# Efficient networks in connections models with heterogeneous nodes and links

Olaizola, Norma and Valenciano, Federico

University of the Basque Country UPV/EHU, University of the Basque Country UPV/EHU

5 October 2022

# Efficient networks in connections models with heterogeneous nodes and links*

By Norma Olaizola† and Federico Valenciano‡

October 5, 2022

**Abstract**

We culminate the extension of the results on efficiency in the seminal connections model of Jackson and Wolinsky (1996), partially addressed in previous papers. In a model where both nodes and links are heterogeneous, we prove that efficiency is reached by networks with a particular type of architecture that we call "hierarchical flower networks". These networks have a unique non-trivial component, within which one of the nodes with a highest value is directly connected with the others in the component, among which some pairs are directly connected. Moreover, the greatest the sum of the values of two nodes, the greatest the strength of their connection, be it direct by a link or indirect by means of two links through the central node.

*JEL* Classification Numbers: A14, C72, D85

*Key words:* Networks, Connections model, Heterogeneity, Efficiency.

---

*This research is supported by Grant PID2019-106146GB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", and by the Basque Government Departamento de Educación, Política Lingüística y Cultura funding for Grupos Consolidados IT1697-22.

†*BRiDGE* group (http://www.bridgebilbao.es), Departamento de Análisis Económico, Universidad del País Vasco UPV/EHU, Avenida Lehendakari Aguirre 83, 48015 Bilbao, Spain; norma.olaizola@ehu.es.

‡*BRiDGE* group (http://www.bridgebilbao.es), Departamento de Métodos Cuantitativos, Universidad del País Vasco UPV/EHU, Avenida Lehendakari Aguirre 83, 48015 Bilbao, Spain; federico.valenciano@ehu.es.

# 1 Introduction

The point of this paper is to culminate the extension of the results on efficiency in the seminal connections models of Jackson and Wolinsky (1996) and Bala and Goyal (2000) partially addressed in several previous papers.[1] In the seminal connections models, a network is the result of creating links between nodes, by bilateral agreements and equal split of the cost of each link in Jackson and Wolinsky's model and unilateral decisions in Bala and Goyal's model. In both models, the cost of a link and its strength or quality (i.e. its decay factor) are exogenously given, giving rise to two-parameter models. This imposes necessarily bilateral formation and compulsory equal share of the fixed cost of each link in Jackson and Wolinsky's model; and unilateral formation and full-covering of its cost by its creator in Bala and Goyal's model, and a fixed level of quality for the resulting link in both. A third parameter is the endowment or value of each node, partially accessed by other nodes, which is the source of revenue. The benefit that each node-player receives through the network is the sum of the fractions of the values of the nodes directly or indirectly connected with it. This fraction is given by the strength of the best path (i.e. the shortest in their models) that connects them. It is assumed that all nodes have the same initial endowment. In other words, both models assume *homogeneity in nodes and links.*[2]

The point of this paper is to culminate the extension of the results about efficiency in a connections model where *both nodes and links are heterogeneous*, i.e. nodes have different values and the strength of a link depends on the amounts invested in it by the two nodes that it connects and is determined by a function of these amounts. Generalizing previous results (Olaizola and Valenciano, 2020a, 2020b, 2021 and 2022), we prove that in these general conditions, under mild conditions about the technology to form links endogenously, efficiency is reached by networks with a particular type of architecture that we call "hierarchical flower networks". These networks have a unique non-trivial component, within which one of the nodes with a highest value is directly connected with the others in the component, among which some pairs are directly connected. Moreover, the greatest the sum of the values of two nodes, the greatest the strength of their connection, be it direct by a link or indirect by means of two links through the central node. Thus in these networks connections are either direct or through the central node. So that only the links with the central node generate externalities, while links between nodes different from the central are only used to see each other by the two nodes it connects.

A feature worth noticing of the proof of the result upon which the result about the efficiency of such networks hinges upon is its algorithmic character, another is the character of the *ad probandum* algorithm. As in Olaizola and Valenciano (2020a and 2021), an algorithm is provided in order to *prove* that any network is dominated by

[1] Olaizola and Valenciano (2020a, 2020b, 2021 and 2022).

[2] Bloch and Dutta (2009) introduce endogenous link strength in a connections model by replacing Jackson and Wolinsky's discrete technology by a *non-decreasing* returns technology.

a particular type of network, a hierarchical flower network in this case, of which the result is a corollary. However, in this more general case a more complex deterministic algorithm with a non-deterministic flavor is needed to prove the result as, by the nature of the problem, a whole tree with possibly many binary branching must be thoroughly covered in order to find the dominant hierarchical flower.

The paper is organized as follows. Section 2 introduces basic notation and terminology. Section 3 introduces the model. Section 4 briefly reviews previous results when only nodes or links are heterogeneous. Section 5 addresses the question of efficiency with heterogeneity in nodes and links. Section 6 concludes pointing out some further work on stability.

# 2 Preliminaries

An *undirected weighted network* (shortened in what follows to a *network*) is a pair $(N, g)$ where $N = \{1, 2, ..., n\}$ with $n \geq 3$ is a set of *nodes* and $g$ is a set of *links* specified by a symmetric *adjacency matrix* $g = (g_{ij})_{i,j \in N}$ of real numbers $g_{ij} \in [0, 1)$, with $g_{ii} = 0$ for all $i$. Alternatively, $g$ can be specified as a map $g : N_2 \to [0, 1)$, where $N_2$ denotes the set of all subsets of $N$ with cardinality 2. When no ambiguity arises we omit $N$ and refer to $g$ as a network. In what follows $\underline{ij}$ stands for $\{i, j\}$ and $g_{\underline{ij}}$ for $g(\{i, j\})$ for any $\{i, j\} \in N_2$. When $g_{\underline{ij}} > 0$ it is said that a *link of strength* $g_{\underline{ij}}$ connects $i$ and $j$. $N_i^d(g) := \{j \in N : g_{\underline{ij}} > 0\}$ denotes the set of *neighbors* of node $i$. A *path* connecting nodes $i$ and $j$ is a sequence of distinct nodes of which the first is $i$, the last is $j$, and every two consecutive nodes are connected by a link. If $i$ and $j$ are two consecutive nodes in a path $p$, we write $ij \in p$ or $\underline{ij} \in p$. $\mathcal{P}_{ij}(g)$ denotes the set of paths in $g$ connecting $i$ and $j$. $N_i(g)$ denotes the set of nodes connected to $i$ by a path. A network is *connected* if any two nodes are connected by a path. A *subnetwork* of a network $(N, g)$ is a network $(N', g')$ s.t. $N' \subseteq N$ and $g' \subseteq g$. A *component* of a network $(N, g)$ is a maximal connected subnetwork. An isolated node (i.e. not connected to any other) is a *trivial component*.

The *empty network* is the one for which $g_{\underline{ij}} = 0$ for all $\underline{ij} \in N_2$. A *complete network* is one where $g_{\underline{ij}} > 0$ for all $\underline{ij} \in N_2$. A *subcomplete network* has only one non-trivial component which is a complete subnetwork, i.e. $g_{\underline{ij}} > 0$ if and only if $\underline{ij} \in M_2$ for some $M \subseteq N$. A *star* is a network with only one non-trivial component with $k$ nodes $(3 \leq k \leq n)$ and $k - 1$ links in which one node (the *center*) is connected by a link with each of the other $k - 1$ nodes. A *flower* is a network with only one non-trivial component with $k$ nodes $(3 \leq k \leq n)$ which consists of a central node connected with each of the other $k-1$ nodes directly, some of which can be directly connected between them.[3] Note that complete networks and stars are particular extreme cases of a flower. Another particular type of flower is the following, which plays a role later.

[3]See e.g. Bala and Goyal (2000) for a different use of the term.

**Definition 1** *A nested split graph network (NSG-network) is a network $g$ such that for all $i, j \in N$ $(i \neq j)$,*

$$\left|N_i^d(g)\right| \leq \left|N_j^d(g)\right| \Rightarrow N_i^d(g) \subseteq N_j^d(g) \cup \{j\}.$$

NSG-networks have a hierarchical architecture, as nodes can be ranked according to their number of neighbors. Note also that NSG-networks are flowers, but not all flowers are NSG-networks.

A star or a flower is said to be *all-encompassing* if $k = n$.

# 3 The model

The point of this paper is the extension of the results about efficiency in the seminal connections models to a model where *both nodes and links are heterogeneous*, i.e. nodes have different values and the strength of a link depends on the amounts invested in it by the two nodes that it connects and is determined by a link-formation technology.

**Definition 2** *A link-formation technology (a technology for short) is a map $\lambda : \mathbb{R}_+^2 \to [0, 1)$ non-decreasing in both arguments, symmetric, i.e. $\lambda(x, y) = \lambda(y, x)$, and s.t. $\lambda(0, 0) = 0$.*

That is, the strength of a link cannot decrease by an increase of the investment in it, the technology does not depend of the identity of its users, and positive strength is costly. We consider the following situation. A set $N = \{1, 2, ..., n\}$ of nodes or players, each of which is endowed with a *value* $v_i > 0$, can form links according to a technology $\lambda$ available to all. A *link-investment* is a $n$-tuple $\mathbf{c} = (\mathbf{c}_i)_{i \in N}$, where $\mathbf{c}_i \in \mathbb{R}_+^N = (c_{ij})_{j \in N}$ with $c_{ii} = 0$, where $c_{ij}$ is the investment of node $i$ in its link with $j$. Link-investment $\mathbf{c}$ yields a weighted network denoted by $g^{\mathbf{c}}$, where $g^{\mathbf{c}}_{\underline{ij}} = \lambda(c_{ij}, c_{ji})$. When $i$ and $j$ invest $c_{ij}$ and $c_{ji}$ in a link connecting them, $\lambda(c_{ij}, c_{ji})$ is the strength of the resulting link, which is the level of fidelity of the transmission of value through it. Flow occurs only through links invested in ($\lambda(0, 0) = 0$), but transmission is never perfect ($0 \leq \lambda(c_{ij}, c_{ji}) < 1$).[4] For a path $p \in \mathcal{P}_{ik}(g^{\mathbf{c}})$, $\delta(p)$ denotes the product of the fidelity levels through each link in that path, i.e. if $p = ii_2i_3...i_mk$, then $\delta(p) = g_{\underline{ii_2}} \times g_{\underline{i_2i_3}} \times ... \times g_{\underline{i_mk}}$. We assume that node $i$ receives from node $k$'s value, $v_k$, the fraction which reaches $i$ through the best path from $k$, that is

$$I_{ik}(g^{\mathbf{c}}) = \max_{p \in \mathcal{P}_{ik}(g^{\mathbf{c}})} v_k \delta(p) = v_k \delta(\overline{p}_{ik}),$$

[4] The value $v_i$ of each node can be interpreted as an endowment of valuable information that can be partially accessed through the network, then the strength of a link is the level of fidelity of the transmission of information through it. Alternatively, $v_i$ can be interpreted as the value of node $i$ as a valuable contact, then the strength of a link is a measure of the quality/intensity/reliability, i.e. the "strength of a tie" (Granovetter, 1973). Under this interpretation, the model can be interpreted as a stylized model of a contact network.

where $\overline{p}_{ik}$ is an *optimal path connecting $i$ and $k$*, i.e. $\overline{p}_{ik} \in \arg\max_{p \in \mathcal{P}_{ik}(g^{\mathbf{c}})} \delta(p)$ (if no path connects $i$ and $k$ we set $\delta(\overline{p}_{ik}) = 0$). Then $i$'s overall revenue from $g^{\mathbf{c}}$ is

$$I_i(g^{\mathbf{c}}) = \sum_{k \in N(i;g^{\mathbf{c}})} I_{ik}(g^{\mathbf{c}}).$$

Thus, $i$'s *payoff* is the value received minus $i$'s investment:

$$\Pi_i^{\delta}(\mathbf{c}) := I_i(g^{\mathbf{c}}) - \sum_{j \neq i} c_{ij} = \sum_{k \in N(i;g^{\mathbf{c}})} v_k \delta(\overline{p}_{ik}) - \sum_{j \neq i} c_{ij}, \tag{1}$$

and the *net value* of the network resulting is the aggregate payoff, i.e. the total value received by the nodes minus the total cost of the network, given by $C(g^{\mathbf{c}}) = \sum_{\underline{ij} \in N_2} c_{\underline{ij}}$, where $c_{\underline{ij}} := c_{ij} + c_{ji}$, i.e.

$$v(g^{\mathbf{c}}) := \sum_{i \in N} \Pi_i^{\delta}(\mathbf{c}) = \sum_{i \in N} I_i(g^{\mathbf{c}}) - C(g^{\mathbf{c}}) = \sum_{\underline{kl} \in N_2} (v_k + v_l)\delta(\overline{p}_{kl}) - C(g^{\mathbf{c}}). \tag{2}$$

In what follows, we denote $s_{kl} := v_k + v_l$, so that

$$v(g^{\mathbf{c}}) = \sum_{\underline{kl} \in N_2} s_{kl} \delta(\overline{p}_{kl}) - C(g^{\mathbf{c}}). \tag{3}$$

Note that $v(g^{\mathbf{c}})$ can be expressed like this:

$$v(g^{\mathbf{c}}) = \sum_{i \in N} v_i \sum_{j \in N \setminus \{i\}} \delta(\overline{p}_{ij}) - C(g^{\mathbf{c}}) = \sum_{i \in N} w_i v_i - C(g^{\mathbf{c}}), \tag{4}$$

where coefficient $w_i := \sum_{j \in N \setminus \{i\}} \delta(\overline{p}_{ij})$ can be interpreted as *node $i$'s visibility* because $w_i v_i$ is the impact of node $i$'s value on the aggregate value of the network. Alternatively $v(g^{\mathbf{c}})$ can also be expressed like this:

$$v(g^{\mathbf{c}}) = \sum_{i \in N} \sum_{j \in N \setminus \{i\}} \delta(\overline{p}_{ij}) v_j - C(g^{\mathbf{c}}) = \sum_{i \in N} V_i - C(g^{\mathbf{c}}), \tag{5}$$

where $V_i := \sum_{j \in N_2 \setminus \{i\}} \delta(\overline{p}_{ij}) v_j$ is node $i$'s *vision*, i.e. the information received by node $i$.

A network $g^{\mathbf{c}}$ *dominates* a network $g^{\mathbf{c}'}$ if $v(g^{\mathbf{c}}) \geq v(g^{\mathbf{c}'})$. A network is *efficient* if it dominates any other. As shown here, efficiency may be reached by a network that leaves some nodes disconnected. A network is *efficient-constrained-to-keep-connectedness* if it is connected and dominates any other connected network.

## 4 Previous results

We begin by briefly reviewing the results in the seminal papers and the above mentioned extensions.

**Homogeneous nodes and homogeneous links**

In terms of the current framework, Jackson and Wolinsky's (1996) discrete model is equivalent to assume homogeneous nodes of value 1 and a technology given by

$$\lambda(c_1, c_2) = \begin{cases} \delta, \text{ if } \min\{c_1, c_2\} \geq \frac{c}{2} \\ 0, \text{ otherwise.} \end{cases} \tag{6}$$

Bala and Goyal's (2000) discrete model is equivalent to assume homogeneous nodes of value 1 and a technology given by

$$\lambda(c_1, c_2) = \begin{cases} \delta, \text{ if } \max\{c_1, c_2\} \geq c \\ 0, \text{ otherwise.} \end{cases} \tag{7}$$

In both cases, parameters $\delta$ $(0 < \delta < 1)$ and $c > 0$ are exogenously given. In both cases, *the only non-empty possibly efficient networks are the complete network and the all-encompassing star.*

**Homogeneous links and heterogeneous nodes**

Olaizola and Valenciano (2021) consider Jackson and Wolinsky's (1996) model with *heterogeneous nodes*, which in the current framework corresponds to discrete technology (6), and assume that each node $i$ has a possibly different value $v_i$. They prove that *efficiency is reached by a nested split graph network where the rank of nodes according to the number of neighbors is consistent with the rank according to their value and some nodes can be disconnected.*

They also prove that efficiency constrained to keep connectedness (i.e. the connected network with greatest value) is also reached by a nested split graph network of this type, which they call *strong* NSG-network, and provide a simple algorithm to obtain it. Formally:

**Definition 3** (*Olaizola and Valenciano,* 2021) *A strong NSG-network* (*SNSG-network*) *of heterogeneous nodes is a nested split graph network $g$ such that for all $i, j \in N$ $(i \neq j)$,*

$$v_i > v_j \Rightarrow \left|N_i^d(g)\right| \geq \left|N_j^d(g)\right|. \tag{8}$$

Thus when nodes have different values, efficiency is achieved through an NSG-network with a double hierarchical structure: the greater the value of a node, the greater the number of neighbors that node has.

**Homogeneous nodes and heterogeneous links**

When links are heterogenous, an obvious prerequisite for a link-investment $\mathbf{c}$ to yield an efficient network is that for each link, if $c > 0$ is the total amount invested in link $ij$, its strength must be

$$\delta(c) := \max_{c_{ij}+c_{ji} \leq c} \lambda(c_{ij}, c_{ji}). \tag{9}$$

Olaizola and Valenciano (2020b) assumes a technology s.t. function $\delta(c)$, given by (9), is non-decreasing and, using the results in Olaizola and Valenciano (2020a), prove that unless a rare condition holds, *the only possibly non-empty efficient networks with homogeneous nodes are the complete network and the all-encompassing star.*[5] They call "supertie" the occurrence of this rare condition. In this case there is a draw between the all-encompassing star, the complete network *and* a whole range of a particular type of nested split graph structures intermediate between them. In Olaizola and Valenciano (2022) it is shown that this cannot occur if $\delta$ is strictly concave, differentiable, and s.t. and $\delta'(c) > 0$, for all $c \geq 0$.

# 5 Heterogeneous nodes and heterogeneous links

We now address the question of efficiency when both nodes and links are heterogeneous. More precisely, we assume that nodes have possibly different values and links are formed according to a technology $\lambda$ s.t. function $\delta$, given by (9), is well-defined and non-decreasing. We refer to such technologies as *regular* technologies. We use a constructive procedure in order to identify the structure of efficient networks, which turns out to be a particular type of flower that we call *hierarchical flower.*

**Definition 4** *A hierarchical flower network is a flower such that: (i) the central node is one of the nodes of greatest value; (ii) the greater the sum of the values of a pair of nodes, the stronger their connection, be it direct or through the central node; in particular, the greater the value of a node, the greater the strength of the link that connects it with the central one.*

We have the following result:

**Theorem 1** *If nodes are possibly heterogeneous and links are formed according to a regular technology, efficiency is reached by a hierarchical flower.*

The proof is based on an algorithm that shows that any network is dominated by a hierarchical flower. In order to precisely describe the algorithm, it is necessary to introduce some notation and a preliminary discussion convenient.

[5] Olaizola and Valenciano (2020a) proves that within a wide class of connections models with homogeneous nodes any network is dominated by a nested split graph network.

Let $g$ be a network with $n$ nodes, $m$ links and a positive net value (otherwise, it is dominated by the empty network of zero net value). Assume that nodes are numbered so that $v_1 \geq v_2 \geq ... \geq v_n$. Let $\delta_1 \geq \delta_2 \geq ... \geq \delta_m$ be the strengths of the $m$ links that form $g$, i.e. $\delta_i = \delta(c_i)$ with $c_1 \geq c_2 \geq ... \geq c_m$. The idea of the proof is to produce a hierarchical flower $g'$ s.t. $v(g') \geq v(g)$ by using all or some of the $m$ available links in $g$. The net value of the network, given by (3), is generated by all the pairs of nodes directly of indirectly connected, namely the contribution of pair $ij$ is $\delta(\overline{p}_{ij})s_{ij}$. Quantities $s_{ij}$ can be arranged in a triangular matrix $\Sigma = (s_{ij})_{i,j \in N, i<j}$, i.e.

$$\Sigma = \begin{bmatrix} s_{12} & s_{13} & s_{14} & \cdots & s_{1n} \\ & s_{23} & s_{24} & \cdots & s_{2n} \\ & & s_{34} & \cdots & s_{3n} \\ & & & \ddots & \vdots \\ & & & & s_{n-1n} \end{bmatrix}, \tag{10}$$

and are preordered by their numerical values. Similarly, the strengths of potential connections, direct or by two-link paths, which could be made by means of the $m$ available links, $\delta_1, \delta_2, ..., \delta_m$, can be arranged in a triangular matrix

$$\Delta = \begin{bmatrix} \delta_1 & \delta_2 & \delta_3 & \delta_4 & \cdots & \delta_m \\ & \delta_1\delta_2 & \delta_1\delta_3 & \delta_1\delta_4 & \cdots & \delta_1\delta_m \\ & & \delta_2\delta_3 & \delta_2\delta_4 & \cdots & \delta_2\delta_m \\ & & & \delta_3\delta_4 & \cdots & \delta_3\delta_m \\ & & & & \ddots & \vdots \\ & & & & & \delta_{m-1}\delta_m \end{bmatrix}, \tag{11}$$

and are preordered by their numerical values. Note that, in both matrices, the entries in rows (in columns) are non-increasing rightwards (downwards). If these $m^2/2$ strengths of potential connections in $\Delta$, already paid for in $g$, were actually independently available, the maximal gross value would be achieved by matching sequentially the best pair of nodes and the best connection. That is, connecting pairs of nodes according to a simple criterion: the greater $s_{ij}$, the greater the strength of the connection of $i$ and $j$ (be it direct or indirect) should be.[6] Unfortunately, this is not possible.[7] After the first two steps, i.e. after connecting 1 and 2 with $\delta_1$, and 1 and 3 with $\delta_2$, connection $\delta_1\delta_2$ is only available for connecting 2 and 3. Now, if $\delta_3 > \delta_1\delta_2$ it can be advantageous to connect 1 and 2 directly with $\delta_3$. However, if there are more nodes, what is most profitable in this case: connecting 1 and 2 with $\delta_3$ or connecting 1 and 4 with $\delta_3$? The problem is that a comparison of the net values of these two three-link networks is *not* reliable, in the sense that connecting 1 and 2 with $\delta_3$ entails eliminating entries

[6] We prove later that the results are not altered if we consider included in $\Delta$ the strengths of all potential connections of any length.

[7] In Olaizola and Valenciano (2021) this is possible because under homogeneous links all direct connections have the same strength, and consequently those by means of paths of length 2 too.

$\delta_3$, $\delta_1\delta_2$, $\delta_1\delta_3$ and $\delta_2\delta_3$ in $\Delta$ because they would not be available any longer. *If there are still disconnected nodes and available links, which of the two options is preferable depends on the subsequent choices of which among the not directly connected pairs of nodes to connect with the subsequently available links.*[8]

We describe an algorithm to produce a network with the available links (i.e. some of those which form $g$) that dominates $g$. Actually, the algorithm generates sequences of hierarchical flowers. At every step, either a new node is connected to node 1, or a pair of spokes is directly connected, although in some cases the algorithm "does nothing" and in some others "branches" or bifurcates and continues in two different directions, producing sequences of hierarchical flowers. The algorithm stops in every sequence when all available links have been used or even before if no pair of nodes remains in the "waiting list" to be directly connected. In this way, a number of hierarchical flower networks is produced among which the one with a greatest net value dominates $g$. In order to describe the algorithm, one must keep account of the links used at each stage, the indirect connections discarded (either because they will remain unused or are not available any more), and the pairs of nodes connected either directly or indirectly if their two-link connection is sure that is not going to be improved in the sequence. Let $\mathcal{S}$ be the set of entries in $\Sigma$, i.e. the set of sums of values of all pairs of nodes to be connected, directly or indirectly; and let $\mathcal{S}^* := \{s_{ij} \in \mathcal{S} : i = 1\}$. Let $\mathcal{D}$ be the set of entries in $\Delta$, i.e. the set of all, in principle feasible one-link and two-link connections; and let $\mathcal{D}^* := \{\delta_j : 1 \leq j \leq m\}$. For a rigorous description of the algorithm it is necessary to provide sets $\mathcal{S}$ and $\mathcal{D}$ with total orders that refine the preorder given by their numerical values.[9] In order to do so proceed in the following way. In the case of $\mathcal{S}$, define the strict total order: $s_{ij} \succ s_{kl}$ if and only if $s_{ij} > s_{kl}$ or $s_{ij} = s_{kl}$ with $ij \neq kl$ and

$$(i \leq k \ \& \ j \leq l) \text{ or } (i > k \ \& \ j < l),$$

whose reflexive closure $\succeq$ is a total order in $\mathcal{S}$.[10]

A similar refinement of the order in set $\mathcal{D}$ is defined breaking ties in the following way. In the case of direct connections just define $\delta_1 \succ \delta_2 \succ ... \succ \delta_m$; in the case of

[8]For instance, if there are 4 nodes with $v_1 = v_2 = v_3 = 100$ and $v_4 = 4$, and 3 available links $\delta_1 = \delta_2 = 0.2$, $\delta_3 = 0.15$. A subcomplete network $g$ that connects 1, 2 and 3 (2 and 3 with a $\delta_3$-link) and leaves 4 disconnected yields a greater net value than an all-encompassing star $g'$ centered at node 1 and connecting the center and node 4 with a $\delta_3$-link. However, if there is a fifth node with $v_5 = 4$ and a fourth available link $\delta_4 = 0.15$, then the network that results from adding a $\delta_4$-link connecting nodes 1 and 4 to network $g$ has a smaller net value than the network that results from adding a $\delta_4$-link connecting nodes 1 and 5 to network $g'$.

[9]Sets $\mathcal{S}$ and $\mathcal{D}$ have been defined as sets of numbers, but strictly formally speaking, these sets should be specified and interpreted as "labelled numbers" because we need to distinguish $s_{ij}$ from $s_{kl}$ ($\delta_i$ from $\delta_j$ or $\delta_k\delta_l$) even if $s_{ij} = s_{kl}$ when $ij \neq kl$ (even if $\delta_i = \delta_j$ when $i \neq j$ or $\delta_j = \delta_k\delta_l$). This can be done by defining them as sets of pairs $(s_{ij}, ij)$ or $(\delta_i, i)$ or $(\delta_i\delta_j, i, j)$. Nevertheless, we avoid this cumbersome notational formality. Just keep in mind that $i < j$ and $v_i \geq v_j$ whenever we write $s_{ij}$, and $\delta_i \geq \delta_j$ whenever we write $\delta_i\delta_j$.

[10]A total or linear order is a reflexive, antisymmetric, transitive and complete or total binary relation.

two-link connections define $\delta_i\delta_j \succ \delta_k\delta_l$ if and only if $\delta_i\delta_j > \delta_k\delta_l$ or $\delta_i\delta_j = \delta_k\delta_l$ with $ij \neq kl$ and

$$(i \leq k \ \& \ j \leq l) \text{ or } (i > k \ \& \ j < l);$$

and in the case of a tie between a two-link connection and a direct connection, define $\delta_i\delta_j \succ \delta_l$ when $\delta_i\delta_j = \delta_l$ only if $j < l$. Then the reflexive closure $\succeq$ of $\succ$ is a total order in $\mathcal{D}$.

**The algorithm**

Let $g$ be a network with positive net value, let $G^q$ denote the set of networks generated at step $q$ from $g$ and let $\mathcal{S}_q$, $\mathcal{S}_q^*$, $\mathcal{D}_q$ and $\mathcal{D}_q^*$ the subsets of $\mathcal{S}$, $\mathcal{S}^*$, $\mathcal{D}$ and $\mathcal{D}^*$ that remain available after step $q$.

Initially, let $G^0 := \{g^\varnothing\}$, where $g^\varnothing$ is the empty network, $\mathcal{S}_0 := \mathcal{S}$, $\mathcal{S}_0^* := \mathcal{S}^*$, $\mathcal{D}_0 := \mathcal{D}$ and $\mathcal{D}_0^* := \mathcal{D}^*$.

*Step* $q$ $(1 \leq q \leq m)$: Let $G^{q-1}$ be the set of networks generated at step $q-1$. For each $g^{q-1} \in G^{q-1}$, proceed as follows. Let $\mathcal{S}_{q-1}, \mathcal{S}_{q-1}^*, \mathcal{D}_{q-1}, \mathcal{D}_{q-1}^*$ be updated sets of sums of pairs' values and of available connections corresponding to the sequence that reached $g^{q-1}$.[11] Take the greatest element in $\mathcal{S}_{q-1}$ (according to the total order defined in $\mathcal{S}$), which must be either some $s_{1j} \in \mathcal{S}_{q-1}^*$ or some $s_{ij} \in \mathcal{S}_{q-1} \setminus \mathcal{S}_{q-1}^*$ (i.e. $i \neq 1$), and the greatest element in $\mathcal{D}_{q-1}$ (according to the total order defined in $\mathcal{D}$), which may be either some $\delta_r \in \mathcal{D}_{q-1}^*$ or some $\delta_k\delta_l \in \mathcal{D}_{q-1} \setminus \mathcal{D}_{q-1}^*$.[12] Four cases are possible depending on which of the four possible combinations forms the greatest product,

$$s_{1j}\delta_r, \ \ s_{ij}\delta_r, \ \ s_{1j}\delta_k\delta_l \text{ or } s_{ij}\delta_k\delta_l.$$

*Case* 1: The greatest product is $s_{1j}\delta_r$. Let $g^q$ be the network that results from adding a $\delta_r$-link connecting 1 and $j$ to $g^{q-1}$, make $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{1j}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_r\}$, and consequently $\mathcal{S}_q^* := \mathcal{S}_{q-1}^* \setminus \{s_{1j}\}$ and $\mathcal{D}_q^* := \mathcal{D}_{q-1}^* \setminus \{\delta_r\}$. If $\mathcal{D}_q^* = \varnothing$, Stop, otherwise proceed to step $q+1$.

*Case* 2: The greatest product is $s_{ij}\delta_r$ with $i \neq 1$. In this case, $i$ and $j$ must be already directly connected with node 1 in $g^{q-1}$, say $s_{1i}\delta_p$ and $s_{1j}\delta_s$ have already been chosen. Then, the algorithm *possibly branches into two paths*:

*Branch* 1: Connect $i$ and $j$ with a $\delta_r$-link, update $g^q$, and make $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{ij}\}$, make $\mathcal{S}_q^* := \mathcal{S}_{q-1}^*$,

$$\mathcal{D}_q := \mathcal{D}_{q-1} \setminus (\{\delta_r, \delta_p\delta_s, \delta_p\delta_r, \delta_s\delta_r\} \cup \{\delta_r\delta_u : \forall u > r\}),$$

[11] A proper but cumbersome notation would be $\mathcal{S}_{q-1}(g^{q-1})$, $\mathcal{S}_{q-1}^*(g^{q-1})$... etc.

[12] The first step selects $s_{12}$ and $\delta_1$, and connects nodes 1 and 2 with a $\delta_1$-link, and the second step selects $s_{13}$ and $\delta_2$, and connects nodes 1 and 3 with a $\delta_2$-link, so that

$$\mathcal{S}_2 = \mathcal{S}_0 \setminus \{s_{12}, s_{13}\}, \mathcal{S}_2^* := \mathcal{S}_0^* \setminus \{s_{12}, s_{13}\},$$

$$\mathcal{D}_2 := \mathcal{D} \setminus \{\delta_1, \delta_2\}; \mathcal{D}_2^* := \mathcal{D}_0^* \setminus \{\delta_1, \delta_2\}.$$

The network $g^2$ that results is a hierarchical flower.

$\mathcal{D}_q^* := \mathcal{D}_{q-1}^* \setminus \{\delta_r\}$. If $\mathcal{D}_q^* = \varnothing$, Stop, otherwise proceed to step $q+1$.

*Branch* 2: If there are still disconnected nodes in $g^{q-1}$, i.e. $\mathcal{S}_{q-1}^* \neq \varnothing$, proceed also as follows. Let $k$ be the node of greatest value still disconnected (i.e. $s_{1k} = \max \mathcal{S}_q^*$). Form $g^q$ by adding a $\delta_r$-link connecting 1 and $k$ to $g^{q-1}$, make $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{1k}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_r\}$, and consequently $\mathcal{S}_q^* := \mathcal{S}_{q-1}^* \setminus \{s_{1k}\}$ and $\mathcal{D}_q^* := \mathcal{D}_{q-1}^* \setminus \{\delta_r\}$. If $\mathcal{D}_q^* = \varnothing$, Stop, otherwise proceed to step $q+1$.

Thus in *Case* 2, the algorithm bifurcates as far as $g^{q-1}$ is not connected, while *only branch* 1 applies whenever $g^{q-1}$ is connected.

*Case* 3: The greatest product is $s_{1j}\delta_k\delta_l$. Then $\delta_k$ and $\delta_l$ must have been previously assigned to connect node 1 with two nodes, say, $p$ and $s$. Let $\delta_r$ be the strength of the link of greatest strength available in $\mathcal{D}_{q-1}^*$, and let $g^q$ be the network that results from $g^{q-1}$ by adding a $\delta_r$-link connecting 1 and $j$, and make $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{1j}, s_{ps}\}$, $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_r, \delta_k\delta_l\}$, and consequently $\mathcal{S}_q^* := \mathcal{S}_{q-1}^* \setminus \{s_{1j}\}$ and $\mathcal{D}_q^* := \mathcal{D}_{q-1}^* \setminus \{\delta_r\}$. If $\mathcal{D}_q^* = \varnothing$, Stop, otherwise proceed to step $q+1$.

*Case* 4: The greatest product is $s_{ij}\delta_k\delta_l$ with $i \neq 1$ and $\mathcal{D}_q^* \neq \varnothing$ and $\mathcal{S}_q \neq \varnothing$. Let $g^q := g^{q-1}$, $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{ij}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_k\delta_l\}$, and consequently $\mathcal{S}_q^* := \mathcal{S}_{q-1}^*$ and $\mathcal{D}_q^* := \mathcal{D}_{q-1}^*$, and proceed to step $q+1$.

The algorithm stops through any possible path after a finite number of steps when all links have been used, i.e. $\mathcal{D}_q^* = \varnothing$, or the waiting list of pairs of nodes is empty, i.e. $\mathcal{S}_q = \varnothing$. As a result, it produces a set of networks. Namely, if $G^q(g^{q-1})$ denotes the network/s (two in case of branching) generated from $g^{q-1}$ as a result of step $q$, then the set of networks generated at step $q$ can be expressed recursively like this

$$G^q := \bigcup_{g^{q-1} \in G^{q-1}} G^q(g^{q-1}).$$

Then we have the following result:

**Proposition 1** *If nodes are possibly heterogeneous and links are formed according to a regular link-formation technology, the network with a greatest net value among the terminal networks that result from applying the algorithm described to any network $g$ with positive net value is a hierarchical flower that dominates $g$.*

We prove first that the algorithm produces a set of hierarchical flowers as terminal networks. In fact, we prove that it produces sequences of such flowers, because at every step, in all cases, the algorithm produces hierarchical flower networks.

**Lemma 1** *At every step the algorithm produces a hierarchical flower.*

**Proof.** We proceed by induction on the number of steps. At steps 1, 2 and 3 this is obvious.[13] Assume that after $q-1$ steps a hierarchical flower $g^{q-1}$ is produced. At

[13] At steps 1 and 2 this is obvious whatever the number of nodes. At step 3 (only possible if $n \geq 3$ and $m \geq 3$), the algorithm yields a star centered at one of the most valuable nodes which connects the four most valuable nodes and/or a triangle connecting the three most valuable nodes (leaving disconnected the others if $n > 3$). In both cases, a flower satisfying condition (*ii*) of Definition 4 is produced.

every step, the algorithm either does nothing, just eliminating a pair and a two-link connection from the set of available ones (Case 4) or makes one of these two actions with the strongest available link: $(a)$ Connects node 1 and the most valuable isolated node (Case 1, Case 2-Branch 1, Case 3), or $(b)$ Connects two spoke nodes not directly connected whose sum of values is the highest (Case 2-Branch 2).

In case $(a)$, the algorithm connects node 1 and the most valuable disconnected node, say $j$, with the best available link, say $\delta_r$. This entails connecting $j$ with all spokes through node 1, and evidently $g^{q-1}$ is a new flower. Moreover, as the link used is (weakly) weaker than any of the links that form $g^{q-1}$, which is a hierarchical flower, it follows immediately that $g^q$ is also hierarchical because the new connections (of all connected nodes in $g^{q-1}$ with $j$) satisfy condition $(ii)$ of Definition 4.

In case $(b)$, the algorithm connects a pair of spokes, say $i$ and $j$, with a highest sum of values with the best available link, say $\delta_r$. In this case, this connection, made with the link of a weakest strength among those that form $g^q$, has no externalities, i.e. only affects $i$ and $j$, and the result is obviously a hierarchical flower. Therefore, the algorithm generates finite sequences of hierarchical flowers.

**Proof of Proposition 1**: Let $g^{q-1}$ be the hierarchical flower after step $q-1$, and $\mathcal{S}_{q-1}$ the set of pairs of not directly connected nodes and $\mathcal{D}_{q-1}$ the set of available connections.

If $s := \max \mathcal{S}_{q-1}$, there are two possibilities: either $s \in \mathcal{S}^*_{q-1}$, i.e. $s = s_{1j}$ for some $j$, which must be an isolated node; or $s \in \mathcal{S}_{q-1} \setminus \mathcal{S}^*_{q-1}$, i.e. $s = s_{ij}$ for some pair $i, j$ of not directly connected spokes that must be connected with the center because $s_{ij} \prec \min\{s_{1i}, s_{1j}\}$.

Consider the two possible best available connections: either $\delta_r = \max \mathcal{D}_{q-1}$ for some $\delta_r \in \mathcal{D}^*_{q-1}$, or $\delta_k\delta_l = \max \mathcal{D}_{q-1}$ for some pair $\delta_k, \delta_l \in \mathcal{D}^*_{q-1}$. In the second case, both must have been chosen previously because $\delta_k\delta_l < \min\{\delta_k, \delta_l\}$. Moreover, both $\delta_k$ and $\delta_l$ must have been chosen to connect node 1 with other two nodes. Otherwise, if $\delta_k$ connects node 1 and node $p$, and $\delta_l$ nodes $p$ and $s$, nodes 1 and $s$ must be connected too because $s_{1s} \succ s_{ps}$, and the links and two-link connections in the triangle should have been eliminated when $\delta_l$ was chosen. A similar reasoning discards the possibility of both $\delta_k$ and $\delta_l$ connecting spokes. Thus, there are only two options: either $\delta_r = \max \mathcal{D}_{q-1}$ for some $\delta_r \in \mathcal{D}^*_{q-1}$ or $\delta_k\delta_l = \max \mathcal{D}_{q-1}$ for some pair $\delta_k, \delta_l \in \mathcal{D}^*_{q-1}$ which connect node 1 with two nodes.

Thus there are four cases corresponding to the four possible ways of matching the best pair of nodes not directly connected and the best available connection.

*Case* 1: The greatest product is $s_{1j}\delta_r$. Since $g^{q-1}$ is a hierarchical flower and $s_{1j}$ is the largest element $\mathcal{S}_{q-1}$, $j$ must be an isolated node in $g^{q-1}$. Therefore, matching the best available link $\delta_r$ and the best available pair $s_{1j}$ is feasible, and $g^q$ is the network that results from adding a $\delta_r$-link connecting 1 and $j$ to $g^{q-1}$. As 1 and $j$ are now connected and $\delta_r$ used, the available pairs and links are now $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{1j}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_r\}$, and consequently $\mathcal{S}^*_q := \mathcal{S}^*_{q-1} \setminus \{s_{1j}\}$ and $\mathcal{D}^*_q := \mathcal{D}^*_{q-1} \setminus \{\delta_r\}$.

*Case* 2: The greatest product is $s_{ij}\delta_r$ with $i \neq 1$. In this case, $i$ and $j$ are already directly connected with node 1 in $g^{q-1}$, by previous selection of, say, $s_{1i}\delta_p$ and $s_{1j}\delta_s$, and the algorithm possibly branches into two paths:

*Branch* 1: Connect $i$ and $j$ with a $\delta_r$-link, update $g^q$, and consequently updating pair's sums results $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{ij}\}$ and $\mathcal{S}_q^* := \mathcal{S}_{q-1}^*$.[14] As to links, this action leads to discard evidently $\delta_r, \delta_p\delta_s, \delta_p\delta_r, \delta_s\delta_r$, but also $\delta_r\delta_u$ for all $u > r$ because it is not possible that $\delta_r\delta_u$ is selected at a posterior step. The reason is the following, whatever the sequence of flowers generated by the algorithm from $g^{q-1}$ and whatever the network in such sequence $g^{q+t}$, $g^{q-1}$ is necessarily a subnetwork of $g^{q+t}$. Thus when $\delta_r\delta_u$ is selected $\delta_u$ must have been previously selected. Then $\delta_r\delta_u$ is feasible only if $\delta_u$ connects $j$ and some spoke $k$, but this is not possible because either $i$ and $k$ are already directly connected or $s_{ik}\delta_u$ should have been selected before $s_{jk}\delta_u$. Thus

$$\mathcal{D}_q := \mathcal{D}_{q-1} \setminus (\{\delta_r, \delta_p\delta_s, \delta_p\delta_r, \delta_s\delta_r\} \cup \{\delta_r\delta_u : \forall u > r\}).$$

If $\mathcal{S}_q^* = \varnothing$, i.e. flower $g^{q-1}$ is all-encompassing (note that $\mathcal{S}_q^* = \mathcal{S}_{q-1}^*$), or $\mathcal{D}_q^* = \varnothing$, i.e. $\delta_r$ is the last available link, this is the best choice at this stage given the preceding sequence of choices. Otherwise, as illustrated by the example of footnote 8, when there are still disconnected nodes in $g^{q-1}$ and available links, this is not necessarily so. For this reason, in this case the algorithm bifurcates:

*Branch* 2: Let $k$ be the node of greatest value still disconnected. Form $g^q$ by adding a $\delta_r$-link connecting 1 and $k$ to network $g^{q-1}$, and update $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{1k}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_r\}$, and consequently $\mathcal{S}_q^* := \mathcal{S}_{q-1}^* \setminus \{s_{1k}\}$ and $\mathcal{D}_q^* := \mathcal{D}_{q-1}^* \setminus \{\delta_r\}$. This branching allows control of the possibility of *regret*, since when $s_{ij}\delta_r$ $(i \neq 1)$ is implemented some "good" paths are disabled to be used in future steps.

*Case* 3: The greatest product is $s_{1j}\delta_k\delta_l$. In this case, as has been shown, $\delta_k$ and $\delta_l$ must have been already assigned to connect two pairs of nodes with node 1, and $j$ must be an isolated node in $g^{q-1}$. Thus matching $s_{1j}$ and $\delta_k\delta_l$ is not feasible. Then the greatest feasible connection is by means of the link of greatest strength available in $\mathcal{D}_{q-1}^*$, say $\delta_r$. Then $g^q$ is the network that results from $g^{q-1}$ by adding a $\delta_r$-link connecting 1 and $j$. Note that in this case $s_{ps}\delta_k\delta_l \geq s_{ps}\delta_r$, in other words, no link available can improve the connection of $p$ and $s$, hence the updating of $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{1j}, s_{ps}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_r, \delta_k\delta_l\}$, and consequently $\mathcal{S}_q^* := \mathcal{S}_{q-1}^* \setminus \{s_{1j}\}$ and $\mathcal{D}_q^* := \mathcal{D}_{q-1}^* \setminus \{\delta_r\}$.

*Case* 4: The greatest product is $s_{ij}\delta_k\delta_l$ with $i \neq 1$. In this case, as has been shown, $\delta_k$ and $\delta_l$ are used in $g^{q-1}$ to directly connect node 1 with two spokes, say $r$ and $s$. Thus the matching is feasible only if $r = i$ and $s = j$, which is necessarily so as we show, i.e. in $g^{q-1}$ nodes 1 and $i$ are already connected by a $\delta_k$-link and nodes 1 and $j$

[14]Note that it cannot be the case that connecting $i$ and $j$ like this does not improve their connection, because if $\delta_p\delta_s \geq \delta_r$, then $\delta_p\delta_s \succ \delta_r$, i.e. $s_{kl}\delta_p\delta_s$ should have been previously chosen for some $kl$, which must be $ij$ and consequently $s_{ij}$ would not be available.

by a $\delta_l$-link.[15] Assume $s_{1r}\delta_k$ and $s_{1s}\delta_l$, with $r \neq i$ or $s \neq j$, were assigned in former steps. Then it must be $s_{rs} \succ s_{ij}$, but then $\delta_k\delta_l$ cannot be in $\mathcal{D}_{q-1}$, because $s_{rs}\delta_k\delta_l$ has been assigned in a previous step and $\delta_k\delta_l$ eliminated as unavailable. Therefore, it must be $r = i$ and $s = j$. In other words, $i$ and $j$ are already connected by the best available connection, so the algorithm "does nothing", i.e. $g^q := g^{q-1}$, only updates $\mathcal{S}_q := \mathcal{S}_{q-1} \setminus \{s_{ij}\}$ and $\mathcal{D}_q := \mathcal{D}_{q-1} \setminus \{\delta_k\delta_l\}$, and $\mathcal{S}^*_q := \mathcal{S}^*_{q-1}$ and $\mathcal{D}^*_q := \mathcal{D}^*_{q-1}$.

The algorithm stops when along a path when $\mathcal{D}^*_q = \varnothing$ or $\mathcal{S}_q = \varnothing$, and generates a set $G^q$ of hierarchical flowers as a result of step $q$ from $g^{q-1}$,

$$G^q := \cup_{g^{q-1} \in G^{q-1}} G^q(g^{q-1}).$$

Evidently, given the finite number of feasible links and pairs of nodes to be connected directly or indirectly, the algorithm, along through all its possible paths, ends necessarily after a finite number of steps.

At every step, in case 1, the algorithm matches the best pair of nodes and the best available link; in case 2, branch 1, matches the best pair, in this case two spokes and the best available link, while in case there are still disconnected nodes, in case 2, branch 2, matches the best available link to connect the central node and the best disconnected one; in case 3, the algorithm matches the best pair, in this case the central node and the best still disconnected node and the best actually available link; in case 4: the algorithm checks the match already done of the best pair and the best two-link connection via node 1.

The impossibility of anticipating which of the two branches can lead to a more valuable network in case 2 would make it necessary to complete all paths in a finite binary tree in order to actually find a network dominating $g$. However, this is not the case as far as we are sure that the best terminal network is sure to dominate $g$, which is the point of the algorithm. And this is so because in all paths, at every step, the algorithm matches the best pair of nodes and the best available connection, with the only exception of case 2 branch 2, which is a caution against the case this can turn out to be a suboptimal choice.

Remains to be discussed the exclusion of connections of length 3 or more. Matrix $\Delta$ and set $\mathcal{D}$ include only paths of length 1 and 2. However, if we had included in $\mathcal{D}$ all possible paths of any length, those with length greater than 2 would never had been chosen. Assume $\mathcal{D}$ includes paths of any length. Let $\delta_p\delta_r\delta_s$ be any 3-link connection available in $\mathcal{D}_{q-1}$ actually connecting a pair of nodes $i$ and $j$ in hierarchical flower $g^{q-1}$. There are three possible cases. Case $(a)$: all three links connect spokes, but in this case $i$ and $j$ are already better connected through the central node. Case $(b)$: $\delta_p$ connects $i$ and 1, $\delta_r$ connects 1 and a spoke, and $\delta_s$ this spoke and $j$, but in this case $i$ and $j$ are

[15] When $s_{1s}\delta_l$ is implemented there is no chance of regret. The possibility of "branching" in a previous step, where $s_{1s}\delta_l$ has been implemented, by alternatively connecting $p$ and $j$ with $\delta_l$ so as to make use of the good connection $\delta_k\delta_l$ to connect 1 and $j$, and later using $\delta_r$ to connect 1 and $s$, it does not work: this network has a strictly smaller net value than $g^q$, and no subsequent addition of available links can reverse this situation.

already better connected through their links with the intermediate spoke. Case $(c)$: $i = 1$, which is better connected with $j$ directly. As for connections of length greater that 3, they contain 3-link connections suboptimal and therefore cannot be optimal.

Then the terminal network with a greatest net value generated by the algorithm dominates $g$. ■

Thus Theorem 1 follows immediately: Efficiency is sure to be reached by a hierarchical flower.

**Comments:**

(i) What is "magic" in the result is that if, for all $i, j \in N$ $(i < j)$, $\delta_{ij}$ denotes the strength of the connection (direct or indirect) of nodes $i$ and $j$ in an efficient flower, we have a triangular matrix:

$$\begin{array}{ccccccccc}
 & 2 & 3 & 4 & 5 & \cdots & j & \cdots & n \\
1 & \delta_{12} & \delta_{13} & \delta_{14} & \delta_{15} & \cdots & \delta_{1j} & \cdots & \delta_{1n} \\
2 & & \delta_{23} & \delta_{24} & \delta_{25} & \cdots & \delta_{2j} & \cdots & \delta_{2n} \\
3 & & & \delta_{34} & \delta_{35} & \cdots & \delta_{3j} & \cdots & \delta_{3n} \\
4 & & & & \delta_{45} & \cdots & & & \vdots \\
\vdots & & & & & \ddots & & & \vdots \\
i & & & & & & \delta_{ij} & & \delta_{in} \\
 & & & & & & & & \vdots \\
n\text{-}1 & & & & & & & & \delta_{n-1,n}
\end{array},$$

where for all $j = 2, ..., n$, $\overline{p}_{1j}$ is one-link path, i.e. $\overline{p}_{1j} = 1j \in g$, so that $\delta_{1j} = \delta(\overline{p}_{1j}) = g_{1j}$; while for all $ij$ $(1 < i < j)$, $\overline{p}_{ij}$ is either one-link path, i.e. $\overline{p}_{ij} = ij \in g$ and $\delta_{ij} = \delta(\overline{p}_{ij}) = g_{ij}$, or $\overline{p}_{ij}$ is a two-link path $\overline{p}_{ij} = i1j$, so that $\delta_{ij} = \delta(\overline{p}_{1i})\delta(\overline{p}_{1j})$, but *in all cases*, i.e. for all $i, j \in N$ $(i < j)$ and all $k, l \in N$ $(k < l)$

$$v_i + v_j > v_k + v_l \Rightarrow \delta_{ij} \geq \delta_{kl}.$$

(ii) Note that, as it could not be otherwise, the result is consistent with previous results obtained in simpler contexts. The question arises about whether it is possible to go further and say something else about the structure of efficient flowers. As has been commented above, in Olaizola and Valenciano (2021) it is proved that under a homogeneous technology (i.e. all links of the same cost and strength) the only possible non-empty efficient networks of heterogeneous nodes are strong nested split graph networks (SNSG-networks, Definition 3). The following counterexample shows that under heterogeneity *in nodes and links* it may be the case that no SNSG-network is efficient.

**Counterexample:** In order to provide a simple example we assume a regular link-formation technology (according to Definition 2) $\lambda : \mathbb{R}^2_+ \to [0, 1)$, s.t. function $\delta$, given by (9), is piecewise linear given by

$$\delta(c) = \begin{cases} 0.5c, \ \text{if } c \leq \frac{2.96}{8.95}, \\ \\ \frac{1}{19.9}c + \frac{2.96}{19.9}, \ \text{if } \frac{2.96}{8.95} \leq c \leq 5, \\ \\ 0.4, \ \text{if } 5 \leq c, \end{cases} \tag{12}$$

which is represented in Figure 1, and denote $\underline{c} := \frac{2.96}{8.95}$ and $\overline{c} := 5$.

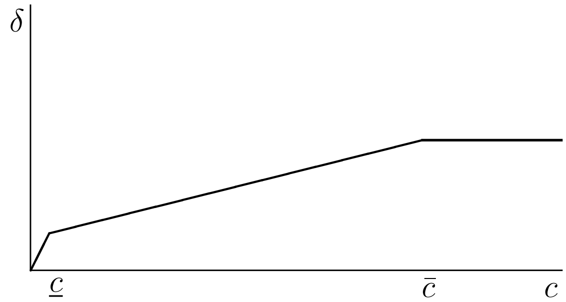


Figure 1: Function $\delta$ for technology $\lambda$

Assume a set of 4 nodes, s.t. $v_1 = v_2 = v_3 = 10$ and $v_4 = 1$. Note first that, if $g^{\Delta}(c)$ denotes the subcomplete network of nodes $1, 2$ and $3$ connected by $c$-links, and $g^{\Lambda}(c)$ any star formed by these three nodes connected by $c$-links, we have

$$v(g^{\Delta}(c)) = 3s\delta(c) - 3c \quad \text{and} \quad v(g^{\Lambda}(c)) = 2s\delta(c) + s\delta(c)^2 - 2c,$$

where $s = 2v_i = 20$ $(i = 1, 2, 3)$.[16] Then we have

$$\begin{aligned} v(g^{\Delta}(\underline{c})) &= 60\delta(\underline{c}) - 3\underline{c} = 60 \times \frac{1.48}{8.95} - 3 \times \frac{2.96}{8.95} = 8.929\,6, \\ v(g^{\Delta}(\overline{c})) &= 60\delta(\overline{c}) - 3\overline{c} = 60 \times 0.4 - 3 \times 5 = 9. \end{aligned}$$

While

$$\begin{aligned} v(g^{\Lambda}(\underline{c})) &= 2s\delta(\underline{c}) + s\delta(\underline{c})^2 - 2\underline{c} = 40 \times \frac{1.48}{8.95} + 20 \times (\frac{1.48}{8.95})^2 - 2 \times \frac{2.96}{8.95} = 6.5, \\ v(g^{\Lambda}(\overline{c})) &= 2s\delta(\overline{c}) + s\delta(\overline{c})^2 - 2\overline{c} = 40 \times 0.4 + 20 \times 0.16 - 2 \times 5 = 9.\,2. \end{aligned}$$

Therefore $g^{\Delta}(\overline{c})$ is the optimal subcomplete network connecting nodes $1, 2$ and $3$, and $g^{\Lambda}(\overline{c})$ the optimal all three nodes encompassing star (up to permutation of the 3 nodes), and

$$v(g^{\Delta}(\overline{c})) = 3s\delta(\overline{c}) - 3\overline{c} = 9 < 9.2 = 2s\delta(\overline{c}) + s\delta(\overline{c})^2 - 2\overline{c} = v(g^{\Lambda}(\overline{c})).$$

The four possible infrastructures underlying SNSG-networks connecting nodes $1, 2, 3$ and 4 are represented in Figure 2. A simple calculation shows that the investment

[16] Both an optimal star or an optimal complete network of homogeneous nodes have all links of the same strength.

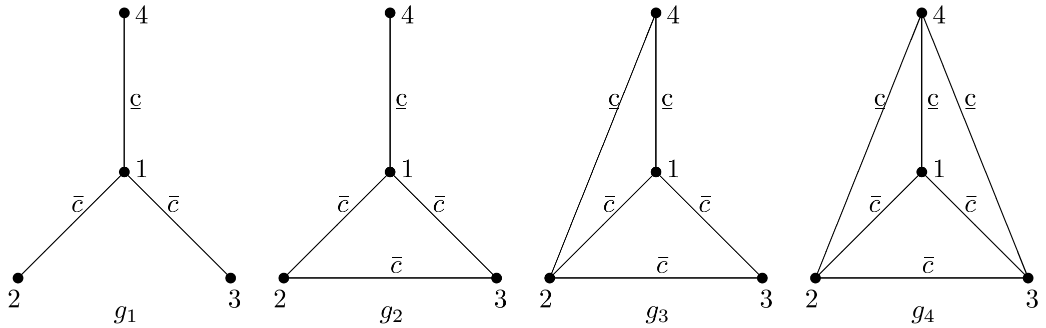


Figure 2: Connected SNSG-infrastructures for $n = 4$

vectors which support each of them optimally, i.e. maximizing the net value, are the following:

| $g_1$ | 2 | 3 | 4 |
|---|---|---|---|
| 1 | $\overline{c}$ | $\overline{c}$ | $\underline{c}$ |
| 2 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 |

,

| $g_2$ | 2 | 3 | 4 |
|---|---|---|---|
| 1 | $\overline{c}$ | $\overline{c}$ | $\underline{c}$ |
| 2 | 0 | $\overline{c}$ | 0 |
| 3 | 0 | 0 | 0 |

,

| $g_3$ | 2 | 3 | 4 |
|---|---|---|---|
| 1 | $\overline{c}$ | $\overline{c}$ | $\underline{c}$ |
| 2 | 0 | $\overline{c}$ | $\underline{c}$ |
| 3 | 0 | 0 | 0 |

and

| $g_4$ | 2 | 3 | 4 |
|---|---|---|---|
| 1 | $\overline{c}$ | $\overline{c}$ | $\underline{c}$ |
| 2 | 0 | $\overline{c}$ | $\underline{c}$ |
| 3 | 0 | 0 | $\underline{c}$ |

,

and their net values are:

$$\begin{aligned} v(g_1) &= g^{\Lambda}(\overline{c}) + 11\delta(\underline{c}) + 2 \times 11 \times \delta(\overline{c})\delta(\underline{c}) - \underline{c} \\ &= 9.2 + 11 \times (\frac{1.48}{8.95}) + 2 \times 11 \times 0.4 \times (\frac{1.48}{8.95}) - \frac{2.96}{8.95} = 12.143. \end{aligned}$$

$$\begin{aligned} v(g_2) &= v(g^{\Delta}(\overline{c})) + 11 \times \delta(\underline{c}) + 2 \times 11 \times \delta(\underline{c}) \times \delta(\overline{c}) - \underline{c} \\ &= 9 + 11 \times \frac{1.48}{8.95} + 2 \times 11 \times \frac{1.48}{8.95} \times 0.4 - \frac{2.96}{8.95} = 11.943. \end{aligned}$$

$$\begin{aligned} v(g_3) &= v(g^{\Delta}(\overline{c})) + 2 \times 11 \times \delta(\underline{c}) + 11 \times \delta(\underline{c}) \times \delta(\overline{c}) - 2\underline{c} \\ &= 9 + 2 \times 11 \times \frac{1.48}{8.95} + 11 \times \frac{1.48}{8.95} \times 0.4 - 2 \times \frac{2.96}{8.95} = 12.704. \end{aligned}$$

$$\begin{aligned} v(g_4) &= v(g^{\Delta}(\overline{c})) + 3 \times 11 \times \delta(\underline{c}) - 3\underline{c} \\ &= 9 + 3 \times 11 \times \frac{1.48}{8.95} - 3 \times \frac{2.96}{8.95} = 13.465. \end{aligned}$$

Now consider the following *not* SNSG-network

| $g$ | 2 | 3 | 4 |
|---|---|---|---|
| 1 | $\overline{c}$ | $\overline{c}$ | $\underline{c}$ |
| 2 | 0 | 0 | $\underline{c}$ |
| 3 | 0 | 0 | $\underline{c}$ |

$$v(g) = g^{\Lambda}(\overline{c}) + 3 \times 11 \times \delta(\underline{c}) - 3 \times \underline{c} = 13.665.$$

Thus $g$ strictly dominates the four optimal SNSG-networks. Therefore, efficiency is *not* reached by any strong nested split graph network.

# 6 Concluding remarks

Efficiency is reached by "flower" networks where the most (one of) valuable node(s) is linked directly with all the rest, and perhaps some additional links connect directly some spokes of this star, so that the greater the sum of the values of any pair of nodes, the greater the strength of their connection (direct or indirect through the central node). Links between spokes are used only to see each other, while a link between a spoke and the center allows to connect indirectly, through the center, that spoke with any other node with which it is not directly connected. We call these networks hierarchical flowers.

A comparison with previous results is pertinent here. Under homogeneity of nodes, i.e. when all nodes have the same value, the possible non-empty efficient hierarchical flowers collapse into only two extreme cases: the all-encompassing star and the complete network (Olaizola and Valenciano, 2020a and 2020b). While when nodes may have different values, but links are homogeneous, the only possible non-empty efficient hierarchical flowers become strong nested split graph networks (Olaizola and Valenciano, 2021).

As in Olaizola and Valenciano (2020a and 2021), an algorithm is provided in order to prove the result, but in this more general case a more complex algorithm is needed as a complete tree must be thoroughly explored. Nevertheless, the point of the algorithm is not actually calculating anything, but *ad probandum*.[17] In fact, no way to calculate the efficient hierarchical flower network has been given. This poses the challenging issue of providing an algorithm to actually calculate or approximate an efficient hierarchical flower, given the weights of a set of nodes and a regular technology. Obtaining a direct, not algorithmic, proof of the hierarchical flower structure of an efficient network is another challenge.

Another line of further work is the question of stability in the connections model with heterogeneous nodes and links, which we have not addressed in this work. Olaizola and Valenciano (2021) consider a notion of pairwise stability "under free cost sharing" in a model with heterogeneous nodes and homogeneous links and show that efficiency and stability in this sense are compatible. Olaizola and Valenciano (2022) introduce a notion of "marginal equilibrium", a weak form of local Nash equilibrium, in a model with homogeneous nodes and heterogeneous links, and show that efficiency and stability in this sense are incompatible. It may be interesting to apply both notions of stability

[17] Unlike the simpler case where only nodes are heterogeneous (Olaizola and Valenciano, 2021), where the algorithm actually calculates an efficient network.

and see whether those results extend in the more general model considered in this paper.

# References

[1] Bala, V., and S. Goyal, 2000, A non-cooperative model of network formation, *Econometrica* 68, 1181-1229.

[2] Bloch, F., and B. Dutta, 2009, Communication networks with endogenous link strength, *Games and Economic Behavior* 66, 39-56.

[3] Granovetter, M.S., 1973, The strength of weak ties, *American Journal of Sociology* 78(6), 1360-1380.

[4] Jackson, M., and A. Wolinsky, 1996, A strategic model of social and economic networks, *Journal of Economic Theory* 71, 44-74.

[5] Olaizola, N., and F. Valenciano, 2020a, "Dominance of weighted nested split graph networks in connections models", *International Journal of Game Theory* 49(6), 75-96.

[6] Olaizola, N., and F. Valenciano, 2020b, Characterization of efficient networks in a generalized connections model with endogenous link strength, *SERIEs* 11(3), 341-367.

[7] Olaizola, N., and F. Valenciano, 2021, Efficiency and stability in the connections model with heterogenous nodes, *Journal of Economic Behavior and Organization*, 189(5), 490-503.

[8] Olaizola, N., and F. Valenciano, 2022, A connections model with decreasing returns link-formation technology, *forthcoming* in *SERIE*s.