sigma \in \Sigma} \gamma_n(p_1, \sigma, \tau).$$

A first extension to the infinite horizon case is provided by the notion of asymptotic value:

Definition 3.7 *The game $\Gamma(p_1)$ has an asymptotic value if the sequence $(v_n(p_1))_{n \geq 1}$ converges to a limit denoted $v(p_1)$.*

A stronger requirement is provided by the notion of uniform value.

Definition 3.8 *Let v be a real number.*

- *Player 1 can guarantee v in $\Gamma(p_1)$ if for all $\varepsilon > 0$ there exists a strategy $\sigma^* \in \Sigma$ of player 1 such that*

$$\liminf_n \inf_{\tau \in \mathcal{T}} \gamma_n(p_1, \sigma^*, \tau) \geq v - \varepsilon.$$

We say that such a strategy σ^ guarantees $v - \varepsilon$ in $\Gamma(p_1)$.*

- *Player 2 can guarantee v in $\Gamma(p_1)$ if for all $\varepsilon > 0$ there exists a strategy $\tau^* \in \mathcal{T}$ of player 2 such that*

$$\limsup_n \sup_{\sigma \in \Sigma} \gamma_n(p_1, \sigma, \tau^*) \leq v + \varepsilon.$$

We say that such a strategy τ^ guarantees $v + \varepsilon$ in $\Gamma(p_1)$.*

- *If both players can guarantee v , then v is called the uniform value of the game $\Gamma(p_1)$.*

Remark 3.9 *If the game $\Gamma(p_1)$ has a uniform value then it has an asymptotic value and both values coincides.*

In the following, we focus on the existence of an uniform value. Note that the existence of the uniform value is a very strong result. Indeed, informally it ensures that both players can guarantee that the mean-average opinion is above (resp. below) a fixed threshold. Moreover, contrary to the asymptotic approach, the optimal strategy does not depend on the length of the game.

In addition, the existence of the uniform value also implies the existence of other solutions concepts. In particular, the discounted value v_ρ also converges to the uniform value when the influencers become patient.

4 Existence of a uniform value

4.1 Existence Result

In the following, we restrict attention to connected social networks and further assume that the matrix M is primitive, that is:

Definition 4.1 A positive matrix M of size $K \times K$ is primitive if there exists $m \geq 1$ such that

$$\forall k \in K, \forall l \in K, (M^m)_{kl} > 0.$$

The assumption holds in particular if the network associated to M is connected and, at least one agent is self influencing (i.e $M_{k,k} > 0$ for some k). Informally, it implies that after m stages of the game, every non-strategic agent $k \in K$ is taking into account the opinion initially held by every other agent.

In this setting, one can show that the opinion game admits a uniform value and, moreover, that this value is independent of the initial distribution of opinions. Namely, one has:

Theorem 4.2 Let Γ be an opinion game such that M is primitive and (C^1, C^{-1}, s) be a signaling structure. For every p_1 , the opinion game $\Gamma(p_1)$ has a uniform value. Moreover there exists $v \in \mathbb{R}$ such that for every $p_1 \in X$,

$$v_\infty(p_1) = v.$$

Remark 4.3 In general, the value v_∞ does depend on the signaling structure (C^1, C^{-1}, s) .

A key insight brought about by Theorem 4.2 is that influencers can design robust targeting strategies that are ϵ -optimal independently of the initial distribution of opinions in the network or of the length of the game (provided it is long enough). In other words, influence strategies can be determined in function of the structure of the network only.

Remark 4.4 The key argument in the proof of Theorem 4.2 is to make appear a contracting property of the transitions. It implies that when considering a long block of actions, the final state almost does not depend on the starting state. Hence, our model shares some similarities with stochastic games with state independent transitions (Thuijsman, 1992) where the transition does not depend on the state variable. Nevertheless, our assumption is weaker since it is only asymptotically true.

4.2 Existence Proof

This subsection provides a proof of Theorem 4.2, i.e. the existence of a uniform value and its independence from the initial vector of opinions. Throughout, we fix an opinion game Γ such that M is primitive and a signaling structure (C^1, C^{-1}, s) . Moreover, we denote by m the smallest integer such that

$$\forall k \in K, \forall l \in K, (M^m)_{kl} > 0.$$

We shall also introduce a specific notation for indexed product of matrices. Namely, given a sequence of indexes $(i_1, j_1, \dots, i_n, j_n)$, we let:

$$\prod_{l=1}^n B(i_l, j_l) := B(i_n, j_n)B(i_{n-1}, j_{n-1}) \dots B(i_2, j_2)B(i_1, j_1).$$

A first lemma proves that the transition function is contracting over block of actions of length $m + 1$.

Lemma 4.5 There exists $\theta \in (0, 1)$ such that for every $h = (i_1, j_1, \dots, i_n, j_n, \dots) \in (I \times J)^\infty$, for every $p \in X$ and every $p' \in X$ then

$$\|q(p, h_{n(m+1)}) - q(p', h_{n(m+1)})\|_\infty \leq \theta^n \|p - p'\|_\infty.$$

Proof. Fix $h = (i_1, j_1, \dots, i_{(m+1)}, j_{(m+1)}) \in H_{n(m+1)}$ then

$$\begin{pmatrix} 1 \\ -1 \\ q(p, h) \end{pmatrix} = \begin{pmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ A_1(h) & A_{-1}(h) & B(h) \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ p \end{pmatrix} \quad (4.1)$$

where

$$B(h) = \prod_{l=1}^{n(m+1)} B(i_l, j_l),$$

$$A_1(h) = \sum_{l=1}^{n(m+1)} \left(\prod_{t=l+1}^{n(m+1)} B(i_t, j_t) \right) A_1(i_l, j_l)$$

and

$$A_{-1}(h) = \sum_{l=1}^{n(m+1)} \left(\prod_{t=l+1}^{n(m+1)} B(i_t, j_t) \right) A_{-1}(i_l, j_l).$$

By construction, for every $(i, j) \in I \times J$ $B(i, j)$ and M have the same positive coefficients. The primitiveness of M then implies that any product of m B -matrices has strictly positive coefficients. It follows that $A_1(h)$ and $A_{-1}(h)$ are strictly positive vectors. Then, since each linear combination computed in Equation 4.1 is convex, there must exist $\theta(h) \in (0, 1)$ such that for every $k, l \in K$,

$$0 \leq \sum_{l=1}^K B(h)_{kl} \leq \theta(h).$$

Let $p, p' \in X$, we then have

$$\begin{aligned} \|q(p, h) - q(p', h)\|_\infty &= \|A_1(h) - A_{-1}(h) + B(h)p - A_1(h) + A_{-1}(h) - B(h)p'\|_\infty, \\ &\leq \max_{k=1, \dots, K} \left| \sum_{l=1}^K B(h)_{kl} (p_l - p'_l) \right| \\ &\leq \theta(h) \|p - p'\|_\infty. \end{aligned}$$

Taking $\theta = \sup_{h \in H_{m+1}} \theta(h)$ and splitting an history $h \in H_{n(m+1)}$ of length $n(m+1)$ into n blocks yields by an immediate induction that

$$\|q(p, h) - q(p', h)\|_\infty \leq \theta^n \|p - p'\|_\infty.$$

■

Therefore, we can deduce that the payoff function is contracting over block of actions of length $m+1$:

Corollary 4.6 *There exists $\theta \in (0, 1)$ such that for every strategy $\sigma \in \Sigma$ and every strategy $\tau \in \mathcal{T}$, for every $p \in X$ and every $p' \in X$ then*

$$\|\gamma_{n(m+1)}(p, \sigma, \tau) - \gamma_{n(m+1)}(p', \sigma, \tau)\| \leq \frac{1}{n} \frac{1 - \theta^n}{1 - \theta} \|p - p'\|_\infty.$$

Proof. Given an history of actions $(i_1, j_1, \dots, i_{n(m+1)-1}, j_{n(m+1)-1})$, we denote by $(p_1, \dots, p_{n(m+1)})$ and by $(p'_1, \dots, p'_{n(m+1)})$ the sequence of states generated by h respectively from $p = p_1$ and $p' = p'_1$. By the previous proposition, we know that for every $t \in \{1, \dots, n\}$,

$$\|p_{1+(t+1)(m+1)} - p'_{1+(t+1)(m+1)}\|_\infty \leq \theta \|p_{1+t(m+1)} - p'_{1+t(m+1)}\|_\infty$$

and for every $l \geq 1$,

$$\|p_{l+1} - p'_{l+1}\|_\infty \leq \|p_l - p'_l\|_\infty.$$

It follows that

$$\begin{aligned}
|\gamma_{n(m+1)}(p, h) - \gamma_{n(m+1)}(p', h)| &\leq \frac{1}{n(m+1)} \sum_{t=1}^{n(m+1)} |g(p_t) - g(p'_t)|, \\
&\leq \frac{1}{n(m+1)} \sum_{l=0}^{n-1} (m+1) \|p_{l(m+1)+1} - p'_{l(m+1)+1}\|_\infty. \\
&\leq \frac{1}{n(m+1)} \sum_{l=0}^{n-1} (m+1) \theta^l \|p_1 - p'_1\|_\infty. \\
&\leq \frac{1}{n} \frac{1 - \theta^n}{1 - \theta} \|p_1 - p'_1\|_\infty.
\end{aligned}$$

We now need to extend this result to strategies. **Since a signaling structure does not depend on the states**, given a pair of strategy (σ, τ) , the distribution over actions is the same starting from p or from p' , hence

$$|\gamma_{n(m+1)}(p, \sigma, \tau) - \gamma_{n(m+1)}(p', \sigma, \tau)| \leq \frac{1}{n} \frac{1 - \theta^n}{1 - \theta} \|p - p'\|_\infty.$$

■

Proposition 4.7 *Let $w = \sup_{p \in X} (\limsup v_n(p))$ then Player 1 can uniformly guarantee the payoff w . More precisely, for every $\varepsilon > 0$ there exists $N \geq 1$ and $\sigma^* \in \Sigma$ such that*

$$\forall p \in X, \forall n \geq N, \forall \tau \in \mathcal{T}, \gamma_n(p, \sigma^*, \tau) \geq w - \varepsilon.$$

Proof. Let $\varepsilon > 0$ and $n_0 = \max\left(\frac{8(m+1)}{\varepsilon}, \frac{4}{\varepsilon(1-\theta)}\right)$. By definition of w , there exists $p^* \in X$ and $n \geq n_0$ such that $v_n(p^*) \geq w - \varepsilon/8$. For every $n, l \geq 1$, for all $p \in X$ we have

$$|v_n(p) - v_l(p)| \leq \frac{|n - l|}{\min(n, l)}.$$

Denote $N = \lceil \frac{n}{m+1} \rceil$, it follows that $n \leq N(m+1) \leq n + (m+1)$ and

$$v_{N(m+1)}(p^*) \geq v_n(p^*) - \frac{(m+1)}{n} \geq w - \varepsilon/8 - \varepsilon/8 = w - \varepsilon/4.$$

Let σ^* be a strategy such that for every $\tau \in \mathcal{T}$,

$$\gamma_{N(m+1)}(p^*, \sigma^*, \tau) \geq v_{N(m+1)}(p^*).$$

By Corollary 4.6, it follows that for every $p \in X$ and every $\tau \in \mathcal{T}$,

$$\begin{aligned}
\gamma_{N(m+1)}(p, \sigma^*, \tau) &\geq \gamma_{N(m+1)}(p^*, \sigma^*, \tau) - \frac{1}{N} \frac{1}{1 - \theta} \|p - p^*\|_\infty, \\
&\geq v_{N(m+1)}(p^*) - \varepsilon/2, \\
&\geq w - 3\varepsilon/4.
\end{aligned}$$

We proved that player 1 has one strategy that can guarantee $w - \varepsilon$ whatever is the initial vector of opinions. It implies that he can guarantee $w - \varepsilon$ in the following modified game: at the initial stage, an opinion vector p is randomly chosen along $\mu \in \Delta_f(X)$. Player -1 learns p whereas Player 1 only knows μ . Hence Player -1 may use different strategies for different vector opinion, whereas Player 1 is restricted to use the same one. A strategy of Player -1 in this new game is therefore a collection of strategies, one for each $p \in X$. Given $\sigma \in \Sigma$ and $\tau \in (\mathcal{T})^X$, we have

$$\gamma_N(\mu, \sigma, \tau) = \sum_{p \in X} \mu(p) \gamma_N(p, \sigma, \tau(p)).$$

It follows that for every $\tau \in (\mathcal{T})^X$,

$$\gamma_N(\mu, \sigma^*, \tau) \geq w - \varepsilon.$$

Hence Player 1 can guarantee $w - \varepsilon$.

The previous case was a worst-case scenario where Player 1 has no information whereas Player -1 knows perfectly the initial vector of opinion. Since Player 1 can guarantee $w - \varepsilon$ in this game, he can also guarantee $w - \varepsilon$ in any N -stage game where he learns more and Player -1 learns less.

We now consider the strategy of Player 1 that repeatedly follow σ^* during N stages. This strategy guarantees $w - 2\varepsilon$ in the long run. \blacksquare

Remark 4.8 *Contrary to the static model of Grabisch et al. (2017), connectivity is not a sufficient condition for our proof. Given a 2-periodic Markov Chain M , one can define the network associated to the Markov chain M^2 . This new network has at least 2 separated components. In Grabisch et al. (2017), this is not a problem since a strategic may not change its target: every non-strategic agent is influenced at every stage in the long run. Since in our model targets can change, for certain strategies, the influence of a player might never reach certain non-strategic. Our proof wouldn't apply in this case.*

5 Characterization of Equilibrium Strategies

In this section, we focus on the characterization of equilibrium strategies in the opinion game Γ . As emphasized in Theorem 4.2, the asymptotic value of the game and asymptotically optimal strategies are independent of the initial distributions of beliefs among non-strategic agents. Therefore, we can focus on the case where initial opinions are uniformly set to 0. In this case, we can write a closed form formula for the payoff.

Given a sequence of pair of actions $h = (i_1, j_1, \dots, i_N, j_N)$ the vector of beliefs after n steps is given by¹:

$$p_n = \sum_{l=1}^n \left(\prod_{m=l+1}^n B(i_m, j_m) \right) (A_1(i_l, j_l) - A_{-1}(i_l, j_l)).$$

Accordingly, the payoff in period n is the mean average over the K non-strategic agents in the network:

$$\frac{1}{K} e' \cdot \sum_{l=1}^n \left(\prod_{m=l+1}^n B(i_m, j_m) \right) (A_1(i_l, j_l) - A_{-1}(i_l, j_l)).$$

And the average payoff in the game of length N is:

$$\gamma_N(h) = \frac{1}{NK} e' \cdot \sum_{n=1}^N \left(\sum_{l=1}^n \left(\prod_{m=l+1}^n B(i_m, j_m) \right) (A_1(i_l, j_l) - A_{-1}(i_l, j_l)) \right) \quad (5.1)$$

5.1 An alternative ball game

Using an approach similar to Acemoglu et al. (2013), our characterization of optimal strategies builds on a correspondence² between the opinion game Γ and the stochastic game where players ought to catch balls that move backwards in time over the network according to the transition probabilities given by M .

More precisely, given an opinion game (K, M, D, λ, μ) , we can define a new game, which we will call the “ball game” and for which the timing of events is every period as follows:

¹Here and below, in order to simplify notation, we let by convention $\prod_{m \in \emptyset} B_m = I$ for any indexed family of matrices (B_m) where I is the identity matrix.

²This correspondence was also put forward by one of the anonymous referee to Grabisch et al. (2017).

- The two strategic players choose nodes in the network where to position themselves, denoted respectively by i and j .
- A new ball is launched uniformly at random (uniform distribution) in the network.
- A ball in node k is caught by player 1 with probability $A_1(i, j)_k$ (resp. by player -1 with probability $A_{-1}(i, j)_k$)
- All balls that have not been caught yet make a move in the network according to the transition probabilities given by the matrix M .

After n stages, the game stops and the winner is the strategic agent that caught more balls. The ball game can be related to the opinion game by inverting the direction of time. Namely, the probability for a strategic player of catching a ball launched in node k after n steps corresponds to the influence he exerts after n steps on the opinion of agent k .

To formalize this relation, let us first note that the ball game can be represented as a stochastic game, which we shall denote $\overleftarrow{\Gamma} = (Y, I, J, f, r)$ and which has the following characteristics.

Definition 5.1 *The game $\overleftarrow{\Gamma} = (Y, I, J, f, r)$ is the stochastic game defined as follow.*

- The set of states is $Y = \mathbb{R}_+^K$ and represents the expected number of balls in each node of the network.
- The set of actions of player 1 is $I = K$, i.e the choice of a position in the network where to catch balls.
- The set of actions of player -1 is $J = K$, i.e the choice of a position in the network where to catch balls.
- The transition $f : Y \times I \times J \rightarrow Z$ is defined for all $y \in Y$, for all $i \in I$ and for all $j \in J$ by

$$f(y, i, j) = yB(i, j) + \frac{1}{K}e'. \quad (5.2)$$

It represents the evolution of the expected number of balls in each node of the network after the already present balls have been caught or have moved over the network according to the Markov chain and a new ball has been added uniformly at random.

- The stage payoff $r : Y \times I \times J \rightarrow [-1, 1]$ is given by the difference between the expected number of balls caught by players 1 and -1 at the current stage:

$$r(y, i, j) = y(A_1(i, j) - A_{-1}(i, j)).$$

One can remark, that the law of evolution of the expected number of balls (Equation 5.2) has an inverse temporal structure with respect to the law of evolution of the opinions (Equation 3.1). More precisely, the expected number of balls that can be caught in node k , given an history of length n h_n , corresponds to the cumulative influence that the choice of k as a target in the current stage would yield in the coming n steps, if both players were then to play the time-inverse of the history h_n . Formally, one has the following proposition.

Proposition 5.2 *Let $n \geq 1$ and $h = (i_1, j_1, \dots, i_n, j_n)$. Denote by $\overleftarrow{h} = (i_n, j_n, \dots, i_1, j_1)$ the time-inverted history, then we have*

$$\gamma_n(\overleftarrow{h}) = \overleftarrow{\gamma}_n(h),$$

where $\overleftarrow{\gamma}$ denote the n -stage payoff in the game $\overleftarrow{\Gamma}$.

Proof. Let us compute the payoff associated to a sequence of actions $h = (i_1, j_1, \dots, i_N, j_N)$ in the game $\overleftarrow{\Gamma}$. We denote by $\overleftarrow{h} = (i_N, j_N, \dots, i_1, j_1) = (i'_1, j'_1, \dots, i'_N, j'_N)$ the inverted history. The expected number of balls in each of the K nodes of the network after l moves is denoted by $y_l \in \mathbb{R}_+^K$. It satisfies the following equations:

- $y_1 = \frac{1}{K}e'$,
- for $l \geq 1$, $y_{l+1} = y_l B(i_l, j_l) + \frac{1}{K}e'$

Hence, we have

$$y_l = \sum_{n=1}^l \frac{1}{K}e' \left(\prod_{m=n}^{l-1} B(i_m, j_m) \right) \quad (5.3)$$

The payoff is then given by the difference between the expected number of balls caught by players 1 and -1, that is

$$\overleftarrow{\gamma}_N(h) := \sum_{l=1}^N y_l [A_1(i_l, j_l) - A_{-1}(i_l, j_l)] \quad (5.4)$$

or equivalently:

$$\overleftarrow{\gamma}_N(h) = \sum_{l=1}^N \left(\sum_{n=1}^l \frac{1}{K}e' \left(\prod_{m=n}^{l-1} B(i_m, j_m) \right) \right) [A_1(i_l, j_l) - A_{-1}(i_l, j_l)] \quad (5.5)$$

$$= \sum_{l=1}^N \left(\sum_{n=1}^{N+1-l} \frac{1}{K}e' \left(\prod_{m=n}^{l-1} B(i'_{N+1-m}, j'_{N+1-m}) \right) \right) [A_1(i'_l, j'_l) - A_{-1}(i'_l, j'_l)] \quad (5.6)$$

$$= \sum_{l=1}^N \left(\sum_{n=1}^{N+1-l} \frac{1}{K}e' \left(\prod_{m=N+1-(l+1)}^{N+1-n} B(i'_m, j'_m) \right) \right) [A_1(i'_l, j'_l) - A_{-1}(i'_l, j'_l)] \quad (5.7)$$

$$= \sum_{l=1}^N \left(\sum_{n=l}^N \frac{1}{K}e' \left(\prod_{m=l+1}^n B(i'_m, j'_m) \right) \right) [A_1(i'_l, j'_l) - A_{-1}(i'_l, j'_l)] \quad (5.8)$$

We obtain the same formula than Equation 5.1 proving that the two payoffs are indeed equal. They are both expressing the time-average of a Markov process. ■

An information structure of the form (C^1, C^{-1}, s) can be associated to the game $\overleftarrow{\Gamma}$ in a similar manner than for the game Γ (see definition 3.4). Hence, strategies in the games Γ and $\overleftarrow{\Gamma}$ formally have exactly the same structure. However, the two games are played in two inverse direction of times. This makes impossible in general to use a strategy in one the two games in the other game. However, a fundamental exception is the state-blind information structure. In this setting, the players do not observe their opponent's actions ex interim and have to best response to the complete history of play. Then, the direction of time is not pertinent anymore and can be discarded. It follows immediately from Proposition 5.2 that, then, the two games are equivalent.

Corollary 5.3 *In the case of a blind structure of information, the two games Γ and $\overleftarrow{\Gamma}$ are equivalent.*

Remark 5.4 *We could have defined a third game played as follows. Each player is choosing a strategy in Γ . The two strategies induce a probability distribution π over finite histories of length n . Then we can compute the payoff by using the "ball game" interpretation: consider the probability distribution $\overleftarrow{\pi}$ induced by π by inverting time. We then compute the payoff as the payoff in the "ball game". By Proposition 5.2, this new game is completely equivalent to the game Γ . It only differs on the interpretation of the payoff function.*

5.2 Forward looking equilibrium strategies

On top of providing a correspondence between the games Γ and $\overleftarrow{\Gamma}$, pure equilibrium strategies for the blind-information structure are also very robust. Indeed, they must provide a best-response to any history of play by the opponent. If such a pure strategy exists, it can in fact be used for any information structure.

Namely, given a pure strategy σ of player 1 for the blind-information structure (i.e an history of play), let us denote by $\sigma_{(C^1, C^{-1}, s)}$ the strategy in the game with information structure (C^1, C^{-1}, s) that plays the sequence of actions σ independently of the actions of the opponent (and respectively for player -1). This was already noticed in a weaker form by Amir (2003) that was comparing strategies in the blind case and strategies in the complete information case in deterministic stochastic games.

Proposition 5.5 *Let (σ, τ) be an equilibrium in pure strategies in the n -stage blind game Γ (resp. $\overleftarrow{\Gamma}$). Then, for any information structure (C^1, C^{-1}, s) , $(\sigma_{(C^1, C^{-1}, s)}, \tau_{(C^1, C^{-1}, s)})$ is an equilibrium in the n -stage game Γ (resp. $\overleftarrow{\Gamma}$). Moreover all the games have the same value.*

Proof. It is a general result that apply to any game. We will prove the result in the case of Γ . Recall that a strategy in the blind game is a distribution over actions. Hence a pure strategy is a sequence of actions. Let $(\sigma, \tau) = ((i_n)_{n \geq 1}, (j_n)_{n \geq 1})$ be an equilibrium in the blind game. We know that for every sequence of actions of player 2, $(j'_n)_{n \geq 1}$,

$$v_n = \gamma_n((i_1, \dots, i_n), (j_1, \dots, j_n)) \leq \gamma_n((i_1, \dots, i_n), (j'_1, \dots, j'_n)).$$

Let (C^1, C^{-1}, s) be a structure of information. Given a strategy τ' for this new structure of information, $(\sigma_{(C^1, C^{-1}, s)}, \tau')$ induces a distribution over actions such that the marginal on $I^{\mathbb{N}}$ is σ . Hence

$$\gamma'_n(\sigma_{(C^1, C^{-1}, s)}, \tau') \geq \gamma_n((i_1, \dots, i_n), (j_1, \dots, j_n)) = v_n.$$

It follows that $\sigma_{(C^1, C^{-1}, s)}$ also guarantees v_n . By symmetry, we can show that $\tau_{(C^1, C^{-1}, s)}$ is guaranteeing v_n . ■

A natural question then is whether we can characterize equilibrium strategies for the opinion game with blind-information structure. In this perspective, we can build on Corollary 5.3 and first characterize equilibrium strategies in the ball game $\overleftarrow{\Gamma}$. Indeed, one obtains as a direct corollary of Proposition 5.5 and Corollary 5.3 that:

Lemma 5.6 *Assume that the blind "ball game" admits a pure Nash equilibrium (σ, τ) . Then for any information structure (C^1, C^{-1}, s) , $(\sigma_{(C^1, C^{-1}, s)}, \tau_{(C^1, C^{-1}, s)})$ is an equilibrium of the opinion game.*

Now, one can note that, in the ball game, the choice of a target has two consequences. First, it determines the probability to catch a ball this period. Second, it influences the future distribution of balls in the network. Yet, at an equilibrium for the blind-information structure, the players can focus on maximizing the probability of catching the ball this period without concern for the future distribution of balls induced. Indeed, on the one hand the value of the game with equal influence is zero and, on the other hand, if at equilibrium the probability of catching a ball wasn't maximized every period, a player could deviate to the strategy that maximizes the probability to catch the ball this period (in order to gain a payoff advantage) and then mimic the strategy of the opponent for all the remaining periods in order to guarantee a zero payoff in future periods. In other words, one has:

Corollary 5.7 *Assume that the blind "ball game" of length N admits a pure Nash equilibrium (σ, τ) then for every $n \in \{1, \dots, N\}$,*

$$\sigma_n \in \{k, y_n(k) = \|y_n\|_\infty\},$$

and

$$\tau_n \in \{k, y_n(k) = \|y_n\|_\infty\},$$

Note that, from the point of view of the opinion game, $y_n(k)$ represents the cumulative influence that the player would exert in future periods (given the future pattern of actions) if he were to target agent k this period. Hence, targeting the agent for which $y_n(k)$ is maximum amounts to adopt a purely forward looking perspective, i.e. to only take into consideration the future propagation of one's influence. In particular, from this perspective, a player is only concerned by the future actions of his opponent and can neglect the current state of opinions in the network. This is to be contrasted with the general case in which, à priori, a player ought to implement a trade-off between maximizing his own future influence and counteracting the influence exerted by his opponent in the previous periods.

Remark 5.8 *Given the inversion of the direction of time between the ball game and the opinion game, maximizing future influence in the opinion game amounts to maximize the probability to catch a ball in the current period while counteracting the influence exerted by the opponent in the previous periods corresponds to minimizing the probability for the opponent to catch a ball in future periods.*

In the blind information structure, the player can not condition his behavior on previous actions of his opponent (because the complete sequence of targets must be chosen ex ante). Hence, he must adopt a purely forward looking behavior at a pure equilibrium. This requires the existence of a target (see Example 5.9) or of a sequence of targets (see Example 5.10) that strongly dominates in terms of influence propagation. If such a dominant sequence of target does not exist, the players ought to add a backward looking perspective to their strategies in order to counteract their opponent's influence (and conversely to prevent their influence from being counteracted by their opponent). In the blind-information game, the only way to account for this backward looking perspective is to play a mixed strategy.

For finer information structures, this backward looking perspective can be implemented by conditioning behavior on the previous action of the opponent or, equivalently, on the current state of opinions in the network. Indeed, in order to react to the current state of opinions, the player must be able to infer it from the previous action of his opponent, i.e. to adopt a backward looking perspective (see 3.5). As highlighted in Example 5.11, this perspective enriches the set of strategies available to the players so that there might exist equilibria in pure strategies with complete information for networks in which there is no equilibrium in pure strategies for the blind information structure.

The introduction of a dynamic perspective also allows to consider a much richer set of behavior than in the static game à la Grabisch et al. (2017) where both players choose a single target that is then fixed throughout the game. As illustrated in Example 5.11, the flexibility brought by the choice of a sequence of targets allows to devise equilibrium in pure strategies for networks in which no pure equilibrium exists in the static game.

Example 5.9 (*Star*) *We consider a star-network with N agents. Agent 1 is the central agents that is linked to each of the other agents. The $N - 1$ other agents are only listening to agent 1. If the lobbying power is small compared to $N - 1$ then, whatever the strategies of the player, the number of ball in the center is far greater than the number of balls on the side and therefore the influencers should focus on the center of the star.*

Example 5.10 *Let us consider the network represented in figure 1, which consists of three layers. The lower layer has one node, which is influencing the n nodes of the second layer, which in turn are each influencing n^2 nodes at the third layer. Moreover, each node of the third layer is influencing equally the node of the first layer. The games with blind information structure are straightforward to analyze in this setting using the ball game metaphor (arrows point in the direction of the movement of balls³). In the first period of the ball game, any target can be chosen, according to corollary 5.7, as the balls are uniformly distributed. In the second period, a node in the second layer must be chosen as it is clearly in the second layer that the expected number of balls is the largest (independently of the action in the first period). From the third period on, the expected number of balls is the largest in the node of the first layer, which must then be chosen as a target.*

³Equivalently an agent has outgoing arrows towards the agents that influence him

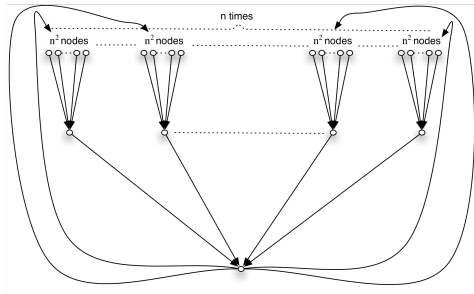


Figure 1: network with layers of increasing size

This example can be generalized by recursion to an arbitrary number K of layers with layer $k + 1$ having n^{a_k} clusters of $n^{a_{k+1}}$ when layer k has n^{a_k} nodes. Then, players shall target the upper layer in the first period and go down layer by layer period after period until they reach the lowest layer, which they shall then systematically target. In the influence game, the timing is inverted, agents should first target the agent in the lower layer (whose influence propagates in the direction opposite to arrows) and, in the last K periods of the game, sequentially target one of the agent in each of the upper layers.

Example 5.11 Let us then consider the the network represented in Figure 2. It is formed on the canvas of a circle among three non-strategic agents, $K = \{\alpha, \beta, \gamma\}$. The circle is however perturbed to ensure the network's matrix is primitive, consistently with our assumption for the existence result (the results below nevertheless apply to the circle as well).

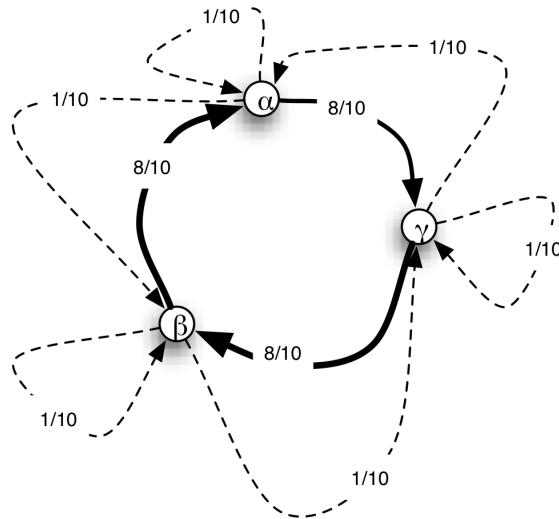


Figure 2: Perturbed Circular Network

Formally, the interactions are captured by the following matrix M .

$$\begin{array}{c}
 \alpha \quad \beta \quad \gamma \\
 \alpha \quad \begin{pmatrix} 1/10 & 8/10 & 1/10 \\ 1/10 & 1/10 & 8/10 \\ 8/10 & 1/10 & 1/10 \end{pmatrix} \\
 \beta \\
 \gamma
 \end{array}$$

In this setting, we assume that both players have equal influence potential $\lambda = \mu = 1$, whereas the local weight are equal to $D(\alpha) = D(\beta) = D(\gamma) = 0$. In other words, the relative influence of strategic players is infinite. We consider three variations of the opinion game in this case: the static case à la Grabisch et al. (2017), the dynamic game with complete and blind information structures respectively. We shall show that there exists equilibrium in pure strategies for the complete information structure whereas in the two other cases, there only exist mixed equilibria.

1. We start by studying the static game à la Grabisch et al. (2017) where both players choose a fixed target in view of maximizing (resp. minimizing) the average asymptotic opinion in the network. One can easily check that, in our setting, the asymptotic opinions of the non-strategic agents for each pair of actions are given by:

$$\begin{pmatrix} (0, 0, 0) & (1, -1, 7/9) & (1, -7/9, -1) \\ (-1, 1, -7/9) & (0, 0, 0) & (7/9, 1, -1) \\ (-1, 7/9, 1) & (-7/9, -1, 1) & (0, 0, 0). \end{pmatrix}$$

Consequently, one obtains the following payoff matrix for the game:

$$\begin{pmatrix} 0 & 7/9 & -7/9 \\ -7/9 & 0 & 7/9 \\ 7/9 & -7/9 & 0 \end{pmatrix}.$$

It is clear that there is no pure equilibrium in the static game.

2. Let us now consider the dynamical version of the game with complete information. We consider the following strategy of Player 1: given a vector of opinion (a, b, c) consider the node with the lowest opinion and target the preceding in the list (α if it is β , β if it is γ and γ if it is α). Let us show that this strategy ensures that starting from the initial vector of opinion $(0, 0, 0)$ then the mean-average opinion is non-negative at every stage in the society.

Starting from $(0, 0, 0)$, there are only two different cases. Either both lobbies target the same nodes, yielding to a new vector of opinions $(0, 0, 0)$, or they target two different nodes yielding a vector of opinions of the form $(-1, 0, 1)$. In both cases, the mean-average is non-negative.

Let us consider now consider opinion vectors of the form $(-1, 0, 1)$. By assumption, Player 1 targets γ and Player -1 has three possibilities:

- If Player -1 targets α , then the new vector of opinion is $(-1, 7/10, 1)$ that is greater coordinate by coordinate than $(-1, 0, 1)$
- If Player -1 targets β , then the new vector of opinion is $(0, -1, 1)$ that is symmetric to $(-1, 0, 1)$
- If Player -1 targets γ , then the new vector of opinion is $(0, 7/10, 0)$ that is greater coordinate by coordinate than $(0, 0, 0)$

All the vector of opinions that appeared in the previous discussion have a non-negative average, hence Player 1 strategy guarantees a non-negative payoffs. Hence, by symmetry, the game has a value of 0. therefore the game has a pure equilibrium.

3. Finally, we consider the dynamical version of the game for the blind information structure. In this framework, let us prove that Player 1 can not guarantee with a pure strategy a non-negative payoff in any N -stage game. In order to prove this result, we will use the correspondence between the opinion game Γ and the Ball game $\bar{\Gamma}$.

Let σ be a pure strategy of Player 1 in the blind opinion game, i.e. a sequence of actions (i_1, \dots, i_N) . We define then the reverse sequence of actions $(i'_1, \dots, i'_N) = (i_N, \dots, i_1)$ and consider the strategy of Player -1 that chooses an action in period t according to the following rule:

- α if $i'_{t+1} = \beta$,
- β if $i'_{t+1} = \gamma$,
- γ if $i'_{t+1} = \alpha$.

That is player -1 chooses the node right before the one previously targeted by player 1, where the probability to find a ball is maximum. As $D(k) = 0$, both players catch the ball with probability 1 if they target the right node. Hence, one can easily check that player -1 hence ensures that Player 1 catch at most $1/3$ of the balls and therefore can't guarantee a non-negative payoff.

5.3 Characterization in the case of small influence

Finally, we focus on the limit case where the influence exerted by the strategic players is vanishingly small. In this setting, the dynamics without external influence provide a useful benchmark: one knows that the vector of beliefs converge to a consensus in which the influence of each initial opinion is proportional to the eigenvector centrality of the corresponding node (see e.g. Golub and Jackson, 2010). Accordingly, if balls move in the network without being caught, they should eventually distribute themselves according to the invariant distribution of M , which is proportional to eigenvector centrality. Both remarks suggest that strategic players with small influence should target nodes with high eigenvector centrality in order to maximize their influence. This is actually the case. Namely, if there is a unique node i^* for which the eigenvector centrality of the matrix M is maximal, one can state the following result.

Theorem 5.12 *Assume that for every $k \in K$, $D(k) > 0$. There exists $N_0 \in \mathbb{N}$ and $\lambda_0 \in \mathbb{R}_+$ such that for every $N \geq N_0$, and for every $\lambda < \lambda_0$ any equilibrium strategy in the game $\overleftarrow{\Gamma}_N(\lambda)$ targets the node with the highest eigenvector centrality from stage N_0 .*

The previous result holds for any information structure, in particular it is true for the blind one. Since we showed that in this case, the game $\overleftarrow{\Gamma}$ and Γ are equivalent up-to the time inversion we obtain the following result.

Theorem 5.13 *Assume that for every $k \in K$, $D(k) > 0$. There exists $N_0 \in \mathbb{N}$ and $\lambda_0 \in \mathbb{R}_+$ such that for every $N \geq N_0$, and for every $\lambda < \lambda_0$ any equilibrium strategy in the game $\Gamma_N(\lambda)$ targets the node with the highest eigenvector centrality between stage 1 and stage $N - N_0$.*

Together with proposition 5.5, this implies that if influence levels are small enough and the game is long enough, it is an equilibrium strategy to target the node with the largest eigenvector centrality independently of the information structure in the game. This is again an instance of forward looking strategy where an agent can restrict attention to the diffusion of his own opinion. This is also adding evidence to the importance of eigenvector centrality as a measure of influence.

5.3.1 Proof of Theorem 5.12

Given $h = (i_1, j_1, i_2, j_2, \dots) \in H_\infty$ an infinite history of actions, we denote by h' the translated history defined by $h' = (i_2, j_2, i_3, j_3, \dots)$. We then define by induction for every $l \geq 2$ the l -translated history by $h^{(l)} = (h')^{(l-1)}$.

Let us denote by $\chi \in \Delta(K)$ the eigenvector centrality of M , and let $i^* \in K$ be the unique node with maximal eigenvector centrality. We then set

$$\epsilon := \frac{1}{4} \min_{i \neq i^*} |\chi_{i^*} - \chi_i|. \quad (5.9)$$

By assumption, one has $\epsilon > 0$.

If M is a primitive matrix, the sequence $(M^n)_{n \geq 1}$ is converging to some matrix M^∞ . Let $N_0 \geq 1$ such that for every $n \geq N_0$, $\|M^n - M^\infty\|_\infty \leq \varepsilon$ and $\|M^n e' - \chi\|_\infty \leq \varepsilon$. Let $N_1 \geq \max(\frac{N_0}{\varepsilon}, 2)$.

Let us now define a model that keep track of every ball in the "Ball game" instead of only considering the expected number of balls at every stage. Informally, for every $l \geq 1$ we define $(X^l)_{t \geq 1}$ a process describing the behavior of the ball introduced at stage l and T_l the time where the ball introduced at time l is caught by one of the two players. For every $l \geq 1$, we define the stochastic process $(X^l)_{t \geq 1}$ (inhomogenous Markov chain) with values in $\{\emptyset\} \cup K$ such that

- for every $t \leq (l-1)$, $X_t^l = \emptyset$,
- for $t = l$, X_t^l is uniformly distributed over K ,
- for every $t \geq l$, if $X_t^l = \emptyset$ then $X_{t+1}^l = \emptyset$. If $X_t^l = k \in K$ then X_{t+1}^l has the following distribution
 - with probability $A_1(i_t, j_t)_k + A_{-1}(i_t, j_t)_k$, it is equal to \emptyset .
 - with probability $B(i_t, j_t)_{k,m}$, it is equal to m .

We assume that conditionally on the sequence h , the stochastic processes are independent. Finally, we define the stopping time $T_l = \inf_{t \geq l} \{X_t^l = \emptyset\}$ corresponding to the stage the ball introduced at stage l is caught by one of the influencer.

For every $k \in K$ and $h \in H_\infty$, one can check that

$$y_m(k) = \mathbb{E}_h (\#\{X_m^l(k), l \geq 1\}) = \sum_{t=1}^{+\infty} \mathbb{P}_h(X_m^l = k).$$

Hence the correspondence between the two models.

Let us now prove that provided the influence potentials are sufficiently small, we can guarantee that the probability that each ball is absorbed in less than N_1 stages is small and therefore at every stage, the probability that the last N_1 Balls introduced in the network have not been caught is close to 1.

We first focus on the probability of absorption of the first ball. We then deduce that at stage N_1 all the balls are still present in the network with high probability and we conclude by an invariance property.

Lemma 5.14 *There exists $\lambda_0 > 0$ such that for every $h \in H_\infty$ and for every $\lambda, \mu \leq \lambda_0$,*

$$\mathbb{P}_h(T_1 \geq N_1 + 1) \geq (1 - \varepsilon)^{\frac{1}{N_1}}.$$

Proof. Fix an history $h \in H_\infty$. Denote by $\underline{D} = \min_{k \in K} D(k) > 0$ and by $\eta = 1 - (1 - \varepsilon)^{(\frac{1}{N_1})^2}$. Let $\lambda_0 = \frac{\eta}{1 - \eta} \underline{D}$. At every stage t , the probability that the ball is caught is

- 0 if the ball is at a node that is not targeted by any player,
- $\frac{\lambda}{\lambda + D(k)}$ or $\frac{\mu}{\mu + D(k)}$ if the node is targeted by one of the two lobbies,
- $\frac{\lambda + \mu}{\lambda + \mu + D(k)}$ if the node is targeted by the two lobbies.

Since the function $x \rightarrow \frac{x}{x + D(k)}$ is increasing in x and decreasing in $D(k)$, we obtain that

$$\mathbb{P}_h(T_1 = t | T_1 \geq t) \leq \frac{2\lambda_0}{2\lambda_0 + \underline{D}}.$$

Hence

$$\mathbb{P}_h(T_1 = t | T_1 \geq t) \leq \eta.$$

It follows by the chain rule that

$$\mathbb{P}_h(T_1 \geq N_1 + 1) \geq (1 - \eta)^{N_1}.$$

By the choice of η , we obtain the result. ■

It follows that at stage N_1 , we are sure that the probability that none of the ball has been caught is

$$\begin{aligned} \mathbb{P}_h(\cap_{l=1}^{N_1} \{T_l \geq N_1\}) &\geq \mathbb{P}_h(\cap_{l=1}^{N_1} \{T_l \geq N_1 + l\}) \\ &\geq \prod_{l=1}^{N_1} \mathbb{P}_h(T_l - l \geq N_1) \\ &\geq \prod_{l=1}^{N_1} \mathbb{P}_{h^{(l)}}(T_l \geq N_1) \\ &\geq (1 - \varepsilon)^{\frac{1}{N_1} N_1} \\ &\geq 1 - \varepsilon. \end{aligned}$$

By assumption given a sequence h of actions, the sequence $(T_i)_{i \geq 1}$ is a sequence of independent variables. Moreover, the law of T_{l+1} under the sequence h is the same as the law of $T_1 + N_1$ under $h^{(l)}$. The previous result being true for every $h \in H_\infty$, we have more generally,

$$\mathbb{P}_h(\cap_{l=N}^{N+N_1} \{T_l \geq N + N_1\}) \geq 1 - \varepsilon.$$

Informally, we are sure that with high probability all the balls that have been introduced in the network between stage N and $N + N_1$ are still there at stage $N + N_1$.

Let us deduce that the distribution of balls at stage $N + N_1$ is very close to the ergodic distribution of M . There are three types of balls:

- Balls that have been caught and that are not counted anymore.
- Balls that have been introduced between stage 1 and $N + N_1 - N_0$ and not caught.
- Balls that have been introduced between stage $N + N_1 - N_0$ and $N + N_1$.

The first case does not concern us. Each of the balls in the second case is distributed close to the invariant measure at stage $N + N_1$. Moreover, we know that with high probability, there are at least N_1 balls in this case. Finally, we know that there are at most N_0 balls in the last case. It follows that the distribution on Balls is mainly driven by the invariant distribution. Formally, we have

$$y_m(k) = \mathbb{E}_h(\#\{l \in [1, m], X_m^l = k\}) \tag{5.10}$$

$$= \sum_{l=1}^m \mathbb{P}_h(X_m^l = k), \tag{5.11}$$

$$= \sum_{l=1}^m \mathbb{P}_h(X_m^l = k | X_m^l \neq \emptyset) \mathbb{P}_h(X_m^l \neq \emptyset), \tag{5.12}$$

$$= \sum_{l=1}^{N+N_1-N_0} \mathbb{P}_h(X_m^l = k | X_m^l \neq \emptyset) \mathbb{P}_h(X_m^l \neq \emptyset) + \sum_{l=N+N_1-N_0}^{N_0} \mathbb{P}_h(X_m^l = k | X_m^l \neq \emptyset) \mathbb{P}_h(X_m^l \neq \emptyset). \tag{5.13}$$

In addition, we know that if $l \leq N + N_1 - N_0$, then

$$|\mathbb{P}_h(X_m^l = k | X_m^l \neq \emptyset) - \chi(k)| = \left| \sum_{p=1}^K \frac{1}{K} M_{pk}^{m-l} - \sum_{p=1}^K \frac{1}{K} M_{pk}^\infty \right| \tag{5.14}$$

$$\leq \varepsilon, \tag{5.15}$$

and

$$\sum_{l=1}^{N+N_1-N_0} \mathbb{P}_h(X_m^l \neq \emptyset) \geq N_1(1-\varepsilon). \quad (5.16)$$

Let us deduce from Equation 5.13, Equation 5.14 and Equation 5.16 that $y_m(i^*) > \max_{i \neq i^*} y_m(i)$. We have

$$\begin{aligned} y_m(i^*) &\geq (\chi(i^*) - \varepsilon) \left(\sum_{l=1}^{N+N_1-N_0} \mathbb{P}_h(X_m^l \neq \emptyset) \right), \\ &\geq (\chi(i^*) - \varepsilon) N_1(1 - \varepsilon), \\ &\geq \chi(i^*) N_1 - N_1 \varepsilon - \varepsilon \chi(i^*). \end{aligned}$$

Moreover, for every $i \in K$, we have

$$\begin{aligned} y_m(i) &\leq (\chi(i) + \varepsilon) \left(\sum_{l=1}^{N+N_1-N_0} \mathbb{P}_h(X_m^l \neq \emptyset) \right) + \sum_{l=N+N_1-N_0}^{N+N_1} \mathbb{P}_h(X_m^l \neq \emptyset), \\ &\leq (\chi(i) + \varepsilon) N_1 + N_0, \\ &\leq \chi(i) N_1 + N_1 \varepsilon + N_0. \end{aligned}$$

Hence the difference between $y_m(i^*)$ and $y_m(i)$ is equal to

$$y_m(i^*) - y_m(i) \geq N_1 \left((\chi(i^*) - \chi(i) - 2\varepsilon - \frac{N_0}{N_1}) - \varepsilon \geq N_1 (4\varepsilon - 2\varepsilon - \varepsilon) - \varepsilon \geq \varepsilon. \right.$$

and we have the strict inequality. Since this is true for every $m \geq 1$ and independently of the actions played by the players, it follows that the optimal strategy for both players is to target i^* at every stage.

6 Conclusion

We consider the problem of maximizing influence through the choice of a sequence of targets when two players compete for predominance over a social network. We show that the problem can be captured as a stochastic game with compact state space. This raises the question of the existence of a uniform value, which we positively answer under the assumption that the adjacency matrix of the network is primitive. Hence strategic players can design optimal targeting strategies independently of the length of the game or of the discount factor (provided it is small enough).

Further analysis of the structure of equilibrium strategies shows that they strongly depend on the level of information each player has about the actions of his opponents and hence about the state of opinions in the network. In absence of information, there exist pure equilibria only if players can devise strategies that are efficient independently of the current state of the network, i.e. they must adopt a purely forward looking perspective on the diffusion of their influence. When players are well informed, they can implement more flexible strategies that account for the state of opinions in the network and implement a trade-off between a forward-looking perspective, according to which they aim at maximizing the future spread of their opinion in the network, and a backward-looking perspective, according to which they aim at counteracting their opponent's previous actions. This flexibility sustains the existence of equilibria in pure strategies in networks where there is no such equilibrium with low level of information or in the static game à la Grabisch et al. (2017) with fixed targets.

These results emphasize the pertinence of adopting a dynamic approach, in particular in view of empirical applications: sequential strategies allow to solve the targeting problem in a wide range of networks and independently of the length of the game. A complete characterization of these strategies seems very

difficult to obtain analytically but it seems to us that an algorithmic approach could be fruitfully applied to the problem in future work.

Acknowledgment: Mandel acknowledges the support of the European Projects Future and Emerging Technologies (FET) SIMPOL (grant no. 610704) and DOLFINs (grant no. 640772). Venel acknowledges the support of the Agence National de la Recherche [ANR CIGNE ANR-15-CE38-0007-01].

References

- D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar. Opinion fluctuations and disagreement in social networks. *Mathematics of Operations Research*, 38(1):1–27, 2013.
- R. Amir. Stochastic games in economics and related fields: an overview. In *Stochastic Games and Applications*, pages 455–470. Springer, 2003.
- P. Amorós and M. S. Puy. Issue convergence or issue divergence in a political campaign? *Public Choice*, 155(3-4):355–371, 2013.
- E. Aragonés, M. Castanheira, and M. Giani. Electoral competition through issue selection. *American journal of political science*, 59(1):71–90, 2015.
- R. J. Aumann, M. Maschler, and R. E. Stearns. *Repeated games with incomplete information*. MIT press, 1995.
- C. Ballester, A. Calvo-Armengol, and Y. Zenou. Who’s who in networks. Wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.
- A. Banerjee, A. Chandrasekhar, E. Dufflo, and M. Jackson. Diffusion of microfinance. *Science*, 341, DOI: 10.1126/science.1236498, 2013.
- F. Beard. Comparative advertising wars: An historical analysis of their causes and consequences. *Journal of Macromarketing*, 30(3):270–286, 2010.
- S. Bharathi, D. Kempe, and M. Salek. Competitive influence maximization in social networks. In *International Workshop on Web and Internet Economics*, pages 306–311. Springer, 2007.
- K. Bimpikis, A. Ozdaglar, and E. Yildiz. Competitive targeted advertising over networks. *Operations Research*, 64(3):705–720, 2016.
- P. Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120, 1972.
- J. M. Box-Steffensmeier and T.-M. Lin. A dynamic model of campaign spending in congressional elections. *Political Analysis*, pages 37–66, 1996.
- S. Chakrabarti and H. Haller. An analysis of advertising wars. *The Manchester School*, 79(1):100–124, 2011.
- D. L. Cogburn and F. K. Espinoza-Vasquez. From networked nominee to networked nation: Examining the impact of web 2.0 and social media on political participation and civic engagement in the 2008 obama campaign. *Journal of Political Marketing*, 10(1-2):189–213, 2011.
- B. De Meyer, P. K. Dubey, and R. Garg. Competing for customers in a social network. *Journal of Dynamics and Games*, 1(3):377–409, 2014.
- M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

- G. Demange and K. Van der Straeten. A communication game on electoral platforms. 2009.
- P. Domingos and M. Richardson. Mining the network value of customers. *Seventh International Conference on Knowledge Discovery and Data Mining*, pages 57–66, 2001.
- U. Doraszelski and S. Markovich. Advertising dynamics and competitive advantage. *The RAND Journal of Economics*, 38(3):557–592, 2007.
- L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.
- B. Golub and M. O. Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, 2010.
- S. Goyal, H. Heidari, and M. Kearns. Competitive contagion in networks. *Games and Economic Behavior*, 2014.
- M. Grabisch, A. Mandel, A. Rusinowska, and E. Tanimura. Strategic influence in social networks. *Mathematics of Operations Research*, forthcoming, 2017.
- J. Granato and M. S. Wong. Political campaign advertising dynamics. *Political Research Quarterly*, 57(3):349–361, 2004.
- J. Hartline, V. Mirrokni, and M. Sundararajan. Optimal marketing strategies over social networks. In *Proceedings of the 17th international conference on World Wide Web*, pages 189–198. ACM, 2008.
- D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. *Proceedings of 9th International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.
- D. Kempe, J. Kleinberg, and E. Tardos. Influential nodes in a diffusion model for social networks. *Proceedings of 32nd International Colloquium on Automata, Languages and Programming*, pages 1127–1138, 2005.
- C. R. Lever. *Strategic Competitions over Networks*. Ph.D. dissertation, Stanford University, 2010.
- J.-F. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, 1981.
- I. Pastine and T. Pastine. Consumption externalities, coordination, and advertising. *International Economic Review*, 43(3):919–943, 2002.
- J. Renault. The value of repeated games with an informed controller. *Mathematics of operations Research*, 37(1):154–179, 2012.
- D. Rosenberg, E. Solan, and N. Vieille. Blackwell optimality in markov decision processes with partial observation. *Annals of statistics*, pages 1178–1193, 2002.
- F. Thuijsman. Optimality and equilibria in stochastic games. *CWI Tracts*, 82:1–107, 1992.
- X. Venel. Commutative stochastic games. *Mathematics of Operations Research*, 40(2):403–428, 2014.
- A. Yadav, E. Kamar, B. Grosz, and M. Tambe. Healer: Pomdp planning for scheduling interventions among homeless youth. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 1504–1506. International Foundation for Autonomous Agents and Multiagent Systems, 2016.
- H. Zhuang, Y. Sun, J. Tang, J. Zhang, and X. Sun. Influence maximization in dynamic social networks. In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*, pages 1313–1318. IEEE, 2013.
- B. Ziliotto. Zero-sum repeated games: Counterexamples to the existence of the asymptotic value and the conjecture $\max \min = \lim v_n$. *The Annals of Probability*, 44(2):1107–1133, 2016.