On a Network Centrality Maximization Game

informs
https://pubsonline.informs.org/journal/moor

MATHEMATICS OF OPERATIONS RESEARCH

Vol. 00, No. 0, Xxxxx 0000, pp. 000–000

Submitted to *Mathematics of Operations Research*

ISSN 0364-765X, EISSN 1526-5471

On a Network Centrality Maximization Game

Costanza Catalano*
Department of Economics, Statistics and Research, Banca d'Italia, Largo Guido Carli 1, 00044 Frascati (Roma), Italy.
costanzacatalano@gmail.com https://sites.google.com/view/costanza-catalano/home

Maria Castaldo
GIPSA-lab, Université de Grenoble Alpes, CNRS, INRIA, Grenoble INP, 11 rue des Mathématiques, F-38000, Grenoble, France.
maria.castaldo@grenoble-inp.fr https://www.mariacastaldo.com

Giacomo Como† and Fabio Fagnani
Department of Mathematical Sciences "G.L. Lagrange," Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy.
giacomo.como@polito.it, fabio.fagnani@polito.it
https://staff.polito.it/giacomo.como/, http://calvino.polito.it/ fagnani/

Abstract. We study a network formation game where n players, identified with the nodes of a directed graph to be formed, choose where to wire their outgoing links in order to maximize their PageRank centrality. Specifically, the action of every player i consists in the wiring of a predetermined number d_i of directed out-links, and her utility is her own PageRank centrality in the network resulting from the actions of all players. We show that this is a potential game and that the best response correspondence always exhibits a local structure in that it is never convenient for a node i to link to other nodes that are at incoming distance more than d_i from her. We then study the equilibria of this game determining necessary conditions for a graph to be a (strict, recurrent) Nash equilibrium. Moreover, in the homogeneous case, where players all have the same number d of out-links, we characterize the structure of the potential maximizing equilibria and, in the special cases $d = 1$ and $d = 2$, we provide a complete classification of the set of (strict, recurrent) Nash equilibria. Our analysis shows in particular that the considered formation mechanism leads to the emergence of undirected and disconnected or loosely connected networks.

Funding: This research was carried out within the framework of the MIUR-funded *Progetto di Eccellenza* of the *Dipartimento di Scienze Matematiche G.L. Lagrange*, Politecnico di Torino, CUP: E11G18000350001. It received partial support from the MIUR-funded project PRIN 2017 "Advanced Network Control of Future Smart Grids" and by the *Compagnia di San Paolo*.

Key words: Network formation games, network centrality, ordinal potential games.

OR/MS subject classification: Primary: 91A43; secondary: 91D30

* The views expressed in this paper are those of the authors and do not involve the responsability of the Bank of Italy.

† Giacomo Como is also with the Department of Automatic Control, Lund University, Sweden.

1. Introduction

The notion of *centrality* is ubiquitous in network science and engineering: it provides a measure of the relative relevance of the nodes in a network. The range of applications is wide [24, 32, 40]. Various definitions of network centrality have appeared in the literature [10], many of which tailored to specific applications. In this paper, we focus on the so-called *PageRank centrality* [12], which is closely related to the notions of centrality introduced by Katz and Bonacich [11, 28, 34].

Understanding how the PageRank centrality measure can be efficiently computed and how it can be modified by perturbing the network is a key problem that has received significant attention in the recent literature [17, 31]. The effect on the PageRank centrality of the nodes caused by the addition or deletion of links in the network is not obvious. While it can be shown [15] that the addition of a directed link always increases the PageRank centrality of its head node, it is less clear how it affects the network centrality of the other nodes. In applications such as citation networks, the World Wide Web, or (on-line) social networks, each node can typically decide where to direct its out-links and often has an interest in gaining visibility, i.e., to increase its own network centrality. A natural question is how such choices modify the PageRank centrality of a node and what is the rewiring that can possibly optimize it. A first analysis in this sense can be found in the works [4, 22], while the article [21] explores computational time issues in these problems. In the works [20] and [6], optimal network intervention problems are studied where an external planner seeks to maximize the sum of Bonacich centralities with either a budget on the links to be created or a cost term in its objective function.

In this paper, we study a network formation model where the nodes of a directed graph to be formed are left free to choose where to place their out-links, with the constraint that each node has a fixed number of links at its disposal (and it has to place them all). We cast the problem into a game-theoretic setting where the utilities of the nodes are exactly their PageRank centralities. We first show that this class of network formation games admits an ordinal potential function and thus these games are potential games. We then study the structure of the best response correspondence and the structure of the equilibria of such games. Our analysis highlights two effects that network formation mechanisms purely based on centrality maximization have on the emerging network structure:

- a preference for local interactions that induces equilibria with a large number of undirected links and short cycles;
- a large number of isolated connected components and a small hierarchical depth.

In particular, in the homogeneous case, where players all have the same number d of out-links, our results provide a characterization of the potential maximizing equilibria, and a complete classification of the Nash equilibria for the two special cases when $d = 1$ and $d = 2$, respectively.

A similar network formation game is considered in the work [30] with the important difference that each player is left free to choose the number of her out-links (that can also be zero). In this framework, the best action of a node is always to link back to nodes in its in-neighborhood. Using this fact, the authors of [30] prove that all Nash equilibria have an undirected graph as a core, with possibly a set of nodes linking to the core and having no in-links. The work [14] later proved that these Nash equilibria explicitly depend on the discount factor of the PageRank centrality, answering a question left open by [30]. Although certain qualitative features of the Nash equilibria are similar in the two models, none of the results in [30] can be extended to our model and various counterexamples are presented in this paper. The recent work [16] is also closely related to ours: there, the authors prove the existence of Nash equilibria for a generalized version of our game. However, their proof is non-constructive and our classification of Nash equilibria cannot be derived from their results.

The problem considered in this paper is an instance of a *network formation game*, where players are identified with nodes and their actions determine the underlying network. The related literature is vast and goes back to the seminal paper [33] proposing a model where the opposite of the utility of a node consists in a discounted sum of distances from all other nodes plus a cost determined by the number of out-links maintained by the node. In their model, undirected links are considered and both incident nodes must agree to create a link between them, while each node can unilaterally prune a link. For this and similar models, the authors of [33] introduce and study a related concept of pairwise stability and of social efficiency. Other works, including [5, 23, 27], have reconsidered the model in [33] and studied network formation games where nodes can autonomously create and delete links. These works focus on the structure of Nash equilibria and on the analysis of the Price of Anarchy, while their contributions differ for the way links are considered, either directed or undirected, and in the form of the utility functions.

The works [36] and [26] consider bounded budget models that are closer in spirit to our framework. In these works, links are not associated to a cost but rather the number of out-links that a node can have is bounded or exactly pre-specified. Two different games are investigated where the nodes' costs are either the sum or the maximum of all distances from the other nodes. In the article [36], the emerging graph is considered directed, and existence of Nash equilibria under uniform budget is proven as well as estimations of the Price of Anarchy are derived. In the work [26], link ownership belongs to just one node, but its effect is bidirectional: the authors prove existence and connectedness of Nash equilibria and establish bounds for the Price of Anarchy.

A related model is the one considered in the article [38], where the authors start with the same link ownership model as in [26], but moves are restricted to swapping one link with another one. Their model is inspired by a similar swapping model proposed by [1] where, instead, links do not have ownership. In the work [38], a related concept of equilibrium is introduced (asymmetric swap-equilibrium) and it is discussed how this generalizes most of the concepts of equilibria previously introduced in network formation games (including [26]). In [38], it is also proved that in every such swap-equilibrium, at most one connected component is 2-edge connected. The work [35] provides an extensive analysis of best response dynamics for various of the network formation games illustrated above. In particular, the swap models in [1] and [38] are shown to be not weakly acyclic so that the convergence of the best response dynamics is not guaranteed. A positive result for the special case when networks are trees is presented: in this case, the swap games are ordinal potential and this ensures finite time convergence of the best response dynamics to a subset of Nash equilibria.

A different network formation model with bounded budget is proposed in the work [25], where each node's utility is either the average or the minimum of the maximal flows to the other nodes and where actions consist in establishing out-links and buying capacity on them. All Nash equilibria are proved to be connected and tight bounds for the Price of Anarchy are provided. These flow games do not enjoy the finite improvement property, hence also in this case convergence of the best response dynamics is not guaranteed.

We remark some striking differences between the model studied in the present paper and the models proposed in the literature on network formation games we just reviewed. Since the PageRank centrality is a normalized measure (i.e., the sum of all centralities is always equal to 1), our model results in a constant-sum game. Yet, as mentioned earlier, we prove that our game is an ordinal potential game: more precisely, the closely related game where the utilities are the logarithm of the PageRank centralities is an exact potential game. This is in contrast with the previously cited network formation games based on distances or flows and has relevant consequences both in terms of the tools used in the analysis and in the type of emerging network structures. E.g., the bounded budget model studied in [26], which is one of the closest to our model, yields Nash equilibria that are connected, while, as we will see, the lack of connectivity of equilibria is the norm in our model.

The remainder of this paper is organized as follows. Section 2 introduces the required graph-theoretic and network centrality background. Section 3 provides a formal definition of the centrality maximization game studied in this paper, along with some of its fundamental properties (including the fact that it is a potential game). In Section 4, we develop an equilibrium analysis of the centrality maximization game for general out-degree profiles. Our main results there are Theorem 1 showing how rational choices yield local connections, Theorem 2 describing the structure of a Nash equilibrium in terms of its connected components, and Theorem 3 providing an alternative characterization of potential-maximizing equilibria for large enough values of the discount factor. We also have a number of secondary results such as Corollary 4 providing a lower bound on the number of undirected links and 3-cycles in Nash equilibria. In Section 5, we focus on homogeneous out-degree profiles: first, we prove a result on the structure of potential-maximizing equilibria (Theorem 4) and then derive a complete classification of Nash equilibria when nodes have all out-degree equal to 1 (Theorem 5) and all out-degree equal to 2 (Theorem 6). Section 6 presents numerical experiments on the role of the game parameters in the network formation. Finally, Section 7 concludes with a summary and the description of some open problems.

A preliminary and incomplete version of our work was presented at the 21st IFAC World Conference [13]. The results reported in [13] are limited to the homogeneous out-degree profiles and correspond to a subset of those presented in Section 5. More specifically, [13, Theorem 6] corresponds to points (i) and (iii) of Theorem 5, while [13, Theorem 8] corresponds Theorem 6(i). On the other hand, [13] did not contain any of the other results reported in this paper.

Throughout the paper, we shall use the following notational convention. The indicator function of a set $\mathcal{A}$ is denoted by $\mathbb{1}_{\mathcal{A}}$, so that $\mathbb{1}_{\mathcal{A}}(a) = 1$ if $a \in \mathcal{A}$ and $\mathbb{1}_{\mathcal{A}}(a) = 0$ if $a \notin \mathcal{A}$. The all-one vector is denoted by $\mathbf{1}$, while δ^i stands for the vector with all zero entries except for entry i which is equal to 1, i.e., $\delta^i_i = 1$ and $\delta^i_j = 0$ for $j \neq i$. For two real-valued functions f and g, the asymptotic notation $f \sim g$ means that $\lim f/g = 1$, whereas $f \asymp g$ means that $0 < \liminf f/g \leq \limsup f/g < +\infty$.

2. Graph-Theoretic Notions and PageRank Centrality In this section, we review some graph-theoretic concepts and introduce the notion of PageRank centrality in a network.

We shall consider finite directed graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V}$ and link set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, containing no self-loops, i.e., such that $(i,i) \notin \mathcal{E}$ for any i in $\mathcal{V}$. The *out-neighborhood* and the *out-degree* of a node i in $\mathcal{V}$ will be denoted by $\mathcal{N}_i = \{j \in \mathcal{G} : (i,j) \in \mathcal{E}\}$ and $d_i = |\mathcal{N}_i|$, respectively. The vector $\mathbf{d} = (d_i)_{i \in \mathcal{V}}$ will be referred to as the *out-degree profile*. We shall always consider cases

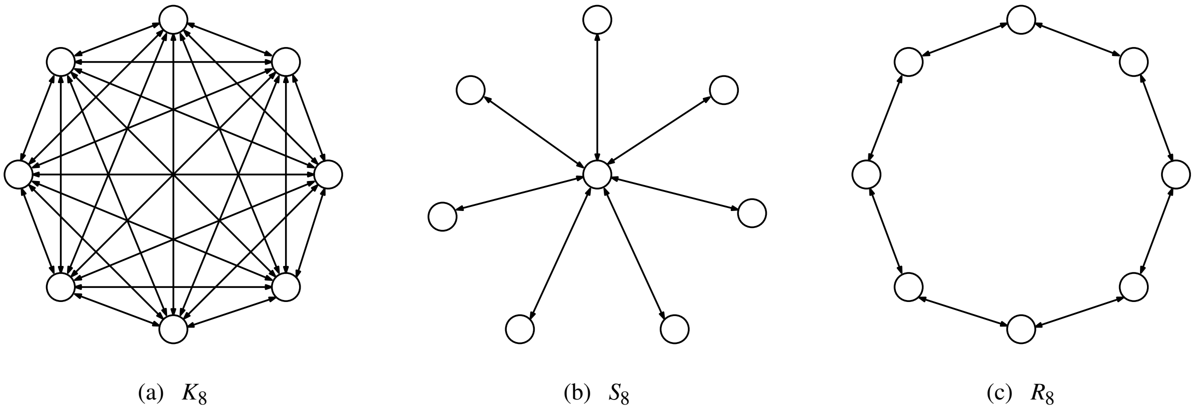


FIGURE 1. The complete, star, and ring graphs with $n = 8$ nodes.

where there are no sink nodes, i.e., $d_i \geq 1$ for every i in $\mathcal{V}$. An *undirected link* in $\mathcal{G}$ is a pair $\{i, j\}$ such that (i, j) and (j, i) are both in $\mathcal{E}$. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is referred to as *undirected* whenever $(i, j) \in \mathcal{E}$ implies that $(j, i) \in \mathcal{E}$. Standard examples of undirected graphs with n nodes include the complete graph K_n (i.e., the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{E} = \{(i, j) : i \neq j \in \mathcal{V}\}$), the star graph S_n (i.e., the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, 2, \ldots, n\}$ and $\mathcal{E} = \{(1, i) : 2 \leq i \leq n\} \cup \{(i, 1) : 2 \leq i \leq n\}$), and the ring graph R_n (i.e., the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, 2, \ldots, n\}$ and $\mathcal{E} = \{(i, j) : |i - j| = 1 \bmod n\}$), all displayed in Figure 1.

An automorphism of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a one-to-one mapping $f : \mathcal{V} \to \mathcal{V}$ such that $(i, j) \in \mathcal{E}$ if and only if $(f(i), f(j)) \in \mathcal{E}$. Two graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\mathcal{G} = (\mathcal{V}', \mathcal{E}')$ are referred to as *isomorphic* if they simply differ by a relabeling of the vertices, i.e., if there exists a one-to-one map $f : \mathcal{V} \to \mathcal{V}'$ such that $(i, j) \in \mathcal{E}$ if and only if $(f(i), f(j)) \in \mathcal{E}'$. The subgraph induced by a subset of nodes $\mathcal{W} \subseteq \mathcal{V}$ is $\mathcal{G}_{\mathcal{W}} = (\mathcal{W}, \mathcal{E} \cap (\mathcal{W} \times \mathcal{W}))$. A k-clique is a subgraph that is isomorphic to K_k.

A length-l *walk* in $\mathcal{G}$ from a node i to a node j is an $(l + 1)$-tuple of nodes $(r_0, r_1, \ldots, r_l)$ such that $r_0 = i$, $r_l = j$, and $(r_{h-1}, r_h) \in \mathcal{E}$ for every $1 \leq h \leq l$. A *path* is a walk $(r_0, r_1, \ldots, r_l)$ such that $r_h \neq r_k$ for every $0 \leq h < k \leq l$, except for possibly $r_0 = r_l$, in which case the path is referred to as *closed*. A closed path of length $l \geq 3$ is called a *cycle*. For two distinct nodes i and j in $\mathcal{V}$, we say that a subset $C \subseteq \mathcal{V} \setminus \{i, j\}$ *cuts* (or is a *cut-set*) between i and j if every path from i to j in $\mathcal{G}$ has at least a node in C. Given a node i in $\mathcal{V}$, we denote by $\mathcal{N}_i^{-l}$ the set of nodes j for which there exists a walk of length at most l from j to i in the graph $\mathcal{G}$. Let also $\mathcal{N}_i^{-\infty} = \bigcup_{l \geq 0} \mathcal{N}_i^{-l}$ be the set of all nodes from which i is reachable in $\mathcal{G}$. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is *strongly connected* if $\mathcal{N}_i^{-\infty} = \mathcal{V}$ for every node i in $\mathcal{V}$. Every graph $\mathcal{G}$ can be decomposed into its *connected components*, i.e., its maximal strongly connected subgraphs. A connected component $\mathcal{G}_{\mathcal{S}}$ of $\mathcal{G}$ is referred to as: a *sink* if there is no link in $\mathcal{G}$ from $\mathcal{S}$ to $\mathcal{V} \setminus \mathcal{S}$; a *source* if $\mathcal{G}_{\mathcal{S}}$ is not a sink and there is no link in $\mathcal{G}$ from

$\mathcal{V} \setminus \mathcal{S}$ to $\mathcal{S}$. Notice that, with this convention, isolated connected components are to be considered as sinks and not as sources.

To a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ we associate its normalized adjacency matrix R in $\mathbb{R}^{\mathcal{V}\times\mathcal{V}}$ with entries

$$R_{ij} = \begin{cases} 1/d_i & \text{if } j \in \mathcal{N}_i \\ 0 & \text{if } j \in \mathcal{V} \setminus \mathcal{N}_i\,. \end{cases} \tag{1}$$

For a parameter β in $(0, 1)$ and a probability vector η in $\mathbb{R}_+^{\mathcal{V}}$, to be referred to respectively as the *discount factor* and the *intrinsic centrality*, the *PageRank centrality* vector is then defined as

$$\pi = (1-\beta)(I - \beta R^\top)^{-1}\eta = (1-\beta)\sum_{k=0}^{+\infty} \beta^k (R^\top)^k \eta\,. \tag{2}$$

The fact that R is a row-stochastic matrix, hence with spectral radius 1, ensures correctness of equation (2) and also implies that π is a probability vector. Indeed, equation (2) implies that $\pi = P^\top \pi$ is the unique invariant probability distribution of the irreducible stochastic matrix

$$P = \beta R + (1-\beta)\mathbf{1}\eta^\top\,. \tag{3}$$

While the PageRank centrality has often been proposed as a measure of the relative relevance of the nodes in network in an axiomatic way [2, 12], the following example illustrates that the PageRank centrality arises a key measure of influence in opinion dynamics models.

Example 1. Consider the Friedkin-Johnsen opinion dynamics [18, 29, 42], whereby nodes i in $\mathcal{V}$ are identified with agents in a social network, each holding a scalar opinion value $y_i(t)$ that gets updated as follows:

$$y_i(t+1) = \beta \sum\nolimits_j R_{ij} y_j(t) + (1-\beta) a_i\,, \qquad t = 0, 1, \dots,$$

where R is a stochastic matrix, β in $(0, 1)$ is a discount factor, and a_i in an anchor opinion value for agent i (originally $a_i = y_i(0)$). It is then easily verified that

$$y^* = \lim_{t\to+\infty} y(t) = (1-\beta)(I - \beta R)^{-1} a\,.$$

For a given vector of weights η in $\mathbb{R}_+^{\mathcal{V}}$ such that $\sum_i \eta_i = 1$, the weighted average equilibrium opinion satisfies

$$\sum\nolimits_i \eta_i y_i^* = \eta^\top (1-\beta)(I-\beta R)^{-1} a = \sum\nolimits_i \pi_i a_i\,,$$

i.e., it coincides with the weighted average of the anchor values weighted by the PageRank centrality vector defined in equation (2). This example illustrates as the PageRank centrality arises a key measure not only in informational networks —such as citation networks or the WWW— but also in social networks. □

In the sequel, it will prove useful to rely on the following probabilistic interpretation of the PageRank centrality. Consider a discrete-time Markov chain $(V_t)_{t\geq 0}$ with finite state space $\mathcal{V}$, and transition probability matrix P as in equation (3), so that, given that its current state $V_t = i$, with probability β the next state V_{t+1} will be chosen uniformly at random among the d_i out-neighbors of i, while with probability $(1-\beta)$ the next state V_{t+1} will be sampled from the probability distribution η. Then, the PageRank centrality vector $\pi = P^\top \pi$ coincides with the unique stationary distribution of V_t. Moreover, for two nodes i and j in $\mathcal{V}$, let $T_i = \min\{t \geq 0 : V_t = i\}$ be the hitting time in i and let $\tau_j^i = \mathbb{E}[T_i : V_0 = j]$ be its conditional expected value when starting from node j. In other terms, τ_j^i represents the time it takes, in expectation, for the Markov chain V_t to go from node j to node i. By [41, Theorem 1.3.5], for every i in $\mathcal{V}$, the vector $(\tau_j^i)_{j\in\mathcal{V}}$ coincides with the unique solution of the following linear system:

$$\tau_i^i = 0, \qquad \tau_j^i = 1 + (1-\beta)\sum_{k\in\mathcal{V}} \eta_k \tau_k^i + \frac{\beta}{d_j}\sum_{k\in\mathcal{N}_j} \tau_k^i, \quad \forall j \neq i\,. \tag{4}$$

Equation (4) can in fact be derived by a conditioning argument and the Markov property: the time to reach i when starting from node i itself is obviously 0, whereas, when starting from a node $j \neq i$, it takes one time step to the Markov chain to make the first move and such a move, with conditional probability $P_{ik} = (1-\beta)\eta_k + \beta \mathbb{1}_{\mathcal{N}_j}(k)/d_j$, will lead to a node k from which the expected time to reach node i is τ_k^i. A related argument can be used to get the following representation of the Page-Rank centrality.

Lemma 1. *In a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the Page-Rank centrality with discount factor β and intrinsic centrality η is given by*

$$\pi_i = \frac{1}{1 + (1-\beta)\tau_\eta^i + \frac{\beta}{d_i}\sum_{j\in\mathcal{N}_i}\tau_j^i}, \qquad \tau_\eta^i = \sum_{j\in\mathcal{V}} \eta_j \tau_j^i, \tag{5}$$

for every node i in $\mathcal{V}$.

Proof Let $T_i^+ = \min\{t \geq 1 \,|\, X_t = i\}$ be the return time in a node i in $\mathcal{V}$. By conditioning on X_1, using the Markov property and equation (4), we have

$$\mathbb{E}[T_i^+ \,|\, X_0 = i] = 1 + \sum_{j\in\mathcal{V}} P_{ij}\tau_j^i = 1 + (1-\beta)\tau_\eta^i + \frac{\beta}{d_i}\sum_{j\in\mathcal{N}_i}\tau_j^i\,.$$

Combining the above with Kac's formula $\pi_i = 1/\mathbb{E}[T_i^+ \,|\, X_0 = i]$ [41, Theorem 1.7.7], we get relation (5). □

We close this section by observing that on the one hand the expected hitting times τ_j^i are independent from node i's out-neighborhood $\mathcal{N}_i$, on the other hand they do depend not only on the structure of the rest of the graph $\mathcal{G}$, but also on the discount factor β and on the intrinsic centrality η.

3. Centrality Maximization Network Formation Games In this section, we introduce a class of directed network formation games based on a competitive centrality maximization mechanism that will be the object of our study. We then prove some fundamental properties of these network formation games.

We consider directed network formation games where a finite set $\mathcal{V}$ of $n \geq 2$ nodes choose where to wire a predetermined number of out-links in order to maximize their own PageRank centrality. Specifically, for an integer vector $\mathbf{d}$ in $\{1, \ldots, n-1\}^{\mathcal{V}}$, a probability vector η in $\mathbb{R}_+^{\mathcal{V}}$, and a scalar parameter β in $(0, 1)$, we assume that each node i in $\mathcal{V}$ is to choose d_i distinct other nodes to link to with the aim of maximizing its own PageRank centrality in the resulting graph $\mathcal{G}$ with intrinsic centrality η and discount factor β. Hence, we model such multi-objective decision problem as a finite strategic game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ where:

- the player set is $\mathcal{V}$;
- the action set $\mathcal{A}_i$ of a player i in $\mathcal{V}$ coincides with the family of all $\binom{n-1}{d_i}$ subsets of $\mathcal{V} \setminus \{i\}$ of cardinality d_i;
- the utility $u_i(x)$ of a player i in $\mathcal{V}$ in a *configuration* $x = (x_1, \ldots, x_n)$ in $\mathcal{X} = \prod_{i \in \mathcal{V}} \mathcal{A}_i$ coincides with the i-th entry of the PageRank centrality vector

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

where $R(x)$ is the normalized adjacency matrix of the graph $\mathcal{G}(x) = (\mathcal{V}, \mathcal{E}(x))$ with node set $\mathcal{V}$ and link set $\mathcal{E}(x) = \{(i, j) \mid i \in \mathcal{V},\ j \in x_i\}$.

In the rest of this paper we shall refer to a game as above as the *centrality game* $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$. We shall use the following standard game-theoretic notions. For a configuration $x = (x_1, \ldots, x_n)$ in $\mathcal{X}$ and a player i in $\mathcal{V}$, the action profile of all players but i is x_{-i} in $\mathcal{X}_{-i} = \prod_{k \neq i} \mathcal{A}_k$, and we write $u_i(x) = u_i(x_i, x_{-i})$ for her utility. The *best response set* of player i to an action profile x_{-i} in $\mathcal{X}_{-i}$ is

$$\mathcal{B}_i(x_{-i}) = \{a \in \mathcal{A}_i \mid u_i(a, x_{-i}) \geq u_i(b, x_{-i}),\ \forall b \in \mathcal{A}_i\}.$$

A (pure strategy) *Nash equilibrium* is a configuration x in $\mathcal{X}$ such that $x_i \in \mathcal{B}_i(x_{-i})$ for every player i in $\mathcal{V}$. A Nash equilibrium x is *strict* if $\{x_i\} = \mathcal{B}_i(x_{-i})$ for every player i in $\mathcal{V}$. Following the terminology introduced in [39], we refer to a game as:

- *ordinal potential* if there exists a function $\Psi : \mathcal{X} \to \mathbb{R}$, to be referred to as an ordinal potential of the game, such that

$$u_i(x) < u_i(y) \;\Leftrightarrow\; \Psi(x) < \Psi(y)\,, \tag{6}$$

for every player i in $\mathcal{V}$ and every two configurations x and y in $\mathcal{X}$ such that $x_{-i} = y_{-i}$.

- *exact potential* if there exists a function $\Psi : \mathcal{X} \to \mathbb{R}$, to be referred to as an exact potential of the game, such that

$$u_i(y) - u_i(x) = \Psi(y) - \Psi(x)\,, \tag{7}$$

for every player i in $\mathcal{V}$ and every two configurations x and y in $\mathcal{X}$ such that $x_{-i} = y_{-i}$.

Clearly, exact potential games are also ordinal potential. It is well known [39] that finite ordinal potential games always admit Nash equilibria. In fact, their set of Nash equilibria can be very large. For the sake of getting a more insightful classification, we introduce the following notions:

- for $l \geq 0$, a $(l+1)$-tuple of configurations $(x^{(0)}, x^{(1)}, \dots, x^{(l)})$ in $\mathcal{X}^{l+1}$ is a length-l *best response path* from a configuration $x^{(0)}$ to a configuration $x^{(l)}$ if, for every $1 \leq k \leq l$, $x^{(k)} \neq x^{(k-1)}$ and there exists i_k in $\mathcal{V}$ such that

$$x^{(k)}_{i_k} \in \mathcal{B}_{i_k}\left(x^{(k-1)}_{-i_k}\right), \qquad x^{(k)}_{-i_k} = x^{(k-1)}_{-i_k},$$

i.e., it is obtained by a sequence of single player modifications choosing best response actions;

- a configuration x in $\mathcal{X}$ is *recursive* if for every other configuration y in $\mathcal{X}$ such that there exists a best response path from x to y, there is also a best response path from y to x;
- a subset $\mathcal{Y} \subseteq \mathcal{X}$ of configurations is *invariant* with respect to best response paths if for every best response path $(x^{(0)}, x^{(1)}, \dots, x^{(l)})$ with initial configuration $x^{(0)}$ in $\mathcal{Y}$, we have that the final configuration $x^{(l)}$ belongs to $\mathcal{Y}$ as well.

For ordinal potential games, recursive configurations enjoy the following remarkable properties.

Lemma 2. *For a finite ordinal potential game:*

(i) every strict Nash equilibrium is recursive;

(ii) every maximizer of the potential function over the configuration space is recursive;

(iii) every recursive configuration is a Nash equilibrium;

(iv) the set of recursive Nash equilibria is invariant with respect to best response paths;

(v) every subset of Nash equilibria that is invariant with respect to best response paths contains only recursive Nash equilibria;

(vi) from every configuration there exists a best response path to a recursive Nash equilibrium.

Proof Let $\Psi : \mathcal{X} \to \mathbb{R}$ be an ordinal potential of the game.

(i) If x^* is a strict Nash equilibrium, then, by definition, the only best response path starting in x^* is the trivial length-0 one. This readily implies that x^* is recursive.

(ii) Let $(x^{(0)}, x^{(1)}, \dots, x^{(l)})$ be a best response path starting in an ordinal potential maximizer $x^{(0)}$ in $\arg\max\{\Psi(x) : x \in \mathcal{X}\}$. By relation (6), we have $\Psi(x^{(0)}) \le \Psi(x^{(1)}) \le \dots \le \Psi(x^{(l)}) \le \Psi(x^{(0)})$. This yields $\Psi(x^{(0)}) = \Psi(x^{(1)}) = \dots = \Psi(x^{(l)})$ and relation (6) implies that $(x^{(l)}, x^{(l-1)}, \dots, x^{(0)})$ is also a best response path, thus proving that $x^{(0)}$ is recursive.

(iii) Let x in $\mathcal{X}$ be a configuration that is not a Nash equilibrium. By definition, there exist a node i in $\mathcal{V}$ and a configuration y in $\mathcal{X}$ with $y_{-i} = x_{-i}$ and $u_i(x) < u_i(y)$. Then, relation (6) implies that $\Psi(x) < \Psi(y)$. Now, if x were recursive, then there would exist a best response path $(y = x^{(0)}, x^{(1)}, \dots, x^{(l)} = x)$ from y to x. But then, relation (6) again would imply that $\Psi(y) = \Psi(x^{(0)}) \le \Psi(x^{(1)}) \le \Psi(x^{(l)}) = \Psi(x)$, thus leading to a contradiction. We have thus shown that, if x is not a Nash equilibrium, then it cannot be recursive.

(iv) Let x in $\mathcal{X}$ be recursive and assume that there exists a best response path γ_1 from x to another configuration y in $\mathcal{X}$. Now, let there be a best response path γ_2 from y to a third configuration z in $\mathcal{X}$. Since the concatenation of γ_1 and γ_2 is itself a best response path from x to z and x is recursive, there should exist a best response path γ_3 from z to x. Then, the concatenation of γ_3 and γ_1 is a best response path from z to y. This proves that y is recursive.

(v) Let $\mathcal{F}$ be a subset of Nash equilibria that is invariant with respect to best response paths. Let $(x^{(0)}, x^{(1)}, \dots, x^{(l)})$ be a best response path starting in some $x^{(0)}$ in $\mathcal{F}$, so that $x^{(k)} \in \mathcal{F}$ is thus a Nash equilibrium for all $0 \le k \le l$. It then follows that, for every $0 \le k < l$ there exists a node i_k in $\mathcal{V}$ such that $x^{(k+1)}_{-i_k} = x^{(k)}_{-i_k}$ and $u_{i_k}(x^{(k+1)}) = u_{i_k}(x^{(k)}_{-i_k})$, so that $(x^{(l)}, x^{(l-1)}, \dots, x^{(0)})$ is also a best response path. This implies that the Nash equilibrium $x^{(0)}$ is recursive.

(vi) Consider the finite directed graph $\mathcal{H} = (\mathcal{X}, \mathcal{F})$ whose node set is the set of configurations $\mathcal{X}$ and where there is a link from x to y is and only if $y_{-i} = x_{-i}$ and $u_i(x) \le u_i(y)$ for some i in $\mathcal{V}$. Let $\mathcal{H}_1 = (\mathcal{X}_1, \mathcal{F}_1), \mathcal{H}_2 = (\mathcal{X}_2, \mathcal{F}_2), \dots, \mathcal{H}_s = (\mathcal{X}_s, \mathcal{F}_s)$, be the sink connected components of $\mathcal{H}$. Then, from every x in $\mathcal{X}$, some y in $\mathcal{Y} = \cup_{h=1}^{s} \mathcal{X}_h$ is reachable. the proof is completed by observing that, by construction, all configurations in $\mathcal{Y} = \cup_{h=1}^{s} \mathcal{X}_h$ are recursive. □

The following example clarifies the difference between Nash equilibria, recurrent Nash equilibria, strict Nash equilibria, and maximizers of the potential function.

	a	b	c	d
a	0	0	0	0
b	0	1	0	0
c	0	0	2	2
d	0	0	2	2

FIGURE 2. Table representation of the 2-player identical interest game in Example 2.

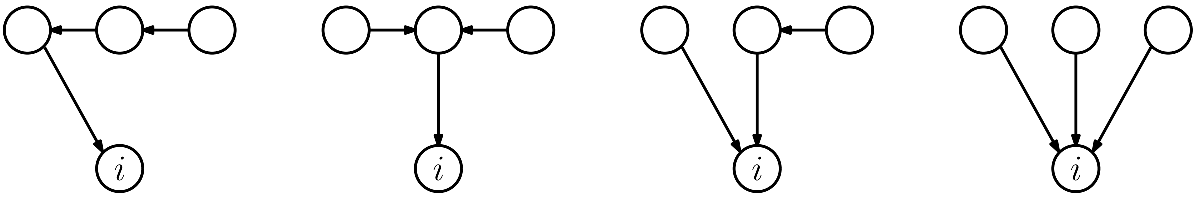


FIGURE 3. For $|\mathcal{V}| = 4$, all four possible —up to isomorphisms— spanning directed rooted trees in a given node i.

EXAMPLE 2. Consider a two-player game with action set $\mathcal{A} = \{a, b, c, d\}$ for both players and identical utilities

$$u_1(x) = u_2(x) = \begin{cases} 2 & \text{if } x_1 \text{ and } x_2 \in \{c, d\} \\ 1 & \text{if } x_1 = x_2 = b \\ 0 & \text{otherwise}. \end{cases}$$

(See also the table in Figure 2.) This is an identical interest game, hence an exact potential game with potential function $\Psi(x) = u_1(x) = u_2(x)$. Notice that: $x^* = (a, a)$ is a Nash equilibrium that is not recurrent; $x^\bullet = (b, b)$ is a strict Nash equilibrium but it is not a maximizer of the potential Ψ; $x^\circ = (c, c)$ is a non-strict recurrent Nash equilibrium and it is a maximizer of the potential Ψ. □

We will now show that centrality games are ordinal potential games. Towards this goal, for a node i in $\mathcal{V}$, consider the set $\mathbb{T}_i$ of *spanning directed rooted trees*, i.e., acyclic graphs $\mathcal{T} = (\mathcal{V}, \mathcal{E}_\mathcal{T})$ with node set $\mathcal{V}$, whereby each node has out-degree 1 except for a node i that has out-degree 0 (see Figure 3 for an illustration). Then, for a configuration x in $\mathcal{X}$, define

$$n_i(x_{-i}) = \sum_{\mathcal{T} \in \mathbb{T}_i} \prod_{(j,k) \in \mathcal{E}_\mathcal{T}} \left(\beta \mathbb{1}_{x_j}(k) + (1 - \beta)\eta_k \right), \tag{8}$$

for every node i in $\mathcal{V}$, and let

$$Z(x) = \sum_{i \in \mathcal{V}} n_i(x_{-i}). \tag{9}$$

Observe that the right-hand side of equation (8) coincides with the sum, over all possible spanning directed rooted trees $\mathcal{T}$ in $\mathbb{T}_i$, of the product of the weights $P_{jk}(x) = \left(\beta \mathbb{1}_{x_j}(k) + (1 - \beta)\eta_k\right)$ of all the $n - 1$ links (j, k) in $\mathcal{E}_\mathcal{T}$. A key observation is then that, since no link (j, k) in $\mathcal{E}_\mathcal{T}$ of any spanning directed rooted tree $\mathcal{T}$ in $\mathbb{T}_i$ is an outgoing link from node i, the right-hand side of equation (8) does not depend on the action x_i of node i. This justifies writing, as we did, the left-hand side of equation (8) as a function of x_{-i} only. We then have the following instrumental result.

Lemma 3. *In a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$,*

$$u_i(x) = \frac{n_i(x_{-i})}{Z(x)}, \tag{10}$$

for every player i in $\mathcal{V}$ and configuration x in $\mathcal{X}$.

Proof By the Markov Chain Tree Theorem [3], the entries of the invariant distribution $\pi(x)$ of the irreducible stochastic matrix $P(x) = \beta R(x) + (1-\beta)\mathbf{1}\eta^\top$ are as in the right-hand side of relation (10). The claim then follows from our previous observation that the PageRank centrality vector coincides with the invariant distribution $\pi(x)$ of $P(x)$. □

Proposition 1. *Consider a centrality game $\Gamma = \Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ and let $\overline{\Gamma}$ be the game with the same player set $\mathcal{V}$ and configuration space $\mathcal{X}$, and utilities*

$$\overline{u}_i(x) = \log u_i(x), \qquad \forall i \in \mathcal{V}, \quad \forall x \in \mathcal{X}.$$

Let $\Psi(x) = -\log Z(x)$. Then:

(i) $\Psi(x)$ is an exact potential for $\overline{\Gamma}$;

(ii) $\Psi(x)$ is an ordinal potential for Γ.

Proof (i) For every node i in $\mathcal{V}$ and configuration x in $\mathcal{X}$, relation (10) implies that

$$\log u_i(x) = -\log Z(x) + \log n_i(x_{-i}).$$

For two configurations x and y in $\mathcal{X}$ such that $x_{-i} = y_{-i}$, the equation above implies that

$$\begin{aligned} \overline{u}_i(y) - \overline{u}_i(x) &= \log u_i(y) - \log u_i(x) \\ &= -\log Z(y) + \log Z(x) + \log n_i(y_{-i}) - \log n_i(x_{-i}) \\ &= -\log Z(y) + \log Z(x) \\ &= \overline{\Psi}(y) - \overline{\Psi}(x), \end{aligned} \tag{11}$$

thus showing that $\overline{\Psi}(x) = -\log Z(x)$ is an exact potential for the game with utilities $\overline{u}_i(x)$.

(ii) This follows from (i) and the fact that

$$\operatorname{sgn}(u_i(x) - u_i(y)) = \operatorname{sgn}(\overline{u}_i(x) - \overline{u}_i(y)),$$

for every configurations x and y in $\mathcal{X}$. □

We shall denote by $\mathcal{X}^*$ the set of all Nash equilibria of a centrality game $\Gamma(n, \beta, \eta, \mathbf{d})$ and with the symbols $\mathcal{X}^\circ$ and $\mathcal{X}^\bullet$, the subsets, respectively, of the recursive and of the strict Nash equilibria. We shall also denote by $\mathcal{X}^{\mathrm{Z}} = \arg\min\{Z(x) : x \in \mathcal{X}\}$ the set of maximizers of the ordinal potential $\Psi(x) = -\log Z(x)$ of the game. It then follows from Lemma 2 that

$$\mathcal{X}^* \supseteq \mathcal{X}^\circ \supseteq \mathcal{X}^\bullet, \qquad \mathcal{X}^* \supseteq \mathcal{X}^\circ \supseteq \mathcal{X}^{\mathrm{Z}}\,.$$

The previous results have direct implications on the asymptotic behavior of some standard learning dynamics based on stochastic strategy revision processes [7, 8], as detailed below. Consider asynchronous dynamics modeled as discrete-time Markov chains with finite state space $\mathcal{X}$ whereby, at every time step $t = 0, 1, 2, \dots$, conditioned on the current configuration $X(t) = x$, one player i is selected uniformly at random from the player set $\mathcal{V}$ and she updates her action to a value $X_i(t+1) = a$ sampled from a conditional distribution $p_i(a|x_{-i})$ while all other players keep playing the same action $X_{-i}(t) = x_{-i}$. In particular, we shall refer to the case when

$$p_i(a|x_{-i}) = p_i^{(0)}(a|x_{-i}) = |\mathcal{B}_i(x_{-i})|^{-1}\,\mathbb{1}_{\mathcal{B}_i(x_{-i})}(a)\,,$$

so that the active player choses her new action uniformly at random from the her best response set $\mathcal{B}_i(x_{-i})$, as the (asynchronous) best response dynamics (c.f. [8]). On the other hand, we shall refer to the case when

$$p_i(a|x_{-i}) = p_i^{(\gamma)}(a|x_{-i}) = \frac{u_i(a, x_{-i}))^{1/\gamma}}{\sum_{b\in\mathcal{A}_i}(u_i(b, x_{-i}))^{1/\gamma}}\,,$$

for some $\gamma > 0$ as the noisy best response dynamics with noise level γ: this can be recognized as the log-linear learning dynamics [7, 37] for the game $\overline{\Gamma}$ with utilities $\overline{u}_i(x) = \log u_i(x)$. Observe that

$$\lim_{\gamma\downarrow 0} p_i^{(\gamma)}(a|x_{-i}) = p_i^{(0)}(a|x_{-i})\,.$$

COROLLARY 1. *Consider a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$. Then,*

(i) for every initial configuration, the best response dynamics gets absorbed in finite time in the set of recursive Nash equilibria $\mathcal{X}^\circ$;

(ii) for every noise level $\gamma \geq 0$, the noisy best response dynamics is reversible with respect to its unique stationary distribution

$$\mu_x^{(\gamma)} = \frac{Z(x)^{-1/\gamma}}{\sum_y Z(y)^{-1/\gamma}}\,, \qquad x \in \mathcal{X}\,,$$

and

$$\mu_x^{(\gamma)} = \lim_{\gamma\downarrow 0} \mu_x^{(\gamma)} = \frac{1}{|\mathcal{X}^{\mathrm{Z}}|}\,\mathbb{1}_{\mathcal{X}^{\mathrm{Z}}}(x)\,, \qquad \forall x \in \mathcal{X}\,.$$

Proof Point (i) of the claim follows from point (i) of Proposition 1 and points (vi) and (iv) of Lemma 2. Point (ii) of the claim follows from point (ii) of Proposition 1 and Section 3.3 in [9]. □

.

4. Equilibrium Analysis for General Out-Degree Profiles In this section, we present our main results on the structure of Nash equilibria in centrality games $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with general out-degree profile $\mathbf{d}$. We recall our standing notation: given a configuration x, we denote by $\mathcal{G}(x)$ the corresponding graph and use a similar functional dependence for every other graph-theoretic concept. E.g., we shall write $\mathcal{N}_i^{-l}(x)$, $\mathcal{N}_i^{-\infty}(x)$ and often, since such sets only depend on x_{-i}, also $\mathcal{N}_i^{-l}(x_{-i})$. Similarly, we shall denote by $\tau_j^i(x_{-i})$ and $\tau_\eta^i(x_{-i})$ the expected hitting times of a node i in $\mathcal{V}$ for the Markov chain with transition probability matrix $P(x) = \beta R(x) + (1-\beta)\mathbf{1}\eta^\top$, since they do not depend on the action x_i of node i, as already observed in Section 2.

4.1. Locality of Best Response Observe that Lemma 1 allows us to rewrite the utility of a player i in $\mathcal{V}$ as

$$u_i(x) = \left(1 + (1-\beta)\tau_\eta^i(x_{-i}) + \frac{\beta}{d_i}\sum_{j\in x_i}\tau_j^i(x_{-i})\right)^{-1}, \qquad x \in \mathcal{X}, \tag{12}$$

leading to the following result.

Proposition 2. *In a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$,*

$$x_i \in \mathcal{B}_i(x_{-i}) \qquad \Longleftrightarrow \qquad \tau_k^i(x_{-i}) \geq \max_{j\in x_i}\tau_j^i(x_{-i}), \quad \forall k \notin x_i, \tag{13}$$

for every node i in $\mathcal{V}$ and strategy profile x_{-i} in $\mathcal{X}_{-i}$.

Proof The only term in the right hand side of relation (12) that depends on the action x_i of player i is the set over which the summation runs. It follows that

$$\mathcal{B}_i(x_{-i}) = \operatorname*{argmin}_{x_i\in\mathcal{A}_i}\sum_{j\in x_i}\tau_j^i(x_{-i})\,.$$

The minimization above is easily solved by choosing any d_i-tuple of nodes in $\mathcal{V}\setminus\{i\}$ that have the smallest hitting times to node i, thus proving the claim. □

Proposition 2 reduces the computation of the best response actions of a node i to finding the d_i nodes that have the smallest expected hitting times of node i. The main difficulty in applying it directly stems from the fact that, as already observed at the end of Section 2, such expected hitting times depend not only on the action profile x_{-i} of the rest of the players, which determines the graph except for node i's out-neighborhood x_i, but also on the discount factor β and on the intrinsic centrality η. We now present a remarkable result asserting that certain fundamental inequalities however always hold, independently from the discount factor β and the intrinsic centrality η.

PROPOSITION 3. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph, β a scalar in $(0, 1)$, and η a probability vector in $\mathbb{R}_+^{\mathcal{V}}$. Then, for every i in $\mathcal{V}$, the expected hitting times solution of equation (4) satisfy the following properties:*

(i) the relative order of $\{\tau_j^i : j \in \mathcal{V}\}$ does not depend on η;

(ii) the relative order of $\{\tau_j^i : j \in \mathcal{N}_i^{-\infty}\}$ only depends on the subgraph $\mathcal{G}_{\mathcal{N}_i^{-\infty}}$;

(iii) for every h, j in $\mathcal{V} \setminus \mathcal{N}_i^{-\infty}$ and k in $\mathcal{N}_i^{-\infty}$,

$$\tau_h^i = \tau_j^i > \tau_k^i; \tag{14}$$

(iv) for every k in $\mathcal{N}_i^{-\infty}$ and cut-set $C \subseteq \mathcal{V} \setminus \{k, i\}$ between k and i,

$$\tau_k^i > \min_{j \in C} \tau_j^i \,. \tag{15}$$

(v) for every three nodes h, i, and j in $\mathcal{V}$, such that $h \neq i \neq j$ and $d_h = d_j$,

$$\mathcal{N}_j \setminus \{h\} = \mathcal{N}_h \setminus \{j\} \qquad \Longrightarrow \qquad \tau_h^i = \tau_j^i \,.$$

Proof We start with a simple but crucial observation. For a node i in $\mathcal{V}$, consider the recursive relations (4) characterizing the expected hitting times τ_j^i as j varies in $\mathcal{V}$. Define

$$\tilde{\tau}_j^i = \frac{\tau_j^i}{1 + (1-\beta)\tau_\eta^i}\,, \qquad i, j \in \mathcal{V}\,, \tag{16}$$

and notice that these quantities satisfy the relations

$$\tilde{\tau}_i^i = 0\,, \qquad \tilde{\tau}_j^i = 1 + \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j} \tilde{\tau}_k^i\,, \qquad j \neq i\,. \tag{17}$$

This yields that the values $\tilde{\tau}_j^i$ coincide with the expected hitting times on node i for the case when $\eta_j = \delta_j^i$. As the transformation (16) preserves the relative order of the expected hitting times, this proves point (i).

Notice that when $\eta = \delta^i$ the equations (4) corresponding to nodes j in $\mathcal{N}_i^{-\infty}$ are completely decoupled from the remaining equations, so that their solutions (hence, their relative order) only depend on the subgraph induced by $\mathcal{N}_i^{-\infty}$. Together with point (i), this proves point (ii).

Thanks again to point (i), it is sufficient to prove points (iii), (iv), and (v) in the special case $\eta_j = \delta^i_j$, as their statements concern only on the relative order of the expected hitting times. In this case, for every j in $\mathcal{V} \setminus \mathcal{N}_i^{-\infty}$, we have that $\mathcal{N}_j \subseteq \mathcal{V} \setminus \mathcal{N}_i^{-\infty}$, so that the equations (4) corresponding to nodes j in $\mathcal{V} \setminus \mathcal{N}_i^{-\infty}$ are decoupled from those corresponding to nodes j in $\mathcal{N}_i^{-\infty}$ and it can be directly verified by substitution that the unique solution $\tilde{\tau}^i$ of equations (17) is such that

$$\tilde{\tau}^i_j = 1/(1-\beta), \qquad \forall j \in \mathcal{V} \setminus \mathcal{N}_i^{-\infty}. \tag{18}$$

Let now $\mathcal{M} = \arg\max\{\tilde{\tau}^i_j : j \in \mathcal{N}_i^{-\infty}\}$, and, for j in $\mathcal{M}$, let $\alpha^i_j = |\mathcal{N}_j \cap \mathcal{N}_i^{-\infty}|/d_j$. Then, relations (17), the fact that $\tilde{\tau}^i_k \leq \tilde{\tau}^i_j$ for all k in $\mathcal{N}_j \cap \mathcal{N}_i^{-\infty}$, and relations (18) imply that

$$\tilde{\tau}^i_j = 1 + \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j} \tilde{\tau}^i_k = 1 + \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j \cap \mathcal{N}_i^{-\infty}} \tilde{\tau}^i_k + \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j \setminus \mathcal{N}_i^{-\infty}} \tilde{\tau}^i_k \leq 1 + \beta \alpha^i_j \tilde{\tau}^i_j + \beta(1-\alpha^i_j)\frac{1}{1-\beta},$$

which implies that

$$\tilde{\tau}^i_j \leq 1/(1-\beta), \qquad \forall j \in \mathcal{N}_i^{-\infty}. \tag{19}$$

To prove (iii), we need to show that the inequality in relation (19) is always strict. Assume by contradiction that $\tilde{\tau}^i_j = 1/(1-\beta)$ for some j in $\mathcal{N}_i^{-\infty}$. Then, for every j in $\mathcal{M}$, relation (19) implies that $\tilde{\tau}^i_j = 1/(1-\beta)$ and $\mathcal{N}_j \cap \mathcal{N}_i^{-\infty} \subseteq \mathcal{M}$. Since $j \in \mathcal{N}_i^{-\infty}$, an iteration of this argument implies that also $i \in \mathcal{M}$, so that $\tilde{\tau}^i_i = 1/(1-\beta)$, which contradicts the first identity in relations (17). Thus point (iii) is proven.

Let $\mathcal{W}$ be the set of nodes in $\mathcal{N}_i^{-\infty} \setminus C$ such that every path connecting them to i hits the cut set C. Suppose by contradiction that $\min\{\tilde{\tau}^i_j : j \in \mathcal{W}\} \leq \min\{\tilde{\tau}^i_j : j \in C\}$, and let j in $\mathcal{W}$ be any minimum point. Notice that $\mathcal{N}_j \subseteq \mathcal{W} \cup C \cup (\mathcal{V} \setminus \mathcal{N}_i^{-\infty})$. Then, by construction,

$$\tilde{\tau}^i_j = 1 + \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j} \tilde{\tau}^i_k \geq 1 + \beta \tilde{\tau}^i_j,$$

so that $\tilde{\tau}^i_j \geq \frac{1}{1-\beta}$. This contradicts point (iii) and proves point (iv).

Finally, to prove point (v), observe that for three nodes h, i, and j in $\mathcal{V}$, such that $h \neq i \neq j$ and $d_h = d_j$, we have $\mathcal{N}_j \setminus \{h\} = \mathcal{N}_h \setminus \{j\}$ if and only if either $\mathcal{N}_j = \mathcal{N}_h$ do not contain neither h nor j, or $h \in \mathcal{N}_j$, $j \in \mathcal{N}_h$, and $\mathcal{N}_j \setminus \{h\} = \mathcal{N}_h \setminus \{j\}$. In the former case, relations (16) and (17) imply that

$$\tilde{\tau}^i_h = 1 + \frac{\beta}{d_h} \sum_{k \in \mathcal{N}_h} \tilde{\tau}^i_k = 1 + \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j} \tilde{\tau}^i_k = \tilde{\tau}^i_j.$$

On the other hand, in the latter case, relations (16) and (17) imply that

$$\tilde{\tau}^i_h - \tilde{\tau}^i_j = \frac{\beta}{d_h} \sum_{k \in \mathcal{N}_h} \tilde{\tau}^i_k - \frac{\beta}{d_j} \sum_{k \in \mathcal{N}_j} \tilde{\tau}^i_k = \frac{\beta}{d_h} (\tilde{\tau}^i_h - \tilde{\tau}^i_j),$$

so that $(1 - \beta/d_h)(\tilde{\tau}^i_h - \tilde{\tau}^i_j) = 0$, i.e., $\tilde{\tau}^i_h = \tilde{\tau}^i_j$. □

A number of important properties of the structure of Nash equilibria of centrality games are a direct consequence of Propositions 2 and 3. Our first main result concerns the structure of the best response sets and shows how the optimal wiring strategy of a node is independent from the intrinsic centrality η and satisfies a locality property. Precisely, given the choices of all the other nodes, the d_i nodes to which a node i needs to link in order to maximize her utility are nodes from which node i itself can be reached in at most d_i hops, provided that there is a sufficient number of such nodes.

Theorem 1 (**Independence from intrinsic centrality and locality of best response**). *Let $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ be a centrality game. Then, for every node i in $\mathcal{V}$ and action profile x_{-i} in $\mathcal{X}_{-i}$:*

(i) the best response set $\mathcal{B}_i(x_{-i})$ does not depend on the intrinsic centrality η;

(ii) if $|\mathcal{N}_i^{-\infty}(x_{-i})| > d_i$ and $x_i \in \mathcal{B}_i(x_{-i})$, then

$$x_i \subseteq \mathcal{N}_i^{-d_i}(x_{-i}) \backslash \{i\}, \tag{20}$$

and, from every node j in x_i there exists a path $(r_0, r_1, \ldots, r_{l-1}, r_l)$ from node $r_0 = j$ to node $r_l = i$ in $\mathcal{G}(x)$ such that $\{r_k : \leq k < l\} \subseteq x_i$ and

$$\tau^i_j(x_{-i}) = \tau^i_{r_0}(x_{-i}) > \tau^i_{r_1}(x_{-i}) > \cdots > \tau^i_{r_{l-1}}(x_{-i}) > \tau^i_{r_l}(x_{-i}) = \tau^i_i(x_{-i}) = 0\,. \tag{21}$$

(iii) if $|\mathcal{N}_i^{-\infty}(x_{-i})| \leq d_i$, then

$$x_i \in \mathcal{B}_i(x_{-i}) \;\Leftrightarrow\; x_i \supseteq \mathcal{N}_i^{-d_i}(x_{-i}) \setminus \{i\}\,. \tag{22}$$

Proof (i) This is a direct consequence of Propositions 2 and 3(i).

(ii) Let x_i in $\mathcal{B}_i(x_{-i})$ be a best response action for node i. Then, Proposition 2 and Proposition 3(iii) imply that $x_i \subseteq \mathcal{N}_i^{-\infty}(x_{-i})$. Given any node j in x_i, suppose by contradiction that there exists no path $(j = r_0, r_1, \ldots, r_l = i)$ from j to i in $\mathcal{G}(x)$ such that $r_k \in x_i$ for all $0 \leq k < l$. Then, the set $C = \mathcal{V} \setminus (x_i \cup \{i, j\})$ would be a cut between j and i and Proposition 3(iv) would imply that

$$\max\{\tau^i_k(x_{-i}) : k \in x_i\} \geq \tau^j_j(x_{-i}) > \min\{\tau^i_k(x_{-i}) : k \in C\},$$

thus violating characterization (13). Hence, there must exist at least one path $(j = r_0, r_1, \dots, r_l = i)$ from j to i in $\mathcal{G}(x)$ such that $r_k \in x_i$ for $0 \le k < l$. For the shortest such path, we have $l \le d_i$, thus implying that $j \in \mathcal{N}_i^{-d_i}(x_{-i})$. By the arbitrariness of j in x_i, we then get that $x_i \subseteq \mathcal{N}_i^{-d_i}(x_{-i})$.

To complete the proof of point (ii), first observe that the last inequality in relation (21) is straightforward since $\tau^i_{r_{l-1}}(x_{-i}) \ge 1 > 0 = \tau^i_i(x_{-i}) = \tau^i_{r_l}(x_{-i})$. On the other hand, for $k = l-2, l-3, \dots, 0$, since $x_{r_k} = \mathcal{N}_{r_k}$ is a cut between r_k and i, Proposition 3(iv) implies that $\tau^i_{r_k}(x_{-i}) > \min\{\tau^i_h(x_{-i}) : h \in x_{r_k}\}$. As $x_i \in \mathcal{B}_i(x_{-i})$, there must exist r_{k+1} in $x_{-i} \cap x_{r_k}$ such that $\tau^i_{r_k}(x_{-i}) > \tau^i_{r_{k+1}}(x_{-i})$. In this way, we can recursively construct a path from j to i satisfying relation (21).

(iii) The claim follows from Propositions 2 and 3(iii), observing that $\mathcal{N}_i^{-\infty}(x_{-i}) = \mathcal{N}_i^{-d_i}(x_{-i})$. □

Notice that Theorem 1(ii) implies that, if $|\mathcal{N}_i^{-\infty}(x_{-i})| > d_i$, i.e., if node i is reachable by at least d_i other nodes, then in any best response action x_i in $\mathcal{B}_i(x_{-i})$ one out-link of node i is necessarily towards a node in its in-neighborhood, a second out-link is towards another node in its in-neighborhood or to a node that is in the in-neighborhood of the previous node, and so on.

An immediate consequence of Theorem 1(i) is the following.

Corollary 2. *In a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$, the set of Nash equilibria $\mathcal{X}^*$ and the subsets of the strict and of the recursive ones, respectively, $\mathcal{X}^\bullet$ and $\mathcal{X}^\circ$, are all independent from η.*

Remark 1. While Corollary 2 ensures their independence from the intrinsic centrality η, the sets $\mathcal{X}^\bullet$, $\mathcal{X}^\circ$, and $\mathcal{X}^*$ may depend on the discount factor β. On the other hand, the set of potential maximizing equilibria $\mathcal{X}^{\mathrm{Z}}$ may also depend on η.

4.2. Examples and Elementary Properties of Nash Equilibria In this subsection, for the sake of simplicity of notation, we often identify a configuration x with the corresponding graph $\mathcal{G}(x)$ and attribute game-theoretic properties directly to the latter (e.g., we will say that a certain graph is a Nash equilibrium).

We start by presenting a number of graphs that turn out to be Nash equilibria for the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ for every possible value of the discount factor β, as can be proved from the following symmetry result.

Corollary 3. *Consider a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ and a configuration x such that $\mathcal{G}(x) = (\mathcal{V}, \mathcal{E})$ is an undirected graph. If for every i in $\mathcal{V}$ and j and k in $\mathcal{N}_i(x)$, there exists an automorphism f of $\mathcal{G}(x)$ such that $f(i) = i$ and $f(j) = k$, then x is a strict Nash equilibrium for every value of β and η.*

Proof The assumption implies that $\tau^i_j(x_{-i}) = \tau^i_k(x_{-i})$ for every i in $\mathcal{V}$, j and k in $\mathcal{N}_i$. For every h in $\mathcal{V} \setminus \mathcal{N}_i(x)$, since $\mathcal{N}_i(x)$ is a cut between h and i, Proposition 3 implies that $\tau^i_h(x_{-i}) > \tau^i_j(x_{-i})$, thus proving that the unique best response of node i is indeed x_i. □

EXAMPLE 3. The complete bipartite graph $K_{l,m}$ is an undirected graph $(\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V} = \mathcal{I} \cup \mathcal{J}$, where $\mathcal{I} \cap \mathcal{J} = \emptyset$, $|\mathcal{I}| = l$, and $|\mathcal{J}| = m$, and link set $\mathcal{E} = \{(i, j), (j, i) : i \in \mathcal{I},\ j \in \mathcal{J}\}$. As a special case, when $l = 1$ and $m = n - 1$, we recover the star graph S_n displayed in Figure 1(b). For $l, m \geq 1$, $K_{l,m}$ satisfies the symmetry assumption of Corollary 3, so that it is a strict Nash equilibrium. □

EXAMPLE 4. Let R_n be the ring graph, as displayed in Figure 1(c). Clearly, R_n satisfies the symmetry assumption of Corollary 3, so that it is a strict Nash equilibrium. □

EXAMPLE 5. All platonic graphs, i.e., graphs that have one of the Platonic solids as their skeleton [43, pp. 263 and 266], satisfy the symmetry assumption of Corollary 3, hence they are all strict Nash equilibria. □

While all the examples above are undirected graphs, general strongly connected Nash equilibria may as well contain directed links, as it will become clear from the examples reported in Section 5.3. The following result provides a lower bound on the number of undirected links i.e., pairs of nodes $\{i, j\}$ such that both directed links (i, j) and (j, i) are links, and 3-cycles in any strongly connected Nash equilibrium graph.

COROLLARY 4. *Let x in $\mathcal{X}^*$ be a Nash equilibrium of a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$. Assume that $\mathcal{G}(x) = (\mathcal{V}, \mathcal{E})$ is strongly connected. Let $n = |\mathcal{V}|$ be the number of nodes, $d_* = \min\{d_i : i \in \mathcal{V}\}$ be the minimum out-degree, c_2, and c_3 be, respectively, the number of undirected links and of 3-cycles in $\mathcal{G}(x)$. Then,*

$$2c_2 \geq n, \qquad d_* c_3 + 2c_2 \geq n d_* . \tag{23}$$

Proof For a node i in $\mathcal{V}$, let $d^{\leftrightarrow}_i = |\{j \in \mathcal{V} : (i, j) \in \mathcal{E}, (j, i) \in \mathcal{V}\}| \leq d_i$ be the number of its outgoing links forming undirected links. Since $\mathcal{G}(x)$ is strongly connected, $|\mathcal{N}^{-\infty}_i(x)| = n > d_i$, so that Theorem 1(ii) implies that $d^{\leftrightarrow}_i \geq 1$. Thus, $2c_2 = \sum_i d^{\leftrightarrow}_i \geq n$, proving the first inequality in relations (23).

If $d^{\leftrightarrow}_i < d_i$, by Theorem 1(ii) there exist $j \neq k$ in $\mathcal{V}$ such that $(i, j), (j, i), (i, k), (k, j) \in \mathcal{E}$, while $(k, i) \notin \mathcal{E}$. Hence, $c_3 \geq \sum_i \min\{d_i - d^{\leftrightarrow}_i, 1\} \geq \sum_i (d_i - d^{\leftrightarrow}_i)/d_i \geq n - \sum_i d^{\leftrightarrow}_i / d_* = n - 2c_2/d_*$, yielding the second inequality in relations (23). □

A useful consequence of Proposition 3 is that the disjoint union of strongly connected graphs that are Nash equilibria remains a Nash equilibrium. Precisely, the following result holds true.

COROLLARY 5. *Consider two disjoint sets $\mathcal{V}_1$ and $\mathcal{V}_2$, centrality games $\Gamma(\mathcal{V}_h, \beta, \eta^{(h)}, \mathbf{d}^{(h)})$ for $h = 1, 2$ and for each of them a (strict) Nash equilibrium $x^{(h)}$. Consider now the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ where $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, $\eta_i = \eta_i^{(h)}$ and $\mathbf{d}_i = \mathbf{d}_i^{(h)}$ for all i in $\mathcal{V}_h$ and $h = 1, 2$. Then, the configuration x such that $x_i = x_i^{(h)}$ for all i in $\mathcal{V}_h$ and $h = 1, 2$ is a (strict) Nash equilibrium.*

Proof Notice that $\mathcal{G}(x)$ is simply the disjoint union of the two graphs $\mathcal{G}(x^{(h)})$ for $h = 1, 2$. Fix h in $\{1, 2\}$ and a node i in $\mathcal{V}_h$. Since, $\mathcal{N}_i^{-\infty}(x) = \mathcal{V}_h$ and $d_i^{(h)} \leq |\mathcal{G}(x^{(h)})| - 1$, it follows from Theorem 1(ii) that all best response sets for node i in the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ are subsets of $\mathcal{V}_h \setminus \{i\}$. Moreover, it follows from Proposition 3(ii) that the set $\{\tau_j^i(x) : j \in \mathcal{V}_h\}$ has the same order as the set $\{\tau_j^i(x^{(h)}) : j \in \mathcal{V}_h\}$. This implies that node i is currently playing a best response action in the union graph and such best response action is unique if it was unique for the centrality game $\Gamma(\mathcal{V}_h, \beta, \eta^{(h)}, \mathbf{d}^{(h)})$, thus proving the claim. □

EXAMPLE 6. Any disjoint union of complete graphs, star graphs, and ring graphs is a strict Nash equilibrium for every value of the discount factor β and of the intrinsic centrality η. □

4.3. Connectivity Structure of Nash Equilibria We now investigate in more generality the structure of the Nash equilibria of the centrality games. Theorem 1 has important consequences on the connectivity structure of Nash equilibria. We start with a preliminary result.

LEMMA 4. *Let x in $\mathcal{X}^*$ be a Nash equilibrium of a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ and let $\mathcal{G}_{\mathcal{K}}(x)$ be a connected component of $\mathcal{G}(x)$. Then:*
(i) if $d_i < |\mathcal{K}|$ for every i in $\mathcal{K}$, then $\mathcal{G}_{\mathcal{K}}(x)$ is a sink;
(ii) if $d_i \geq |\mathcal{K}|$ for some i in $\mathcal{K}$, then $\mathcal{G}_{\mathcal{K}}(x)$ is a source.

Proof Fix some i in $\mathcal{K}$. Since $\mathcal{G}_{\mathcal{K}}(x)$ is a connected component,

$$\mathcal{N}_i \cap \mathcal{N}_i^{-\infty} \subseteq \mathcal{K} \subseteq \mathcal{N}_i^{-\infty} \tag{24}$$

so that in particular $|\mathcal{K}| \leq |\mathcal{N}_i^{-\infty}|$.

(i) If $d_i < |\mathcal{K}|$, then $d_i < |\mathcal{N}_i^{-\infty}|$, so that Theorem 1(ii) implies that $\mathcal{N}_i \subseteq \mathcal{N}_i^{-d_i}$. This, by the double inclusion (24), yields $\mathcal{N}_i \subseteq \mathcal{K}$. If this is true for every i in $\mathcal{K}$, then $\mathcal{G}_{\mathcal{K}}(x)$ is necessarily a sink.

(ii) If $d_i \geq |\mathcal{K}|$, then $\mathcal{N}_i \nsubseteq \mathcal{K} \setminus \{i\}$, so that $\mathcal{G}_{\mathcal{K}}(x)$ cannot be a sink. Condition (24) now implies that $\mathcal{N}_i \nsubseteq \mathcal{N}_i^{-\infty}$, so that $d_i \geq |\mathcal{N}_i^{-\infty}|$ by Theorem 1(ii) and, consequently, $\mathcal{N}_i^{-\infty} \setminus \{i\} \subseteq \mathcal{N}_i$ by Theorem 1(iii).

Finally, using again relation (24), we obtain that $\mathcal{N}_i^{-\infty} = \mathcal{K}$. This implies that $\mathcal{G}_{\mathcal{K}}(x)$ is necessarily a source. □

Remark 2. In the special case of homogeneous out-degree profiles, i.e., $\mathbf{d} = d\mathbf{1}$, Lemma 4 implies that a connected component $\mathcal{G}_{\mathcal{K}}(x)$ that is a source is necessarily such that $d \geq |\mathcal{K}|$. This in turn implies that $\mathcal{G}_{\mathcal{K}}(x)$ is a $|\mathcal{K}|$-clique and every node within it has $d - |\mathcal{K}| + 1$ out-links towards other connected components that are sinks.

We conclude this section with the following result providing necessary conditions on the connectivity properties for a graph to be a Nash equilibrium of a centrality game. Before stating it, we remind the reader that, by convention, isolated components of a graph are classified as sinks but not as sources.

Theorem 2 (**Connectivity of Nash equilibria**). *Let x in $\mathcal{X}^*$ be a Nash equilibrium of a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$. Then,*

(i) every connected component of $\mathcal{G}(x)$ is either a sink or a source.

Moreover:

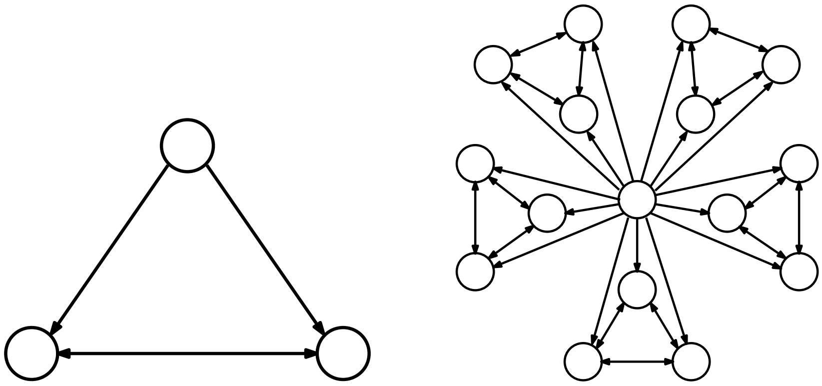
(ii) if x is recursive, then at most one of the connected components of $\mathcal{G}(x)$ is a source;

(iii) if x is strict and $\max_i d_i < n - 1$, then all the connected components of $\mathcal{G}(x)$ are isolated.

Proof (i) This follows directly from Lemma 4.

(ii) By contradiction, let $\mathcal{G}_{\mathcal{I}}(x)$ and $\mathcal{G}_{\mathcal{K}}(x)$ be two distinct connected components of $\mathcal{G}(x)$ that are sources. Observe that, since $\mathcal{G}_{\mathcal{I}}(x)$ is not a sink, there must exist some node i in $\mathcal{I}$ such that $x_i \not\subseteq \mathcal{I}$. Then, by point (i), there exist a connected component $\mathcal{G}_{\mathcal{J}}(x)$ that is a sink and node j in $x_i \cap \mathcal{J}$. In particular, $j \notin \mathcal{I} \cup \mathcal{K}$. It follows from Theorem 1(iii) that, for every k in $\mathcal{K}$, $y_i = (x_i \setminus \{j\}) \cup \{k\}$ is a best response action for player i. The graph $\mathcal{G}(y)$ possesses a connected component $\mathcal{G}_{\mathcal{K}}(y)$ that is neither a source nor a since. Thus, by item (i), y is not a Nash equilibrium, hence Lemma 2(iii) implies that y is not recursive. Since (x, y) is a length-1 best response path and y is not recursive, the Nash equilibrium x is not recursive.

(iii) By contradiction, let $\mathcal{G}_{\mathcal{I}}(x)$ be a connected component of $\mathcal{G}(x)$ that is a source. Since $\mathcal{G}_{\mathcal{I}}(x)$ is not a sink, there exists a node i in $\mathcal{I}$ such that $x_i \not\subseteq \mathcal{I}$. Moreover, since $\mathcal{G}_{\mathcal{I}}(x)$ is a source, we have that $\mathcal{N}_i^{-\infty}(x_{-i}) = \mathcal{I}$. Hence, $|\mathcal{I}| = |\mathcal{N}_i^{-\infty}(x_{-i})| \leq d_i$, for otherwise Theorem 1(ii) would imply that $x_i \subseteq \mathcal{I}$. Then, Theorem 1(iii) implies that every $y_i \subseteq \mathcal{V} \setminus \{i\}$ such that $|y_i| = d_i$ and $\mathcal{N}_i^{-\infty}(x_{-i}) \setminus \{i\} \subseteq y_i$ is a possible best response action for player i. There are $\binom{n-|\mathcal{I}|}{d_i+1-|\mathcal{I}|}$ such subsets. By assumption, $n > d_i + 1$, so that $\binom{n-|\mathcal{V}'|}{d_i+1-|\mathcal{I}|} > 1$, so that x cannot be a strict Nash equilibrium. □



(a) $n = 3$ and $d = (1, 1, 2)$. (b) $n = 16$ and $d = (3, 3, \ldots, 3, 15)$.

FIGURE 4. Graphs of Example 7 in two special cases.

REMARK 3. The assumption $\max_i d_i < n - 1$ in Theorem 2(iii) cannot be removed. Indeed, e.g., for $n = 3$ and $\mathbf{d} = (2, 1, 1)$, the configuration $x = (\{2, 3\}, \{3\}, \{2\})$ is a strict Nash equilibrium, yet $\mathcal{G}(x)$ admits a connected component $\mathcal{G}_{\{1\}}(x)$ that is a source (see Figure 4(a)).

4.4. Analysis of Potential Maximizing Equilibria In this subsection, we focus on the subset $\mathcal{X}^{\mathrm{Z}}$ of Nash equilibria where the potential $\Psi(x) = -\log Z(x)$ achieves its maximum value over the configuration space $\mathcal{X}$. Recall that, by Lemma 2(ii), the set of potential maximizers $\mathcal{X}^{\mathrm{Z}}$ is a nonempty subset of the set $\mathcal{X}^{\circ}$ of recursive Nash equilibria. Moreover, by Corollary 1, the stationary distribution of the noisy best response dynamics in a centrality game approaches a uniform distribution on $\mathcal{X}^{\mathrm{Z}}$ in the vanishing noise limit, motivating the interest in characterizing this specific subset of Nash equilibria. In particular, we characterize the structure of such potential maximizing equilibria for values of the discount factor β close to 1. Notice that the limit as β approaches 1 is especially relevant, as this is when the weight of the network structure on determining the PageRank centrality is maximized with respect to that of the intrinsic centrality.

For a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$, define the function $m : \mathcal{X} \to \mathbb{N}$ by setting $m(x)$ equal to the number of connected components of $\mathcal{G}(x)$ that are sinks. The key fact is the following result characterizing the asymptotic behavior of the function $Z(x)$ as β approaches 1.

PROPOSITION 4. *In a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$, for every configuration x in $\mathcal{X}$,*

$$Z(x) \asymp (1 - \beta)^{m(x)-1}, \tag{25}$$

as $\beta \to 1$.

Proof For every configuration x in $\mathcal{X}$, player i in $\mathcal{V}$, and spanning directed rooted tree $\mathcal{T} = (\mathcal{V}, \mathcal{E}_\mathcal{T})$ in $\mathbb{T}_i$, let

$$w_\mathcal{T}(x_{-i}) = \prod_{(j,k)\in\mathcal{E}_\mathcal{T}} \left(\beta \mathbb{1}_{x_j}(k) + (1-\beta)\eta_k\right),$$

be the weight of $\mathcal{T}$ in x. Let also $\mathcal{E}^x_\mathcal{T} = \mathcal{E}_\mathcal{T} \cap \mathcal{E}(x)$ be the set of links of $\mathcal{T}$ that are also links of $\mathcal{G}(x)$, and let $\overline{\mathcal{E}}^x_\mathcal{T} = \mathcal{E}_\mathcal{T} \setminus \mathcal{E}(x)$ be the set of links of $\mathcal{T}$ that are not links of $\mathcal{G}(x)$. Then, define

$$B_\mathcal{T}(\beta) = \prod_{(j,k)\in\mathcal{E}^x_\mathcal{T}} (1-(1-\beta)(1-\eta_k)), \qquad C_\mathcal{T} = \prod_{(j,k)\in\overline{\mathcal{E}}^x_\mathcal{T}} \eta_k,$$

and observe that the weight of $\mathcal{T}$ in x satisfies

$$w_\mathcal{T}(x_{-i}) = \prod_{(j,k)\in\mathcal{E}^x_\mathcal{T}} (1-(1-\beta)(1-\eta_k)) \prod_{(j,k)\in\overline{\mathcal{E}}^x_\mathcal{T}} (1-\beta)\eta_k = B_\mathcal{T}(\beta) C_\mathcal{T} (1-\beta)^{|\overline{\mathcal{E}}^x_\mathcal{T}|}.$$

Since $B_\mathcal{T}(\beta) \xrightarrow{\beta\to 1} 1$, it follows that

$$w_\mathcal{T}(x_{-i}) \sim C_\mathcal{T}(1-\beta)^{|\overline{\mathcal{E}}^x_\mathcal{T}|}, \tag{26}$$

as $\beta \to 1$. Notice that for every nonempty $\mathcal{K} \subseteq \mathcal{V} \setminus \{i\}$, there must exist at least one link (k,h) in $\mathcal{E}_\mathcal{T}$ from a node k in $\mathcal{K}$ to a node h in $\mathcal{V} \setminus \mathcal{K}$. If $\mathcal{G}_\mathcal{K}(x)$ is a connected component of $\mathcal{G}(x)$ that is a sink, such link is necessarily in $\overline{\mathcal{E}}^x_\mathcal{T}$. This implies that for every connected component of $\mathcal{G}_\mathcal{K}(x)$ that is a sink and that does not contain i, there exists a link (k,h) in $\overline{\mathcal{E}}^x_\mathcal{T}$ from some k in $\mathcal{K}$. Hence, in particular,

$$|\overline{\mathcal{E}}^x_\mathcal{T}| \geq m(x) - 1. \tag{27}$$

We now show that there indeed exist spanning directed rooted trees with exactly $m(x)-1$ links in $\overline{\mathcal{E}}^x_\mathcal{T}$. Assume that $\mathcal{G}_{\mathcal{K}_1}(x), \ldots, \mathcal{G}_{\mathcal{K}_{m(x)}}(x)$ are all the connected components of $\mathcal{G}(x)$ that are sinks and choose a node i in $\mathcal{K}_1$. Define $\overline{\mathcal{K}}_1 = \mathcal{N}^{-\infty}_i(x_{-i})$ and let $\mathcal{T}_1 = (\overline{\mathcal{K}}_1, \mathcal{E}_{\mathcal{T}_1})$ be a spanning directed tree in $\mathcal{G}_{\overline{\mathcal{K}}_1}(x)$ rooted in i. We now fix, arbitrarily, nodes i_m in $\mathcal{K}_m$ for $m = 2, \ldots, m(x)$ and, recursively, we construct spanning directed trees $\mathcal{T}_m = (\overline{\mathcal{K}}_m, \mathcal{E}_{\mathcal{T}_m})$ in $\mathcal{G}_{\overline{\mathcal{K}}_m}(x)$ rooted in i_m where

$$\overline{\mathcal{K}}_m = \mathcal{N}^{-\infty}_{i_m}(x_{-i_m}) \setminus \bigcup_{m'<m} \overline{\mathcal{K}}_{m'}.$$

Then, $\mathcal{T} = (\mathcal{V}, \mathcal{E}_\mathcal{T})$ where

$$\mathcal{E}_\mathcal{T} = \bigcup_{m=1}^{m(x)} \mathcal{E}_{\mathcal{T}_m} \cup \{(i_m, i_1) \mid m = 2, \ldots, m(x)\})$$

is a spanning directed rooted tree in $\mathbb{T}_i$ with $|\overline{\mathcal{E}}^x_\mathcal{T}| = m(x) - 1$. By the way $Z(x)$ is defined, this together with relations (26) and (27) proves the result. □

Proposition 4 allows us to prove the following result, stating that, for large enough values of the discount factor β, all configurations maximizing the potential of a centrality game are maximizers of the positive-integer valued function $m(x)$.

THEOREM 3. *In a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$, there exists $\overline{\beta} < 1$ such that*

$$\mathcal{X}^{\mathrm{Z}} \subseteq \underset{x \in \mathcal{X}}{\arg\max}\, m(x)\,, \tag{28}$$

for every β in $(\overline{\beta}, 1)$.

Proof It follows from Proposition 4 and the fact that $\mathcal{X}$ is a finite set that there exist two positive numbers $0 < c_1 < c_2$ such that

$$c_1 \leq \frac{Z(x)}{(1-\beta)^{m(x)-1}} \leq c_2 \tag{29}$$

for every x in $\mathcal{X}$ and β in $(0,1)$. Put $\overline{\beta} = 1 - c_1/c_2$ and take $\beta > \overline{\beta}$. Let x^* in $\mathcal{X}^{\mathrm{Z}}$ be a potential maximizer and assume by contradiction that $x^* \notin \arg\max\{m(x) : x \in \mathcal{X}\}$. This implies that there exists x^{**} in $\mathcal{X}$ such that $m(x^{**}) \geq m(x^*) + 1$. From relation (29) we then obtain

$$Z(x^*) \geq c_1(1-\beta)^{m(x^*)-1} \geq c_1(1-\beta)^{m(x^{**})-2} \geq \frac{c_1}{c_2(1-\beta)} Z(x^{**}) \geq \frac{c_1}{c_2(1-\beta)} Z(x^*)$$

This yields $\beta \leq 1 - c_1/c_2 = \overline{\beta}$ contradicting the assumption $\beta > \overline{\beta}$. Hence, the result follows. □

We now analyze how the maximum value $\max\{m(x) : x \in \mathcal{X}\}$ of the number of connected components of $\mathcal{G}(x)$ that are sinks, among all possible configurations x, is determined by the out-degree profile $\mathbf{d}$ of the centrality game. We start by relabeling nodes in such a way that the out-degree is non-decreasing, i.e.,

$$d_1 \leq d_2 \leq \cdots \leq d_n\,. \tag{30}$$

We now partition the node set into groups in such a way that the degree of every node is strictly less than the cardinality of the group it belongs to. This will ensure that it is possible for every node to choose its out-neighbors within the group, thus creating a sink connected component. In this way, we will be able to determine the maximum number of sink connected components that a configuration may achieve. To this aim, we define $\nu_0 = 0$ and introduce the sequences of positive integers $\nu_1, \ldots, \nu_{m^*(\mathbf{d})}$ and $n_1, \ldots, n_{m^*(\mathbf{d})}$ iteratively by the recursion

$$n_k = \min\left\{1 \leq j \leq n - \nu_{k-1} \mid d_{j+\nu_{k-1}} \leq j-1\right\}, \qquad \nu_k = \sum_{1 \leq h \leq k} n_h\,, \tag{31}$$

for $k = 1, \ldots, m^*(\mathbf{d})$, where $m^*(\mathbf{d})$ is the largest index k for which the set in the right-hand side of the first equation in (31) is nonempty. Then, put $\nu = \nu_{m^*(\mathbf{d})}$ and observe that $\nu \leq n$. Then, the following result holds true.

LEMMA 5. *In a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$, we have that*

$$\max\{m(x) : x \in \mathcal{X}\} = m^*(\mathbf{d}) .$$

Proof We prove the result by first showing that there exists a configuration x^* in $\mathcal{X}$ for which $\mathcal{G}(x^*)$ has exactly $m^*(\mathbf{d})$ connected components that are all sinks and have order, respectively, $n_1, \dots, n_{m^*(\mathbf{d})}$. We will then show that $m(x) \le m^*(\mathbf{d})$ for every configuration x in $\mathcal{X}$.

Let n_k and ν_k be defined as in relations (31). For $1 \le k \le m^*(\mathbf{d})$, consider the set $\mathcal{S}_k = \{\nu_{k-1} + 1, \dots, \nu_k\}$. Notice that, by construction, we have that

$$d_i \le d_{\nu_k} \le n_k - 1 = |\mathcal{S}_k| - 1 , \qquad \forall i \in \mathcal{S}_k .$$

Notice also that, if $n_k > 2$, then

$$d_i \ge d_{\nu_{k-1}+2} > 2 - 1 = 1 , \qquad \forall i \in \mathcal{S}_k \setminus \{\nu_{k-1} + 1\} .$$

This implies that we can find a configuration x^* in $\mathcal{X}$ such that, for every k, $x_i^* \subseteq \mathcal{S}_k$ for every i in $\mathcal{S}_k$ and $\mathcal{G}_{\mathcal{S}_k}$ is connected. It is sufficient for instance to impose that

$$x^*_{\nu_{k-1}+1} \supseteq \{\nu_{k-1} + 2\}, \quad x^*_{\nu_k} \supseteq \{\nu_k - 1\}, \quad x^*_{\nu_{k-1}+j} \supseteq \{\nu_{k-1} + j - 1, \nu_{k-1} + j + 1\}, \quad \forall 1 < j < n_k .$$

Therefore, $\mathcal{G}_{\mathcal{S}_k}(x^*)$ is a connected component that is a sink for every $k = 1, \dots, m^*(\mathbf{d})$. The set $\mathcal{R} = \{\nu + 1, \dots, n\}$ of remaining nodes with the highest degrees cannot contain any further connected component that is a sink: indeed, if $\mathcal{S} \subseteq \mathcal{R}$ is nonempty and $s = \max \mathcal{S}$, we have that $d_s > s - \nu - 1 \ge |\mathcal{S}| - 1$ so that $\mathcal{G}_{\mathcal{S}}(x^*)$ cannot be a sink.

Consider now any configuration x in $\mathcal{X}$ and let $\mathcal{G}_{\mathcal{S}_1}(x), \dots, \mathcal{G}_{\mathcal{S}_{m(x)}}(x)$ be its connected components that are sinks, having order $\overline{n}_k = |\mathcal{S}_k|$ for $k = 1, \dots, m(x)$. For $0 \le k \le m(x)$, put $\overline{\nu}_k = \sum_{0 \le h \le k} \overline{n}_h$ and $\mu_k = \max \mathcal{S}_k$. Assume without loss of generality that $\mu_1 < \mu_2 < \dots < \mu_{m(x)}$. Assume by contradiction that $m(x) > m^*(\mathbf{d})$. Then, in particular

$$d_{\mu_k} < \overline{n}_k , \qquad 1 \le k \le m^*(\mathbf{d}) + 1 . \tag{32}$$

We now show by induction on k that

$$\overline{\nu}_k \ge \nu_k , \tag{33}$$

for every $0 \le k \le m^*(\mathbf{d})$. As $\nu_0 = \overline{\nu}_0 = 0$, the inequality (33) is automatically verified for $k = 0$. On the other hand, assume that inequality (33) holds true for some $0 \le k < m^*(\mathbf{d})$. Then,

$$\mu_{k+1} \ge |\mathcal{S}_1 \cup \dots \cup \mathcal{S}_{k+1}| = \overline{\nu}_{k+1} = \overline{\nu}_k + \overline{n}_{k+1} \ge \nu_k + \overline{n}_{k+1} . \tag{34}$$

If $\mu_{k+1} < \nu_{k+1}$, then relations (31), (34), and (32) lead to the contradiction

$$d_{\mu_{k+1}} \geq \mu_{k+1} - \nu_k \geq \overline{n}_{k+1} > d_{\mu_{k+1}} .$$

Then, necessarily $\mu_{k+1} \geq \nu_{k+1}$, so that relations (32), (30), and (31) imply that

$$\overline{n}_{k+1} \geq d_{\mu_{k+1}} - 1 \geq d_{\nu_{k+1}} - 1 = n_{k+1} .$$

Together with inequality (33), the above implies that

$$\overline{\nu}_{k+1} = \overline{n}_{k+1} + \overline{\nu}_k \geq n_{k+1} + \nu_k = \nu_{k+1} .$$

We have thus proved by induction that inequality (33) holds true for every $0 \leq k \leq m^*(\mathbf{d})$. Notice that relations (31) imply that $d_i > i - \nu_{m^*(\mathbf{d})} - 1$, for every $i > \nu_{m^*(\mathbf{d})}$. If we apply this to $i = \mu_{m^*(\mathbf{d})+1}$ and use relation (34) with $k = m^*(\mathbf{d})$ and relation (32) with $k = m^*(\mathbf{d}) + 1$, we obtain the contradiction

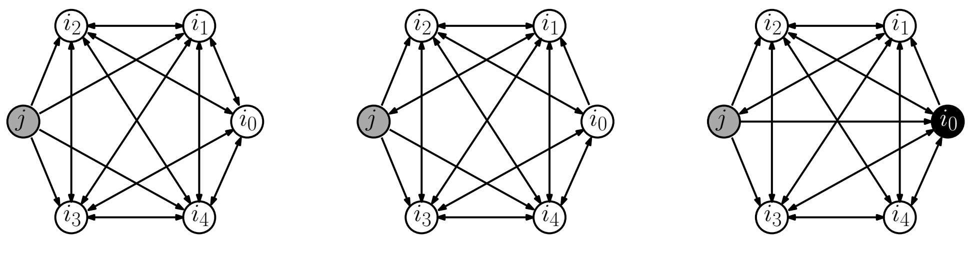
$$d_{\mu_{m^*(\mathbf{d})+1}} > \mu_{m^*(\mathbf{d})+1} - \nu_{m^*(\mathbf{d})} - 1 \geq \overline{n}_{m^*(\mathbf{d})+1} - 1 \geq d_{\mu_{m^*(\mathbf{d})+1}} .$$

Hence, $\mathcal{G}_{\mathcal{S}_k}(x)$ cannot be a sink component for every $1 \leq k \leq m^*(\mathbf{d}) + 1$. This implies that $m(x) \leq m^*(\mathbf{d})$. □

The characterization of the index $m^*(\mathbf{d})$ provided by Lemma 5 allows us to perform explicit computations, as in the example below.

Example 7. Let n and d be positive integers such that $n - 1$ is a multiple of $d + 1$. Consider an out-degree-profile such that $n - 1$ nodes have the same out-degree $d_i = d$ for $i = 1, 2, \ldots, n - 1$, and a hub node n has out-degree $d_n = n - 1$. From relations (31) we immediately obtain that $n_k = d + 1$ for every $1 \leq k \leq m^*(\mathbf{d})$ where $m^*(\mathbf{d}) = \frac{n-1}{d+1}$. Since there cannot be sink components with less than $d + 1$ elements, we conclude that $m^*(\mathbf{d})$ is achieved by all those configurations x in $\mathcal{X}$ such that $\mathcal{G}(x)$ consists of the union of exactly $m^*(\mathbf{d})$ $(d + 1)$-cliques plus the hub node pointing to all other nodes (see Figure 4 for an illustration in the special cases $n = 3$ and $d = 1$, and $n = 16$ and $d = 2$, respectively). It then follows from Theorem 3 that for values of the discount factor β sufficiently close to 1, all potential maximizers x in $\mathcal{X}^{\mathrm{Z}}$ are of this form. □

5. Equilibrium Analysis for Homogeneous Out-Degree Profiles In this section, we refine the equilibrium analysis in centrality games in the special case of homogeneous out-degree profiles. First, we analyze the structure of the potential maximizers x^* in $\mathcal{X}^{\mathrm{Z}}$ when $\mathbf{d} = d \cdot \mathbf{1}$ for arbitrary positive integer d. Then, we focus on two special cases of homogeneous out-degree profiles: $\mathbf{d} = \mathbf{1}$ and $\mathbf{d} = 2 \cdot \mathbf{1}$, respectively. For these cases, we are able to fully characterize the sets of Nash equilibria $\mathcal{X}^*$, recursive Nash equilibria $\mathcal{X}^\circ$, and strict Nash equilibria $\mathcal{X}^\bullet$.



(a) $\mathcal{G}(x)$ (b) $\mathcal{G}(y)$ (c) $\mathcal{G}(z)$

FIGURE 5. Graphs associated to the best response path (x, y, z) in the proof of Lemma 6 in the special case $d = 4$. Configuration z is not a Nash equilibrium since node i_0 strictly prefers linking to j instead of i_1.

5.1. Potential Maximizing Equilibria for Homogeneous Out-Degree Profiles We start with the following result, proving that certain configurations can never be recursive.

LEMMA 6. *Let x in $\mathcal{X}^*$ be a Nash equilibrium of a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$, with $d > 1$. If $\mathcal{G}(x)$ contains a $(d+1)$-clique $\mathcal{G}_{\mathcal{K}}(x)$ that is a sink component and a source node j in $\mathcal{V} \setminus \mathcal{K}$ with $d_j = d$ and $x_j \subseteq \mathcal{K}$, then x is not recursive.*

Proof We are going to construct a length-2 best response path from x to a configuration z in $\mathcal{X}$ that is not a Nash equilibrium (see Figure 5 for an illustration in the special case $d = 4$). By Lemma 2(iii)-(iv), this will imply that x is not recursive. We start by labeling the nodes in the clique as $\mathcal{K} = \{i_0, i_1, \dots, i_d\}$ in such a way that $x_j = \{i_1, \dots, i_d\}$. We then consider a configuration y such that $y_{-i_1} = x_{-i_1}$ and $y_{i_1} = (x_{i_1} \setminus \{i_0\}) \cup \{j\} = \{i_2, \dots, i_d, j\}$. Observe that since $x_{i_1} \in \mathcal{B}_{i_1}(x_{-i_1})$ (because x is a Nash equilibrium) and $x_j = \{i_1, \dots, i_d\} = x_{i_0}$, Proposition 3(v) and Proposition 2 imply that also $y_{i_1} \in \mathcal{B}_{i_1}(x_{-i_1})$. Let then z be a configuration such that $z_{-j} = y_{-j}$ and $z_j = (y_j \setminus \{i_d\}) \cup \{i_0\}$ (notice that $z \neq y$ since $d > 1$). Observe that, since $y_{i_0} \setminus \{i_d\} = \{i_1, \dots, i_{d-1}\} = y_{i_d} \setminus \{i_0\}$ and $y_j \in \mathcal{B}_j(y_{-j})$, Proposition 3(v) implies that also $z_j \in \mathcal{B}_j(y_{-j})$. This proves that (x, y, z) is a best response path. To prove that z is not a Nash equilibrium we show that $z_{i_0} \notin \mathcal{B}_{i_0}(z_{-i_0})$. Indeed, since $z_{i_1} = \{j\} \cup \{i_2, \dots, i_d\}$ and $z_j = \{i_0, i_1, \dots, i_{d-1}\}$, from relations (4) we get

$$\tau_{i_1}^{i_0}(z_{-i_0}) - \tau_j^{i_0}(z_{-i_0}) = \frac{\beta}{d} \sum_{h \in z_{i_1}} \tau_h^{i_0}(z_{-i_0}) - \frac{\beta}{d} \sum_{h \in z_j} \tau_h^{i_0}(z_{-i_0}) = \frac{\beta}{d} \left(\tau_j^{i_0}(z_{-i_0}) - \tau_{i_1}^{i_0}(z_{-i_0}) \right) + \frac{\beta}{d} \tau_{i_d}^{i_0}(z_{-i_0}) \,.$$

This yields

$$\tau_{i_1}^{i_0}(z_{-i_0}) - \tau_j^{i_0}(z_{-i_0}) = \frac{\beta}{d - \beta} \tau_{i_d}^{i_0}(z_{-i_0}) > 0 \,.$$

It then follows from Proposition 2 that $z_{i_0} \notin \mathcal{B}_{i_0}(z_{-i_0})$, so that z is not a Nash equilibrium. By Lemma 2(iii) that z is not recursive. Finally, Lemma 2(iv) implies that x is not recursive. □

THEOREM 4. *Consider a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with homogeneous out-degree profile $\mathbf{d} = d \cdot \mathbf{1}$ with $1 \leq d < n = |\mathcal{V}|$. Then, there exists $\overline{\beta} < 1$ such that, for every $\overline{\beta} < \beta < 1$ and x^* in $\mathcal{X}^{\mathrm{Z}}$:*

(i) if n is a multiple of $d+1$, then $\mathcal{G}(x^*)$ is the union of $n/(d+1)$ isolated $(d+1)$-cliques;

(ii) if $d > 1$ and $n > (d+1)d$, then $\mathcal{G}(x^*)$ is the union of $\lfloor n/(d+1) \rfloor$ isolated connected components.

Proof From Lemma 2(ii), we know that x^* is a recursive Nash equilibrium. Theorem 2(i)-(ii) then guarantees that all connected components of $\mathcal{G}(x^*)$ are sinks, except for possibly one source. From relations (31) and Lemma 5 we obtain that $m^*(\mathbf{d}) = \lfloor \frac{n}{d+1} \rfloor$ and $n_k = d+1$ for every $1 \leq k \leq \lfloor \frac{n}{d+1} \rfloor$. Write $n = \lfloor \frac{n}{d+1} \rfloor (d+1) + r$ for some $r < d+1$. Since $m(x^*) = \lfloor \frac{n}{d+1} \rfloor$, considering there cannot be sink components with less than $d+1$ elements, $\mathcal{G}(x^*)$ must consist of $\lfloor \frac{n}{d+1} \rfloor$ sink components each of size at least $d+1$ and possibly of a single source component with at most r vertices.

If n is a multiple of $d+1$, then $r = 0$ and thus $\mathcal{G}(x^*)$ consists of the union of exactly $\lfloor \frac{n}{d+1} \rfloor$ isolated $(d+1)$-cliques.

Suppose now that $n > (d+1)d$ and $d > 1$. Notice that $\lfloor \frac{n}{d+1} \rfloor \geq d \geq r$. If, by contradiction, a source component was present, it would follow that there would necessarily exist a sink component $\mathcal{G}_{\mathcal{K}}(x^*)$ with exactly $d+1$ nodes. Let j be a node in the source component. By Theorem 1(iii), there exists a length-1 best response path, consisting in the rewiring of all the out-links of node j towards $\mathcal{K}$, leading to a configuration x^{**} where $x_j^{**} \subseteq \mathcal{K}$. Since we are assuming that $d > 1$, Lemma 6 implies that x^{**} is not recursive. By Lemma 2(iii), also $x^* \notin \mathcal{X}^{\circ}$. Then, Lemma 2(ii) implies that $x^* \notin \mathcal{X}^{\mathrm{Z}}$, leading to a contradiction. This implies that $\mathcal{G}(x^*)$ is the union of $\lfloor \frac{n}{d+1} \rfloor$ isolated components. □

REMARK 4. It is worth noting that, if n is not a multiple of $d+1$ and either $n \leq d(d+1)$ or $d = 1$, then we cannot rule out the possibility that, for some potential maximizer x^* in $\mathcal{X}^{\mathrm{Z}}$, the associated graph $\mathcal{G}(x^*)$ contains a connected component that is a source. E.g., for $n = 5$ and $d = 2$, a configuration with associated graph as the one displayed in Figure 6(b) is recursive and trivially maximizes $m(x)$. On the other hand, for the case when $d = 1$, Theorem 5 below yields a complete classification of the Nash equilibria showing in particular that, whenever n is odd, recursive Nash equilibria always contain a singleton source component.

5.2. The Case $\mathbf{d} = \mathbf{1}$

In this subsection, we consider centrality games with homogeneous out-degree profiles $\mathbf{d} = \mathbf{1}$, i.e., when every node has a single out-link. In this case, the best response presented in Theorem 1 takes the following special form.

PROPOSITION 5. *Consider a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{1})$. Then,*

$$\mathcal{B}_i(x_{-i}) = \begin{cases} \mathcal{N}_i^{-1}(x_{-i}) \setminus \{i\} & \text{if } \mathcal{N}_i^{-1}(x_{-i}) \neq \{i\} \\ \mathcal{V} \setminus \{i\} & \text{if } \mathcal{N}_i^{-1}(x_{-i}) = \{i\} \end{cases} \tag{35}$$

for every player i in $\mathcal{V}$ and action profile x_{-i} in $\mathcal{X}_{-i}$.

Proof Since $\mathbf{d} = \mathbf{1}$, for every j in $\mathcal{N}_i^{-1}(x_{-i}) \setminus \{i\}$ we have that $x_j = \{i\}$, so that relations (4) imply that $\tau_j^i = 1 + (1-\beta) \sum_{v \in \mathcal{V}} \eta_v \tau_v^i$. Hence, if $\mathcal{N}_i^{-1}(x_{-i}) \neq \{i\}$, then Theorem 1(ii) implies the result. If $\mathcal{N}_i^{-1}(x_{-i}) = \{i\}$, then the result follows directly from Theorem 1(iii). □

We now introduce the family of graphs $\mathcal{K}_2^r$ obtained by adding to a finite set of disjoint 2-cliques, r extra nodes, each of which having exactly one out-link pointing towards an arbitrary node belonging to any of the 2-cliques. The following result provides a complete characterization of the sets of Nash equilibria, recursive Nash equilibra, and strict Nash equilibria, respectively, for centrality games with homogeneous out-degree profile $\mathbf{d} = \mathbf{1}$.

THEOREM 5 (**NASH EQUILIBRIA FOR $\mathbf{d} = \mathbf{1}$**). *For a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{1})$ and a configuration x:*

(i) x is a Nash equilibrium if and only if $\mathcal{G}(x) \in \bigcup_{r \geq 0} \mathcal{K}_2^r$;

(ii) x is a strict Nash equilibrium if and only if $\mathcal{G}(x) \in \mathcal{K}_2^0$;

(iii) x is a recursive Nash equilibrium if and only if $\mathcal{G}(x) \in \mathcal{K}_2^0 \cup \mathcal{K}_2^1$.

Proof (i) If x in $\mathcal{X}$ is such that $\mathcal{G}(x) \in \mathcal{K}_2^r$, then both nodes that belong to a 2-clique as well nodes that link to a 2-clique are playing a best response action according to relations (35), so that x is a Nash equilibrium. Conversely, if x is a Nash equilibrium, then for every node i in $\mathcal{V}$, Proposition 6 guarantees that either there is another node j such that both $(j,i) \in \mathcal{E}$ and $(i,j) \in \mathcal{E}$, or $\mathcal{N}_i^{-1} = \{i\}$. In the former case, $\mathcal{G}_{\{i,j\}}(x)$ is a connected component that is a sink. In the latter case, $\mathcal{G}_{\{i\}}(x)$ is a connected component that is a source. Hence, $\mathcal{G}(x)$ is necessarily of type $\mathcal{K}_2^r$.

(ii) It follows from Theorem 2(iii) that if x is a strict Nash equilibrium, then $\mathcal{G}(x) \in \mathcal{K}_2^0$. On the other hand, if x in $\mathcal{X}$ is such that $\mathcal{G}(x) \in \mathcal{K}_2^0$, then it follows from relations (35) that every node has just one incoming link and is thus playing its unique best response action. This implies that $x \in \mathcal{X}^\bullet$.

(iii) It follows from Theorem 2(ii) that every recursive Nash equilibrium x is such that all connected components of $\mathcal{G}(x)$ are sinks except for possibly one source. Since $\mathbf{d} = \mathbf{1}$, necessarily the sinks have order 2 and the source has order 1, so that $\mathcal{G}(x) \in \mathcal{K}_2^0 \cup \mathcal{K}_2^1$. Since $\mathcal{G}(x) \in \mathcal{K}_2^0$ implies that $x \in \mathcal{X}^\bullet$ by point (ii), it remains to show that $\mathcal{G}(x) \in \mathcal{K}_2^1$ implies that $x \in \mathcal{X}^\circ$. For that, let x in $\mathcal{X}$ be such that $\mathcal{G}(x) \in \mathcal{K}_2^1$. Denote by s the unique source node and let i, j be the nodes in the 2-clique such that $(s, i) \in \mathcal{E}(x)$. From relations (35), all nodes in $\mathcal{V} \setminus \{s, i\}$ are currently playing their unique best response and no transition is thus possible. Also, from relations (35), we deduce that $\mathcal{B}_s(x) = \mathcal{V} \setminus \{s\}$ and $\mathcal{B}_i(x) = \{j, s\}$. Therefore, every best response path starting from x makes a first step y such that either $y_{-s} = x_{-s}$ and $y_s \in \mathcal{V} \setminus \{i, s\}$, or $y_{-i} = x_{-i}$ and $y_i = s$. In both cases $\mathcal{G}(y) \in \mathcal{K}_2^1$. We now apply Lemma 2(v) to the set of configurations x such that $\mathcal{G}(x) \in \mathcal{K}_2^1$ and conclude that they are all recursive. □

REMARK 5. Notice that, if the number of players n is even, then Theorem 5 implies that the set of recursive Nash equilibria of the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{1})$ coincides with the set of its strict Nash equilibra as they are both given by $\mathcal{K}_2^0$. Instead, if n is odd, the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{1})$ admits no strict Nash equilibria, while the set of recursive Nash equilibra coincide with $\mathcal{K}_2^1$.

5.3. The Case $\mathbf{d} = 2 \cdot \mathbf{1}$

In this subsection, we focus on the special case of the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with homogeneous out-degree profile $\mathbf{d} = 2 \cdot \mathbf{1}$, i.e., when every node has to place exactly two out-links. We shall first present a full classification of recursive and strict Nash equilibria and then provide a complete classification of the Nash equilibria of the centrality game.

We start with following result.

PROPOSITION 6. *Consider the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with $\mathbf{d} = 2 \cdot \mathbf{1}$. Then,*

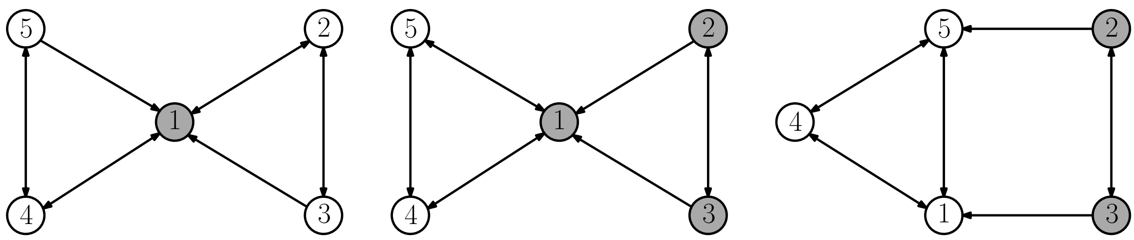
$$\mathcal{B}_i(x_{-i}) = \begin{cases} \{\{j,k\} \colon j, k \in \mathcal{V},\ i \neq j \neq k \neq i\} & \text{if } \mathcal{N}_i^{-\infty}(x_{-i}) = \{i\} \\ \{\{j,k\} \colon k \in \mathcal{V},\ i \neq k \neq j\} & \text{if } \mathcal{N}_i^{-\infty}(x_{-i}) = \{i, j\}, \end{cases} \tag{36}$$

for every player i in $\mathcal{V}$ and strategy profile x_{-i} in $\mathcal{X}_{-i}$. Moreover, if $|\mathcal{N}_i^{-\infty}(x_{-i})| > 2$, then

$$\mathcal{B}_i(x_{-i}) \subseteq \left\{\{j,k\} \colon j,k \in \mathcal{N}_i^{-1}(x_{-i}),\ i \neq j \neq k \neq i\right\} \cup \left\{\{j,k\} \colon j \in \mathcal{N}_i^{-1}(x_{-i}),\ k \in \mathcal{N}_j^{-1}(x_{-j}),\ i \neq j \neq k \neq i\right\}. \tag{37}$$

Proof Relations (36) follow from Theorem 1(iii). On the other hand, relation (37) follows from Theorem 1(ii): if node i is reachable in configuration x by at least two nodes besides itself, then either its best response is a pair of nodes $j \neq k$ that both point directly towards it, or it consists of a node j that points directly towards it and of a node k that points directly towards j. □

As discussed in Example 4, every ring graph R_n is a Nash equilibrium of the centrality game $\Gamma(\mathcal{V}, \beta, \eta, 2 \cdot \mathbf{1})$. Notice that ring graphs are the only connected undirected graphs where all nodes have degree 2. A remarkable fact is that there exists a recursive strongly connected Nash equilibrium



(a) B_5 (b) B_5' (c) B_5''

FIGURE 6. The Butterfly graphs. Grey nodes have more than one best response action.

that is not undirected. This is displayed in Figure 6 (a) and will be referred to as the *butterfly graph* B_5. Figures 6 (b) and 6 (c) display two more graphs that turn out to be best response evolutions of B_5, as for the result below.

LEMMA 7. *Consider the centrality game* $\Gamma(\mathcal{V}, \beta, \eta, 2 \cdot \mathbf{1})$ *and let* x *in* $\mathcal{X}$ *be such that* $\mathcal{G}(x) = B_5$. *Then:*

(i) x is a non-strict Nash equilibrium;

(ii) x is a recursive Nash equilibrium and the configurations reachable from x by a best response path are exclusively configurations y such that $\mathcal{G}(y)$ is isomorphic to some of the graphs B_5, B_5', and B_5'' of Figure 6.

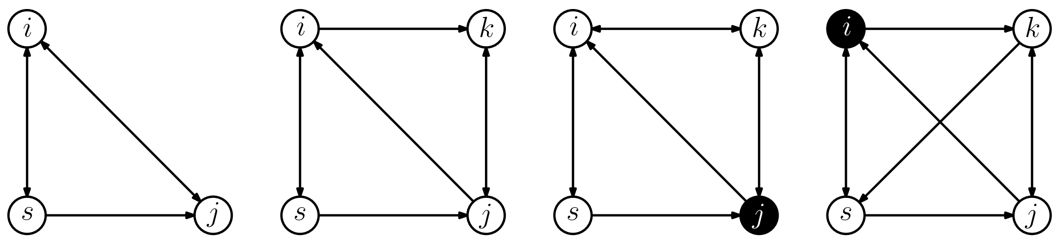
Proof (i) We analyze the best response set of every node in configuration x, starting with node 1. Proposition 3(v) applied to the two triples $\{1,4,5\}$ and $\{1,2,3\}$ implies that $\tau_4^1(x) = \tau_5^1(x)$ and $\tau_2^1(x) = \tau_3^1(x)$. By symmetry, we then conclude that all four hitting times are equal to each other. This implies that every pair of nodes is a best response for node 1 in configuration x. Moving to node 4, we can see that it has two in-neighbors: node 4 and node 1. From system (4) and Proposition 3(iv) (observing that $\{1\}$ is a cut set between node 2 and node 4) we obtain that $\tau_1^4(x) - \tau_5^4(x) = \frac{\beta}{2}(\tau_2^4(x) - \tau_1^4(x)) > 0$. Therefore, by Proposition 2, the unique best response for node 4 is $\{5, 1\}$, namely the action currently played. Finally, node 5 has only one in-neighbor, node 4, and node 4 has only one in-neighbor different from node 5, that is node 1. Therefore, from Theorem 1(ii), we deduce that the unique best response of node 5 is $\{4, 1\}$. By symmetry, also nodes 2 and 3 are playing their unique best response. This says that x is a Nash equilibrium. It is not strict because the best response of node 1 contains six different possible pairs.

(ii) We first notice that four out of the six best response actions of player 1 lead to configurations x' such that $\mathcal{G}(x')$ is isomorphic to B_5, while two of them lead to configurations y such that $\mathcal{G}(y)$ is equal (or isomorphic) to B_5'. This graph consists of a sink that is a 3-clique and of a source that is a 2-clique with both nodes of it out-linking to one of the nodes of the 3-clique (see Figure 6 (b)). We now show that such configurations y are Nash equilibria. For simplicity, we assume that $\mathcal{G}(y)$ coincides with B'_5 as in Figure 6 (b). Regarding node 5, Proposition 3(iv)-(v) (using a cut argument with cut set $\{1\}$) yield $\tau_2^5(y) = \tau_3^5(y) > \tau_1^5(y) = \tau_4^5(y)$. Hence, the only best response for node 5 is $\{1, 4\}$. By symmetry, the only best response for node 4 is $\{1, 5\}$. Finally, because of relations (36), the best response of node 2 is any pair $\{3, s\}$ with s in $\{1, 4, 5\}$ and the best response of node 3 is any pair $\{2, s\}$ with s in $\{1, 4, 5\}$. This implies that y is a Nash equilibrium. The inverse transition is a best response for 1 in y and leads back to x. The only other possible transitions from y are through a modification of one of the out-links of either node 2 or node 3 and lead to configurations z such that $\mathcal{G}(z)$ is isomorphic to B''_5 (see Figure 6 (c)). We show that also such configurations z are Nash equilibria (for simplicity we assume that $\mathcal{G}(z)$ coincides with B''_5 as in Figure 6 (c)). An argument completely analogous to the one used in configuration y yields that node 5 is playing its unique best response. Regarding node 1, notice that $\tau_4^1(z) = \tau_5^1(z)$ thanks again to Proposition 3(v). Moreover, from equations (4), we obtain that

$$\tau_5^1(z) - \tau_2^1(z) = \frac{\beta}{2}(\tau_4^1(z) - \tau_3^1(z)), \quad \tau_4^1(z) - \tau_3^1(z) = \frac{\beta}{2}(\tau_5^1(z) - \tau_2^1(z) - \tau_4^1(z)) = -\frac{\beta}{2}\tau_2^1(z),$$

and this implies that both $\tau_2^1(z)$ and $\tau_3^1(z)$ are both strictly greater than $\tau_4^1(z) = \tau_5^1(z)$. As a consequence, also node 1 (and, by symmetry, node 4) is currently playing its unique best response. Finally, relations (36) show that node 2 and node 3 are still playing a best response in configuration z. This proves that z is a Nash equilibrium. Finally notice that in configuration z, only nodes 2 and 3 can modify their action and that any modification will lead to a configuration whose graph is either isomorphic to B'_5 or to B''_5. We have thus proven that all the configurations whose graph is isomorphic to one of the three graphs B_5, B'_5, or B''_5 are Nash equilibria and that such set is closed by best response paths. By Lemma 2(v), this implies that all such Nash equilibria are recursive. □

We now go more in depth with our analysis of Nash equilibria proving a necessary condition on the connected components of Nash equilibria of the centrality game $\Gamma(\mathcal{V}, \beta, \eta, 2 \cdot \mathbf{1})$. In particular, we show that ring graphs R_n and butterfly graphs B_5 are, in this context, the only possible strongly connected Nash equilibria. Towards this goal, we first introduce the graph $T_{(s,j),i}$ as the directed graph on the node set $\{i, j, s\}$ having one directed link (s, j) and all the other links undirected (see Figure 7 (a)). Notice that node s and node i have out-degree 2 and thus they have no choice but to connect to the remaining two nodes. Instead, node j has out-degree 1 and its best response is either



(a) $T_{(s,j),i}$ (b) (c) (d)

FIGURE 7. The directed graph $T_{(s,j),i}$ and other explanatory graphs for the proof of Lemma 8.

linking to i or to node s. If node j moves its out-link from i to s, we obtain the isomorphic graph $T_{(j,i),s}$. This says that $T_{(s,j),i}$ is a non-strict, though recursive, Nash equilibrium of the centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with $\mathcal{V} = \{s, i, j\}$ and $\mathbf{d} = (2, 2, 1)$. The following result illustrates the role played by the graph $T_{(s,j),i}$ in a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with $\mathbf{d} = 2 \cdot \mathbf{1}$.

LEMMA 8. *Let x in $\mathcal{X}^*$ be a Nash equilibrium for a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with $\mathbf{d} = 2 \cdot \mathbf{1}$. Let $\mathcal{G}_{\mathcal{K}}(x) = (\mathcal{K}, \mathcal{E}')$ be a connected component of $\mathcal{G}(x)$ that is a sink. If there exists a link (s, j) in $\mathcal{E}'$ such that $(j, s) \notin \mathcal{E}'$, then $\mathcal{G}_{\mathcal{K}}(x)$ contains a subgraph of type $T_{(s,j),i}$.*

Proof Since $\mathcal{G}_{\mathcal{K}}(x) = (\mathcal{K}, \mathcal{E}')$ is a sink connected component of $\mathcal{G}(x)$ and every node has out-degree 2, it follows that $|\mathcal{K}| \geq 3$. In particular, $|\mathcal{N}_s^{-\infty}(x)| \geq |\mathcal{K}| \geq 3$ so that Theorem 1(ii) implies that $x_s \subseteq \mathcal{N}_s^{-2}(x)$. By assumption, $j \notin \mathcal{N}_s^{-1}(x)$, hence there exists a node i in $\mathcal{K}$ such that $(j, i) \in \mathcal{E}'$ and $(i, s) \in \mathcal{E}'$. Moreover, the second claim in Theorem 1(ii) implies that $(s, i) \in \mathcal{E}'$. We are left to prove that also $(i, j) \in \mathcal{E}'$. By contradiction, suppose this is not the case. The argument just used for the pair of nodes s and j, can now be applied to the pair of nodes j and i and deduce the existence of a fourth node k in $\mathcal{V}' \setminus \{i, j\}$ such that $(i, k), (k, j), (j, k) \in \mathcal{E}'$. Notice that, as a consequence, $k \neq s$. The graph depicted in Figure 7 (b) would thus be a subgraph of $\mathcal{G}_{\mathcal{K}}(x)$. Now there are two possibilities: either $(k, i) \in \mathcal{E}'$ or $(k, i) \notin \mathcal{E}'$. In the former case, we are in the situation of Figure 7 (c) and we claim that j is not playing a best response. Indeed, subtracting the expected hitting times from nodes i and s to node j and using relations (4), we obtain that $(1 + (\beta/2))(\tau_i^j(x) - \tau_s^j(x)) = (\beta/2)\tau_k^j(x) > 0$ and so $\{s, k\}$ gives j a strictly better utility than $x_j = \{i, k\}$. Consider finally the case when $(k, i) \notin \mathcal{E}'$. Arguing for the pair of nodes i and k as we did before we deduce that $(k, s) \in \mathcal{E}'$, i.e., we are in the situation of Figure 7 (d). Using again relations (4), we obtain $(1 + (\beta/2))(\tau_k^i(x) - \tau_j^i(x)) = (\beta/2)\tau_s^i(x) > 0$ and so i would not be playing best response, for $\{j, s\}$ would give i a strictly better utility than $x_i = \{s, k\}$. This completes the proof. □

We can now prove the following result.

PROPOSITION 7. *Consider a centrality game $\Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with $\mathbf{d} = 2 \cdot \mathbf{1}$. Let x in $\mathcal{X}^*$ be a Nash equilibrium and let $\mathcal{G}_{\mathcal{K}}(x) = (\mathcal{K}, \mathcal{E}')$ be a connected component of $\mathcal{G}(x)$. Then:*

(i) if $\mathcal{G}_{\mathcal{K}}(x)$ is a source component, then it is either a singleton or a 2-clique;

(ii) if $\mathcal{G}_{\mathcal{K}}(x)$ is a sink component, then it is either a ring graph R_k or the Butterfly graph B_5.

Proof (i) This follows directly from Remark 2.

(ii) Since $\mathcal{G}_{\mathcal{K}}(x)$ is a connected component of a graph $\mathcal{G}(x)$ that is a sink and since all its nodes have out-degree exactly 2, if $\mathcal{G}_{\mathcal{K}}(x)$ is undirected then it must be a ring graph. Otherwise, if $\mathcal{G}_{\mathcal{K}}(x)$ is not undirected, then there must exist at least two directed links in $\mathcal{E}'$, say (s, j) and (r, k), so that Lemma 8 implies the existence of two subgraphs $T_{(s,j),i}$ and $T_{(r,k),t}$.

First, suppose that the subgraphs $T_{(s,j),i}$ and $T_{(r,k),t}$ share one or two nodes: the only way this can occur is if either (a) $k = i$ and $t = j$ or (b) $j = k$ and $\{s, i\} \cap \{r, t\} = \emptyset$. In case (a) we obtain that $\mathcal{G}_{\mathcal{K}}(x)$ coincides with the graph in Figure 8(a) and, thanks to Proposition 3(v), $\tau^i_j(x) = \tau^i_r(x)$. This implies that if $\{j, s\}$ is a best response for node i, then so is $\{s, r\}$. If node i chooses action $y_i = \{s, r\}$, however, we get a configuration y whose connected component $\mathcal{G}_{\mathcal{K}}(y)$ coincides with the graph in Figure 7(c) that in the proof of Lemma 8 was shown not to correspond to a Nash equilibrium. In case (b), the obtained graph corresponds to the Butterfly graph B_5 in Figure 6(a) and since in B_5 every node has out-degree equal to 2, it necessarily coincides with the sink component $\mathcal{G}_{\mathcal{K}}(x)$.

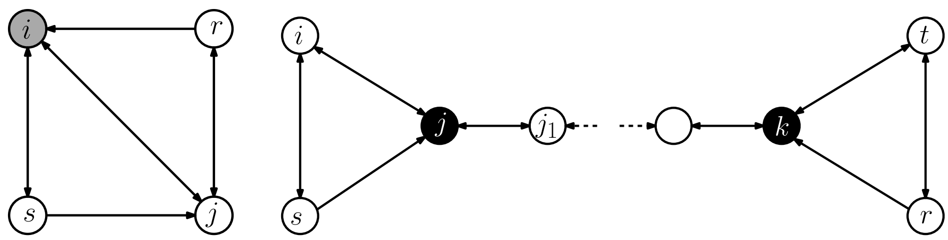
Now, suppose instead that the two subgraphs $T_{(s,j),i}$ and $T_{(r,k),t}$ do not share any node. Since in $T_{(s,j),i}$ node j has out-degree 1, and $d_j = 2$, there must exist a fourth node, say j_1, such that both $(j, j_1) \in \mathcal{E}'$ and $(j_1, j) \in \mathcal{E}'$. A recursive argument based on the finiteness of the graph now shows that there must exist a sequence of distinct nodes $j = j_0, j_1, \ldots, j_l$, with $l \geq 2$, such that $\{j_a, j_{a+1}\}$ are 2-cliques in $\mathcal{G}_{\mathcal{K}}(x)$ for $a = 1, \ldots, l-1$ and from j_l there is a directed link to some node k in $\{s, j, i, j_1, \ldots j_{l-1}\}$. Lemma 8 then implies that the only possibility is that $k = j_{l-2}$, so that $l \geq 3$ and we obtain the graph depicted in Figure 8(b). Since again every node has out-degree equal to 2, it necessarily coincides with the sink component $\mathcal{G}_{\mathcal{K}}(x)$. However, relations (4) imply that

$$\tau^j_{j_1}(x_{-j}) - \tau^j_s(x_{-j}) = \frac{\beta}{2}\left(\tau^j_{j_2}(x_{-j}) - \tau^j_i(x_{-j})\right) = \frac{\beta^2}{4}\left(\tau^j_{j_3}(x_{-j}) + \tau^j_{j_1}(x_{-j}) - \tau^j_i(x_{-j})\right)$$

so that

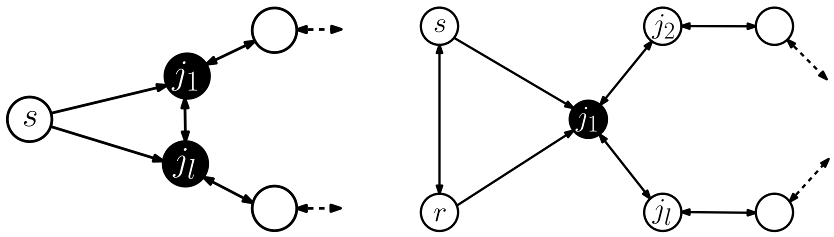
$$\tau^j_{j_1}(x_{-j}) - \tau^j_s(x_{-j}) = \frac{\beta^2/4}{1-\beta^2/4}\tau^j_{j_3}(x_{-j}) > 0,$$

and thus node j is not playing a best in response in configuration x since $y_j = \{i, s\}$ would give it a strictly better utility than $x_j = \{i, j_1\}$. This completes the proof. □



(a) (b)

Figure 8. Explanatory graph for the proof of Proposition 7. Gray nodes have multiple best response and black nodes are not playing best response.



(a) (b)

Figure 9. (a) Singleton source linking to two adjacent nodes in a ring graph. (b) 2-clique source linking to a single node in a ring graph. Black nodes are not playing a best response.

In particular, Proposition 7(ii) provides a complete classification of the graphs corresponding to Nash equilibria that are strongly connected or consist of isolated connected components. We are now ready to classify all strict and recursive Nash equilibria of the centrality game $\Gamma(\mathcal{V}, \beta, \eta, 2 \cdot \mathbf{1})$ in terms of the following graph families:

- $\mathcal{R}$ is the family of graphs obtained by taking an arbitrary disjoint union of ring graphs R_k with $k \geq 3$;
- $\mathcal{K}_{3,2}$ is the family of graphs that consist of a disjoint union of 3-ring graphs R_3 and of a unique 2-clique source component with two outgoing links each pointing to any of the nodes belonging to the ring graphs;
- $\mathcal{K}_{3,\mathcal{B}}$ is the family of graphs that are a disjoint union of 3-ring graphs R_3 and of a unique butterfly graph B_5.

Theorem 6 (**Strict and recursive Nash equilibria with $\mathbf{d} = \mathbf{2}$**). *For a centrality game $\Gamma(\mathcal{V}, \beta, \eta, 2 \cdot \mathbf{1})$ and a configuration x in $\mathcal{X}$:*

(i) x is a strict Nash equilibrium if and only if $\mathcal{G}(x) \in \mathcal{R}$;

(ii) x is a recursive Nash equilibrium if and only if $\mathcal{G}(x) \in \mathcal{R} \cup \mathcal{K}_{3,2} \cup \mathcal{K}_{3,\mathcal{B}}$.

Proof (i) If $\mathcal{G}(x)$ is a disjoint union of ring graphs, then x is a strict Nash equilibria thanks to Corollary 5. On the other hand, it follows from Proposition 7, Lemma 7, and Theorem 2(iii) that there cannot be other strict Nash equilibria.

(ii) We first show that if $\mathcal{G}(x) \in \mathcal{K}_{3,2}$, then x is a Nash equilibrium. Let $\mathcal{G}_{\mathcal{K}}(x)$, with $\mathcal{K} = \{r, s\}$, be the unique 2-clique source component. If both remaining out-links of r and s are directed to one or two nodes of the same 3-ring component $\mathcal{G}_{\mathcal{T}}(x)$, the subgraph induced by $\mathcal{K} \cup \mathcal{T}$ is isomorphic to either B_5' or B_5'' (see Figure 6). The graph $\mathcal{G}(x)$ is in this case the disjoint union of 3-rings and of a graph isomorphic to either B_5' or B_5''. Consequently, x is a Nash equilibrium because of Lemma 7 and Corollary 5. Suppose instead that the remaining out-links of r and s are directed to nodes of two different disjoint 3-rings $\mathcal{G}_{\mathcal{T}_1}(x)$ and $\mathcal{G}_{\mathcal{T}_2}(x)$. Suppose that $\mathcal{T}_1 = \{i, j, k\}$ and that r links to k. Then i (and analogously j) is playing its unique best response by noticing that $\tau^i_j(x) = \tau^i_k(x)$ because of Proposition 3(v) and the fact that k is a cut set with respect to the remaining nodes in the graph (see Proposition 3(iv)). Regarding node k, we have that $\tau^k_j(x) = \tau^k_i(x)$ because of Proposition 3(v) and from we obtain that $\tau^k_r(x) - \tau^k_j(x) = \frac{\beta}{2}(\tau^k_s(x) - \tau^k_i(x)) > \frac{\beta}{2}(\tau^k_r(x) - \tau^k_j(x))$ that implies $\tau^k_s > \tau^k_r(x) > \tau^k_j(x)$. Since $\mathcal{N}^{-2}_k(x) = \{k, i, j, r, s\}$, it follows from Proposition 6 that k is playing its unique best response. Since $\mathcal{N}^{-2}_r(x) = \{r, s\}$ also r is playing a best response and by symmetry, also s and the nodes in the other 3-ring are playing a best response. $\mathcal{G}(x)$ is thus a Nash equilibrium again because of Corollary 5. Similarly, any graph in $\mathcal{K}_{3,\mathcal{B}}$ is a Nash equilibrium because of Lemma 7 and Corollary 5. Previous considerations and Lemma 7(ii) also show that, from any $\mathcal{G}$ in $\mathcal{K}_{3,2} \cup \mathcal{K}_{3,\mathcal{B}}$, the graphs reachable in a best response path are all graphs in $\mathcal{K}_{3,2} \cup \mathcal{K}_{3,\mathcal{B}}$. By Lemma 2(v), this implies that all configurations x such that $\mathcal{G}(x) \in \mathcal{K}_{3,2} \cup \mathcal{K}_{3,\mathcal{B}}$ are recursive.

We are left with proving that if x is recursive, then $\mathcal{G}(x) \in \mathcal{R} \cup \mathcal{K}_{3,2} \cup \mathcal{K}_{3,\mathcal{B}}$. Suppose that $\mathcal{G}(x) \notin \mathcal{R}$. By Proposition 7, there are three possibilities to analyze: (a) $\mathcal{G}(x)$ contains a butterfly B_5 as a sink component; (b) $\mathcal{G}(x)$ contains a 2-clique K_2 as a source component; (c) $\mathcal{G}(x)$ contains a singleton as a source component.

In case (a), Lemma 7(ii) implies that there exist best response transitions that generate a source component from B_5. For this reason, Theorem 2(ii) forbids the presence of source components in $\mathcal{G}(x)$ and also of other sink components isomorphic to a B_5. Moreover, Lemma 7(ii) also yields that there exists a best response path from x to a configuration y with associated graph $\mathcal{G}(y)$ whose connected components are all sinks consisting of ring graphs except for a 2-clique $\{r, s\}$ that is the unique source with both r and s linking to a node j_1 in a ring of maximal length l. If $l > 3$, we can argue as follows. Let j_1 be the node in the ring to which both r and s point to, and let j_2 and j_l be its adjacent nodes in the ring (see Figure 9(b)). By system (4), $\tau_s^{j_1}(y)-\tau_{j_2}^{j_1}(y)=(\beta/2)(\tau_s^{j_1}(y)-\tau_{j_3}^{j_1}(y))$. Since $\tau_{j_3}^{j_1}(y)>\min\{\tau_{j_2}^{j_1}(y),\tau_{j_l}^{j_1}(y)\}=\tau_{j_2}^{j_1}(y)$ (where the inequality follows from Proposition 3(iv), since $C=\{j_2,j_l\}$ is a cut between j_3 and j_1, and the equality by symmetry), it follows that $\tau_{j_2}^{j_1}(y)>\tau_s^{j_1}(y)$. This says that j_1 is not playing a best response action. Consequently y is not a Nash equilibrium and thus x is not a recursive Nash equilibrium. Therefore, if $\mathcal{G}(x)$ contains a butterfly graph B_5, then $\mathcal{G}(x)\in\mathcal{K}_{3,\mathcal{B}}$. A completely analogous argument, in case (b), shows that if $\mathcal{G}(x)$ contains a 2-clique K_2 as a source component, then $\mathcal{G}(x)\in\mathcal{K}_{3,2}$.

Finally, suppose we are in case (c): $\mathcal{G}(x)$ contains a singleton source node s. From x, there exists a best response path leading to a configuration y where node s is linking to two adjacent nodes j_1 and j_l in the same ring of maximal length $l\geq 3$, as in Figure 9(a). By labeling the nodes in the length-l ring as $j_1,j_2,\dots,j_l$ so that j_h and j_k are adjacent if and only if $|h-k|=1$ modulo l, and using relations (4), we get that

$$\tau_s^{j_1}(y)-\tau_{j_2}^{j_1}(y)=\frac{\beta}{2}(\tau_{j_l}^{j_1}(y)-\tau_{j_3}^{j_1}(y)). \tag{38}$$

If $l>3$, a cut and a symmetry argument again implies that $\tau_{j_l}^{j_1}(y)-\tau_{j_3}^{j_1}(y)<0$. This again shows that j_1 is not playing a best response. Finally, if $l=3$, the nodes j_3 and j_l coincide. In this case, equation (38) and Proposition 3 (v) yield $\tau_s^{j_1}(y)=\tau_{j_2}^{j_1}(y)=\tau_{j_3}^{j_1}(y)$ and, thus, the pair $\{j_2,s\}$ is a best response for node j_1. This choice leads to a configuration z whose graph $\mathcal{G}(z)$ is isomorphic to the one in Figure 7 (c) that cannot be a Nash because of Lemma 8. Therefore, x is not recursive and this completes the proof. □

REMARK 6. In both the homogeneous cases analyzed, with one and two out-links, the set of strict Nash equilibria $\mathcal{X}^\bullet$ and that of recursive Nash equilibria $\mathcal{X}^\circ$ are both independent from the discount factor β. As we will see later on, however, in the case of two links the set of Nash equilibria $\mathcal{X}^*$ generally depends on the value of β.

REMARK 7. Notice that the disjoint union of two recursive Nash equilibria may not be a recursive Nash equilibrium in general. As an example, consider two disjoint butterfly graphs B_5. Individually, such configurations are recursive according to Lemma 7. However, Theorem 6(ii) implies that a recursive equilibrium cannot admit the simultaneous presence of two disjoint subgraphs isomorphic to B_5.

6. Numerical Results In this section, we present some numerical results that corroborate our theoretical findings and highlight further features of the equilibrium structure of the considered centrality games.

Our numerical experiments consider centrality games $\Gamma = \Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ and simulate the asynchronous best response dynamics introduced in Section 3: an initial configuration $X(0)$ is chosen at random from $\mathcal{X}$ and subsequent configurations are iteratively generated at times $t = 0, 1, 2, \ldots$ by sampling a node i uniformly at random from the set $\mathcal{V}$ and moving to a new configuration $X(t+1)$ such that $X_{-i}(t+1) = X_{-i}(t)$, while $X_i(t+1)$ is chosen uniformly from the best response set $\mathcal{B}_i(X_{-i}(t))$. By Corollary 1(i), with probability one $X(t)$ gets absorbed in finite time in the set $\mathcal{X}^{\circ}$ of recursive Nash equilibria. In our simulations, we stop the dynamics at some time T and for the final configuration $x = X(T)$ we compute the number of connected components $c(x)$.

At every time t, given the current configuration $X(t) = x$, best response actions for the sampled node i are computed from their characterization (13) in terms of expected hitting times. We make use of Theorem 1 in order to reduce the computational burden and restrict the search of nodes j with the lowest expected hitting times on i to the set $\mathcal{N}_i^{-d_i}(x_{-i})$ of nodes from which node i can be reached in at most d_i hops. Building on Proposition 3(i)-(ii), we compute the expected hitting times $\tilde{\tau}_j^i$ for the case $\eta = \delta^i$: using relation (18), this requires solving the linear system (17) in the unknowns $(\tilde{\tau}_j^i)_{j \in \mathcal{N}_i^{-\infty}(x_{-i})}$. The dimension $|\mathcal{N}_i^{-\infty}(x_{-i})|$ of this linear system depends of course on the configuration x, but it often turns out to be much lower than the network order n: in particular, this tends to occur when the number $m(x)$ of sink components in $\mathcal{G}(x)$ is large. As illustrated in Section 4.4, for values of the discount factor β close to 1, this corresponds to high values of the potential function $\Psi(x)$, which is non-decreasing along the best-response dynamics.

Figure 10 shows the evolution of the asynchronous best response dynamics starting from the network on the left for $T = 2000$ time steps. The order of the network is $n = 50$, the degree profile $\mathbf{d} = 7 \cdot \mathbf{1}$, and the discount factor is chosen as $\beta = 0.7$. The network images represent the configurations reached by the dynamics at times $t = 900, 1040, 1120, 1280$, while the plot underneath shows the evolution of the number of connected components of the reached configuration at each time step.

6.1. The Homogeneous Case The first set of simulations is carried out for centrality games $\Gamma = \Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ with homogeneous degree profiles, namely $\mathbf{d} = d \cdot \mathbf{1}$.

Notice that, for such centrality games, Theorem 4 characterizes the class of potential maximizing equilibria (that are a subclass of recursive Nash equilibria) when the discount factor β is sufficiently

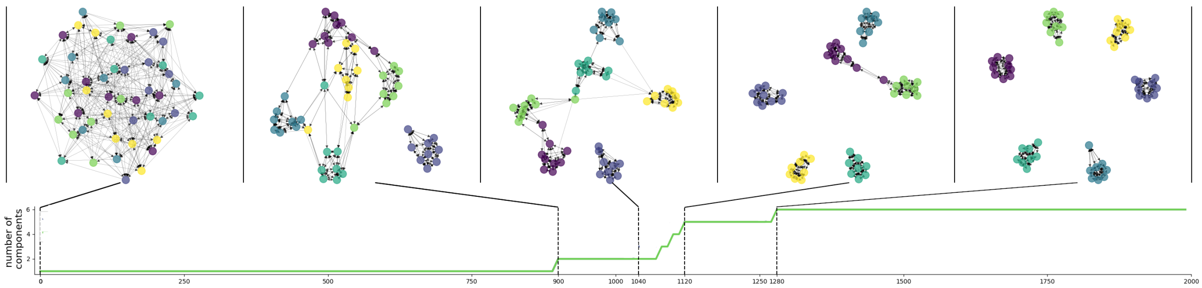


FIGURE 10. From left to right, evolution of the asynchronous best response dynamics for a centrality game with $n = 50$, $d = 7 \cdot \mathbf{1}$ and $\beta = 0.7$ and times between 0 and 2000. Colors are used to represent the final connected component the nodes belong to. In the bottom part, in green, we represent the number of connected components as function of time.

close to 1 and $n \geq d(d+1)$. All such potential maximizing equilibria are composed of isolated components of size $(d+1)$ and $(d+2)$ and thus, for such equilibria x, it holds that

$$\frac{n}{d+2} \leq c(x) \leq \frac{n}{d+1}\,. \tag{39}$$

Notice that any recursive equilibrium is composed of a number of sink connected components that must have size at least $d+1$ and of at most a source connected component of size at most d. This implies that, for every recursive equilibrium x (not necessarily a potential maximizer), we have the bound

$$c(x) \leq 1 + \frac{n-1}{d+1},$$

that for large n is asymptotically equivalent to the upper bound in relation (39). However, there exist recursive equilibria for which instead the left inequality in relation (39) is not satisfied: e.g., configurations x such that $\mathcal{G}(x)$ is the ring graph R_n are strongly connected strict Nash equilibria for $d = 2$.

For every $n \in \{100, 150, 200, 250\}$, $d \in \{3, 4, 5, 6, 7, 8, 9\}$, and $\beta \in \{0.1, 0.2, \dots, 0.9\}$, we randomly generate 50 networks with n nodes and out-degree d by choosing, independently for each node i, the set of its d out-neighbors uniformly at random in $\mathcal{V} \setminus \{i\}$. We then numerically simulate the asynchronous best response dynamics for $T = 100000$ time steps and compute $c(x)$ in the final configuration $X(T) = x$. We plot on the y-axis always the normalized index $C(x) = c(x)(d+1)/n$.

Figure 11 visually summarizes some statistics on this index with the use of boxplots. The top figure shows, for each combination of β and d, the boxplots of $C(x)$ as a function of n. Precisely, the boxes represent the distributions of $C(x)$ between the first and the third quartile, with the middle bars indicating the median, as n varies in $\{100, 150, 200, 250\}$. The whiskers extending from the box indicate the maximum and minimum values of the data sample, excluding outliers. We notice

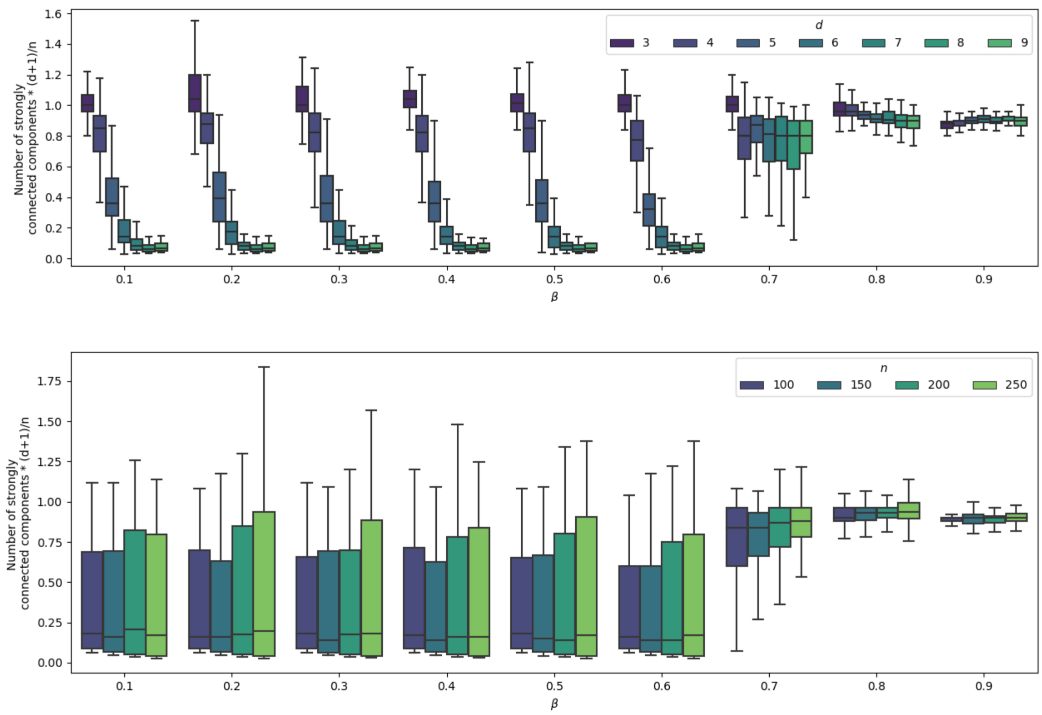
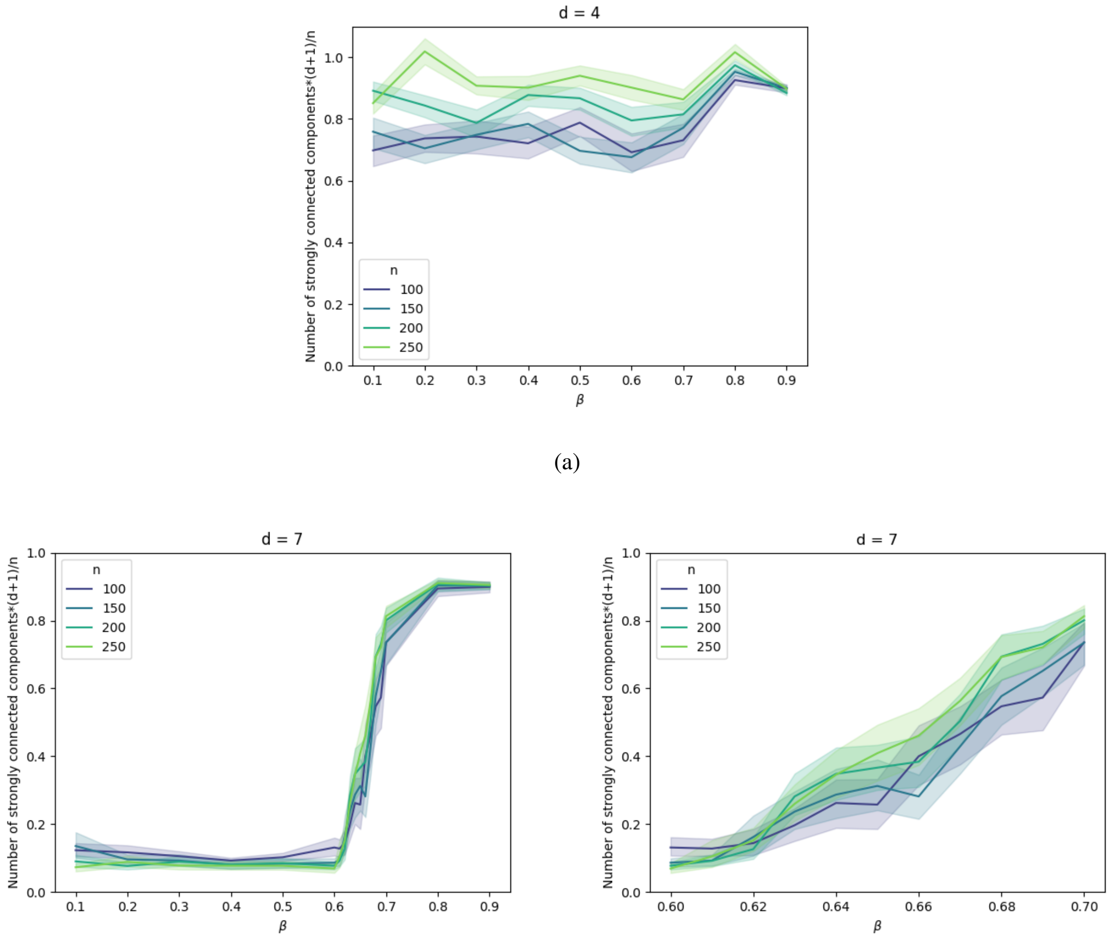


Figure 11. Distributions of the normalized number of connected components $C(x)$ in the configuration x reached by the asynchronous best response dynamics after $T = 100000$ time steps, starting from random homogeneous graphs with $n \in \{100, 150, 200, 250\}$ nodes, out-degrees $d \in \{3, 4, 5, 6, 7, 8, 9\}$, and $\beta \in [0.1, 0.9]$. Top: boxplots show the distribution of $C(x)$ as n varies for various fixed β and d. Bottom: boxplots show the distribution of $C(x)$ as d varies for various fixed β and n.

that for small d, the number of connected components is always quite high, and the value of the discount factor β does not really play a role in it. In contrast, for higher values of d, the number of connected components seems to be strongly influenced by the value of β: namely, for $\beta \leq 0.6$ the index $C(x)$ is close to 0, indicating a highly connected network, while for $\beta \geq 0.7$ the index $C(x)$ appears to be close to 1, indicating a highly fragmented network. The bottom figure shows instead the boxplots of the distribution of $C(x)$ as d varies in $\{3, 4, 5, 6, 7, 8, 9\}$, for each combination of β and n. Here, we can observe that the same phenomenon arises, where the number of connected components seems to consistently increase when $\beta > 0.6$, regardless of the value of n.

Figure 12 proposes a different visualization of the simulation for two values of d. Here, we plot $C(x)$ as a function of β for different values of n and (a) $d = 4$ or (b) $d = 7$. Solid lines indicate the averages while shaded areas indicate the corresponding variances. We notice that for $d = 4$, $C(x)$ is always above 0.6 suggesting that the recursive Nash equilibria obtained are largely disconnected for every value of β. Instead, for $d = 7$ a transition phase phenomenon is suggested to happen around $\beta^* = 0.66$: for this reason numerical experiments have been carried out for as many values of β in



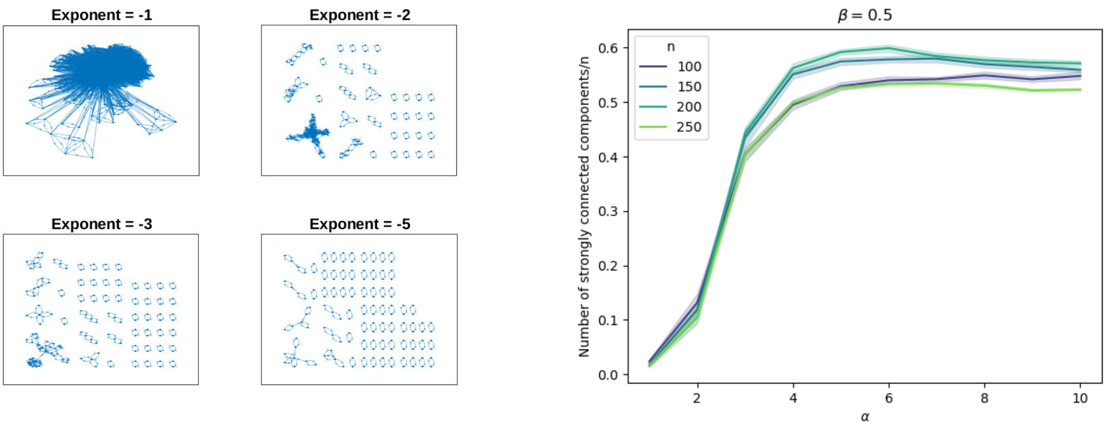
(a)

(b) (c)

Figure 12. Normalized number of connected components $C(x)$ in the configuration x reached by the asynchronous best response dynamics after $T = 100000$ time steps, starting from random homogeneous graphs with different values of n, β and d. Plots show $C(x)$ as function of β for various values of n in the cases (a) $d = 4$ and (b) $d = 7$, while (c) is a zoom of plot (b) considering values $\beta \in \{0.6, 0.61, 0.62, \ldots, 0.7\}$. The solid lines refer to the average normalized number of connected components over 50 initially generated networks against β for fixed n, d, while shaded areas indicate the corresponding variances.

the range $[0.6, 0.7]$ as in the rest of the range $[0, 1]$. Figure 12 (c) reports a zoom for $\beta \in [0.6, 0.7]$ of Figure 12 (b). Above this value β^*, we obtain again maximally disconnected networks close to the potential maximizers investigated in Theorem 4, while below such value more connected networks emerge, confirming what we have already observed.

6.2. Power-Law Out-Degree Distributions A second set of numerical experiments have been carried out for centrality games $\Gamma = \Gamma(\mathcal{V}, \beta, \eta, \mathbf{d})$ whose out-degree profile has been randomly



(a) (b)

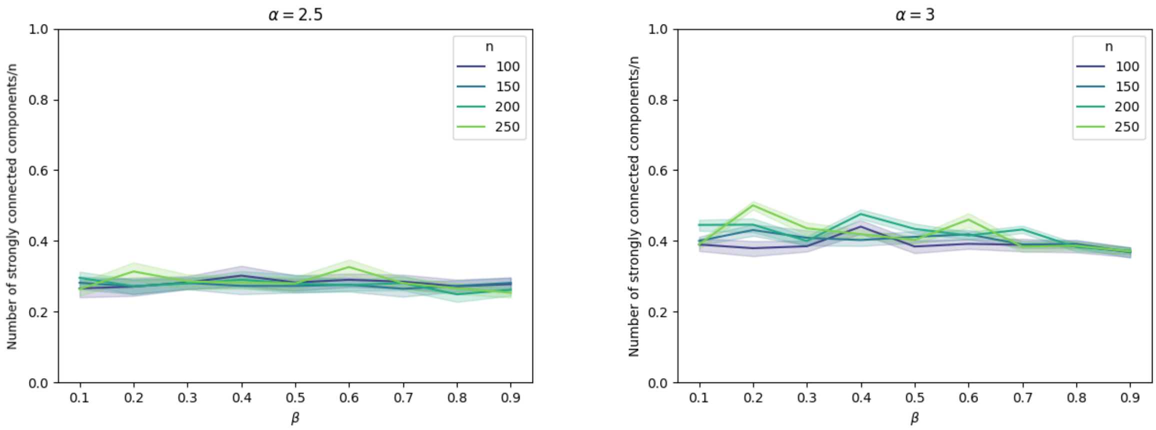
Figure 13. Results on the configuration reached by the asynchronous best response dynamics after $T = 100000$ time steps, starting from networks generated by the power law (40). (a) Final configurations reached with $n = 150$, $\beta = 0.5$, and various values of α. (b) Normalized number of connected components in final configurations against α for $\beta = 0.5$ and various n. Solid lines indicate averages while shaded areas indicate the corresponding variances.

sampled from a truncated power-law distribution:

$$\mathbb{P}(d_i = d) = \begin{cases} \dfrac{d^{-\alpha}}{\sum_{k=1}^{n-1} k^{-\alpha}} & \text{if } 1 \le d < n \\ 0 & \text{if } d \ge n, \end{cases} \tag{40}$$

where $\alpha > 0$ is a parameter. A peculiar feature of graphs with such power law distributions (particularly when α is small) is the presence of *hubs*, meant as nodes with very large out-degree.

We have carried out several numerical experiments with various values of n, β, and α. In all cases, the initial configuration has been randomly generated by first sampling the out-degree distribution $\mathbf{d}$ with probability distribution (40) and then choosing, independently for each node i, the set of its d_i out-neighbors uniformly at random in $\mathcal{V} \setminus \{i\}$. As in the previous case, the best response dynamics has been simulated for $T = 100000$ time steps. Figure 13(a) shows four final configurations reached by the asynchronous best response dynamics started from a random initial configuration, with $n = 150$, $\beta = 0.5$, and $\alpha \in \{1, 2, 3, 5\}$. We notice the presence of a hub for $\alpha = 1$, while higher values of α yield more and more isolated components as maximal out-degree gets smaller and smaller. For the same value of β, we performed more extensive simulations in the style followed for the homogeneous case, by generating 50 networks for each $n \in \{100, 150, 200, 250\}$ and $\alpha \in \{1, 2, \ldots, 9, 10\}$. In Figure 13(b), we plot average and variance of the normalized index $c(x)/n$



(a) (b)

FIGURE 14. Normalized number of connected components $c(x)/n$ in the configuration x reached by the asynchronous best response dynamics after $T = 100000$ time steps, starting from networks generated by the power law (40). The index $c(x)/n$ is plotted against β for for various n and with (a) $\alpha = 2.5$, (b) $\alpha = 3$. Solid lines indicate averages while shaded areas indicate the corresponding variances.

for the final configuration x against α for the different values of n. This plot confirms that for small values of α the index $c(x)/n$ is small indicating the presence of large connected components, presumably because of the presence of hubs. On the other hand, for high values of α the index $c(x)/n$ is close to 0.5, which means that we have highly disconnected networks where most of the components are just 2-cliques. This is consistent with the fact that, for high values of α, the maximal out-degree gets very small and most of the vertices will have out-degree equal to 1.

In consideration of the fact that typically the power-law distribution is considered for $\alpha > 2$ (otherwise the average degree is unbounded in n), we also have carried out more extensive results for the values $\alpha = 2.5$ and $\alpha = 3$, both values for which the second moment of the degree distribution diverges as n grows large. In Figure 14 we have plotted the average (over 50 instances) and variance of the normalized index $c(x)/n$ for the final configuration x against β for $n \in \{100, 150, 200, 250\}$ and $\alpha = 2.5$ or $\alpha = 3$. These plots suggest that the final network remains largely disconnected with the number of connected components that seems to increase linearly in n. We can also notice that the parameter β does not appear to have a significant effect on the structure of the emerging equilibria in either the chosen values of α.

7. Conclusion

In this paper, we have proposed and analyzed a family of network formation games in which every node i, equipped with a fixed number of out-links d_i, is free to choose how to direct them in order to maximize its PageRank centrality. We have first shown that the considered model is a potential game. Our results show that best responses are essentially local: a player i tends to link to nodes from which node i can be reached in at most d_i steps. This fact yields fundamental information on the structure of networks that are Nash equilibria: connected components can only be sources or sinks and at most one source can show up in the class of recursive Nash equilibria, where best response dynamics is known to get absorbed in finite time. This implies that typically equilibria are largely disconnected, with several undirected links and small cycles. For the special case of homogeneous out-degree profiles with $d_i = d$ for every node i, the analysis of the potential function allows us to reach further insight on the subclass of recursive Nash equilibria that maximize the potential (when the discount factor of the centrality is sufficiently high): they are all composed of isolated components of size $d+1$ and $d+2$. For the case of $d = 1, 2$ we have a complete classification of the recursive Nash equilibria.

In the final section, we carry out some initial numerical studies. Besides corroborating our theoretical results, our numerical results suggest the possible presence of phenomena not yet investigated, such as for instance the possible presence of threshold type behaviors with respect to the discount parameter of the PageRank centrality. This is left for future investigation.

Fragmentation and lack of connectivity seem to be the norm in this network formation process and we might question how realistic this can be. Indeed, in some real-world networks this may occur: notable examples are the citation graphs, the World Wide Web, or other social networks like the sentimental relation graph reported in [24, Figure 2.7]. Nevertheless, there are contexts where being part of a larger community may bring an advantage to the individuals. Some preliminary work in this direction is reported in [19] where a community is engaged in an inferential task whose performance depends on the size of the community. A challenging problem (and new at the best of our knowledge) is to combine the two mechanisms and consider games where players have to trade off between trying to be central in their community and, at the same time, being part of a community large enough to well perform some collective task.

Acknowledgments

The authors wish to thank the anonymous reviewers for several valuable suggestions that improved their manuscript.

References

[1] Alon N, Demaine ED, Hajiaghayi M, , Leighton T (2010) Basic network creation games. *Proceedings of the 22nd ACM symposium on Parallelism in algorithms and architectures*, volume SPAA '10, 106 – 113 (ACM, New York).

[2] Altman A, Tennenholtz M (2005) Ranking systems: The PageRank axioms. *Proceedings of the 6th ACM Conference on Electronic Commerce*, 1–8, EC ’05 (New York, NY, USA: Association for Computing Machinery).

[3] Anantharam V, Tsoucas P (1989) A proof of the Markov chain tree theorem. *Statistics & Probability Letters* 8:189–192.

[4] Avrachenkov K, Litvak N (2006) The effect of new links on Google PageRank. *Stochastic Models* 22(2):319–331.

[5] Bala V, Goyal S (2000) A noncooperative model of network formation. *Econometrica* 68(5):1181 – 1229.

[6] Belhaj M, Bervoets S, Frédéric Deroïan F (2016) Efficient networks in games with local complementarities. *Theoretical Economics* 11:357 – 380.

[7] Blume L (1993) The statistical mechanics of strategic interaction. *Games and Economic Behavior* 5:387–424.

[8] Blume L (1995) The statistical mechanics of best response strategy revision. *Games and Economic Behavior* 11(2):111–145.

[9] Blume L (1997) *The Economy As An Evolving Complex System II*, chapter Population Games (CRC Press).

[10] Boldi P, Vigna S (2014) Axioms for centrality. *Internet Mathematics* 3-4(10):222–262.

[11] Bonacich P (1987) Power and centrality: A family of measures. *American Journal of Sociology* 92(5):1170–1182.

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

[13] Castaldo M, Catalano C, Como G, Fagnani F (2020) On a centrality maximization game. *IFAC* 53(2):2844–2849.

[14] Chen W, Teng SH, Wang Y, Zhou Y (2009) On the α-sensitivity of Nash equilibria in PageRank-based network reputation games. *Frontiers in Algorithmics*, 63–73.

[15] Chien S, Dwork C, Kumar R, Simon DR, Sivakumar D (2004) Link evolution: Analysis and algorithms. *Internet Mathematics* 1(3):277–304.

[16] Cominetti R, Quattropani M, Scarsini M (2022) The buck-passing game. *Mathematics of Operations Research* 47(3):1731–1756.

[17] Como G, Fagnani F (2015) Robustness of large-scale stochastic matrices to localized perturbations. *IEEE Transactions on Network Science and Engineering* 2(2):53–64.

[18] Como G, Fagnani F (2016) From local averaging to emergent global behaviors: The fundamental role of network interconnections. *Systems and Control Letters* 95:70–76.

[19] Como G, Fagnani F, Proskurnikov A (2022) Reaching optimal distributed estimation through myopic self-confidence adaptation. *IFAC PapersOnLine* .

[20] Corbo J, Calvo-Armengol A, Parkes D (2006) A study of Nash equilibrium in contribution games for peer-to-peer networks. *Operating Systems Review* 40(3):61 – 66.

[21] Csáji BC, Jungers RM, Blondel VD (2010) PageRank optimization in polynomial time by stochastic shortest path reformulation. *Algorithmic Learning Theory*, 89–103.

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

[23] Demaine ED, Hajiaghayi M, Mahini H, Zadimoghaddam M (2007) The price of anarchy in network creation games. *Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing*, volume PODC '07, 292 – 298 (ACM, New York).

[24] Easley D, Kleinberg J (2010) *Networks, Crowds, and Markets: Reasoning About a Highly Connected World* (Cambridge University Press).

[25] Echzell H, Friedrich T, Lenzner P, A M (2021) Flow-based network creation games. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, volume IJCAI'20, 139 – 145 (ACM, New York).

[26] Ehsani S, Fadaee S, Fazli M, Mehrabian A, Sadeghabad S, Safari M, Saghafian M (2011) On a bounded budget network creation game. *Proceedings of the twenty-third annual ACM symposium on Parallelism in algorithms and architectures*, volume SPAA '11, 207 – 214 (ACM, New York).

[27] Fabrikant A, Luthra A, Maneva E, Papadimitriou CH, Shenker S (2003) On a network creation game. *Proceedings of the twenty-second annual symposium on Principles of distributed computing*, volume PODC '03, 347 – 351 (ACM, New York).

[28] Friedkin N (1991) Theoretical foundations for centrality measures. *American Journal of Sociology* 96:1478–1504.

[29] Friedkin NE, Johnsen EC (1990) Social influence and opinions. *Journal of Mathematical Sociology* 15(3-4):193–206.

[30] Hopcroft J, Sheldon D (2008) Network Reputation Games. Technical report, Cornell University.

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

[32] Jackson M, ed. (2008) *Social and Economic Networks* (Princeton and Oxford: Princeton Univ. Press).

[33] Jackson MO, Wolinsky A (1996) A strategic model of social and economic networks. *Journal of Economic Theory* 71(1):44 – 74.

[34] Katz L (1953) A new status index derived from sociometric analysis. *Psychometrika* 18:39–43.

[35] Kawald B, Lenzner P (2013) On dynamics in selfish network creation. *Proceedings of the twenty-fifth annual ACM symposium on Parallelism in algorithms and architectures*, volume SPAA '13, 83 – 92 (ACM, New York).

[36] Laoutaris N, Poplawski LJ, Rajaraman R, Sundaram R, Teng SH (2008) Bounded budget connection (bbc) games or how to make friends and influence people, on a budget. *Proceedings of the twenty-seventh ACM symposium on Principles of distributed computing*, volume PODC '08, 165 – 174 (ACM, New York).

[37] Marden JR, Shamma JS (2012) Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior* 75(2):788–808.

[38] Mihalák M, Schlegel J (2012) Asymmetric swap-equilibrium: A unifying equilibrium concept for network creation games. *Mathematical Foundations of Computer Science 2012*, volume Lecture Notes in Computer Science, Vol. 7464, 693 – 704 (Springer, Berlin).

[39] Monderer D, Shapley LS (1996) Potential games. *Games and Economic Behavior* 14(1):124 – 143.

[40] Newman M (2003) The structure and function of complex networks. *SIAM Review* 45(2):167–256.

[41] Norris JR (1997) *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press).

[42] Proskurnikov AV, Tempo R (2017) A tutorial on modeling and analysis of dynamic social networks. part i. *Annual Reviews in Control* 43:65–79.

[43] Read RC, Wilson RJ (1998) *An Atlas of Graphs* (Oxford University Press).