

TI 2025-038/II  
Tinbergen Institute Discussion Paper

# Efficient Information Aggregation: Optimal Structure Of Signal Networks

*Bernd Heidergott<sup>1</sup>*

*Frank den Hollander<sup>2</sup>*

*Ines Lindner<sup>3</sup>*

*Azadeh Parvanesh<sup>4</sup>*

<sup>1</sup> Vrije Universiteit Amsterdam, Tinbergen Institute

<sup>2</sup> Leiden University

<sup>3</sup> Vrije Universiteit Amsterdam, Tinbergen Institute

<sup>4</sup> Leiden University

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900

# EFFICIENT INFORMATION AGGREGATION: OPTIMAL STRUCTURE OF SIGNAL NETWORKS

BERND HEIDERGOTT

*VU University Amsterdam, Department of Operations Analytics,  
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*  
`bheidergott@feweb.vu.nl`

FRANK DEN HOLLANDER

*Mathematical Institute, Leiden University,  
Einsteinweg 55, 2333 CC Leiden, The Netherlands*  
`denholla@math.leidenuniv.nl`

INES LINDNER

*VU University Amsterdam, Department of Economics,  
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*  
`i.d.lindner@vu.nl`

AZADEH PARVANEH

*Mathematical Institute, Leiden University,  
Einsteinweg 55, 2333 CC Leiden, The Netherlands*  
`s.a.parvaneh.ziabari@math.leidenuniv.nl`

**ABSTRACT.** This paper develops a mathematical framework to study signal networks, in which nodes can be active or inactive, and their activation or deactivation is driven by external signals and the states of the nodes to which they are connected via links. The focus is on determining the optimal number of key nodes (= highly connected and structurally important nodes) required to represent the global activation state of the network accurately. Motivated by neuroscience, medical science, and social science examples, we describe the node dynamics as a continuous-time inhomogeneous Markov process. Under mean-field and homogeneity assumptions, appropriate for large scale-free and disassortative signal networks, we derive differential equations characterising the global activation behaviour and compute the expected hitting time to network triggering. Analytical and numerical results show that two or three key nodes are typically sufficient to approximate the overall network state well, balancing sensitivity and robustness. Our findings provide insight into how natural systems can efficiently aggregate information by exploiting minimal structural components.

*Key words:* Signal network, aggregated nodes, key nodes, active versus inactive nodes, triggering.

*MSC2020:* 60B20, 05C80, 46L54.

*Acknowledgment:* FdH and AP were supported by the Netherlands Organisation for Scientific Research (NWO) through NETWORKS Gravitation Grant no. 024.002.003. AP was also supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101034253.



---

*Date:* May 26, 2025.

## 1. INTRODUCTION AND BACKGROUND

This paper investigates how to efficiently aggregate nodes in a signal network to accurately and concisely represent the network’s overall state. Our analysis is motivated by the observation that, in many cases, a very small number of nodes suffices to capture the essential behaviour of the entire network, even when the network itself is large. Section 1.1 introduces the concept of signal networks and outlines the study’s main objectives. Sections 1.2 through 1.4 illustrates our framework’s relevance through three key examples: brain networks, aging networks, and friendship networks.

**1.1. Signal networks.** Consider a network with a large number of nodes. The state of a node is modeled with the help of a strictly decreasing *activation function*  $\phi(u)$ ,  $u \geq 0$ , as follows. If at time  $t$  a node is triggered by an exogenous signal, then its state at time  $t' > t$  is  $\phi(t' - t)$ . As long as the state is larger than or equal to a threshold value, the node is *active*, while as soon as the state drops below the threshold value, the node is *inactive*. The prior information gets overridden when the node is triggered again at a later point in time, and the triggering process restarts independently of the past.

We represent large sets of nodes as an *aggregated node*. We assume that the time instances of signals at individual nodes form a general renewal process. Under this mild assumption, the process of active versus inactive states for the aggregated nodes is well approximated by a *Poisson point process* with exponentially distributed holding times between active states due to the *superposition principle* for a large number of sparse point processes (see [12, Theorem 3.1]). Our primary focus will be on the subnetwork formed by the aggregated nodes, which we refer to as the *signal network*. The network is encoded as ‘triggered’ when a fraction at least  $\gamma \in (0, 1)$  of the aggregated nodes is active. The time-lapse during which an aggregated node stays active or inactive depends on the state of its neighboring aggregated nodes. We assume that the signal network is *scale-free*, meaning that its empirical degree distribution is approximately a power law, and *disassortative*, meaning that node degrees tend to be negatively correlated. We assume that the signal network contains  $n \gg 1$  aggregated nodes, of which  $1 \leq k \ll n$  are *key nodes*, i.e., ‘hubs’ in the signal network that play a critical role in the functioning of the network, in the sense that when they are active the entire network is triggered. See Fig. 1 for an illustration.

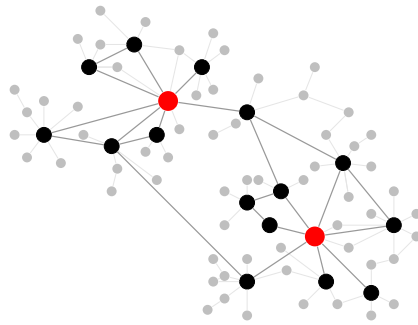


FIGURE 1. An example of a signal network. The aggregated nodes are black and the key nodes are red. The grey nodes form the background of the aggregated nodes.

Think of the bodily reaction of creating ‘goosebumps’ after an emotional experience. For this reaction to occur, the network of aggregated nodes has to be triggered across a certain threshold. Our claim is that this triggering is best organised by aggregating the overall state of the network into a monitoring of the state of the key nodes. Our main question is: What is the optimal value for  $k$  to do this efficiently?

- **Claim:** If the network has  $k = 2$  or  $3$  key nodes, then the states of these key nodes capture the overall state of the network sufficiently well. In contrast, for  $k = 1$ , the key node may already be active when far less than  $\gamma$  of the aggregated nodes are active, while for  $k \geq 4$ , the key nodes may still be inactive when far more than  $\gamma$  of the aggregated nodes are active.

In other words, it suffices to take a *snapshot* of a very small number of key nodes' activity state to get a fair impression of the *overall* activity state of a large signal network. (In order for this claim to be valid, the signal network must avoid absorbing states, for which the distributions of the active and inactive times have to be chosen appropriately: the more neighbours are active, the shorter the time of inactivity and the longer is the time of activity.)

**1.2. Brain network as signal network.** A brain network in cognitive neuroscience is an important example of a signal network that reflects how the brain processes information. A brain network can be described as a graph, with nodes representing neural elements (e.g., neurons or brain regions) and links representing anatomical connectivity (e.g., synapses or axonal projections). Such a network is called a *structural brain network*. If the links describe dynamic interactions or statistical dependencies between the nodes, it is called a *functional brain network*. Signals received by the body are transmitted to the brain directly via cranial nerves or the spinal cord. Unlike graph theory, which emphasises only connectivity patterns, structural brain networks are influenced by physical and topological distances. Neurons and brain regions that are spatially close are more likely to be connected, whereas neurons or brain regions that are spatially distant are less likely to be connected [10]. In [25], it is argued that the key nodes of a brain network are highly connected and central (the 'hubs' in the network) and that these nodes serve as primary points for the overall functioning of the network. Several empirical studies have shown that *small-world* architectures are present in both structural and functional brain networks, in humans and other animals, across a wide range of space and time scales [26, 6, 24, 27, 10]. The small-world architecture ensures short distances for quick signal processing, while *sufficient clustering* ensures stability (when a node falls out). In addition, the brain has a strong synchronisation capacity, which the brain modules need in order to be functional, similar to the firefly synchronisation phenomenon [9].

Regarding the *scale-freeness* of brain networks, the reports differ depending on the data sets and the approach taken. Due to physical constraints and the cost of adding connections in the brain, there are strict upper limits on the number and the density of connections at any given node. As a result, structural brain networks (even those at a large scale) are unlikely to exhibit scale-free degree distributions across a broad range of node degrees [2, 25]. However, an analysis of the degree distribution reveals deviations from a Gaussian or an exponential profile [25]. Truncated power-law degree distributions have also been reported for humans [17].

Functional brain networks show different degree distributions depending on the scale of observation. Scale-freeness evidence has been reported in voxel-level analyses [11, 16], while exponentially truncated power-law distributions have been reported in region-level analyses [1, 5]. Moreover, truncated power-law degree distributions have been reported for cats [28]. Disassortativity is often considered a natural characteristic of biological networks [22]. However, studies show that the structural brain network exhibits an *assortative degree organisation* [14, 8], while the functional brain network is disassortative [7]. These findings are further supported by a [20]. Disassortativity helps to ensure that if one of the key nodes is damaged, then the impact is restricted to only a small part of the network, which is crucial for overall stability and functionality.

The *centrality of hubs* means that they are often part of the shortest paths between other nodes, which makes their role important. Hubs play a crucial role in the brain by connecting different signals and controlling the flow of information between separate areas in the brain.

Because much of the communication between brain regions passes through these hubs, their function has a significant impact on the overall brain performance. Any perturbation in the state of a hub may quickly spread throughout the network. Moreover, hubs help conserve wiring length and volume because they allow information to travel efficiently without needing long-range connections. As a result, if hubs are damaged or malfunctioning, this can substantially impact the overall function of the brain [25]. Understanding the optimal number of hubs in a brain network is therefore crucial, as it provides valuable insight into the overall brain function. Our main interest will be functional brain networks.

**1.3. Aging network as signal network.** The body can be seen as a network, where nodes represent functional units (organs, cells, or health measurements) and links represent interactions or dependencies between them. Aging can be viewed as the accumulation of damage across this network over time. Damage can spread when one part of the system fails due to internal degradation or external signals. This failure increases the risk of damage in neighboring parts. Some nodes, such as vital organs in the human body, are critical for the overall function, and we can, therefore, call them the key nodes of the aging network. When these critical nodes fail, the entire system collapses, leading to death. In 1825, Benjamin Gompertz [13] stated an empirical relationship for the mortality rate  $m(t)$  at time  $t$ , given by

$$m(t) \approx \alpha e^{\beta t}, \quad t \geq 0,$$

with parameters  $\alpha, \beta > 0$  (and with  $\approx$  meaning approximately). This relationship is since referred to as *Gompertz law*. It shows that, as we age, our death risk increases exponentially. The Gompertz law works relatively well for adult ages, roughly from 40 to 90 years, but not at very young or very old ages.

[12] offers a causal mathematical model that explains aging and mortality through network theory, which offers a novel and mathematical derivation of the Gompertz law. The body is modelled as a network where each node represents a health indicator, such as blood pressure, sugar, cholesterol levels, heart rate, physical mobility, etc., and each link represents a channel for interaction. In this network, two nodes are designated as *mortality nodes* (most vital nodes), while the others are designated *aging nodes*. The network is assumed to be scale-free and disassortative. Each node has a state that is either healthy or damaged. The health network evolves over time via a Markov process, where nodes can switch between healthy and damaged. Damage to nodes spreads via the links between them. An individual is considered dead when both mortality nodes become damaged. With the help of a mean-field assumption and a homogeneity assumption, the mortality rate is shown to be the product of the solutions of two non-linear differential equations, and this product approximates the Gompertz law in the age range of 40 to 80 years.

The aging network studied in [12] closely aligns with the signal network considered in the present paper, where active nodes correspond to damaged nodes and passive nodes correspond to healthy nodes. In [12], the number of mortality nodes is set to two, following the network description of mortality proposed in [21], which is based on a suitable empirical fit to data. This choice lacks a theoretical justification, and the preference for two mortality nodes over other values is primarily based on empirical performance. Investigating the optimal number of key nodes, as is done in the present paper, provides a deeper insight into the structural and functional robustness of aging networks.

**1.4. Friendship network as signal network.** Another example of a signal network is the friendship network, where nodes represent individuals in the community under consideration, and links represent their (mutual) friendships. In this context, signals can be understood as the spread of information, such as news, rumours, or opinions. A node is considered passive when it is not affected by the incoming information and active when it is and can respond to it.

Empirically, social networks often follow a scale-free degree distribution. A widely used model to describe such networks is the preferential attachment model in [3]. This dynamic random network model assumes that new nodes are more likely to attach to already highly connected nodes (i.e., nodes with a high degree), which leads to negative degree correlation and a scale-free topology [3]. While the classical preferential attachment model tends to produce disassortative networks based purely on degree, [18] proposes more socially grounded mechanisms, which can result in assortative network structures where individuals tend to connect with other individuals similar to themselves. However, under certain conditions, particularly when relationships are hierarchical or are driven by influence rather than mutual friendships, the network may also exhibit disassortative characteristics. Examples are social media networks (like X-follower networks) or advisor-student relationships. Moreover, [23] and [4] find that disassortativity is a result of an environment in which the growth mechanism depends on the fitness of a node in competitive dynamics and weight-driven dynamics, respectively.

The analogy to the body evolving from single-celled organisms, incrementally adding body parts based on fitness criteria, is tempting. With an environment that is changing in time, different body parts and functionalities become less or more important, and therefore, the fitness of a node may change with time. During an ice age, for example, functionalities that keep the body warm would increase fitness, and so, as the body evolves, new nodes that are added will likely attach to nodes that aid survival in the cold, such as fat cells. The result is a *hierarchical organization*, meaning the occurrence of many small densely connected clusters, which combine to form larger, less densely connected groups, which again combine to form even larger and even less densely connected groups. Our body also exhibits hierarchical organisation with cells making up tissue, tissues making up organs, organs making up organ systems, and organ systems ultimately making up higher-level functionality such as the ability to read. The analogy with the social context is that the body needs to be efficiently organised.

With the above considerations, our assumption of disassortativity in signal networks is *purely degree-based*. Therefore, a friendship network can also be a good signal network example. This idea is further supported by the well-known friendship paradox, which states that, on average, our friends have more friends than we do. A mathematical analysis of the friendship paradox [15] shows that only a small fraction of the nodes may have more connections than their friends do in large preferential attachment networks, such as worldwide friendship networks. Finding the optimal number of key nodes (i.e., highly connected and influential individuals) in such networks is an interesting direction for future research in this context.

## 2. MODEL AND KEY ASSUMPTIONS

Section 2.1 formulates the model. Sections 2.2–2.3 state two *key assumptions* under which the model can be analysed in closed form.

**2.1. Model.** Recall Figure 1. Labelling the *aggregated nodes* by  $1, \dots, n$ , we describe the state of the signal network as a continuous-time stochastic process

$$\mathbf{X}(t) = (X_1(t), \dots, X_n(t)), \quad t \geq 0,$$

where  $X_i(t) = 1$  when node  $i$  is active at time  $t$  and  $X_i(t) = 0$  when node  $i$  is passive at time  $t$ . We let  $\mathbf{X}(0) = (0, \dots, 0)$ . The stochastic process  $\mathbf{X} = (\mathbf{X}(t))_{t \geq 0}$  is a time-inhomogeneous Markov chain.

We denote the set of neighbours of node  $i$  by  $N(i)$ . For each node  $i$ , we assume that the set  $N(i)$  is non-empty. We assume that the rates at time  $t$  at which node  $i$  transitions to the active state or the passive state, respectively, are given by

$$\Gamma_+(i, t) = \lambda(a_i(t)), \quad \Gamma_-(i, t) = \mu(a_i(t)),$$

where  $\lambda(\cdot)$  and  $\mu(\cdot)$  are decreasing, respectively, increasing functions, and  $a_i(t)$  is the fraction of active neighbours of node  $i$  at time  $t$ , i.e.,

$$a_i(t) = \frac{1}{|N(i)|} \sum_{j \in N(i)} \mathbb{1}_{\{X_j(t)=1\}}.$$

We define the average fraction of active nodes at time  $t$  by

$$a(t) = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{X_j(t)=1\}}.$$

Labelling the *key nodes* by  $1, \dots, k$ , we define

$$\tau^\gamma = \inf \{t \geq 0: a(t) \geq \gamma\}$$

to be the first time when the signal network gets triggered and

$$\tau_k = \inf \{t \geq 0: (X_1(t), \dots, X_k(t)) \in (1, \dots, 1)\}$$

be the first time when all  $k$  key nodes are active. Due to the central role played by the key nodes, the activation of all key nodes should indicate that most aggregated nodes are active (i.e., the system is triggered). Therefore, we should expect that  $\mathbb{E}[\tau^\gamma] \approx \mathbb{E}[\tau_k]$  ( $\approx$  means approximately) for suitable values of  $k$ . For this purpose, we define

$$k_c^-(\gamma) = \sup \{1 \leq k \leq n: \mathbb{E}[\tau^\gamma] \geq \mathbb{E}[\tau_k]\} \vee 1,$$

$$k_c^+(\gamma) = \inf \{1 \leq k \leq n: \mathbb{E}[\tau^\gamma] \leq \mathbb{E}[\tau_k]\} \wedge n,$$

where we use the convention that  $\sup \emptyset = -\infty$  and  $\inf \emptyset = +\infty$ . Since  $\tau_k$  is strictly increasing in  $k$ , the optimal value of  $k$  is defined to be

$$k_c(\gamma) = \begin{cases} k_c^-(\gamma), & \text{if } |\mathbb{E}[\tau^\gamma] - \mathbb{E}[\tau_{k_c^-(\gamma)}]| \leq |\mathbb{E}[\tau^\gamma] - \mathbb{E}[\tau_{k_c^+(\gamma)}]|, \\ k_c^+(\gamma), & \text{otherwise,} \end{cases}$$

i.e., the value of  $k$  for which  $\mathbb{E}[\tau_k]$  is closest to  $\mathbb{E}[\tau^\gamma]$ .

**Remark 2.1. [Activation rate]** A key object in analysing  $k_c(\gamma)$  is the activation rate at time  $t$  for  $k$  key nodes, given by

$$m_k(t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathbb{P}\{\tau_k \leq t + \Delta \mid \tau_k \geq t\} = -\frac{1}{\mathbb{P}\{\tau_k \geq t\}} \frac{d}{dt} \mathbb{P}\{\tau_k \geq t\}.$$

We will also look at  $m_k(t)$  later. ♠

Under a *mean-field assumption* and a *homogeneity assumption* (see below), we estimate  $k_c(\gamma)$  for different choices of  $\lambda$  and  $\mu$  when the number  $n$  of aggregated nodes is large. These two assumptions are reasonable due to the scale-free and disassortative nature of the signal network. The Pareto principle, popularised by Richard Koch [19], suggests that roughly 80% of effects result from 20% of causes. Therefore, we consider  $\gamma = 0.4$  as a conservative estimate of the fraction of nodes that need to be active to trigger the entire network, *regardless* of the specific nature of the signal network. The theory presented in Sections 2-3 works for every choice of  $\gamma \in (0, 1)$ , but for the numerics in Section 2 we use the choice  $\gamma = 0.4$ .

**2.2. Mean-field assumption.** The mean-field assumption is the approximation

$$a_i(t) \approx a(t) \approx \hat{a}(t) = \mathbb{E}[a(t)],$$

in combination with the fact that rates for the transition to the active state or the passive state are independent of the individual nodes are given by

$$\Gamma_+(i, t) \approx \lambda(\hat{a}(t)), \quad \Gamma_-(i, t) \approx \mu(\hat{a}(t)).$$

We will refer to  $\hat{a}(t)$  as the *activation fraction* at time  $t$ .



In a large network, the properties of scale-freeness and disassortativity ensure the presence of many high-degree nodes distributed across the network. This leads to the nodes' near independence and results in them behaving approximately in the same way. It is, therefore, reasonable to apply the law of large numbers, which suggests that the average state of the nodes is a good approximation for the state of the entire network in support of the above mean-field assumption.

**2.3. Homogeneity assumption.** For  $i \in \{0, 1, \dots, k\}$ , we define

$$S_i = \left\{ (z_1, \dots, z_n) \in \{0, 1\}^n : \#\{1 \leq j \leq k : z_j = 1\} = i \right\}.$$

We assume that the states of the signal network in  $S_0, S_1, \dots, S_k$  are *further aggregated* into  $k + 1$  distinct states. See Fig. 2 for an illustration. In that case  $(k - i)\lambda(\hat{a}(t))$  and  $i\mu(\hat{a}(t))$  are the transition rates out of state  $S_i$  and into state  $S_{i+1}$ , respectively,  $S_{i-1}$ . These transition rates are depicted in Figure 2. In fact, if  $Q(t) = (q_{i,j}(t))_{0 \leq i,j \leq k-1}$  denotes the infinitesimal generator (transition rate matrix) of the inhomogeneous continuous-time Markov chain  $\mathbf{X}$  at time  $t$ , then

$$q_{i,j}(t) = \lim_{\Delta \downarrow 0} \frac{\mathbb{P}\{\mathbf{X}(t + \Delta) \in S_j \mid \mathbf{X}(t) \in S_i\} - \delta_{i,j}}{\Delta}$$

$$= \begin{cases} (k - i)\lambda(\hat{a}(t)), & \text{if } 0 \leq i \leq k - 1, j = i + 1, \\ i\mu(\hat{a}(t)), & \text{if } 1 \leq i \leq k, j = i - 1, \\ -[(k - i)\lambda(\hat{a}(t)) + i\mu(\hat{a}(t))], & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases}$$

where  $\delta_{i,j}$  is the Kronecker delta (which is 1 if  $i = j$  and 0 otherwise).

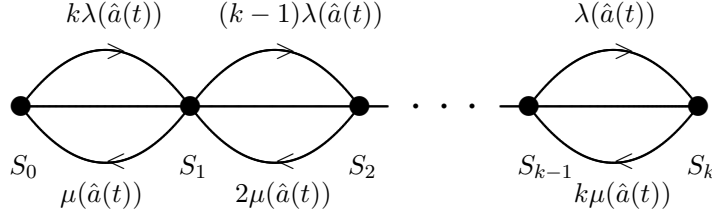


FIGURE 2. The transition rates between distinct states in the aggregated time-inhomogeneous Markov chain.

### 3. MATHEMATICAL RESULTS

In Section 3.1, with the help of the mean-field assumption and following a similar approach as in [12], we derive a mathematical expression for the activation fraction  $\hat{a}(t)$  at time  $t$  as the solution to an autonomous differential equation, and for the hitting time  $\tau^\gamma$ , representing the first time when the network is triggered. In Section 3.2, with the help of the mean-field and homogeneity assumptions, we derive a mathematical expression for the expectation of the hitting time  $\tau_k$ , defined as the first time when all  $k$  key nodes get active, and in Section 3.3 for the activation rate of the  $k$  key nodes. We use the symbol  $\approx$  to emphasise that the expressions are obtained under the two assumptions. Our key results are Theorems 3.3 and 3.6 below.

### 3.1. Evolution of the activation fraction.

**Lemma 3.1.** [Evolution of the activation fraction] *The activation fraction at time  $t$  is the solution to an autonomous differential equation*

$$\frac{d}{dt}\hat{a}(t) \approx (1 - \hat{a}(t))\lambda(\hat{a}(t)) - \hat{a}(t)\mu(\hat{a}(t)), \quad \hat{a}(0) = 0. \quad (3.1)$$

*Proof.* Since all nodes are initially passive, we have  $\hat{a}(0) = 0$ . Moreover, the rate at which one of the  $n(1 - \hat{a}(t))$  passive nodes at time  $t$  becomes active equals  $n(1 - \hat{a}(t))\lambda(\hat{a}(t))$ , while the rate at which one of the  $n\hat{a}(t)$  active nodes at time  $t$  becomes passive equals  $n\hat{a}(t)\mu(\hat{a}(t))$ . Therefore,  $n(1 - \hat{a}(t))\lambda(\hat{a}(t)) - n\hat{a}(t)\mu(\hat{a}(t))$  describes how the number of active nodes evolves (increases or decreases) at time  $t$ . Divide this value by  $n$  to get the fraction of active nodes at time  $t$ .  $\square$

**Remark 3.2.** [Link to the trigger time] Under the mean-field assumption,  $\tau^\gamma$  is deterministic and is equal to

$$\hat{\tau}^\gamma = \inf \{t \geq 0: \hat{a}(t) \geq \gamma\},$$

i.e., the first time the system reaches the threshold level  $\gamma$ . Therefore, we analyse  $\tau^\gamma$  by solving the equation in (3.1) and then looking at the first time  $t$  at which

$$\hat{a}(t) \geq \gamma.$$

♠

### 3.2. Expected activation time of the key nodes.

**Theorem 3.3.** [Expected time for key nodes activation] *For  $k \geq 1$ ,*

$$\mathbb{E}[\tau_k] \approx \int_0^\infty (1 - p_k(t)) dt,$$

where  $p_k(t)$  is obtained from the following system of coupled ordinary differential equations:

$$\begin{aligned} \frac{d}{dt} p_0(t) &= p_1(t)\mu(\hat{a}(t)) - p_0(t)k\lambda(\hat{a}(t)), \\ \frac{d}{dt} p_j(t) &= p_{j-1}(t)(k - j + 1)\lambda(\hat{a}(t)) + p_{j+1}(t)(j + 1)\mu(\hat{a}(t)) \\ &\quad - p_j(t)[(k - j)\lambda(\hat{a}(t)) + j\mu(\hat{a}(t))], \quad 1 \leq j \leq k - 2, \\ \frac{d}{dt} p_{k-1}(t) &= p_{k-2}(t)2\lambda(\hat{a}(t)) - p_{k-1}(t)[\lambda(\hat{a}(t)) + (k - 1)\mu(\hat{a}(t))], \quad k > 1, \\ \frac{d}{dt} p_k(t) &= p_{k-1}(t)\lambda(\hat{a}(t)), \end{aligned}$$

with the following initial conditions:

$$p_0(0) = 1 \quad \text{and} \quad p_1(0) = \dots = p_k(0) = 0.$$

Here, for  $t \geq 0$  and  $j = 0, 1, \dots, k$ ,

$$p_j(t) \in [0, 1] \quad \text{and} \quad \sum_{j=0}^k p_j(t) = 1.$$

*Proof.* Let  $\tilde{\mathbf{X}}$  be a time-inhomogeneous Markov chain on the state space  $\{S_0, S_1, \dots, S_k\}$ , with transition rates identical to those of  $\mathbf{X}$ , except that the transition from  $S_k$  to  $S_{k-1}$  is removed and  $S_k$  is considered as an absorbing state. Let  $p_j(t) = \mathbb{P}\{\tilde{\mathbf{X}}(t) \in S_j\}$  for  $0 \leq j \leq k$ . Since all nodes are initially passive, we have  $p_0(0) = 1$  and  $p_1(0) = \dots = p_k(0) = 0$ .

For  $\Delta \downarrow 0$  and  $0 \leq j \leq k-1$  we have

$$\begin{aligned} p_j(t + \Delta) &= \sum_{i=0}^k p_i(t) \mathbb{P}\{\tilde{\mathbf{X}}(t + \Delta) \in S_j \mid \tilde{\mathbf{X}}(t) \in S_i\} \\ &= \sum_{i=0}^{k-1} p_i(t) (q_{i,j}(t)\Delta + \delta_{i,j} + o(\Delta)) \\ &= p_j(t) + \sum_{i=0}^{k-1} p_i(t) q_{i,j}(t)\Delta + o(\Delta). \end{aligned}$$

Similarly, for  $\Delta \downarrow 0$  we have

$$\begin{aligned} p_k(t + \Delta) &= \sum_{i=0}^k p_i(t) \mathbb{P}\{\tilde{\mathbf{X}}(t + \Delta) \in S_k \mid \tilde{\mathbf{X}}(t) \in S_i\} \\ &= p_k(t) + \sum_{i=0}^{k-1} p_i(t) q_{i,k}(t)\Delta + o(\Delta). \end{aligned}$$

This implies that, for  $0 \leq j \leq k$ ,

$$\frac{d}{dt} p_j(t) = \sum_{i=0}^{k-1} p_i(t) q_{i,j}(t).$$

Hence we have

$$\begin{aligned} \frac{d}{dt} p_0(t) &= p_0(t) q_{0,0}(t) + p_1(t) q_{1,0}(t) = p_1(t) \mu(\hat{a}(t)) - p_0(t) k \lambda(\hat{a}(t)), \\ \frac{d}{dt} p_k(t) &= p_{k-1}(t) q_{k-1,k}(t) = p_{k-1}(t) \lambda(\hat{a}(t)), \end{aligned}$$

and, for  $k > 1$ ,

$$\begin{aligned} \frac{d}{dt} p_{k-1}(t) &= p_{k-2}(t) q_{k-2,k-1}(t) + p_{k-1}(t) q_{k-1,k-1}(t) \\ &= p_{k-2}(t) 2\lambda(\hat{a}(t)) - p_{k-1}(t) [\lambda(\hat{a}(t)) + (k-1)\mu(\hat{a}(t))], \end{aligned}$$

and, when  $k > 2$ , for  $1 \leq j \leq k-2$ ,

$$\begin{aligned} \frac{d}{dt} p_j(t) &= p_{j-1}(t) q_{j-1,j}(t) + p_j(t) q_{j,j}(t) + p_{j+1}(t) q_{j+1,j}(t) \\ &= p_{j-1}(t)(k-j+1)\lambda(\hat{a}(t)) + p_{j+1}(t)(j+1)\mu(\hat{a}(t)) - p_j(t)[(k-j)\lambda(\hat{a}(t)) + j\mu(\hat{a}(t))]. \end{aligned}$$

Finally,

$$\mathbb{E}[\tau_k] = \int_0^\infty \mathbb{P}\{\tau_k > t\} dt \approx \int_0^\infty \mathbb{P}\{\tilde{\mathbf{X}}(t) \notin S_k\} dt = \int_0^\infty (1 - p_k(t)) dt.$$

□

### 3.3. Activation rates for the key nodes.

**Lemma 3.4.** [Representation of the activation rate in terms of the pre-trigger probability] For every  $t \geq 0$ ,

$$m_k(t) \approx \lambda(\hat{a}(t)) \mathbb{P}\{\mathbf{X}(t) \in S_{k-1} \mid \mathbf{X}(u) \notin S_k, 0 \leq u \leq t\}.$$

*Proof.* In a short time length,  $\mathbf{X}$  can only reach  $S_k$  by going from  $S_{k-1}$  to  $S_k$  in a single jump. Hence, taking

$$\begin{aligned} A_t &= \{\mathbf{X}(u) \notin S_k, 0 \leq u \leq t\} = \{\tau_k \geq t\}, \\ B_t &= \{\mathbf{X}(t) \in S_{k-1}\}, \\ C_{t,\Delta} &= \{\exists 0 \leq u \leq \Delta: \mathbf{X}(t+u) \in S_k\}, \end{aligned}$$

we have, for  $\Delta \downarrow 0$ ,

$$\begin{aligned} \mathbb{P}\{\tau_k \leq t + \Delta \mid \tau_k \geq t\} &= \frac{\mathbb{P}\{A_t \cap B_t \cap C_{t,\Delta}\}}{\mathbb{P}\{A_t\}} + o(\Delta) = \mathbb{P}\{C_{t,\Delta} \mid A_t \cap B_t\} \mathbb{P}\{B_t \mid A_t\} + o(\Delta) \\ &= \mathbb{P}\{C_{t,\Delta} \mid B_t\} \mathbb{P}\{B_t \mid A_t\} + o(\Delta), \end{aligned}$$

where the last equality follows from the Markov property at time  $t$ . Noting that the rate of jumping from  $S_{k-1}$  to  $S_k$  is  $\lambda(\hat{a}(t))$ , we have

$$\mathbb{P}\{C_{t,\Delta} \mid B_t\} \approx \lambda(\hat{a}(t)) \Delta + o(\Delta), \quad \Delta \downarrow 0,$$

and hence

$$m_k(t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathbb{P}\{\tau_k \leq t + \Delta \mid \tau_k \geq t\} \approx \lambda(\hat{a}(t)) \mathbb{P}\{B_t \mid A_t\}.$$

□

**Lemma 3.5. [Computation of the pre-trigger probability]** For every  $k \geq 1$ ,

$$\mathbb{P}\{\mathbf{X}(t) \in S_{k-1} \mid \mathbf{X}(u) \notin S_k, 0 \leq u \leq t\} \approx \frac{p_{k-1}(t)}{1 - p_k(t)},$$

where  $p_{k-1}(t)$  and  $p_k(t)$  are defined as in Theorem 3.3.

*Proof.* Define  $\tilde{\mathbf{X}}$  as in the proof of Theorem 3.3. Also, let  $p_j(t) = \mathbb{P}\{\tilde{\mathbf{X}}(t) \in S_j\}$  for  $0 \leq j \leq k$ . Then

$$\begin{aligned} \mathbb{P}\{\mathbf{X}(t) \in S_{k-1} \mid \mathbf{X}(u) \notin S_k, 0 \leq u \leq t\} &= \frac{\mathbb{P}\{\{\mathbf{X}(t) \in S_{k-1}\} \cap \{\mathbf{X}(u) \notin S_k, 0 \leq u \leq t\}\}}{\mathbb{P}\{\mathbf{X}(u) \notin S_k, 0 \leq u \leq t\}} \\ &\approx \frac{\mathbb{P}\{\tilde{\mathbf{X}}(t) \in S_{k-1}\}}{\mathbb{P}\{\tilde{\mathbf{X}}(t) \notin S_k\}} = \frac{p_{k-1}(t)}{1 - p_k(t)}. \end{aligned}$$

□

**Theorem 3.6. [Activation rate for key nodes]** For every  $k \geq 1$ ,

$$m_k(t) \approx \lambda(\hat{a}(t)) \frac{p_{k-1}(t)}{1 - p_k(t)},$$

where  $p_{k-1}(t)$  and  $p_k(t)$  are defined as in Theorem 3.3.

*Proof.* Combine Lemmas 3.4–3.5. □

#### 4. OPTIMAL NUMBER OF KEY NODES

In Section 4.1, we investigate the optimal number of key nodes  $k_c(\gamma)$  for five specific choices of the rate functions  $\lambda(\cdot)$  and  $\mu(\cdot)$  governing the transitions to the active and to the passive state, respectively. In Section 4.2, we do the same for the aging network and the rate functions simulated in [21]. The numerics is based on the approach mentioned in Remark 3.2 and the choice  $\gamma = 0.4$  mentioned below Remark 2.1.

**4.1. Optimal number of key nodes for various rate functions.** We assume that  $\lambda(a) < \mu(a)$  for all  $a \in (0, 1]$ , which ensures that the system does not exhibit unbounded growth in activation and, instead, tends to a stable and predominantly passive state. The corresponding results are presented in Table [1](#) and Figure [3](#).

**Constant rate functions.** As observed in Table [1](#), numerical estimates for the constant rate functions  $\lambda(a) = 0.45$  and  $\mu(a) = 0.65$  suggest that the optimal number of key nodes is  $k_c(\gamma) = 2$ . This finding is further supported by the numerical plot of the activation rate for key nodes  $m_k$  shown in Figure [3](#)(A). The activation rate  $m_k(t)$  increases over time for each value of  $k$  and eventually reaches a steady state.

**Exponential rate functions.** We examine two scenarios: (i) the exponential-constant rate functions  $\lambda(a) = e^{-a}$  and  $\mu(a) = 1$ ; (ii) constant-exponential rate functions  $\lambda(a) = 1$  and  $\mu(a) = e^a$ . As shown in Table [1](#), numerical estimates suggest that the optimal number of key nodes is  $k_c(\gamma) = 2$  in both scenarios. Figures [3](#)(B)-(C) illustrate the activation rate  $m_k(t)$  in these cases, which further support the conclusion that  $k = 2$  is optimal.

**Power rate functions.** As observed in Table [1](#), numerical estimates for the power-constant rate functions  $\lambda(a) = (a + 1)^{-1}$  and  $\mu(a) = 1$  suggest that the optimal number of key nodes is  $k_c(\gamma) = 1$ . However, due to the closeness of  $\hat{\tau}^\gamma$  to the numerical estimate of  $\mathbb{E}[\tau_2]$ , along with the potential impact of numerical errors, the value  $k = 2$  might also be considered a reasonable estimate of  $k_c(\gamma)$ . This observation is further supported by the numerical plot of the activation rate  $m_k(t)$  shown in Figure [3](#)(D).

**Logarithmic rate functions** As shown in Table [1](#), numerical estimates for the logarithmic rate functions  $\lambda(a) = (\log(3 + a))^{-1}$  and  $\mu(a) = \log(3 + a)$  suggest that the optimal number of key nodes is  $k_c(\gamma) = 2$ . Figure [3](#)(E) shows the dynamics of the activation rate  $m_k(t)$ , which further supports that  $k = 2$  is a reasonable choice for the optimal number of key nodes in this case.

TABLE 1. Numerical estimates of the expected hitting time  $\mathbb{E}[\tau_k]$  and the trigger time  $\hat{\tau}^\gamma$  for various choices of the rate functions and  $\gamma = 0.4$ . The optimal number of key nodes  $k_c(\gamma)$ , for which  $\mathbb{E}[\tau_k]$  is closest to  $\hat{\tau}^\gamma$ , corresponds to the  $k$ -values indicated in boldface. Recall Remark [3.2](#).

$k$	$\lambda(a) = 0.45$ $\mu(a) = 0.65$	$\lambda(a) = e^{-a}$ $\mu(a) = 1$	$\lambda(a) = 1$ $\mu(a) = e^a$	$\lambda(a) = (a + 1)^{-1}$ $\mu(a) = 1$	$\lambda(a) = (\log(3 + a))^{-1}$ $\mu(a) = \log(3 + a)$
1	1.37	0.79	0.62	<b>0.78</b>	0.72
2	<b>4.94</b>	<b>3.14</b>	<b>2.22</b>	2.95	<b>2.69</b>
3	9.36	6.19	4.27	5.66	5.22
4	17.48	11.94	8.12	10.61	9.97
5	33.35	23.45	15.84	20.19	19.49
6	65.08	47.34	31.88	39.49	39.30
$\hat{\tau}^\gamma$	3.46	2.79	2.00	1.61	3.11

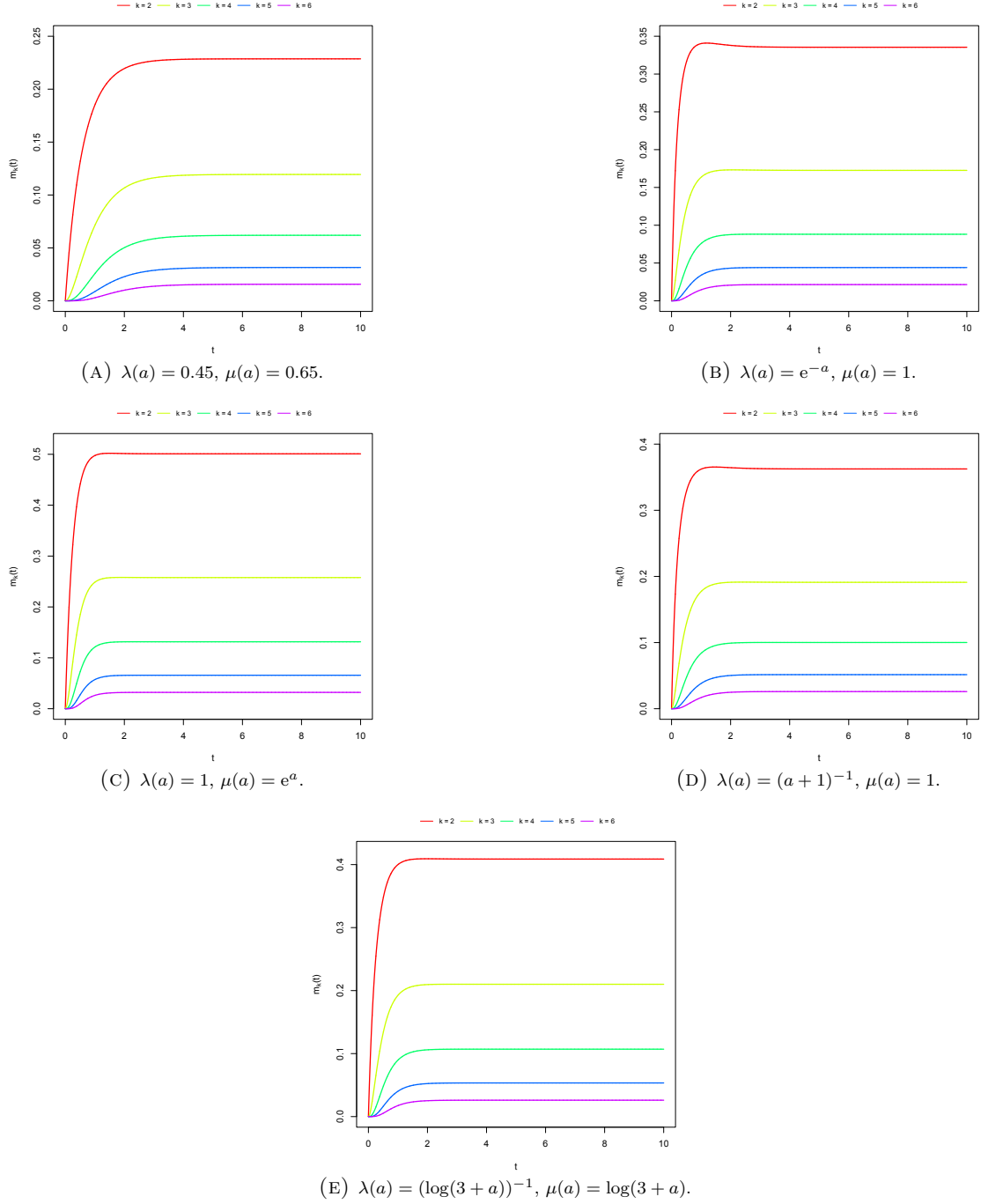


FIGURE 3. Comparison of the activation rate dynamics for the key nodes under different choices of the rate functions  $\lambda(a)$  and  $\mu(a)$  for the choice  $\gamma = 0.4$ . The colours indicate the number of key nodes. In all panels,  $m_k(t)$  starts from zero and increases initially, but the rate and level of this growth depend on how the activation and deactivation rates vary. When  $\lambda(a)$  is constant or increasing (e.g., panels A and C), activation remains effective as  $a$  grows, and  $m_k(t)$  reaches higher plateaus (compare panels B and C). In contrast, when  $\lambda(a)$  decreases with  $a$  (e.g., panels B, D and E), activation weakens over time, resulting in lower plateaus. In all cases, increasing  $k$  slows down the growth of  $m_k(t)$  and reduces its plateau, since the same activation rate is shared among more key nodes.

**4.2. Optimal number of key nodes in an aging network.** Consider the rate functions chosen in [21] (and analysed in [12]),

$$\lambda(a) = \Gamma_0 e^{r_+ a}, \quad \mu(a) = \frac{\Gamma_0}{R} e^{-r_- a}, \quad a \in [0, 1], \quad (4.1)$$

where the parameters are chosen to be  $r_+ = 10.27$ ,  $r_- = 6.5$ ,  $R = 1.5$  and  $\Gamma_0 = 0.00113$ .

TABLE 2. Numerical estimates of the expected hitting times  $\mathbb{E}[\tau_k]$  for the rate functions in (4.1).

$k$	$\mathbb{E}[\tau_k]$
1	86.94
2	95.49
3	97.14
4	97.62
5	97.80
6	97.89
7	97.93

TABLE 3. Numerical estimates of  $\hat{\tau}^\gamma$  for various values of  $\gamma$  for the rate functions in (4.1).

$\gamma$	$\hat{\tau}^\gamma$
0.2	81.79
0.3	91.19
0.4	95.05
0.5	96.69
0.6	97.40

A comparison of Tables 2–3 shows that  $k_c(\gamma) = 1$  for  $\gamma = 0.2, 0.3$ ,  $k_c(\gamma) = 2$  for  $\gamma = 0.4$ , and  $k_c(\gamma) = 3$  for  $\gamma = 0.5$ . For  $\gamma = 0.6$ , although the numerical estimates suggest that  $k_c(\gamma) = 4$ , the small difference of  $\hat{\tau}^{0.6}$  from the estimate of  $\mathbb{E}[\tau_3]$ , along with the potential impact of numerical errors, indicates that  $k_c(\gamma) = 3$  is also a reasonable and efficient choice. Note that for the parameter values used in (4.1), the threshold  $\gamma = 0.4$  is particularly appropriate, as death can occur when approximately half of the aggregated nodes are damaged. Therefore, the optimal number of key nodes is  $k_c(\gamma) = 2$ . Even for a conservative choice of  $\gamma \in [0.4, 0.6]$ , it is reasonable to take  $k_c(\gamma) \in \{2, 3\}$ . This provides a mathematical justification for using two mortality nodes in a health network, as adopted in [21, 12].

We also note that the rate functions in (4.1) fail to satisfy the inequality  $\lambda(a) < \mu(a)$  for all  $a \in (0, 1]$ , which makes it somewhat different from the five choices used in Section 4.1.

## 5. CONCLUSION

We have shown that, for large scale-free and disassortative signal networks, only a very small number of nodes need to be monitored to get a fair impression of the overall state of the network. This dramatic form of *data aggregation* is important because for large networks it is difficult, if not impossible, to monitor the overall state. We have further shown that the reduction

is *universal* and is not related to the underlying structure of the network. Three examples serve as an illustration: brain networks, aging networks, and friendship networks. Based on two assumptions – mean-field and homogeneity – we have shown how the optimal number of key nodes can be identified analytically and computed numerically for different activation and deactivation rates.

## REFERENCES

- [1] S. Achard, R. Salvador, B. Whitcher, J. Suckling, E.T. Bullmore, A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *Journal of Neuroscience* 26 (2006), 63–72.
- [2] L.A.N. Amaral, A. Scala, M. Barthelemy, H.E. Stanley, Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America* 97 (2000), 11149–11152.
- [3] A.-L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* 286 (1999), 509–512.
- [4] A. Barrat, M. Barthélemy, A. Vespignani, Modeling the evolution of weighted networks. *Physical Review E* 70 (2004), 066149.
- [5] D.S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke, E.T. Bullmore, Adaptive reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National Academy of Sciences of the United States of America* 103 (2006), 19518–19523.
- [6] D.S. Bassett, E.T. Bullmore, Small world brain networks. *Neuroscientist* 12 (2006), 512–523.
- [7] L.M. Bettencourt, G.J. Stephens, M.I. Ham, G.W. Gross, Functional structure of cortical neuronal networks grown in vitro. *Physical review E: Statistical, nonlinear, and soft matter physics* 75 (2007), 021915.
- [8] U. Braun, M.M. Plichta, C. Esslinger, C. Sauer, L. Haddad, O. Grimm, Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *Neuroimage* 16 (2012), 1404–1412.
- [9] M. Buchanan, *Small Worlds and the Groundbreaking Science of Networks*. W.W. Norton & Company, New York London, 2002.
- [10] E. Bullmore, O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10 (2009), 186–198.
- [11] V.M. Eguíluz, D.R. Chialvo, G.A. Cecchi, M. Baliki, A.V. Apkarian, Scale-free brain functional networks. *Physical Review Letters* 94 (2005), 018102.
- [12] V. Flitner, B. Heidergott, F. den Hollander, I. Lindner, A. Parvaneh, H. Strulik, A unifying theory of aging and mortality. <https://doi.org/10.48550/arXiv.2412.12815>
- [13] B. Gompertz, On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London* 115 (1825), 513–583.
- [14] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C.J. Honey, V.J. Wedeen, O. Sporns, Mapping the structural core of human cerebral cortex. *PLOS Biology* 6 (2008), e159.
- [15] R.S. Hazra, F. den Hollander, A. Parvaneh, The friendship paradox for sparse random graphs. *Probability Theory and Related Fields* (2025). <https://doi.org/10.1007/s00440-025-01365-w>
- [16] M.P. van den Heuvel, C.J. Stam, M. Boersma, H.E. Hulshoff Pol, Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage* 43 (2008), 528–539.
- [17] Y. Iturria-Medina, R.C. Sotero, E.J. Canales-Rodriguez, Y. Aleman-Gomez, L. Melie-Garcia, Studying the human brain anatomical network via diffusionweighted MRI and graph theory. *Neuroimage* 40 (2008), 1064–1076.
- [18] M.O. Jackson, *Social and Economic Networks*. Princeton University, 2010.
- [19] R. Koch, *The 80/20 Principle: The Secret to Achieving More with Less*. 3rd ed., Crown, 2011.
- [20] S. Lim, F. Radicchi, M.P. van den Heuvel, O. Sporns, Discordant attributes of structural and functional brain connectivity in a two-layer multiplex network. *Scientific Reports* 9 (2019), 2885.
- [21] A.B. Mitnitski, A.D. Rutenber, S. Farrell, K. Rockwood, Aging, frailty and complex networks. *Biogerontology* 18 (2017) 433–446.
- [22] M.E. Newman, The structure and function of complex networks. *SIAM Review* 45 (2003), 167–256.
- [23] R. Pastor-Satorras, A. Vazquez, A. Vespignani, Dynamical and correlation properties of the Internet. *Physical Review Letters* 87 (2001), 258701.
- [24] J.C. Reijneveld, S.C. Ponten, H.W. Berendse, C.J. Stam, The application of graph theoretical analysis to complex networks in the brain. *Clinical Neurophysiology* 118 (2007), 2317–2331.
- [25] O. Sporns, *Networks of the Brain*. MIT Press, Cambridge, MA, 2011.
- [26] O. Sporns, D. Chialvo, M. Kaiser, C.C. Hilgetag, Organization, development and function of complex brain networks. *Trends in Cognitive Sciences* 8 (2004), 418–425.



- [27] C.J. Stam, J.C. Reijneveld, Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomedical Physics* 1 (2007), 3.
- [28] S. Yu, D. Huang, W. Singer, D. Nikolic, A small world of neuronal synchrony. *Cerebral Cortex* 18 (2008), 2891–2901.