Efficient structure of noisy communication networks

Breitmoser, Yves and Vorjohann, Pauline

EUV Frankfurt (Oder)

26 November 2012

Efficient structure of noisy communication networks

Yves Breitmoser* EUV Frankfurt (Oder)

Pauline Vorjohann HU Berlin

November 26, 2012

Abstract

In the canonical network model, the connections model, only three specific network structures are generically efficient: complete, empty, and star networks. This renders many plausible network structures inefficient. We show that requiring robustness with respect to stochastic transmission failures rehabilitates incomplete, circular network structures. Specifically, we show that near the "bifurcation" where both star and complete network are efficient in the standard connections model, star and complete network are generally inefficient as transmission failures become possible. As for four-player networks, we additionally show that the circle network is uniquely efficient and robust near this bifurcation.

JEL–Codes: D85, C70

Keywords: communication network, information flow, stochastics, robustness, efficiency, connections model

*Corresponding author. Address: Europa–Universität Viadrina, Postfach 1786, 15207 Frankfurt(Oder), Germany, email: breitmoser@euv-frankfurt-o.de, Telephone/Fax: +3355534 2291/2390

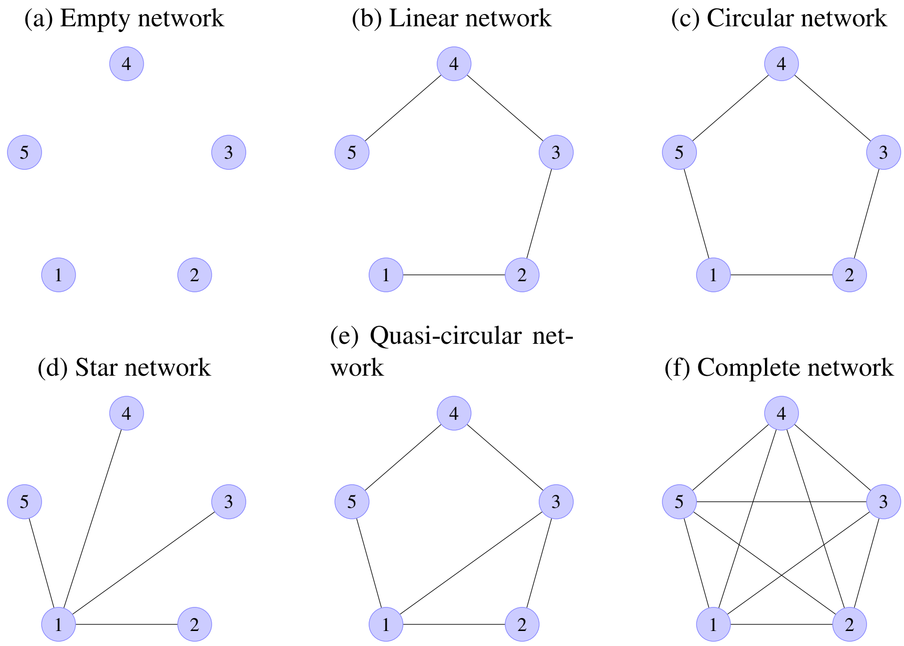
1 Introduction

Most forms of communication take place within networks. This includes direct communication within families and between friends, technical communication between computers, and impersonal communication such as television and radio. For this reason, communication networks have been analyzed fairly comprehensively, starting with Jackson and Wolinsky (1996)'s seminal analysis of the "connections model". Jackson and Wolinsky showed that in networks with two-way information flow, where the value of information decays over time, efficient (non-empty) networks are either star-shaped or complete. The efficiency of stars explains the existence of TV stations for nation-wide communication, and the efficiency of completeness explains communication within families. Subsequent research established this dichotomy of economically efficient communication structures in many related circumstances. For example, only star or complete networks are efficient if only the link initiator bears the link costs (Bala and Goyal, 2000a; Hojman and Szeidl, 2008), if players are farsighted (Dutta et al., 2005), if link strength is endogenous (Bloch and Dutta, 2009), if transfer payments between players are possible (Bloch and Jackson, 2007), and in a slightly weakened form, the result holds true even if link costs are heterogenous between players (Jackson and Rogers, 2005; Galeotti et al., 2006).

These results seem to assert that intermediate network structures are inefficient, and in particular that incomplete, circular networks are wasteful. A well-known counterexample to this assertion is the shape of computer networks. Computer networks need to be "robust" to transmission failure and server downtime, in order to ensure connectivity even if specific links or nodes disappear temporarily. In computer science and combinatorics, this led to the analysis of expander graphs (for a survey, see Hoory et al., 2006), which trade off robustness and number of links rather effectively, although not "efficiently" in the economic sense. In general, the resulting networks tend to be incomplete, which ensures cost efficiency, but also redundant and partially circular, which ensures existence of alternative communication paths in case specific links are unavailable temporarily.

In this paper, we argue that efficient friendship networks are redundant, circular, and incomplete for largely the same reason that computer networks are: complete linking is too costly (i.e. too time-consuming, contrary to say within families),

Figure 1: Elementary network structures



but robustness is important as communication between friends is relatively sporadic (contrary to say TV programs). We set up a simple model of networks with stochastic information flow and show that redundancy and circularity indeed follow if link costs are “intermediate”.

Our model extends the connections model by introducing stochastic link failure—friends meet stochastically. The connections model with decay/discounting remains included as a special case. The intuition underlying our result is that in the absence of noise, utilities simply depend on the distances between players. With noise, expected utilities also depend on the number of alternative paths between players, i.e. on the degree of redundancy of links in the network. We derive expected utilities in closed form and analyze welfare in various network structures. We analyze efficiency near the “bifurcation” where both star and complete network are efficient in the connections model. In four-player networks, the circle becomes uniquely efficient near this bifurcation as noise sets in, and for the case of more than four players, we show that neither star nor complete network are efficient. Then, circular, redundant networks such as the “wheel” (see Figure 4 below) improve welfare.

To our knowledge, the present paper is the first to analyze a network with two-way information flow exhibiting both noise and decay. This extends the existing literature, which has established efficiency of redundant, circular network structures only in networks without decay, i.e. with perfectly patient players. In models of two-way information flow, redundancy is efficient if communication is noisy (Bala and Goyal, 2000b; Haller and Sarangi, 2005) or if link strength is endogenous (Deroïan, 2009; Bloch and Dutta, 2009). Circular structures also obtain in networks with one-way information flow (Bala and Goyal, 2000a, and Kim and Wong, 2007). Another difference to the earlier stochastic models is that Bala and Goyal (2000b) and Haller and Sarangi (2005) analyze networks where links may fail to exist globally with a certain probability and agents maximize a utility function based on the expected number of persons they will be connected to in the resulting network, whereas we explicitly analyze information flow. Similarly to the aforementioned papers, we abstain from modeling costs of information acquisition (Kannan et al., 2007), from discussing non-cooperative implementation (Haller et al., 2007; Harrison and Muñoz, 2008) aside from a brief discussion in the conclusion, and from modeling heterogeneous players (Galeotti, 2006; Billand et al., 2008).

Section 2 introduces notation and elementary terms. Section 3 introduces our model of stochastic information flow. Sections 4 and 5 analyze the four-player and n-player networks. Section 6 concludes.

2 Basic notation and definitions

The set of network nodes (players) is denoted as $N = \{1, 2, \ldots, n\}$, with $0 < n < \infty$, and players are denoted as $i, j \in N$. The existence of a link between players $i, j \in N$ is indicated through $g_{ij} \in \{0, 1\}$, where $g_{ij} = g_{ji} = 1$ indicates existence and $g_{ij} = g_{ji} = 0$ indicates non-existence. All links are undirected. The link matrix $G = (g_{ij})_{i,j \in N}$ defines the "network" and $\mathcal{G}$ denotes the set of such networks. We write $G \subseteq G'$ if $g_{ij} \leq g'_{ij}$ for all $i, j \in N$, and $G \subset G'$ if additionally $G \neq G'$. The network resulting by adding the link ij (if not already existent) to network G is denoted as $G \cup \{ij\}$.

The *degree* $d_i(G) = \sum_{j \neq i} g_{ij}$ of $i \in N$ is the number of links in G involving i. A node with degree d bears costs $C(d) \in \mathbb{R}$. Costs are increasing in d and satisfy $C(0) = 0$. For all $G \in \mathcal{G}$, $u_i(G) \in \mathbb{R}$ is i's expected utility in G. Utility less costs is $u_i(G) - C(d_i(G))$, and a network is *efficient* if it maximizes the welfare $\sum_{i \in N} \big[u_i(G) -$

$C(d_i(G))\big]$ over all $G \in \mathcal{G}$.

A path between i and j ($i \neq j$) in G is $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ such that $p_1 = i$, $p_k = j$, $p_l \in N$ for all $l \leq k$, $g_{p_l, p_{l+1}} = 1$ for all $l < k$, and $p_{l'} \neq p_{l''}$ if $l' \neq l''$ for all $l', l'' < k$. The set of paths between i and j in G is $\mathbf{P}_{ij}(G)$, and for all $\mathbf{p} \in \mathbf{P}_{ij}(G)$, the length of $\mathbf{p}$ is $l(\mathbf{p})$. The *distance* $l_{i,j}(G)$ between i and j is the length of the shortest path connecting them in G, i.e. $l_{i,j}(G) = \infty$ if $\mathbf{P}_{ij} = \emptyset$, and $l_{i,j}(G) = \min_{\mathbf{p} \in \mathbf{P}_{ij}(G)} l(\mathbf{p})$ otherwise. A network G is *connected* if $l_{i,j}(G) < \infty$ for all $i \neq j$.

Figure 1 reviews the elementary network structures in case $n = 5$. If the number of players is $n > 5$, the structures generalize as follows. A network G is *empty* if $g_{ij} = 0$ for all $i, j \in N$. It is *complete* if $g_{ij} = 1$ for all $i \neq j$. The network is a *star* if there exists $k \in N$ such that $g_{ij} = 1$ if and only if either $i = k$ or $j = k$. It is linear if there exists a bijection $o : N \to N$ such that $g_{ij} = 1$ if and only if $|o(i) - o(j)| = 1$. It is called *circle* if there exists a bijection $o : N \to N$ such that $g_{ij} = 1$ if and only if $o(i) - o(j) \equiv 1 \pmod{n}$ or $o(i) - o(j) \equiv -1 \pmod{n}$. Finally, it is called *quasi-circle* if it is incomplete and contains a circle.

3 A model of stochastic information flow

Definition

Consider a model of information flow where players communicate via emails. They read and send emails in rounds (e.g. at night), they send them stochastically to a selection of their contacts, and the emails sent contain all information that they have. The choice whether any given contact is sent an email on a given day is random and i.i.d. across contacts.

Definition 1 (Email model)**.** The interaction proceeds in rounds. In round $t = 0$, a random, non-empty selection of nodes $N' \subset N$ is provided with a piece of information. For all $N' \subset N$, $\Pr(N')$ denotes the probability that the information originates in N'. In each round $t \geq 1$, any $i \neq j$ exchange information with probability $g_{ij} \cdot \alpha$, with $\alpha \in (0, 1]$, and if they do, they exchange all information they possessed at the beginning of the round. The value of the information that reaches $i \in N$ in round $t \geq 0$ is $v_i(t)$.

Define $\pi_i\big(t | N', G\big)$, $i \in N$ and $t \geq 0$, as the probability that it takes exactly t

rounds until the information reaches i in network G. For example, $\pi_i(0|N') = 1$ if $i \in N'$, $\pi_i(0|N') = 0$ if $i \in N \setminus N'$, and $\pi_i(t|N') = 0$ for all t if $d_i(G) = 0$ and $i \in N \setminus N'$. The expected value of information that originated in $N' \subset N$ is for any $i \in N$

$$u_i(G|N') = \sum_{t \geq 0} \pi_i(t|N', G) \cdot v_i(t), \tag{1}$$

and overall i's expected utility in G (not considering link costs) is

$$Eu_i(G) = \sum_{\substack{N' \subset N \\ N' \neq \emptyset}} \Pr\left(N'\right) \cdot u_i(G|N'). \tag{2}$$

The email model contains the connections model (Jackson and Wolinsky, 1996) as a special case.

Example 2 (Connections model)**.** Given some $a > 0$, the expected utility in the connections model is $u_i(G) = a \cdot \sum_{j \neq i} v_i\left(l_{i,j}(G)\right)$. This corresponds with the email model if $\alpha = 1$ and $\Pr(N') = 1/n$ for all singletons N', $\Pr(N') = 0$ otherwise.

In relation to the connections model, the email model therefore allows for the probability of communication between connected players to be smaller than one. Within either email or connections model, one may consider exponential discounting, $v_i(t) = \delta^t$, $\delta \in [0, 1]$, hyperbolic discounting, $v_i(t) = \kappa/(\kappa + t)$, $\kappa > 0$, quasi-linear discounting, $v_i(t) = \max\{0, 1 - t/\eta\}$, $\eta > 0$, or any other value function. After establishing a few basic properties, we will focus on exponential discounting.

The email model is not the only way to model stochastic information flow in networks. As an alternative, let us mention a "telephone model" where players meet sequentially (in random order) and exchange all information they accumulated up to the respective point in time. This telephone model does not contain the connections model as a special case, however, and we are interested especially in the robustness of efficienct structures as noise is added to the connections model.

Basic properties

Figure 2 defines two three-player networks that allow us to illustrate the basic properties. We look at these networks from the perspective of player 1, assuming he discounts exponentially, $v_i(t) = \delta^t$, and that links cost $c = 0$. Fix $\alpha \in (0,1]$ and assume that the piece of information is initially known by either of the three players (determined by a uniform draw). In the star network N_S, player 1 knows the information in $t = 0$ with probability $1/3$, and otherwise (if any other player learns the information in $t = 0$) he learns it in $t > 0$ with probability $\alpha(1-\alpha)^{t-1}$. In aggregate, 1's expected utility in the star network N_S is

$$u_i(N_S) = \left(1 \cdot \delta^0 + 2 \cdot \sum_{t=1}^{\infty} \alpha(1-\alpha)^{t-1}\delta^t\right)/3 = \frac{(3\alpha-1)\,\delta+1}{3\,((\alpha-1)\,\delta+1)}. \tag{3}$$

In the eyes of player 1, N_C adds a seemingly redundant link between 2 and 3 to the star, but if players communicate stochastically, this link establishes an alternative route for the information to reach 1. Thus, it increases 1's expected utility in relation to the star. The expected utility of 1 in N_C is

$$\begin{aligned} u_i(N_C) &= \frac{\delta^0}{3} + \frac{2}{3}\left(\alpha\delta + \sum_{t=2}^{\infty}\delta^t\left((1-\alpha)^{2t-3}\,\alpha\left(2\alpha-\alpha^2\right)(t-1) + (1-\alpha)^{2t-2}\,\alpha\right)\right), \\ &= \frac{\alpha\delta\left((\alpha-1)\left(\alpha^2-3\alpha+1\right)\delta+1\right)}{\left((\alpha-1)^2\,\delta-1\right)^2}. \end{aligned} \tag{4}$$

The utilities are equivalent only if $\alpha = 0$ or $\alpha = 1$, i.e. when players exchange information each round with probability 1 (as in the connections model) or not at all. If $\alpha \in (0,1)$, in turn, the link between 2 and 3 induces a positive externality for player 1, and this externality, albeit low in magnitude (see Figure 2c), will prove relevant. Lemma 4 generalizes this observation and shows that links in the email model induce positive externalities in general (i.e. under Assumption 3). Note the contrast to the connections model ($\alpha = 1$), where a link provides a positive externality to third players only if it shortens the shortest path ("distance") to or between them.

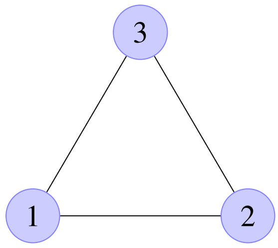
Assumption 3. $\alpha \in (0,1]$, G is connected, and for all $i \in N : \Pr\left(\{i\}\right) > 0$ and $v_i(t)$ is decreasing in t.

Lemma 4 (Positive externality)**.** *If* $\alpha \in (0,1)$*, then* $G \subset G' \Rightarrow u_i(G) < u_i(G')$.

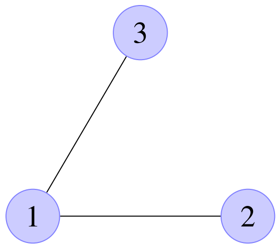
Proof. Fix any $G \subset G'$, define $j', j'' \in N : j' \neq j''$ such that $g_{j'j''} = 0$ and $g'_{j'j''} = 1$, and define $G'' := G \cup \{j'j''\}$. As $\Pr\left(\{j'\}\right) > 0$, this implies $u_i(G''|\{j'\}) > u_i(G|\{j'\})$.

Figure 2: The two connected three-player networks

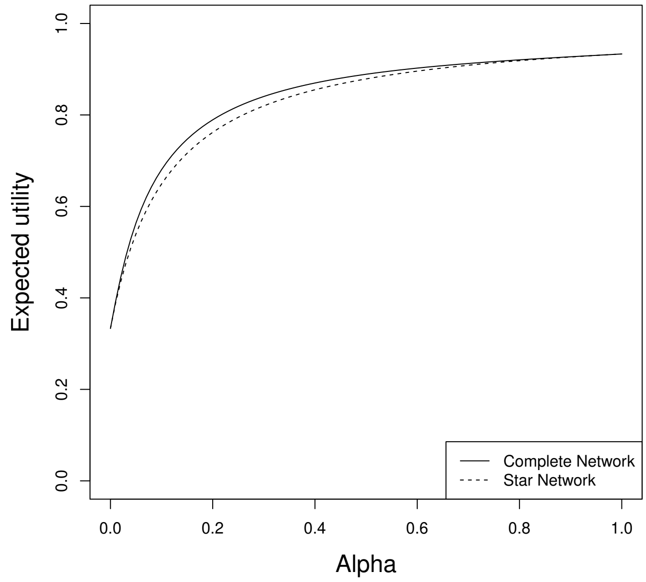
(a) Complete N_C



(b) Star N_S



(c) Exp. uility of player 1



Since $u_i(G''|N') \geq u_i(G|N')$ follows for all $N' \subseteq N$ by stochastic dominance, this yields $u_i(G) < u_i(G'')$ and by induction $u_i(G) < u_i(G')$. $\square$

Now pick any network G satisfying Assumption 3 and consider the expected number of rounds that it takes until a piece of information that originated in $N' = \{j\}$ reaches i. Let $Et(i,j,G) := \sum_{t \geq 0} t \cdot \pi_i(t|\{j\}, G)$ denote this expectation. If $\alpha = 1$, then $Et(i,j,G) = l_{i,j}(G)$; the duration has zero variance and it is independent of any link in the network but those on the shortest path between i and j. In case $\alpha \in (0,1)$, however, $Et(i,j,G)$ depends on all links in G, as implied by Lemma 4. Specifically, for any $\alpha > 0$, $Et(i,j,G)$ is bounded above by $\alpha^{-1} \cdot l_{i,j}(G)$, which results if there is only one path between i and j (e.g. if G is linear), and it is bounded below by $l_{i,j}(G)$, which results if there are infinitely many paths between i and j. The following lemma shows that the expected value of information originating in $\{j\}$ is bounded correspondingly. Thus, given α, the upper bound of utility over all email networks G is equal to the utility in the connections model, while the lower bound illustrates the range of utilities that may result by varying G even when $l_{i,j}(G)$ is held constant. Thus, the difference between these bounds is the maximal value of redundancy given the distance $l_{i,j}(G)$.

Lemma 5 (Bounds)**.** *For all $l > 0$ and $\alpha \in (0,1]$, the expected value of information in all $G \in \mathcal{G}$ satisfying $l_{i,j}(G) = l$ satisfies*

$$\sum_{t \geq l} \binom{l-1}{t-1} \alpha^t (1-\alpha)^{t-l} \cdot v_i(t) \leq u_i(G|\{j\}) \leq v_i(l), \tag{5}$$

and both bounds are tight.

Proof. By Lemma 4, the lower bound obtains in any minimal network G where i and j have distance l (i.e. in any G such that $\nexists G' \subset G : l_{i,j}(G') = l$). Any such minimal network G is "linear" in the sense that $\mathbf{P}_{ij}(G)$ is a singleton, and i's expected utility in G is equal to the lower bound above. Also by Lemma 4, the upper bound is approximated in any maximal network G where i and j have distance l (i.e. in any G such that $\nexists G' \supset G : l_{i,j}(G') = l$), as the number of nodes n tends to infinity. For any $n < \infty$, any maximal network G with $l_{i,j}(G) = l$ has the following structure. There exists a mapping o from N onto $\{1,2,\ldots,l-1\}$ such that for all $k \notin \{i,j\} : g_{ik} = 1$ iff $o(k) = 1$ and $g_{jk} = 1$ iff $o(k) = l-1$, and such that for all $k,k' \notin \{i,j\} : g_{kk'} = 1$ iff $|o(k) - o(k')| = 1$. It is easy to verify that any network violating this structure is either not maximal or does not satisfy $l_{i,j}(G) = l$, and that any network with this structure satisfies $l_{i,j}(G) = l$. The supremum of the expected utility is the limit of $u_i(G|\{j\})$ in such networks as n tends to infinity and is equal to the claimed bound. □

Lower and upper bound equate in case $\alpha = 1$, i.e. in the connections model. In all other cases, the network structure matters beyond mere distances $l_{i,j}(G)$, which will be shown to affect network efficiency in the next section.

4 Efficiency in four-player networks

From now on, we assume for simplicity that all players are equally likely to originate information, and that only single players may originate new information. This assumption is standard in the sense that any $i \in N$ thus weighs all opponents $j \neq i$ equally. In conjunction with standard linearity and symmetry assumptions, this implies the utility assumed by Jackson and Wolinsky (1996) if we set $\alpha = 1$.

Assumption 6. $\Pr(N') = 1/n$ for all singleton sets $N' \subseteq N$ and $\Pr(N') = 0$ otherwise.

Costs are linear, $C(d) = c \cdot d$ with $c \in \mathbb{R}_+$, and value functions are symmetric between players, $v_i = v_j$ for all $i \neq j$.

Note that we normalize utilities differently than Jackson and Wolinsky do. We consider expected utilities, which essentially normalizes them by dividing through $1/n$. If $n = 4$, this implies that the costs thresholds discussed below are $1/4$ of Jackson and Wolinsky's. Now, recall the general result for the symmetric connections model.

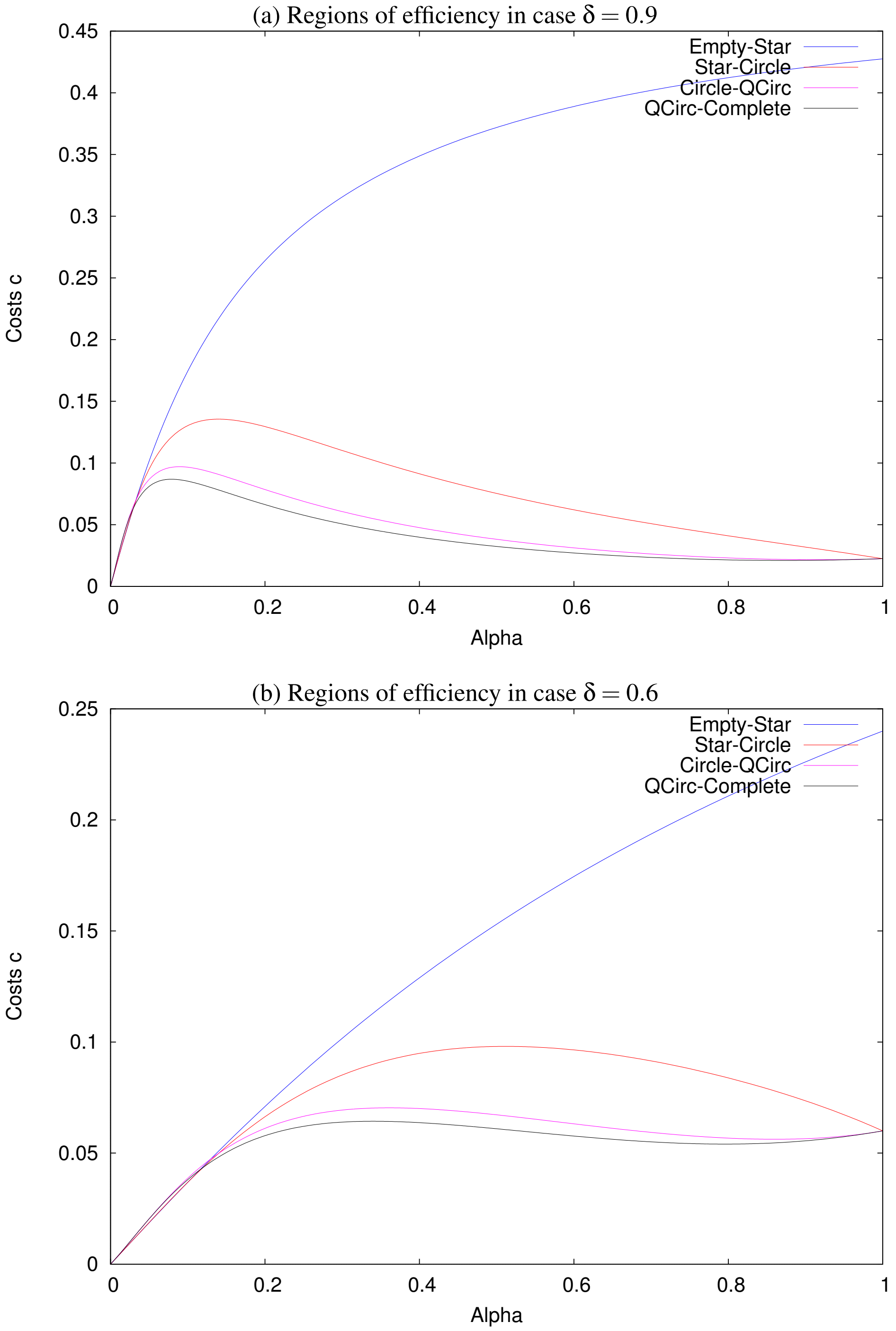
Proposition 7 (Jackson and Wolinsky, 1996)**.** *If* $\alpha = 1$*, the efficient network is either empty, complete, or star for almost all* $c \in \mathbb{R}_+$.

Both star and complete network are efficient if $c = v(1) - v(2)$, i.e. if the gain from abbreviating an indirect connection through establishing a direct one equates with the costs of this connection. At this "bifurcation", actually all networks with diameter one or two are efficient. Besides stars and complete networks, this contains circles and quasi-circles (Figure 1) in the four-player case. In this section, we derive the expected utilities of the players in all these networks, see Figure 3 for a first glance at the results, and then we evaluate efficiency in case $\alpha \approx 1$. Near the bifurcation, the circle will be uniquely efficient, and since it does not extend the star by adding connections, this shows that qualitatively different structures emerge if robustness with respect to even infinitesimal noise is required.

Proposition 8 (Four players)**.** *Fix* $n = 4$*,* $\delta > 1/6$*, and* $v_i(t) = \delta^t$ *for all* i*. If costs* c *are such that complete network and star network induce the same welfare at* $\alpha = 1$*, then there exists* $\underline{\alpha} < 1$ *such that the circle is uniquely efficient for all* $\alpha \in \left(\underline{\alpha}, 1\right)$.

Proposition 8 is proved through a sequence of lemmas that characterize welfare in the various candidate networks and by a subsequent analysis of the neighborhood of $\alpha = 1$ to obtain. The four networks relevant in the efficiency analysis are the star network (Fig. 1d), the circular network (Fig. 1c), the quasi-circular network (Fig. 1e), and the complete network (Fig. 1f). It is straightforward to show that the only remaining network that can be efficient (in case $n = 4$) is the empty network, in which the aggregate expected utility is 1. We start with the star (briefly allowing for any $n \geq 4$, which will be useful in the next section) and the complete network. The proofs of these lemmas are straightforward but tedious, due to the large number of case distinctions and indefinite sums to be evaluated, and therefore relegated to the appendix.

Figure 3: Regions of efficiency



Note: The lines mark the boundaries of the (α, c)-regions where specific network structures are efficient. Above the upmost line, the empty network is efficient, between upmost and second line, the star is efficient, and subsequently, circle, quasi-circle, and complete network are efficient (in this order).

Lemma 9. *In the n-player star network S^n, the expected utility of the central player is $u_n(S^n)$ and that of any peripheral player $i \neq n$ is $u_i(S^n)$ as defined next.*

$$u_n(S^n) = \frac{\delta\,(\alpha n - 1) + 1}{((\alpha - 1)\,\delta + 1)\,n} \qquad u_i(S^n) = \frac{\delta^2\,\left(\alpha^2 n - 3\,\alpha + 1\right) + (3\,\alpha - 2)\,\delta + 1}{((\alpha - 1)\,\delta + 1)^2\,n}$$

Lemma 10. *In the four-player complete network C^4, the expected utility of $i \in N$ is*

$$u_i(C^4) = \frac{\begin{array}{c}(\alpha-1)^4\left(4\alpha^6-24\alpha^5+60\alpha^4-74\alpha^3+36\alpha^2-9\alpha+1\right)\delta^3+\\ (\alpha-1)\left(2\alpha^6-10\alpha^5+35\alpha^4-71\alpha^3+67\alpha^2-23\alpha+3\right)\delta^2+\left(\alpha^4-6\alpha^3+12\alpha^2-13\alpha+3\right)\delta-1\end{array}}{4\left((\alpha-1)^3\delta+1\right)^2\left((\alpha-1)^4\delta-1\right)}.$$

Now let us compare the social welfares of these two networks. If $\delta > 1/2$ and we decrease α starting at the bifurcation where star and complete network are equally efficient, the welfare of the complete network starts to exceed that of the star.

$$\left.\frac{d\sum_{i\in N} u_i(C^4)}{d\alpha} - \frac{d\sum_{i\in N} u_i(S^4)}{d\alpha}\right|_{\alpha=1} = \frac{3\,(\delta-1)\,\delta\,(2\,\delta-1)}{2} \underset{\delta>1/2}{<} 0. \tag{6}$$

That is, the aggregate utility in the complete network decreases more slowly than that in the star network if $\delta > 1/2$. This is surprising, as network links lose in value (in absolute terms) as α decreases, and yet the complete network gains in efficiency on the star. Thus, the welfare gain of adding links is not generally increasing in α, nor is it generally decreasing in α, as can be checked easily for $\delta < 1/2$.

Corollary 11. *$G' \subset G''$ does not imply that $\sum_{i\in N} u_i(G') - \sum_{i\in N} u_i(G'')$ is monotonic in α.*

This kind of non-monotonicity obstructs general analyses substantially. Further, Eq. (6) may also suggest that the complete network (for $\delta > 1/2$) also induces higher welfare than any other incomplete network as we decrease α starting in the bifurcation. This would imply that the complete network is uniquely efficient in this neighborhood. This is not the case, however, as we show by looking at the quasi-circle.

Lemma 12. *In the four-player quasi-circular network Q^4, the expected utilities of the* *degree-3 players is $u_1(Q^4)$ and those of the degree-2 players is $u_2(Q^4)$, with*

$$u_1(Q^4) = \frac{\begin{array}{c}(\alpha-1)^6\left(4\alpha^5-20\alpha^4+38\alpha^3-26\alpha^2+8\alpha-1\right)\delta^4-(\alpha-1)^3\left(2\alpha^6-16\alpha^5+60\alpha^4-107\alpha^3+86\alpha^2-30\alpha+4\right)\delta^3-\\(\alpha-1)\left(2\alpha^5-17\alpha^4+55\alpha^3-73\alpha^2+36\alpha-6\right)\delta^2-\left(3\alpha^3-10\alpha^2+14\alpha-4\right)\delta-1\end{array}}{4\left((\alpha-1)^2\delta-1\right)\left((\alpha-1)^3\delta+1\right)^3}$$

$$u_2(Q^4) = \frac{\begin{array}{c}(\alpha-1)^5\left(4\alpha^5-20\alpha^4+36\alpha^3-22\alpha^2+7\alpha-1\right)\delta^4-2(\alpha-1)^2\left(2\alpha^6-13\alpha^5+39\alpha^4-58\alpha^3+41\alpha^2-14\alpha+2\right)\delta^3+\\\left(\alpha^6-10\alpha^5+41\alpha^4-84\alpha^3+83\alpha^2-36\alpha+6\right)\delta^2+2\left(\alpha^3-4\alpha^2+6\alpha-2\right)\delta+1\end{array}}{4\left((\alpha-1)^2\delta-1\right)^2\left((\alpha-1)^3\delta+1\right)^2}.$$

Again, look at the derivatives of the aggregate utilities.

$$\left.\frac{d\sum_{i\in N}u_i(Q^4)}{d\alpha}-\frac{d\sum_{i\in N}u_i(C^4)}{d\alpha}\right|_{\alpha=1} = \frac{(\delta-1)\,\delta}{2} < 0. \tag{7}$$

For all δ, decreasing α increases the aggregate utility in the quasi-circular network in relation to that of the complete network. Thus, the efficient network in the neighborhood is incomplete. In conjunction with the above result that the welfare gain of adding all links in $C^4 \setminus S^4$ to the star is decreasing in α, this implies that the last link's value is increasing in α, while the other links' values are decreasing in α.

Corollary 13. *$G' \subset G''$ with $g'_{ij} = g''_{ij} = 0$ does not imply $\sum_{i\in N}u_i\big(G'\cup\{ij\}\big)-\sum_{i\in N}u_i(G') > \sum_{i\in N}u_i\big(G''\cup\{ij\}\big)-\sum_{i\in N}u_i(G'')$ or vice versa.*

Hence, network links are complementary in terms of the value they create. An additional link may be welfare improving only if specific other links are present. This is intuitive, but it poses another obstacle to general analyses, as it shows that the aggregate structure needs to be analyzed as a whole. To conclude our analysis, let us now look at the circle.

Lemma 14. *In the four-player circular network O^4, the expected utility of $i \in N$ is*

$$u_i(O^4) = \frac{(\alpha-1)^2\left(4\alpha^4-16\alpha^3+14\alpha^2-6\alpha+1\right)\delta^3-\left(2\alpha^4-18\alpha^3+30\alpha^2-16\alpha+3\right)\delta^2+\left(3\alpha^2-8\alpha+3\right)\delta-1}{4\left((\alpha-1)^2\delta-1\right)^3}.$$

Now, the derivatives of the aggregate utilities are

$$\left.\frac{d\sum_{i\in N}u_i(O^4)}{d\alpha}-\frac{d\sum_{i\in N}u_i(Q^4)}{d\alpha}\right|_{\alpha=1} = \frac{(\delta-1)\,\delta}{2} < 0. \tag{8}$$

In relation to Eq. (7), this shows that both diagonal links are equally useful at $\alpha \approx 1$. For all $\delta \in (0,1)$, the circle therefore becomes more efficient relative to the quasi-circle as α decreases. Since the costs remain constant, this implies that the circle induces higher welfare than the quasi-circle near the bifurcation. The following concludes the proof of the proposition.

Proof of Proposition 8. Social welfares in star and complete networks are equal at $\alpha = 1$ if $c = (1-\delta)\,\delta/4$. Using these costs and $\alpha = 1$, the social welfare in all considered networks (i.e. in all four-player networks with diameter 1 or 2) is $3\,\delta^2 + 1$. Since the expected utilities in all cases are continuous in α, the welfare in all considered networks remains greater than 1 in the neighborhood of $c = (1-\delta)\,\delta/4$ and $\alpha = 1$. Since the empty network generally induces welfare 1, it is therefore not efficient in any such neighborhood. Further, as shown above, Eq. (8), the circle is more efficient than the quasi-circular network in the neighborhood of $\alpha = 1$ for all $\delta \in (0,1)$. By Eq. (7) and transitivity, this applies with respect to the complete network as well, and the relation to the star is

$$\left.\frac{d\sum_{i\in N} u_i(O^4)}{d\alpha} - \frac{d\sum_{i\in N} u_i(S^4)}{d\alpha}\right|_{\alpha=1} = \frac{(\delta-1)\,\delta\,(6\,\delta-1)}{2} \underset{\delta<1/6}{<} 0. \tag{9}$$

Thus, the circle is the efficient network for all $\delta \geq 1/6$. □

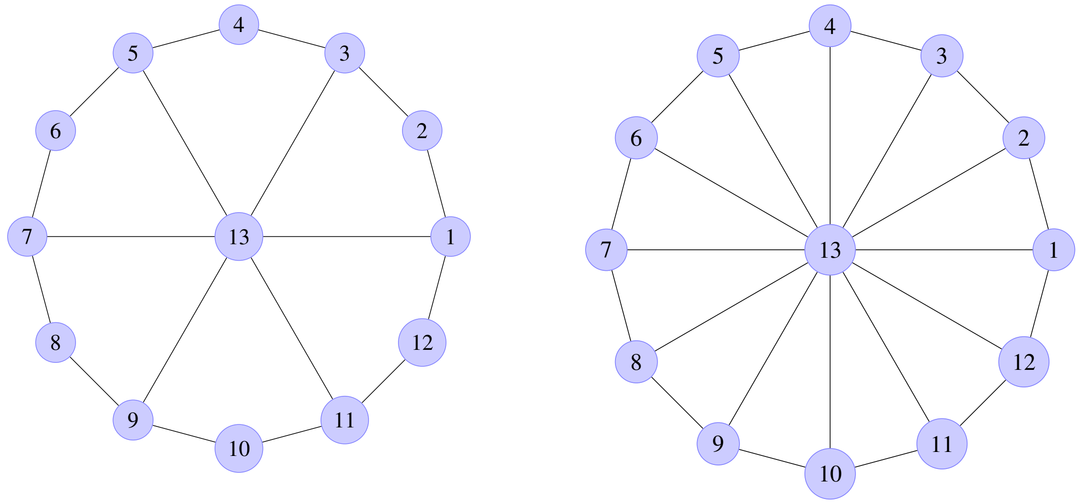
Thus, structures differing qualitatively from star and complete networks emerge if we require robustness to communication failures in the connections model. The following section shows that this result obtains similarly in n-player networks.

5 Efficiency in large networks

In networks with $n > 4$ players, the circular network is not generally efficient anymore. To see this, recall that at the bifurcation discussed above, all networks with diameter 1 or 2 are efficient, while the circle has a diameter of at least 3 if $n > 5$. Further, many relevant intermediate network structures exist for general n, i.e. many plausible networks on the transition from star to complete network. We obtain a general result, however, by showing that a full wheel (Figure 4) is more efficient than both star and complete network near the bifurcation, which obviously implies that

Figure 4: Two “Wheel” networks

(a) A sparse wheel (b) The full wheel



neither star nor complete network are efficient. Note that we use the term wheel to refer to a circle with “spokes” while the literature following Bala and Goyal (2000a) refers to directed circles as wheels.

Definition 15 (Wheels)**.** A network G is called wheel if there exists $k \in N$ such that the network on $N \setminus \{k\}$ is a circle and $g_{ik} = 1$ for at least one $i \neq k$. It is a full wheel if $g_{ik} = g_{ki} = 1$ for all $i \neq k$.

Proposition 16. *Fix $n > 4$ and $\delta > 1/3$. There exist $\alpha \approx 1$ and $c \approx \delta(1-\delta)/n$ such that the full wheel induces a strictly higher welfare than both star and complete network.*

Proposition 16 implies that incomplete, redundant structures generally emerge if robustness with respect to noise is required, i.e. not just in case $n = 4$. Its proof differs from the proof of Proposition 8 in that closed-form representations of the expected utilities for general numbers of players $n > 4$ and arbitrary probabilities $\alpha \in (0,1)$ are not available. These payoffs can be characterized by Taylor expansion in the neighborhood of $\alpha = 1$, however, which we exploit. The expected utilities in stars have already been derived in Lemma 9. We proceed by characterizing the payoffs in complete networks (for arbitrary n).

Lemma 17. *Let $u_i(C^n)$ denote the expected utility of $i \in N$ in the n-player complete network C^n. For any $n > 3$ and any $\varepsilon \geq 0$ there exists $\varepsilon' \geq 0$ such that*

$$\left|u_i(C^n) - \tilde{u}_i(C^n)\right| \leq \varepsilon \qquad \text{and} \qquad \left|du_i(C^n)/d\alpha - d\tilde{u}_i(C^n)/d\alpha\right| \leq \varepsilon$$

for all $\alpha \in [1-\varepsilon', 1]$, with

$$\tilde{u}_i(C^n) = \frac{\left((1-\alpha)\,\alpha^{2n-3}\,\delta^2 + \alpha\,\delta\right)\,(n-1) + 1}{n}.$$

Essentially, Lemma 17 exactly characterizes the expected utility, in terms of value and first derivative with respect to α, in $\alpha = 1$. In relation to the expected utilities in the n-player star network, provided by Lemma 9, it implies that the observation in Eq. (6) that the star is inferior to the complete network (if $\delta > 1/2$) in the neighborhood of the bifurcation generalizes to $n > 4$.

$$\left.\frac{d\sum_{i\in N} u_i(C^n)}{d\alpha} - \frac{d\sum_{i\in N} u_i(S^n)}{d\alpha}\right|_{\alpha=1} = \frac{(\delta-1)\,\delta\,(2\delta-1)\,(n-2)\,(n-1)}{n} \underset{\delta>1/2}{<} 0 \quad (10)$$

The next lemma provides a similar characterization for the payoffs in the full wheel. Here, we focus on the expected payoffs in wheels of at least eight players. For, a large number of alternative paths along which the information can spread through the wheel have to be distinguished, and in case $n \geq 8$, all such paths can be analyzed in a unified manner. The cases $n = 5, 6, 7$ can be analogously to $n \geq 8$, but the case distinctions need to be adapted. We verified that the results reported for $n \geq 8$ continue to hold similarly if $n = 5, 6, 7$ (details are available from the authors).

Lemma 18. *Let $u_i(W^n)$ denote the expected utility of $i \in N$ in the n-player full wheel. For any $n \geq 8$ and any $\varepsilon \geq 0$ there exists $\varepsilon' \geq 0$ such that*

$$\left|u_i(W^n) - \tilde{u}_i(W^n)\right| \leq \varepsilon \qquad \text{and} \qquad \left|du_i(W^n)/d\alpha - d\tilde{u}_i(W^n)/d\alpha\right| \leq \varepsilon$$

for all $\alpha \in [1-\varepsilon', 1]$, with

$$\tilde{u}_i(W^n) = \begin{array}{c} -\Big[(\alpha-1)\,\alpha^2\,\delta^3\,(n-1)\,\big(\alpha^4 n - 3\alpha^3 n + 3\alpha^2 n + n - 2\alpha^6 + 10\alpha^5 - 26\alpha^4 + 34\alpha^3 - 20\alpha^2 - 8\big) \\ -\alpha^2\,\delta^2\,(n-1)\,\big(n - 2\alpha^4 + 8\alpha^3 - 12\alpha^2 + 2\big) - n - 4\alpha\,\delta\,(n-1)\Big]/n^2 \end{array}.$$

Now, if we look at the derivatives of the aggregate utilities,

$$\frac{d\sum_{i\in N} u_i(C^n)}{d\alpha} - \frac{d\sum_{i\in N} u_i(W^n)}{d\alpha}\bigg|_{\alpha=1} = \frac{(\delta-1)\,\delta\,(n-1)\,((2\delta-1)\,n-4\,(3\delta-1))}{n},$$

we find that the difference is not generally positive or negative. The sign depends on $(2\delta-1)\,n \gtrless 4\,(3\delta-1)$. If n is large, then $(2\delta-1)\,n > 4\,(3\delta-1)$, the difference is negative, and thus complete networks gain on wheels as α decreases (i.e. if $\alpha < 1$ and c held constant). In turn, if n is small or δ intermediate, e.g. $\delta = 0.5$, then the term is positive and the wheel gains on the complete network if α decreases. Since social welfare induced in wheels and complete networks equate at the bifurcation (for $\alpha = 1$), the effect of decreasing α thus requires further analysis. By additionally taking cost variations into account, we can show that there are (α, c) near the bifurcation such that the wheel improves upon both complete network and star.

Proof of Proposition 16. At $\alpha = 1$ and $c = \delta(1-\delta)/n$, all networks with diameter no more than 2 induce the same social welfare. Their welfare is also strictly greater than 1, which is the welfare induced by the empty network. As expected utilities are continuous in α and c, all networks with diameter 1 or 2 thus induce strictly greater welfare than the empty network in the neighborhood of $\alpha = 1$ and $c = \delta(1-\delta)/n$. It remains to show that there is a trajectory $(d\alpha, dc)$ along which the differences in welfare between wheel and star, on the one hand, and between wheel and complete network, on the other, increases. Using

$$k_{W,S} = -\frac{d\sum_{i\in N}\big(u_i(W^n)-u_i(S^n)\big)/d\alpha}{d\sum_{i\in N}\big(u_i(W^n)-u_i(S^n)\big)/dc} = \frac{(\delta-1)\,\delta\,(4\delta-1)}{n}$$

$$k_{W,C} = -\frac{d\sum_{i\in N}\big(u_i(W^n)-u_i(C^n)\big)/d\alpha}{d\sum_{i\in N}\big(u_i(W^n)-u_i(C^n)\big)/dc} = \frac{(\delta-1)\,\delta\,((2\delta-1)\,n-4\,(3\delta-1))}{(n-4)\,n},$$

the proposition follows from

$$k_{W,C} - k_{W,S} = \frac{2\,(1-\delta)\,\delta^2\,(n-2)}{(n-4)\,n} > 0. \tag{11}$$

For, along the trajectory $(d\alpha, dc) = (-1, -k)$ both welfare differences (wheel in relation to star and wheel in relation to complete network) increase for any $k \in$

$(k_{W,S}, k_{W,C})$, and by Eq. (11) the interval $(k_{W,S}, k_{W,C})$ is generally not empty. $\square$

6 Conclusion

The paper analyzed efficiency of networks with stochastic information transmission. We set up a simple model that contains the connections model (Jackson and Wolinsky, 1996) as a special case and evaluated network efficiency as noise is small. We found that the standard result that either star or complete networks are efficient, if links are formed at all, no longer holds true even for small levels of noise in meeting/transmission stochastics. Instead, incomplete, redundant networks become efficient. Figures 3 and 5 illustrate that these results generalize directly to case with substantial noise and to alternative valuation functions. Our analysis thus rehabilitates efficiency of a variety of intermediate network structures between the minimal star and the maximal complete network on which the existing literature has focused.

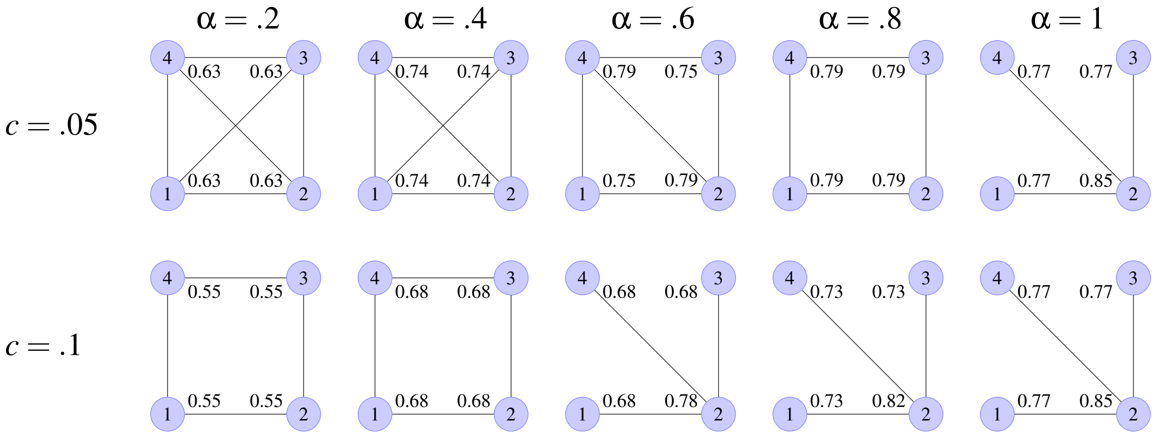
We derived expected utilities in networks with stochastic temporary link unavailability based on exact analyses of arrival times and information flow in the network. Previous work of Bala and Goyal (2000b) and Haller and Sarangi (2005) analyzed connectivity in networks with stochastic, but permanent node unavailability. We did not touch strategic stability, but for the circle in four-player networks, as an example, pairwise stability follows immediately. Cutting a link in the circle yields a linear network, which has diameter 3 and is therefore not stable near the bifurcation. In turn, adding a link yields a quasi-circle (see Figure 1), which was shown to reduce welfare, and by the existence of positive externalities of additional links (Lemma 4), this implies that the players adding the link suffer. Hence, there are no gains from pairwise deviations.

The approach underlying our analysis, the analysis of stochastic information flow, can be applied similarly to networks with large numbers of players and links, at the very least by Monte Carlo simulation. We have seen that link complementarities obtain (see e.g. Cor. 13), which is intuitive and strongly suggests that the set of efficient network structures will adhere to certain regularities. A better understanding of these complementarities may thus yield a precise characterization of the "intermediate" network structures that are efficient in the presence of noise even if n is large, which seems to be an interesting task for future research.

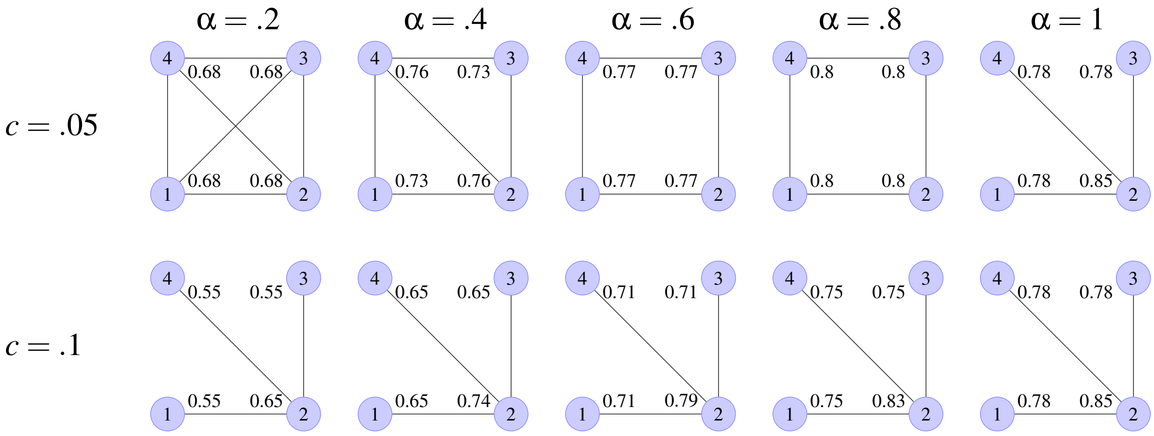
Figure 5: Efficient structures for various unavailability rates α

Note: The valuation $v_i(t)$ is the value of information that is learned after t rounds, α is the probability that a link is unavailable in a given round (i.i.d. across links and across rounds), and c are the costs of link formation. The expected utilities of the players are given next to their nodes.

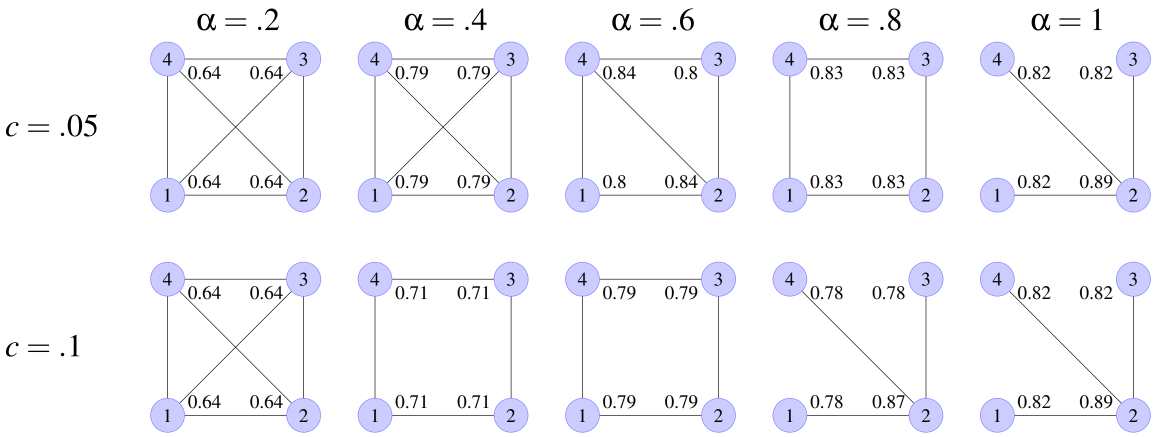
(a) Exponential utility $v_i(t) = \delta^t$ with $\delta = 0.8$



(b) Hyperbolic utility $v_i(t) = \kappa/(\kappa + t)$ with $\kappa = 4$



(c) Quasi-linear utility $v_i(t) = \max\{0, 1 - t/\eta\}$ with $\eta = 7$



A Relegated proofs

Throughout this section, let $u_i(j)$ denote i's expected utility in a given network G if only j currently has the information, and let $\pi_i(t|j)$ denote the probability that the information reaches i in round t. Further, define the random variable $D_i(j)$, for all $i, j \in N$, as the number of rounds it takes until the information, originally known by j, reaches i.

Proof of Lemma 9 (Star, n nodes) Let player n be the central player and players $1, \ldots, n-1$ be the peripheral players. Thus, the expected utility of n if only $i \neq n$ has the information is

$$u_n(i) = \alpha \sum_{t=1}^{\infty} (1-\alpha)^{t-1} \delta^t = \frac{\alpha \delta}{1-(1-\alpha)\delta}.$$

Using $u_n = \big(1+(n-1)\, u_n(i)\big)/n$ and rearranging terms yields the claimed expression. Second, for all $i \neq n$, u_i is defined as $u_i = \big(1+u_i(n)+(n-2)\, u_i(j)\big)/n$, with $j \in N \setminus \{i,n\}$, where $u_i(n) \equiv u_n(i)$ and

$$u_i(j) = \alpha^2 \sum_{t=1}^{\infty} (1-\alpha)^{t-2} \delta^t \,(t-1) = \frac{\alpha^2 \delta^2}{(\alpha\delta - \delta + 1)^2}.$$

$\square$

Proof of Lemma 10 (Complete, 4 nodes) The expected utility is derived by distinguishing four cases. **Case A:** $D_i(j) \leq D_k(j)$ for all $i \neq j, i \neq k, j \neq k$.

$$u_i(j|A) = \alpha \sum_{t=1}^{\infty} (1-\alpha)^{3t-3} \delta^t = \frac{\alpha \delta}{(\alpha-1)^3 \delta + 1}.$$

Case B: $D_k(j) < D_i(j)$ for exactly one $k \neq i, j$ (i.e. one player other than i learns the information before i). The probability that i learns it in $t \geq 2$ is

$$\pi_i(t|j, \text{Case B}) = \alpha \left(1-(1-\alpha)^2\right) \sum_{d=1}^{t-1} (1-\alpha)^{4t-d-5} = \frac{(1-\alpha)^{3t} (\alpha-2)\, \alpha \left(\alpha + (1-\alpha)^t - 1\right)}{(1-\alpha)^5},$$

and i expected utility conditional on case B is

$$u_i(j|B) = \sum_{t=2}^{\infty} \pi_i(t|j, \text{Case B}) \cdot \delta^t = \frac{(\alpha-2)\,(\alpha-1)^2\,\alpha^2\,\delta^2}{\left((\alpha-1)^3\,\delta+1\right)\,\left((\alpha-1)^4\,\delta-1\right)}.$$

Case C: $D_k(j) = D_{k'}(j) < D_i(j)$ for $k \neq k'$ and both different from i and j. The expected utility in this case is

$$u_i(j|C) = \left(1-(1-\alpha)^3\right)\,\alpha^2 \sum_{t=2}^{\infty} (1-\alpha)^{3t-5}\,\delta^t\,(t-1) = \frac{(1-\alpha)\,\alpha^3\,\left(\alpha^2-3\alpha+3\right)\,\delta^2}{\left((\alpha-1)^3\,\delta+1\right)^2}.$$

Case D: $D_k(j) < D_{k'}(j) < D_i(j)$. The probability that i learns it in $t \geq 3$ this way is

$$\begin{aligned}\pi_i(t|j, \text{Case D}) &= \left(1-(1-\alpha)^2\right)\,\left(1-(1-\alpha)^3\right)\,\alpha \sum_{\mathrm{d}_1=1}^{t-2}\ \sum_{\mathrm{d}_2=\mathrm{d}_1+1}^{t-1} (1-\alpha)^{3t+\mathrm{d}_2-\mathrm{d}_1-6} \\ &= \frac{(1-\alpha)^{3t}\,(\alpha-2)\,\alpha\,\left(\alpha^2-3\alpha+3\right)\,\left((\alpha-1)\,\alpha t-\alpha^2-(1-\alpha)^t+1\right)}{(\alpha-1)^6}\end{aligned}$$

and consequently the expected utility in case D is

$$u_i(j|D) = \sum_{t=3}^{\infty} \pi_i(t|j, \text{Case D}) \cdot \delta^t = \frac{(\alpha-2)\,(\alpha-1)^4\,\alpha^3\,\left(\alpha^2-3\alpha+3\right)\,\delta^3}{\left((\alpha-1)^3\,\delta+1\right)^2\,\left((\alpha-1)^4\,\delta-1\right)}.$$

Overall, i's utility is $u_i = \left\{1+(n-1)\left[u_i(j|A)+2\,u_i(j|B)+u_i(j|C)+2\,u_i(j|D)\right]\right\}/n$.

$\square$

Proof of Lemma 12 (Quasi-circle, 4 nodes) Label players similar to Figure 1e, i.e. players 1 and 3 have degree 3, players 2 and 4 have degree 2. We begin with determining the expected utility of 1 if the information is initially with 2 (or symmetrically 4). Three cases are distinguished here. *Case A:* $D_1(2) \leq D_3(2)$.

$$u_1(2|A) = \alpha \sum_{t=1}^{\infty} (1-\alpha)^{2t-2}\,\delta^t = -\frac{\alpha\,\delta}{(\alpha-1)^2\,\delta-1}$$

Case B: $D_3(2) < D_1(2) \leq D_4(2)$, i.e. 3 learns it before 1 (and may therefore tell him), but 4 does not.

$$u_1(2|B) = \left(1-(1-\alpha)^2\right)\alpha\sum_{t=2}^{\infty}\delta^t\sum_{d=1}^{t-1}(1-\alpha)^{3t-d-4}$$
$$= -\frac{(\alpha-2)(\alpha-1)\alpha^2\delta^2}{\left((\alpha-1)^2\delta-1\right)\left((\alpha-1)^3\delta+1\right)}$$

Case C: $D_3(2) < D_4(2) < D_1(2)$, i.e. both 3 and 4 learn it before 1 and may tell him.

$$u_1(2|C) = \left(1-(1-\alpha)^3\right)\alpha^2\sum_{t=3}^{\infty}\delta^t\sum_{\mathrm{d}_3=1}^{t-2}(1-\alpha)^{3t-\mathrm{d}_3-5}(t-\mathrm{d}_3-1)$$
$$= \frac{(\alpha-1)^3\alpha^3\left(\alpha^2-3\alpha+3\right)\delta^3}{\left((\alpha-1)^2\delta-1\right)\left((\alpha-1)^3\delta+1\right)^2}$$

In aggregate, the expected utility of 1 if 2 gets the information first is the sum of the expected utilities in these basic cases, and rearranging a little, this yields

$$u_1(2) = \frac{\alpha\delta\left((\alpha-1)^3\left(\alpha^4-5\alpha^3+9\alpha^2-5\alpha+1\right)\delta^2-(\alpha-1)\left(3\alpha^2-6\alpha+2\right)\delta-1\right)}{\left((\alpha-1)^2\delta-1\right)\left((\alpha-1)^3\delta+1\right)^2}.$$

The expected utility of 1 conditional on 3 getting the information first is determined similarly. Three cases are distinguished implicitly, namely (i) $D_1 \leq D_2, D_4$, (ii) $D_2 < D_1 \leq D_4$ or $D_4 < D_1 \leq D_2$, (iii) $D_2, D_4 < D_1$, and in aggregate, the following results.

$$u_1(3) = \alpha\sum_{t=1}^{\infty}(1-\alpha)^{3t-3}\delta^t + 2\alpha\left(1-(1-\alpha)^2\right)\sum_{t=2}^{\infty}(1-\alpha)^{3t-4}\delta^t(t-1)$$
$$+\left(1-(1-\alpha)^3\right)\alpha^2\sum_{t=2}^{\infty}(1-\alpha)^{3t-5}\delta^t(t-1)^2$$
$$= \frac{\alpha\delta\left((\alpha-1)^4\left(\alpha^4-5\alpha^3+10\alpha^2-6\alpha+1\right)\delta^2-(\alpha-1)\left(\alpha^4-\alpha^3-5\alpha^2+8\alpha-2\right)\delta+1\right)}{\left((\alpha-1)^3\delta+1\right)^3}$$

Aggregate over all three cases, the expected utility of 1 is $u_1 = \left[1+2u_1(2)+u_1(3)\right]/n$, and rearranging a little, the claimed term results.

Next, let us consider the expected utility of 2 (symmetrically that of 4). First, assume 1 gets the information first, and as before, distinguish three cases, A, B, and C. *Case A:* $D_2 \leq D_3$. Conditional on A, the probability that 2 learns it in $t \geq 1$ is

$$\pi_2(t|1,A) = (1-\alpha)^{3t-4}\,\alpha^2\,(t-1) + (1-\alpha)^{3t-3}\,\alpha.$$

Case B: $D_3 \leq D_4$ and $D_3 < D_2$. The probability of $t \geq 2$ conditional on B is

$$\pi_2(t|1,B) = \alpha\left(1-(1-\alpha)^2\right)\sum_{\mathrm{d}_3=1}^{t-1}(1-\alpha)^{2t+\mathrm{d}_3-4} = \frac{(1-\alpha)^{2t}\,(\alpha-2)\,\alpha\,\left(\alpha+(1-\alpha)^t-1\right)}{(\alpha-1)^4}.$$

Case C: $D_4 < D_3 < D_2$. The probability of $t \geq 3$ conditional on C is

$$\begin{aligned}\pi_2(t|1,C) &= \left(1-(1-\alpha)^2\right)^2\alpha\sum_{\mathrm{d}_3=2}^{t-1}(1-\alpha)^{2t+\mathrm{d}_3-5}\,(\mathrm{d}_3-1)\\ &= \frac{(1-\alpha)^{2t}\,(\alpha-2)^2\,\alpha\,\left((1-\alpha)^t\,\alpha t-\alpha^2-2\left((1-\alpha)^t-1\right)\alpha+(1-\alpha)^t-1\right)}{(\alpha-1)^5}\end{aligned}$$

Now, aggregating over these three cases, the expected utility of 2 if 1 gets the information first is

$$\begin{aligned}u_2(1) &= \sum_{t=1}^{\infty}\delta^t\,\pi_2(t|1,A)+\sum_{t=2}^{\infty}\delta^t\,\pi_2(t|1,B)+\sum_{t=3}^{\infty}\delta^t\,\pi_2(t|1,C)\\ &= \frac{\alpha\delta\left((\alpha-1)^3\left(\alpha^4-5\alpha^3+9\alpha^2-5\alpha+1\right)\delta^2-(\alpha-1)\left(3\alpha^2-6\alpha+2\right)\delta-1\right)}{\left((\alpha-1)^2\,\delta-1\right)\left((\alpha-1)^3\,\delta+1\right)^2}\end{aligned}$$

Finally, we proceed similarly to determine the expected utility of 2 if 4 gets the information first. *Case A:* $D_1 = D_3$, i.e. they cannot have learned it from one another, and one of them will tell 2. The probability of the information arriving in round $t \geq 2$ conditional on A is

$$\pi_2(t|4,A) = (1-\alpha)^{2t-4}\,\alpha^2\left(2\alpha-\alpha^2\right)(t-1).$$

Case B: $D_1 < D_2 \leq D_3$ or $D_3 < D_2 \leq D_1$. The probability of $t \geq 2$ conditional on B

is

$$\pi_2(t|4,B) = \alpha^2 \sum_{d=1}^{t-1} (1-\alpha)^{3t-d-4} = -\frac{(1-\alpha)^{2t}\,\alpha\left(\alpha+(1-\alpha)^t-1\right)}{(\alpha-1)^4}$$

Case C: $D_1, D_3 < D_2$. The probability of $t \geq 3$ conditional on C is

$$\begin{aligned}\pi_2(t|4,C) &= \left(1-(1-\alpha)^2\right)^2 \alpha \sum_{\mathrm{d}_1=1}^{t-2} \sum_{\mathrm{d}_3=\mathrm{d}_1+1}^{t-1} (1-\alpha)^{2t+\mathrm{d}_3-\mathrm{d}_1-5} \\ &= \frac{(1-\alpha)^{2t}\,(\alpha-2)^2\,\alpha\left((\alpha-1)\,\alpha t-\alpha^2-(1-\alpha)^t+1\right)}{(\alpha-1)^5}\end{aligned}$$

Aggregating again, we obtain the expected utility of 2 in case 4 gets the information first.

$$\begin{aligned}u_2(4) &= \sum_{t=2}^{\infty} \delta^t \pi_2(t|4,A) + 2\sum_{t=2}^{\infty} \delta^t \pi_2(t|4,B) + 2\sum_{t=3}^{\infty} \delta^t \pi_2(t|4,C) \\ &= \frac{\alpha^2\delta^2\left((\alpha-1)^2\left(\alpha^3-5\alpha^2+8\alpha-2\right)\delta-\alpha^2+2\right)}{\left((\alpha-1)^2\delta-1\right)^2\left((\alpha-1)^3\delta+1\right)}\end{aligned}$$

Thus, the expected utility of 2 overall is $u_2 = \left[1+2u_2(1)+u_2(4)\right]/n$, which yields the claimed term. $\square$

Proof of Lemma 14 (Circle, 4 nodes) Consider a circular network similar to Figure 1c, i.e. 1 is linked with 2 and 4, 2 is linked with 1 and 3, and so on. First consider the expected utility of 1 if the information is initially with 2 (or symmetrically with 4, i.e. his neighbor).

$$\begin{aligned}u_1(2) = \alpha\delta + (1-\alpha)\,\alpha\delta^2 + \sum_{t=3}^{\infty} \delta^t &\left((1-\alpha)^{2t-3}\,\alpha^2\,(t-1)+(1-\alpha)^{2t-2}\,\alpha\right) \\ &+\alpha^2\left(2\alpha-\alpha^2\right)\sum_{t=3}^{\infty}(1-\alpha)^{2t-4}\,\delta^t\binom{t-1}{2}\end{aligned}$$

Simplifying the sums and rearranging terms, we obtain

$$u_1(2) = \frac{\alpha\delta\left((\alpha-1)^2\left(\alpha^3-4\alpha^2+3\alpha-1\right)\delta^2+(\alpha-1)(3\alpha-2)\delta-1\right)}{\left((\alpha-1)^2\delta-1\right)^3}.$$

Next, the expected utility of 1 if 3 (the opposite player) gets the information first is

$$\begin{aligned} u_1(3) &= \alpha^2\left(2\alpha-\alpha^2\right)\sum_{t=2}^{\infty}(1-\alpha)^{2(t-2)}\delta^t(t-1)^2+2\alpha^2\sum_{t=2}^{\infty}(1-\alpha)^{2t-3}\delta^t(t-1) \\ &= \frac{\alpha^2\delta^2\left((\alpha-1)^2\left(\alpha^2-4\alpha+2\right)\delta+\alpha^2-2\right)}{\left((\alpha-1)^2\delta-1\right)^3}, \end{aligned}$$

and the expected utility of 1 results as $u_1 = \left[1+2*u_1(2)+u_1(3)\right]/4$. By symmetry, $u_1 = u_i$ for all $i \in N$. $\square$

Proof of Lemma 17 (Complete, *n* nodes) Given the continuity of both $u_i(C^N)$ and $\tilde{u}_i(C^N)$ in the neighborhood of $\alpha = 1$, it suffices to show that

$$u_i(C^n)-\tilde{u}_i(C^n)\Big|_{\alpha=1} = 0 \qquad \text{and} \qquad \frac{du_i(C^n)}{d\alpha}-\frac{d\tilde{u}_i(C^n)}{d\alpha}\Big|_{\alpha=1} = 0. \tag{12}$$

In general, it is possible to rearrange u_i into sums of products of the form

$$u_i = \sum_{n=0}^{\infty}(1-\alpha)^n f_n(\alpha) \qquad\qquad \tilde{u}_i = \sum_{n=0}^{\infty}(1-\alpha)^n \tilde{f}_n(\alpha)$$

such that no $f_n(\alpha)$, $n \geq 0$, is still divisible by $(1-\alpha)$, nor any $\tilde{f}_n(\alpha)$. Eq. (12) holds for any $\tilde{u}_i$ that satisfies $f_0 = \tilde{f}_0$ and $f_1 = \tilde{f}_1$. That is, we have to characterize all paths of information flow that require 0 or 1 failed meeting in total.

In complete networks, there are exactly two such paths of information flow between any i and j, $i \neq j$. On the one hand, with probability α, i meets j in round t, and on the other hand, with probability $(1-\alpha)$, i and j do not meet in round 1, but all other possible meetings take place (i.e. i meets every $k \neq j$, and all of them meet j

in round 2). Thus,

$$\tilde{u}_i(j|C^n) = \alpha * \delta + (1-\alpha)\,\alpha^{n-2}\,\alpha^{n-1}\,\delta^2,$$

and $\tilde{u}_i$ as claimed results as $\tilde{u}_i(C^n) = \left[1 + (n-1)\,\tilde{u}_i(j|C^n)\right]/n$. □

Proof of Lemma 18 (Wheel, *n* nodes) The basic idea is the same as in the proof of Lemma 17, i.e. we define $\tilde{u}_i$ such that $f_0 = \tilde{f}_0$ and $f_1 = \tilde{f}_1$ as defined above. The number of paths that require 0 or 1 failed meeting in total is larger than above, however.

Let n denote the central player and let $1, \ldots, n-1$ denote the peripheral players. First, consider the expected utility of the central player n. There are two relevant paths from any $i \neq n$ to n, namely the immediate one, with probability α, and the delayed one, with the possibility of going via either peripheral neighbor of i. There are several ways of defining a function $\tilde{u}_n$ with the required properties; we choose the following one.

$$\tilde{u}_n = \left[1 + (n-1)\big(\alpha\,\delta + (1-\alpha)\,\alpha^2\,(1-(1-\alpha)^3)\,\delta^2\big)\right]/n$$

Next, fix any peripheral player $i \neq n$. First, in case n gets the information first, there are two relevant paths (either i and n meet in the first round or not), and $\tilde{u}_i(n)$ can be characterized as follows.

$$\tilde{u}_i(n) = \left(1-(1-\alpha)^3\right)\,(1-\alpha)\,\alpha^2\,\delta^2 + \alpha\,\delta$$

Second, in case a peripheral player with distance 1 to i gets the information first, the expected utility can be characterized as

$$\tilde{u}_i(1) = \left(1-(1-\alpha)^2\right)\,(1-\alpha)\,\alpha\,\delta^2 + \alpha\,\delta.$$

Similarly, the expected utilities in the remaining cases, where distances along the periphery are 2, 3, or 4+, are characterizable as (in case $n \geq 8$)

$$\tilde{u}_i(2) = 2\,(1-\alpha)\,\alpha^2\,\delta^2 + \left(1-(1-\alpha)^2\right)\alpha^2\,\delta^2$$
$$\tilde{u}_i(3) = \left(\left(1-(1-\alpha)^2\right)\left(1-(1-\alpha)^3\right)(1-\alpha)\,\alpha^3 + \left(1-(1-\alpha)^2\right)(1-\alpha)\,\alpha^3\right)\delta^3 + \alpha^2\,\delta^2$$
$$\tilde{u}_i(4+) = \left(\left(1-(1-\alpha)^3\right)(1-\alpha)\,\alpha^3 + (1-\alpha)\,\alpha^2\right)\delta^3 + \alpha^2\,\delta^2.$$

The expected utility overall results as

$$\tilde{u}_i = \left[\tilde{u}_i(n) + 1 + 2\,\tilde{u}_i(1) + 2\,\tilde{u}_i(2) + 2\,\tilde{u}_i(3) + (n-8)\,\tilde{u}_i(4+)\right]/n.$$

$\square$

References

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

Bala, V. and Goyal, S. (2000b). A strategic analysis of network reliability. *Review of Economic Design*, 5(3):205–228.

Billand, P., Bravard, C., and Sarangi, S. (2008). Existence of nash networks in one-way flow models. *Economic Theory*, 37(3):491–507.

Bloch, F. and Dutta, B. (2009). Communication networks with endogenous link strength. *Games and Economic Behavior*, 66(1):39–56.

Bloch, F. and Jackson, M. (2007). The formation of networks with transfers among players. *Journal of Economic Theory*, 133(1):83–110.

Deroïan, F. (2009). Endogenous link strength in directed communication networks. *Mathematical Social Sciences*, 57(1):110–116.

Dutta, B., Ghosal, S., and Ray, D. (2005). Farsighted network formation. *Journal of Economic Theory*, 122(2):143–164.

Galeotti, A. (2006). One-way flow networks: the role of heterogeneity. *Economic Theory*, 29(1):163–179.

Galeotti, A., Goyal, S., and Kamphorst, J. (2006). Network formation with heterogeneous players. *Games and Economic Behavior*, 54(2):353–372.

Haller, H., Kamphorst, J., and Sarangi, S. (2007). (Non-) existence and scope of Nash networks. *Economic Theory*, 31(3):597–604.

Haller, H. and Sarangi, S. (2005). Nash networks with heterogeneous links. *Mathematical Social Sciences*, 50(2):181–201.

Harrison, R. and Muñoz, R. (2008). Stability and equilibrium selection in a link formation game. *Economic Theory*, 37(2):335–345.

Hojman, D. and Szeidl, A. (2008). Core and periphery in networks. *Journal of Economic Theory*, 139(1):295–309.

Hoory, S., Linial, N., and Wigderson, A. (2006). Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561.

Jackson, M. and Rogers, B. (2005). The economics of small worlds. *Journal of the European Economic Association*, 3(2-3):617–627.

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

Kannan, R., Ray, L., and Sarangi, S. (2007). The structure of information networks. *Economic Theory*, 30(1):119–134.

Kim, C. and Wong, K. (2007). Network formation and stable equilibrium. *Journal of Economic Theory*, 133(1):536–549.