Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

On a Centrality Maximization Game

Maria Castaldo * **Costanza Catalano** ** **Giacomo Como** **
Fabio Fagnani **

* *Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, GIPSA-lab, F-38000 Grenoble, France (e-mail: Maria.Castaldo@gipsa-lab.grenoble-inp.fr).*
** *Department of Mathematical Sciences "G.L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy (e-mail: {costanza.catalano,giacomo.como,fabio.fagnani}@polito.it).*

Abstract: The Bonacich centrality is a well-known measure of the relative importance of nodes in a network. This notion is, for example, at the core of Google's PageRank algorithm. In this paper we study a network formation game where each player corresponds to a node in the network to be formed. The action of a player consists in the assignment of m out-links and his utility is his own Bonacich centrality. We study the Nash equilibria (NE) and the best response dynamics of this game. In particular, we provide a complete classification of the set of NE when $m = 1$ and a fairly complete classification of the NE when $m = 2$. Our analysis shows that the centrality maximization performed by each node tends to create undirected and disconnected or loosely connected networks, namely 2-cliques for $m = 1$ and rings or a special "Butterfly"-shaped graph when $m = 2$. Our results build on locality property of the best response function in such game that we formalize and prove in the paper.

Keywords: Network centrality, network formation, Bonacich centrality, PageRank, game theory, social networks.

1. INTRODUCTION

Centrality is a key issue in network science. It aims at ranking the relevance of nodes in a network and its applications are ubiquitous. It permits to individuate the nodes in infrastructural networks where shocks can potentially trigger disruptive cascade effects (Ballester and Zenou (2006)) or rather the nodes in a socio-economic network that have more influence in the opinion formation and diffusion (Kempe et al. (2015)). Nodes with high centrality are the natural target of intervention strategies (Galeotti and Goyal (2009), Galeotti et al. (2017)) that aim to maximally enhance or depress the network performance.

In the literature, different definitions of centrality can be found, such as the *degree* centrality or the *eigenvalue* centrality (see for references Latora et al. (2017), Section 2.3). In this paper, we focus on the so-called *Bonacich* centrality measure, introduced in a seminal paper by the American sociologist Bonacich (1987). Formally, the Bonacich centrality π_i of a node i in a directed unweighted network is defined as

$$\pi_i = \beta \sum_{j \in N_i^-} \frac{\pi_j}{d_j} + (1-\beta)\eta_i \,, \qquad (1)$$

where N_i^- is the in-neighborhood of node i in the network, d_j is the out-degree of node j, η_i can be interpreted as the a-priori centrality of i (possibly the same for all nodes), and $\beta \in (0, 1)$ is some fixed parameter. Notice that by (1), the centrality of node i depends on the centrality of the nodes j linking at i (discounted by the number of their out-links) and on its intrinsic centrality. The centrality of a node is then somewhat inherited by the nodes connected to it: a node is important in the measure that important nodes have a link to it.

The Bonacich centrality have found wide applications in many contexts, as in social networks (e.g. representing citations among scientists), in describing Nash equilibria in networked quadratic games (Ballester and Zenou (2006)), in production networks among firms (Acemoglu et al. (2012)), and in opinion dynamics models as the Friedkin-Johnsen model (Friedkin and Johnsen (1990)). A famous instance of the Bonacich centrality is the so-called *PageRank centrality* for web pages, introduced by Brin and Page (1998), which is at the core of modern search engines like Google. Any search query on the web leads indeed to a set of possible related web pages that are sorted and presented by the engine according to their centrality ranking. Due to the relevance of the PageRank centrality for the external visibility of a web page, the problem of understanding how this measure can be efficiently computed and how it can be modified by perturbing the network has recently become very popular; see for example Ishii and Tempo (2014), Como and Fagnani (2015). The effect on the centrality caused by adding or deleting links in the network is not obvious from the recursive definition (1). It is not difficult to see that the addition of a link (i, j) always increases the centrality of the node j; less clear is how it affects the

* Giacomo Como is also with the Department of Automatic Control, Lund University, Sweden. This work was supported by MITI CNRS via 80 PRIME grant DOOM Disorders of Online Media, MIUR grant Dipartimenti di Eccellenza 2018–2022 [CUP: E11G18000350001], the Swedish Research Council, and by the Compagnia di San Paolo.

centrality of node i or, possibly, of all the other nodes in the network. In a context like that of web pages, where each node can decide only where to point its out-links and the aim is to gain visibility (that is, to increase its centrality in the network), the question of how such choice modifies its centrality and what is the rewiring that can possibly optimize it, turns out to be a natural relevant question. A first analysis in this sense can be found in Avrachenkov and Litvak (2006) and de Kerchove et al. (2008), while Csáji et al. (2010) explore computational time issues of these problems.

In this paper, we take this point of view by assuming that nodes are left free to choose their out-links and we cast the problem into a game-theoretic setting where rewards of nodes are exactly their centralities. We investigate the shapes that the network assumes when maximazing the centrality is the only driving force: we study the Nash equilibria of our game, i.e. configurations of the network in which every node is playing its optimal action, and the behavior of the best response dynamics, i.e. a discrete dynamics in which, at every time step, a random player plays an optimal action (see Section 3 for formal definitions). We can see our problem as an instance of a *network formation game*, where the actions of the players (the nodes of the network) are the ones defining the underlying network structure; we refer the reader to Jackson (2005) for a survey on network formation games and their applications in economy and sociology.

We study the problem under the assumption that all nodes are allowed to place the same number m of out-links. We obtain a complete classification of the Nash equilibria in the case $m = 1$, and a fairly complete classification of Nash equilibria in the case $m = 2$. Moreover, we carry on a detailed analysis on the Nash equilibria that are limit points of the best response dynamics. The main message that comes from this analysis is that the centrality maximization performed by each node tends to create undirected and disconnected or loosely connected networks: the components are 2-cliques for $m = 1$, rings and a special *Butterfly* graph for $m = 2$.

While completing this research, we discovered that a similar game-theoretic formulation was considered in Cominetti et al. (2018), Section 7, where authors prove the existence of Nash equilibria for a generalized version of our game. However, the proof in Cominetti et al. (2018) is non-constructive and our classification of Nash equilibria, we believe, cannot be derived from their results.

The paper is structured as follows. In Section 2 we present the game theoretical setting; Section 3 recalls classical results and definitions of game theory, while Section 4 describes the main results of the paper. All intermediate technical results are in Section 5; due to length restrictions all the proofs have been omitted and can be found in Castaldo et al. (2019). Section 6 concludes with a summary and some open problems.

2. THE MODEL

In this section, we formally define the centrality maximization game and we state the problems we want to address. Consider a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, \ldots, n\}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of (directed) edges. We denote by $(i, j) \in \mathcal{E}$ a directed edge from node i to node j. We assume throughout the paper that $\mathcal{G}$ does not contain self-loops. In- and out- neighborhoods of a node i are indicated, respectively, by N_i^- and N_i. Their cardinalities $d_i^- = |N_i^-|$ and $d_i = |N_i|$ are, respectively, the in- and the out-degree of node i. Under the assumption that $d_i > 0$ for every $i \in \mathcal{V}$, we equip $\mathcal{G}$ with the normalized weight matrix R whose entries R_{ij} are defined as

$$R_{ij} = \frac{1}{d_i} \mathbb{1}_{\{(i,j)\in\mathcal{E}\}},$$

where $\mathbb{1}$ is the characteristic function. The entry R_{ij} represents the weight attributed to the link (i, j). The Bonacich centrality $\pi = (\pi_1, \ldots, \pi_n)$ of $\mathcal{G}$ in Eq. (1) can be more compactly written as

$$\pi = (1 - \beta)(I - \beta R^\top)^{-1}\eta \tag{2}$$

where I is the identity matrix, $\beta \in (0, 1)$, $\eta \in \mathbb{R}^n$ is a fixed probability vector [1] and $R^\top$ denotes the transpose of the matrix R. A direct check shows that π is a probability vector. Expanding (2) in a power series, we can write the Bonacich centrality of node i as

$$\pi_i = (1-\beta)\left[\eta_i + \beta \sum_j \eta_j R_{ji} + \beta^2 \sum_{j,l} \eta_j R_{jl} R_{li} + \cdots\right]. \tag{3}$$

Interpreting η as a vector assigning an a-priori centrality (not depending on the graph) to each node (possibly the uniform one $\eta_i = n^{-1}$ for all i), formula (3) says that the Bonacich centrality of a node in the graph $\mathcal{G}$ is the discounted sum of its own centrality η_i and of the centrality of the other nodes discounted by the weight of the paths connecting to i through the constant β. Notice that the constant $(1 - \beta)$ appears just to normalize π to a probability vector.

In our setting, we start with the set of nodes $\mathcal{V} = \{1, \ldots, n\}$ and we suppose that each node i is a player that assigns m directed edges from i to m other distinct elements in $\mathcal{V}$. This construction results in a graph $\mathcal{G}$ and the Bonacich centrality of node i in $\mathcal{G}$ represents its utility. This can be thought as a classical game where

- $\mathcal{V}$ is the set of players;
- given $i \in \mathcal{V}$, the corresponding set of actions $\mathcal{A}_i$ is the family of all subsets of $\mathcal{V}\backslash\{i\}$ of cardinality m;
- let $\mathcal{A} = \prod_i \mathcal{A}_i$ and $x = (x_1, \ldots, x_n) \in \mathcal{A}$ a strategy profile (or *configuration*). We define the graph $\mathcal{G}(x) = (\mathcal{V}, \mathcal{E}(x))$ where $\mathcal{E}(x) = \{(i, j) \mid i \in \mathcal{V},\ j \in x_i\}$. Notice that by construction $\mathcal{G}(x)$ has constant out-degree equal to m. We denote by $R(x)$ the normalized weight matrix of $\mathcal{G}(x)$ [2]. Given $\beta \in (0, 1)$ and $\eta \in \mathbb{R}^n$ a probability vector such that $\eta_i > 0$ for all i, we define the utility vector $u(x) = (u_1(x), \ldots, u_n(x))$ as the Bonacich centrality of $\mathcal{G}(x)$:
$$u(x) = (1 - \beta)(I - \beta R(x)^\top)^{-1}\eta.$$

The game we have introduced is denoted by $\Gamma(\mathcal{V}, \beta, \eta, m)$ to recall all the parameters entering in the construction.

The main goal of this paper is to analyze the structure of Nash equilibria for the game $\Gamma(\mathcal{V}, \beta, \eta, m)$ and to in-

[1] v is a *probability* vector if $\sum_i v_i = 1$ and $v_i \geqslant 0$ for all i.
[2] That is, $R_{ij}(x) = m^{-1}$ if $(i, j) \in \mathcal{E}(x)$, $R_{ij}(x) = 0$ otherwise.

vestigate the asymptotic behavior of its best response dynamics, which is defined in the next section. The game is homogeneous in the sense that we give every node the chance to place the same number m of out-links in the network. A natural generalization of this problem would be to consider a different number m_i of out-links for each node; we leave this to future work.

3. PRELIMINARIES

In this section we recall some fundamental definitions and classical results in game theory that will be used in the next sections.

Given $x \in \mathcal{A}$ and $i \in \mathcal{V}$, we adopt the usual convention to indicate with $x_{-i} \in \mathcal{A}_{-i} = \prod_{k \neq i} \mathcal{A}_k$ the vector x restricted to the components in $\mathcal{V}\backslash\{i\}$ and to use the notation $x = (x_i, x_{-i})$.

Definition 1. Let $i \in \mathcal{V}$ and $x_{-i} \in \mathcal{A}_{-i}$. We define the *best response set* $\mathcal{B}_i(x_{-i})$ of node i given the strategy x_{-i} as

$$\mathcal{B}_i(x_{-i}) = \text{argmax}_{x_i \in \mathcal{A}_i} u_i(x_i, x_{-i}).$$

The best response set represents the set of actions of player i that maximize his utility u_i, given the strategy x_{-i} played by all the other players. We now recall the definition of (strict) Nash Equilibria and best response dynamics.

Definition 2. Let $x \in \mathcal{A}$ be a strategy profile. If for all $i \in \mathcal{V}$, $x_i \in \mathcal{B}_i(x_{-i})$, then x a *Nash equilibrium.* If for all $i \in \mathcal{V}$, $\mathcal{B}_i(x_{-i}) = \{x_i\}$, then x a *strict Nash equilibrium.* We denote by $\mathcal{N}$ and $\mathcal{N}^{\text{st}}$ the set of, respectively, Nash equilibria and strict Nash equilibria.

Definition 3. The *(asynchronous) best response dynamics* is a discrete time dynamics Y_t on the state space $\mathcal{A}$ in which at every time $t \in \mathbb{N}$, a player i is chosen uniformly at random and he revises his action by picking an element y in $\mathcal{B}_i\big((Y_{t-1})_{-i}\big)$ uniformly at random.

A classical result of Monderer and Shapley (1996) states that if a game is ordinal potential [3], then its best response dynamics converges in finite time with probability one to (a subset of) Nash equilibria, independently on the initial condition. Cominetti et al. (2018) (Proposition 7.5 and Section 7.2) proved that our game is ordinal potential, which let us formulate the following result:

Proposition 4. The best response dynamics on the game $\Gamma(\mathcal{V}, \beta, \eta, m)$ always converges in finite time with probability one to a set $\mathcal{N}^* \subseteq \mathcal{N}$ of Nash equilibria.

Typically $\mathcal{N}^*$ is a proper subset of $\mathcal{N}$. Moreover, as strict Nash equilibria are absorbing points of the best response dynamics, it holds that $\mathcal{N}^{\text{st}} \subseteq \mathcal{N}^*$; however, in general they are not equal. If we consider the transition graph on the configuration set $\mathcal{A}$ induced by the best response dynamics Y_t, the set $\mathcal{N}^*$ can be described as its smallest trapping set (no edge leading out of $\mathcal{N}^*$) that is globally reachable (from every configuration in $\mathcal{A}$ there is a path leading inside $\mathcal{N}^*$). Nash equilibria in $\mathcal{N}^*$ play a crucial role in games as they are those the best response dynamics will eventually converge to, while Nash equilibria in $\mathcal{N}\backslash\mathcal{N}^*$ will only show up in the transient behavior.

[3] A game is ordinal potential if there exists a function $\Psi : \mathcal{A} \to \mathbb{R}$ s.t. $u_i(x_i, x_{-i}) < u_i(x_i', x_{-i}) \Leftrightarrow \Psi(x_i, x_{-i}) < \Psi(x_i', x_{-i})$.

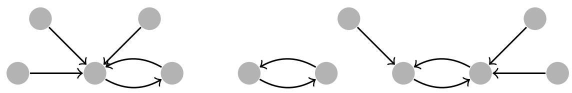

Fig. 1. An example of a graph of type $C_2^{3,6}$.

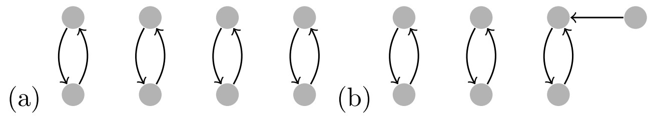

Fig. 2. (a) A graph of type $C_2^{n/2,0}$ with $n = 8$; (b) A graph of type $C_2^{(n-1)/2,1}$ with $n = 7$.

Our aim is to investigate the structure of these three sets $\mathcal{N}^{\text{st}} \subseteq \mathcal{N}^* \subseteq \mathcal{N}$ for the game $\Gamma(\mathcal{V}, \beta, \eta, m)$ that we have introduced in the previous section.

4. MAIN RESULTS

In this paper we focus on the case when $m = 1$ and $m = 2$, namely when nodes are allowed to set, respectively, one or two out-links towards other nodes. Through a characterization of the best response set $\mathcal{B}_i(x_{-i})$, we are capable of giving a full description of the three sets $\mathcal{N}^{\text{st}}$, $\mathcal{N}^*$ and $\mathcal{N}$ of Nash equilibria for $m = 1$, and a full description of $\mathcal{N}^{\text{st}}$ and $\mathcal{N}^*$ for $m = 2$, together with a necessary condition for $\mathcal{N}$. The case $m = 2$ presents a much more complex behavior and, for certain aspects, as complex as the general case.

4.1 The case of out-degree $m = 1$

In order to describe our results, it is convenient to introduce a particular family of graphs.

Definition 5. We call a 2-*clique* the complete directed graph (without self-loops) with two nodes and we indicate it by C_2; we call a *singleton* a node with zero in-degree. Given $l, r \in \mathbb{N}$, we define $C_2^{l,r}$ as the directed graph obtained by taking the disjoint union of l copies of C_2 plus r extra singletons, each of them having exactly one out-link towards a node in any of the 2-cliques.

Notice that $C_2^{l,r}$ has exactly $n = 2l + r$ nodes and all nodes have out-degree equal to one. Figure 1 is an example of graph of type $C_2^{l,r}$ for $l = 3$ and $r = 6$. The following theorem is our first main result for the case $m = 1$.

Theorem 6. For any choice of the parameters β and η, the game $\Gamma(\mathcal{V}, \beta, \eta, 1)$ has the following properties:

(1) the set of Nash equilibria $\mathcal{N}$ coincides with all the configurations $x \in \mathcal{A}$ for which $\mathcal{G}(x)$ is of type $C_2^{l,r}$ with $2l + r = n$;
(2) the set of strict Nash equilibria $\mathcal{N}^{st}$ is empty when n is odd and it coincides with all the configurations $x \in \mathcal{A}$ for which $\mathcal{G}(x)$ is of type $C_2^{n/2,0}$ when n is even.

Figure 2(a) represents a strict Nash equilibrium for $\Gamma(\mathcal{V}, \beta, \eta, 1)$ with $n = 8$, while Fig. 2(b) shows a non-strict Nash equilibrium for $n = 7$. The following corollary completely captures the asymptotic behavior of the best response dynamics of $\Gamma(\mathcal{V}, \beta, \eta, 1)$; in particular it shows that the Nash equilibrium of Fig. 2(b) belongs to $\mathcal{N}^*$.

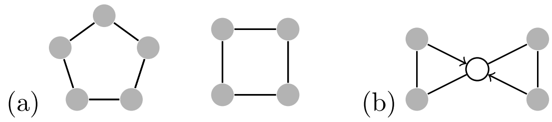

Fig. 3. (a) Example of strict Nash equilibrium for the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$ with $n = 9$. (b) The Butterfly graph. White nodes do not have unique best response.

Corollary 7. Consider the best response dynamics for the game $\Gamma(\mathcal{V}, \beta, \eta, 1)$. For any choice of β and η, it holds that:

- if n is even, the limit set $\mathcal{N}^*$ coincides with $\mathcal{N}^{\mathrm{st}}$;
- if n is odd, the limit set $\mathcal{N}^*$ coincides with those $x \in \mathcal{A}$ for which $\mathcal{G}(x)$ is of type $C_2^{(n-1)/2,1}$.

Notice that when $n = 2k$, the best response dynamics will eventually be absorbed in any of the $|\mathcal{N}^*| = n!2^{-k}(k!)^{-1}$ strict Nash equilibria with probability one. On the other hand, when $n = 2k + 1$ the best response dynamics will eventually reach the (unique) trapping set consisting of $|\mathcal{N}^*| = (n-1)n!2^{-k}(k!)^{-1}$ configurations of type $C_2^{(n-1)/2,1}$. In this case, it can be shown that the best response dynamics will keep fluctuating ergodically in the set $\mathcal{N}^*$ with uniform equilibrium probability.

4.2 The case of out-degree $m = 2$

We call *ring* graph an undirected graph whose vertices are arranged in a ring so that each vertex has exactly two neighbors (see for example Fig. 3(a), where each connected component is a ring graph). The *length* of a ring graph is the number of its vertices. From now on we say that an edge (i, j) in $\mathcal{G}$ is *undirected* if also (j, i) is an edge of $\mathcal{G}$, otherwise we call it *directed*. We say that a graph is *undirected* if all its edges are undirected. In figures, we represent directed edges with arrows and undirected edges with simple lines.

The first main result of this section is a complete characterization of the set of strict Nash equilibria.

Theorem 8. For any choice of β and η, the set of strict Nash equilibria $\mathcal{N}^{st}$ of the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$ consists of all the configurations $x \in \mathcal{A}$ for which $\mathcal{G}(x)$ is the union of ring graphs.

A consequence of this fact is that for any $n \geqslant 3$ there always exists a strict Nash equilibrium, as the ring graph of length n is always one of these. Figure 3(a) provides an example of strict Nash equilibrium with $n = 9$.

We now investigate the structure of all Nash equilibria. Given a Nash equilibrium $x \in \mathcal{A}$, let $\{\mathcal{G}_\lambda(x)\}_{\lambda=1,\dots,\Lambda}$ be the decomposition of $\mathcal{G}(x)$ in terms of its strongly connected components. The *condensation graph* of $\mathcal{G}(x)$ is defined as the graph $\mathcal{H}(x)$ whose nodes are the components $\{\mathcal{G}_\lambda(x)\}_\lambda$ and where there is an edge from $\mathcal{G}_{\lambda_1}(x)$ to $\mathcal{G}_{\lambda_2}(x)$ if there exists an edge in $\mathcal{G}(x)$ from a node in $\mathcal{G}_{\lambda_1}(x)$ to a node in $\mathcal{G}_{\lambda_2}(x)$. The condensation graph $\mathcal{H}(x)$ is directed and acyclic. The following theorem describes the topology of $\mathcal{H}(x)$ when $x \in \mathcal{N}$, thus characterizing the structure of the Nash equilibria of the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$. We remind that a vertex is called a *sink* if it has zero out-degree and it is called a *source* if it has zero in-degree.

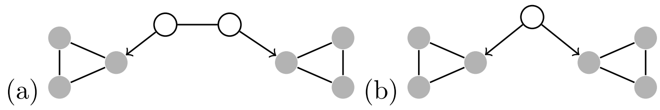

Fig. 4. Ex. of nonstrict Nash equilibria for $\Gamma(\mathcal{V}, \beta, \eta, 2)$. White nodes do not have unique best response.

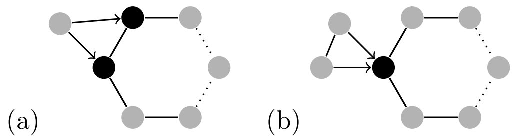

Fig. 5. (a) Singleton linking to two adjacent nodes in a ring. (b) 2-clique linking to a single node in a ring. Black nodes are not in best response.

Theorem 9. Let $x \in \mathcal{A}$ be a Nash equilibrium for the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$ and $\mathcal{H}(x)$ be its condensation graph on the components $\{\mathcal{G}_\lambda(x)\}_\lambda$. For any choice of β and η, the following facts hold:

(1) every component $\mathcal{G}_\lambda(x)$ is either a sink or a source in $\mathcal{H}(x)$ (or both if isolated);
(2) every source component is either a single vertex (singleton) or a 2-clique;
(3) every sink component is either a ring or the Butterfly graph in Fig. 3(b).

Notice that the Butterfly graph is a nonstrict Nash equilibrium as the best response of the node in the center is not unique, i.e. it can change action while maintaining the same utility. Figure 4 provides other two examples of nonstrict Nash equilibria: in both structures we can identify either a singleton or a 2-clique linking to rings; the nodes in white have not unique best response.

Remark 10. Not all the configurations $x \in \mathcal{A}$ that satisfy conditions (1), (2) and (3) of Theorem 9 are Nash equilibria. Indeed, by direct computation it is easy to see that the following examples are not Nash equilibria:

(1) a singleton linking to two adjacent nodes in a ring of length greater or equal than four (see Fig. 5(a));
(2) a 2-clique linking to a single node in a ring of length greater or equal than four (see Fig. 5(b)).

We are now ready to characterize the limit set $\mathcal{N}^* \subseteq \mathcal{N}$ for the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$, i.e. the absorbing points of its best response dynamics.

Corollary 11. Consider the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$ and let i s.t. $i = n \mod 3$. Then for any choice of β and η, it holds that:

- if $i = 0, 1$, the limit set $\mathcal{N}^*$ coincides with $\mathcal{N}^{\mathrm{st}}$;
- if $i = 2$, the limit set $\mathcal{N}^*$ coincides with $\mathcal{N}^{\mathrm{st}} \cup \mathcal{G}_b^3$, where $\mathcal{G}_b^3$ is the set of all graphs that are unions of rings of length three and a Butterfly graph or unions of rings of length three and a 2-clique linking to any nodes in the rings (see for example Fig. 4(a)).

5. PROOFS OF THE RESULTS

The proofs of our results are based on a probabilistic interpretation of the game in terms of Markov chains. We first recall some preliminary notions and then in Subsections 5.1 and 5.2 we present all the technical results that let us prove the main ones, respectively for the case $m = 1$ and $m = 2$. Due to length restrictions, all the proofs are omitted and can be found in Castaldo et al. (2019).

A (discrete-time) Markov chain X_t on a finite state space $\mathcal{V} = \{1, \dots, n\}$ and with transition matrix $P \in \mathbb{R}^{n\times n}$, P stochastic [4], is a sequence of random variables $X_1, X_2, \dots$ with values in $\mathcal{V}$ such that $\mathbb{P}(X_{t+1} = i|X_1 = j_1, \dots, X_t = j_t) = \mathbb{P}(X_{t+1} = i|X_t = j_t) = P_{j_t i}$. Given $s \in \mathcal{V}$, we define $T_s := \inf\{t \geqslant 0 : X_t = s\}$ the *hitting* time on s and $T_s^+ := \inf\{t \geqslant 1 : X_t = s\}$ the *return* time to s. Given $i, s \in \mathcal{V}$, we define $\tau_i^s := \mathbb{E}_i[T_s]$ the *expected* hitting time on s of the Markov chain X_t with initial state i. It is known that if P is an irreducible matrix, then the Markov chain admits a unique invariant distribution, that is a probability vector π s.t. $\pi = P^\top \pi$. The invariant distribution π can be written in terms of hitting times:

Proposition 12. (Norris (1997)). Let X_t be a Markov chain on the finite state space $\mathcal{V}$ and with irreducible transition matrix P, and let π be its (unique) invariant distribution. Then it holds that

$$\pi_s = \left(1 + \sum_{i\in\mathcal{V}} P_{si}\tau_i^s\right)^{-1}, \tag{4}$$

where the expected hitting times τ_i^s, $i \in \mathcal{V}$, are the only family of values satisfying the following system:

$$\begin{cases} \tau_i^s = 0 & \text{if } i = s, \\ \tau_i^s = 1 + \sum_{j\in\mathcal{V}} P_{ij}\tau_j^s & \text{if } i \neq s. \end{cases} \tag{5}$$

Manipulating (2) and using the fact that $\mathbf{1}^\top\pi = 1$ with $\mathbf{1}$ the all-ones vector, we can see that the Bonacich centrality π satisfies the relation $\pi = (\beta R^\top + (1-\beta)\eta\mathbf{1}^\top)\pi$. Since $P = \beta R + (1-\beta)\mathbf{1}\eta^\top$ is an irreducible stochastic matrix, it means that π is the (unique) invariant distribution of the Markov chain having P as transition matrix. We now use this characterization in the context of our game. Given a configuration $x \in \mathcal{A}$, we write

$$P(x) = \beta R(x) + (1-\beta)\mathbf{1}\eta^\top \tag{6}$$

and we denote by $\tau_i^s(x)$ the hitting time on s of the Markov chain having $P(x)$ as transition matrix and starting from i. When the configuration x is clear from the context, sometimes we write τ_i^s instead of $\tau_i^s(x)$ to ease the notation. The utility vector $u(x)$ can be written in terms of the formula (4) as $u_s(x_s, x_{-s}) = (1 + \sum_{i\in\mathcal{V}} P_{si}(x)\tau_i^s(x))^{-1}$. Since the terms $P_{si}(x)$ only depend on x_s (the out-links from s), while the hitting times $\tau_i^s(x)$ only depend on x_{-s}, with slight abuse of notation we rewrite the utility function as

$$u_s(x_s, x_{-s}) = \left(1 + \sum_{i\in\mathcal{V}} P_{si}(x_s)\tau_i^s(x_{-s})\right)^{-1}. \tag{7}$$

A consequence of (7) is an explicit formula describing the best response set, as shown in the following remark.

Remark 13. Consider the game $\Gamma(\mathcal{V}, \beta, \eta, m)$, a node $s \in \mathcal{V}$ and $x_{-s} \in \mathcal{A}_{-s}$. Then the set $\mathcal{B}_s(x_{-s})$ can be written as:

$$\mathcal{B}_s(x_{-s}) = \operatorname*{argmin}_{x_s\in\mathcal{A}_s} \sum_{i\in\mathcal{V}} R_{si}(x_s)\tau_i^s(x_{-s}). \tag{8}$$

In the following, given $x \in \mathcal{A}$ we denote by $N_s^-(x)$ the in-neighborhood of the vertex s in the graph $\mathcal{G}(x)$, i.e. $i \in N_s^-(x)$ if and only if $s \in x_i$ (or equivalently, if and only if $R_{is}(x) > 0$). Notice that $N_s^-(x)$ depends just on x_{-s} so with a slight abuse of notation we can write $N_s^-(x_{-s})$.

[4] A matrix P is stochastic if each row is a probability vector.

5.1 The case of out-degree $m = 1$

We first characterize the best response set of a player. An important observation is the following:

Remark 14. If $m = 1$, then for any $s \in \mathcal{V}$ and $x_s \in \mathcal{A}_s$ it holds that $R_{sx_s}(x_s) = 1$, and for all $i \neq x_s$, $R_{si}(x_s) = 0$. So (8) takes the form $\mathcal{B}_s(x_{-s}) = \operatorname{argmin}_{i\in\mathcal{V}\setminus\{s\}} \tau_i^s(x_{-s})$.

The following proposition shows that the best response action of a player in the game $\Gamma(\mathcal{V}, \beta, \eta, 1)$ takes always place in his in-neighborhood, as long as it is nonempty.

Proposition 15. Consider the game $\Gamma(\mathcal{V}, \beta, \eta, 1)$ and let $s \in \mathcal{V}$ and $x_{-s} \in \mathcal{A}_{-s}$. It holds that:

(1) If $N_s^-(x_{-s}) \neq \varnothing$, then $\mathcal{B}_s(x_{-s}) = N_s^-(x_{-s})$;
(2) If $N_s^-(x_{-s}) = \varnothing$, then $\mathcal{B}_s(x_{-s}) = \mathcal{V}\backslash\{s\}$.

The proof makes use of Proposition 12 to show that two nodes in $N_s^-(x_{-s})$ have the same hitting times on s, and that this time is strictly smaller than the one from any node not in $N_s^-(x_{-s})$. In addition it can be proved that τ_j^s does not depend on j when $|N_s^-(x_{-s})| = \varnothing$.

When studying Nash equilibria, since every node with a nonempty in-neigbourhood has to link at one of his in-neighbors (Proposition 15, item (2)), the only possible configurations for the network are those described by Theorem 6, item (1). Moreover, since the best response of a node is unique only when it has exactly one in-neighbor, strict Nash equilibria assume the configurations presented in Theorem 6, item (2). See Castaldo et al. (2019) for a detailed proof of Theorem 6 and Corollary 7.

5.2 The case of out-degree $m = 2$

As for $m = 1$, we want to better characterize the best response set of a player. The following two lemmas are needed to prove the subsequent Proposition 18, where we show that the best response actions of a node s are always towards nodes that are at most at in-distance two from it.

Lemma 16. Consider the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$, and let $x \in \mathcal{A}$ and $s \in \mathcal{V}$. It holds that:

(1) for every $i \neq s$, $\tau_i^s(x) \leqslant \eta_s^{-1}(1-\beta)^{-1}$;
(2) if there exists $i \neq s$ such that $\tau_i^s(x) = \eta_s^{-1}(1-\beta)^{-1}$, then $N_s^-(x) = \varnothing$.

The next lemma provides a different upper bound on the return times $\tau_i^s(x)$ when $|N_s^-(x)| \geqslant 1$. We denote by $N_s^{-2}(x)$ the set $N_s^-(x) \cup \{N_t^-(x) : t \in N_s^-(x)\}$, that is the in-neighborhood of s in $\mathcal{G}(x)$ at distance at most two.

Lemma 17. Consider the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$, and let $x \in \mathcal{A}$ and $s \in \mathcal{V}$ such that $|N_s^-(x)| \geqslant 1$. Let $k \in N_s^-(x)$ and set $T_1 = (1-\frac{\beta}{2})(1-\beta)^{-1}(\eta_s + \frac{\beta}{2}\eta_k)^{-1}$ and $T_2 = (1-\beta)^{-1}(\eta_s + \frac{\beta}{2}\eta_k)^{-1}$. Then it holds that:

(1) $\tau_k^s(x) \leqslant T_1$ and for all $i \neq k$, $\tau_i^s(x) \leqslant T_2$;
(2) if $\tau_k^s(x) = T_1$ and for all $i \neq k, s$, $\tau_i^s(x) = T_2$, then $|N_s^{-2}(x)| = 1$.

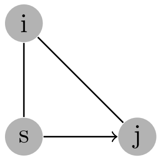

Fig. 6. The directed graph $T_{(s,j),i}$.

The following proposition characterizes the best response set of a player in the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$.

Proposition 18. Consider the game $\Gamma(\mathcal{V}, \beta, \eta, 2)$, and let $x \in \mathcal{A}$ and $s \in \mathcal{V}$. It holds that:

(1) if $N_s^{-2}(x) = \varnothing$, then $\mathcal{B}_s(x_{-s}) = \{\{v, w\} : v, w \in \mathcal{V} \backslash \{s\}, v \neq w\}$;
(2) if $|N_s^{-2}(x)| = 1$, then $\mathcal{B}_s(x_{-s}) = \{\{r, v\} : v \in \mathcal{V} \backslash \{s, r\}\}$, where $\{r\} = N_s^{-2}(x) = N_s^{-}(x)$;
(3) if $|N_s^{-2}(x)| \geqslant 2$, then $\mathcal{B}_s(x_{-s}) \subseteq \{\{v, w\} : v, w \in N_s^{-}(x), v \neq w\} \cup \{\{v, w\} : v \in N_s^{-}(x) \text{ and } w \in N_v^{-}(x)\}$.

A sketch of the proof can be summarized as follows. We prove item (2) and (3) of Proposition 18 by contradiction. Infact, by assuming that items (2) and (3) do not hold true, we end up with some nodes having hitting times equal to the upper bound in item (1) of Lemma 16 in the first case, or with some nodes having hitting times equal to T_1 and T_2 in item (1) of Lemma 17 in the second case. Then items (2) of Lemmas 16 and 17 lead, respectively, to a contradiction of the ipothesis of items (2) and (3) of Proposition 18.

Remark 19. Suppose that $|N_s^{-2}(x)| \geqslant 2$ for $x \in \mathcal{A}$ and $s \in \mathcal{V}$ and let $x_s = \{i, j\} \in \mathcal{B}_s(x_{-s})$. Item (3) of Proposition 18 implies that, if $j \notin N_s^{-}(x)$, then $i \in N_s^{-}(x)$ and $j \in N_i^{-}(x)$. In other words, (j, i) and (i, s) must be edges of $\mathcal{G}(x)$, as well as (s, i) and (s, j) since s is playing $\{i, j\}$.

Theorems 8 mainly follows from Proposition 18 and the observation that the best response of a node s can be unique only in the case $|N_s^{-2}(x)| \geqslant 2$.

Definition 20. We denote by $T_{(s,j),i}$ the directed graph on the vertices $\{i, j, s\}$ having one directed edge (s, j) and all the other edges undirected (see Fig. 6).

Lemma 21. Let $x \in \mathcal{A}$ be a Nash eq. of $\Gamma(\mathcal{V}, \beta, \eta, 2)$, $\mathcal{H}(x)$ be the condensation graph of $\mathcal{G}(x)$ and $\mathcal{G}_\lambda(x) = (\mathcal{V}_\lambda, \mathcal{E}_\lambda)$ be a sink in $\mathcal{H}(x)$. If there exists a directed edge $(s, j) \in \mathcal{E}_\lambda$, then $\mathcal{G}_\lambda(x)$ contains a structure of type $T_{(s,j),i}$.

In view of Remark 19, to prove the above Lemma it suffices to show that $(i, j) \in \mathcal{E}(x)$; this (nontrivially) follows from item (3) of Proposition 18, see Castaldo et al. (2019). Proposition 18 and Lemma 21 are the main building blocks for the proofs of Theorem 9 and Corollary 11.

6. CONCLUSIONS

In this paper we proposed a game in which every node of a network aims at maximizing its Bonacich centrality by choosing where to direct its out-links, whose number is fixed to be equal to m. We have completely characterized the sets $\mathcal{N}^{\text{st}}$, $\mathcal{N}^*$ and $\mathcal{N}$ of Nash equilibria when $m = 1$ and the sets $\mathcal{N}^{\text{st}}$ and $\mathcal{N}^*$ when $m = 2$, providing also necessary conditions for $\mathcal{N}$. Our results show that the centrality maximization performed by each node tends to create disconnected and undirected networks, partially due to the locality property of the best response actions. In particular, both for $m = 1, 2$ all the m-regular undirected networks result to be (strict) Nash equilibria. A natural follow-up of our work would be the analysis of Nash equilibria of the game for $m \geqslant 3$, possibly in an heterogeneous setting where m is different for each node. Preliminary numerical experiments show that this tendency to create disconnected networks show up also for bigger m, and that not all m-regular undirected networks are Nash equilibria.

REFERENCES

Acemoglu, D., Carvalho, V.M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggergate fluctuations. *Econometrica*, 80(5), 1977–2016.

Avrachenkov, K. and Litvak, N. (2006). The effect of new links on google pagerank. *Stoch. Models*, 22, 2006.

Ballester, C. and Zenou, Y. (2006). Who's who in networks. *Econometrica*, 74, 1403–1417.

Bonacich, P. (1987). Power and centrality: A family of measures. *American J. of Sociology*, 92(5), 1170–1182.

Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks*, 30, 107–117.

Castaldo, M., Catalano, C., Como, G., and Fagnani, F. (2019). On a centrality maximization game. URL `https://arxiv.org/abs/1911.06737v2`. Unpublished.

Cominetti, R., Quattropani, M., and Scarsini, M. (2018). The buck-passing game. URL `https://arxiv.org/abs/1808.03206`. Unpublished.

Como, G. and Fagnani, F. (2015). Robustness of large-scale stochastic matrices to localized perturbations. *IEEE Trans. on Network Sci. and Eng.*, 2(2), 53–64.

Csáji, B.C., Jungers, R.M., and Blondel, V.D. (2010). Pagerank optimization in polynomial time by stochastic shortest path reformulation. In *Algorithmic Learning Theory*, 89–103.

de Kerchove, C., Ninove, N., and van Dooren, P. (2008). Maximizing pagerank via outlinks. *Linear Algebra and its Applications*, 429(5), 1254 – 1276.

Friedkin, N.E. and Johnsen, E.C. (1990). Social influence and opinions. *J. of Math. Sociology*, 15(3-4), 193–206.

Galeotti, A., Golub, B., and Goyal, S. (2017). Targeting interventions in networks. URL `https://arxiv.org/abs/1710.06026`. Cambridge Working Papers in Economics 1744.

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

Ishii, H. and Tempo, R. (2014). The pagerank problem, multiagent consensus, and web aggregation: A systems and control viewpoint. *IEEE Control Systems Magazine*, 34(3), 34–53.

Jackson, M.O. (2005). *A Survey of Models of Network Formation: Stability and Efficiency*, chapter 1, 11–57. Cambridge University Press.

Kempe, D., Kleinberg, J., and Tardos, E. (2015). Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4), 105–147.

Latora, V., Nicosia, V., and Russo, G. (2017). *Complex Networks: Principles, Methods and Applications*. Cambridge University Press.

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

Norris, J.R. (1997). *Markov Chains*. Cambridge University Press.