



UvA-DARE (Digital Academic Repository)

Information processing in complex networks

Quax, R.

Publication date
2013

[Link to publication](#)

Citation for published version (APA):

Quax, R. (2013). *Information processing in complex networks*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 2

The diminishing role of highly connected units in the dynamical behavior of complex systems

Material from this chapter is submitted as a Research Article to Journal of the Royal Society Interface. In addition, the presented concept is published as an editorial letter in Journal of Computational Science. Lastly, material from this chapter and the introduction is published in the European Physics Journal in the Special Topics section.

2.1 Introduction

Many non-equilibrium systems consist of dynamical units which interact through a network to produce complex behavior as a whole. In a wide variety of such systems, each unit has an internal microstate that quasi-equilibrates to the distribution of states of the units it interacts with, or ‘interaction potential’, which results in its next observable macrostate. This assumption is also known as the local thermodynamic equilibrium (LTE), originally formulated to describe radiative transfer inside stars (99, 100). Examples of systems of coupled units that have been described in this manner include brain networks (12, 83, 101, 102), cellular regulatory networks (103–107), immune networks (108, 109), social interaction networks (75, 77, 110–112), and financial trading markets (23, 110, 113). A state change of one unit may subsequently cause a neighbor unit to change its state, which may in turn cause other units to change, and so on. The core problem of understanding the system’s behavior is that the topology of

interactions mixes cause and effect of units in a complex manner, making it hard to tell which units drive the system dynamics.

The main goal of complex systems research is to understand how the dynamics of individual units combine to produce the behavior of the system as a whole. A common method to dissect the collective behavior into its individual components is to remove a unit and observe the effect (25, 29–31, 114–119). In this manner it has been shown, for instance, that highly connected units, or hubs, are crucial for the structural integrity of many real systems (116), i.e., removing only a few hubs disconnects the system into subnetworks which can no longer interact. On the other hand, Tanaka et al. find that sparsely connected units are crucial for the dynamical integrity of systems where the remaining (active) units must compensate for the removed (failed) units (25). Less attention has been paid to study the interplay of the unit dynamics and network topology, from which the system's behavior emerges, in a unified manner.

Here we introduce an information-theoretical approach to quantify to what extent the system's state is actually a representation of an instantaneous state of an individual unit, using the definitions of entropy and mutual information. We will define the information dissipation time (IDT) of a unit as the time it takes for the information about the unit to be lost from the network. It is shown analytically that the IDT of a unit is a decreasing function of its connectivity, in the limit. It is also calculated numerically, and compared to three empirical datasets of the dynamical importance of units in real networks.

We consider large static networks of identical units whose dynamics can be described by the Gibbs measure. The Gibbs measure describes how a unit changes its state subject to the combined potential of its interacting neighbors, in case the LTE is appropriate and using the maximum-entropy principle (120, 121) to avoid assuming any additional structure. In other words, we focus on systems where each unit can actually be in a large number of (unobservable) microstates which translate many-to-one to the

(observable) macrostates. In fact, each unit may even be a subsystem in its own right in a multi-scale setting, such as a cell in a tissue or a person in a social network. At a small timescale, each unit probabilistically chooses its next macrostate depending on the current states of its neighbors, termed Markov networks (122). We consider random interaction networks with a given degree distribution $p(k)$, which denotes the probability that a randomly selected unit has k interactions with other units, and which have a maximum degree k_{\max} that grows less than linear in the network size N . Self-loops are not allowed. No additional topological features are imposed, such as degree-degree correlations or community structure. An important consequence of this for our purposes is that the network is ‘locally tree-like’ (42, 123), i.e., link cycles are exceedingly long.

We validate our analytical predictions qualitatively using numerical experiments of random networks of 6000 ferromagnetic Ising spins where the number of interactions k of a spin is distributed as a power-law $p(k) \propto k^{-\gamma}$. Ising-spin dynamics are extensively studied and are often used as a first approximation of the dynamics of a wide variety of complex physical phenomena (42). We find qualitative evidence in empirical data of the dynamical importance of units as function of their degree in three different domains, namely viral marketing in social networks (124), evolutionary conservation of human proteins (125), and the transmission of a neuron’s activity in neural networks (126).

2.2 Results

2.2.1 Information dissipation time of a unit

As a measure of the dynamical importance of a unit s we calculate its information dissipation time (IDT), denoted $D(s)$. In words, it is the time it takes for the information about the state of the unit s to disappear from the network’s state. As another way of describing it, it is the time it takes for the network as a whole to forget a particular state of a single unit.

In this subsection we derive analytically a relation between the number of interactions of a unit and the IDT of its state.

Terminology

A system S consists of units s_1, s_2, \dots among which some pairs of units $E = (s_i, s_j), (s_k, s_l), \dots$ interact with each other. Each interaction is undirected and the number of interactions that involve unit s_i is denoted by k_i , called its *degree*, which equals d with probability $p(d)$, called the degree distribution. The set of k_i units that s_i interacts with directly is denoted by $h_i = \{x : (s_i, x) \in E\}$. The state of unit s_i at time t is denoted by s_i^t , and the collection $S^t = s_1^t, s_2^t, \dots, s_N^t$ forms the state of the system. Each unit probabilistically chooses its next state based on its current state and the current state of each of its nearest neighbors in the interaction network. Unit s_i chooses the next state x denoted by the conditional probability distribution $p(s_i^{t+1} = x | h_i^t)$. This is also known as a Markov network.

Information as a measure of dynamical impact

The instantaneous state of a system S^t consists of $H(S^t)$ bits of Shannon information. In other words, $H(S^t)$ answers to unique yes/no questions (bits) which must be specified in order to determine a unique state S^t . As a consequence, the more bits about S^t are determined by the instantaneous state $s_i^{t_0}$ of a unit s_i at time $t_0 \leq t$, the more the system state S^t depends on the unit's state $s_i^{t_0}$.

The impact of a unit's state $s_i^{t_0}$ on the system state S^t at a particular time t is their mutual information $I(S^t | s_i^{t_0})$. In the extreme case that $s_i^{t_0}$ fully determines the state S^t , the entropy of the system state coincides with the entropy of the unit state, and the dynamical impact is maximum at $H(S^t) = H(s_i^{t_0}) = I(S^t | s_i^{t_0})$. In the other extreme case that the unit state $s_i^{t_0}$

is completely irrelevant to the system state S^t , the information is minimum at $I(S^t | s_i^{t_0}) = 0$.

This mutual information integrated over time (as $t \rightarrow \infty$) is then a measure of the extent that the system's state trajectory is affected by an instantaneous state of the unit. In other words, it measures the 'dynamical importance' of the unit. Instead of the integration we choose to calculate the IDT of the unit, which in our model not only fully characterizes the integral mutual information but is also a measure of how this dynamical importance decreases over time. It is therefore a more informative measure; also, it may be more general, as the integration may be undefined for other systems.

Defining the IDT of a unit

At each time step, the information stored in a unit's state s_i^t is partially transmitted to the next states of its nearest neighbors (93, 95), which in turn transmit it to their nearest neighbours, and so on. The state of unit s at time t dictates the system state at the same time t at the amount

$$I_0^k \equiv I(S^t | s^t) = I(s^t | S^t) = H(s^t),$$

with the understanding that unit s has k interactions. At time $t+1$ the system state is still influenced by the unit's state s^t , now at the amount

$$I_1^k \equiv I(S^{t+1} | s^t) = I(h^{t+1}, s^{t+1} | s^t).$$

As a result, a unit with k connections locally dissipates its information at a rate I_1^k / I_0^k . (Note the analog with the quantity f from the example of coin flips in the Introduction, in Section 1.1.3.) Here we utilize the observation that the information about a unit's state s^t , which is at first present at the unit itself at the maximum amount $H(s^t)$, can be only transferred at time $t+1$ to the direct neighbors h of s through the nearest-neighbor interactions.

Since we are particularly interested in the *non-local* dynamical importance of a unit, we compute only the amount of information that a unit transfers to its neighbors and not to its own new state, so instead of I_1^k we will use

$$I_1^k \equiv I(h^{t+1} | s^t).$$

It greatly simplifies the subsequent expressions. In addition, the change makes sense: we are interested in how a unit influences the rest of the network, not its own state. Even if a portion of the unit's information would be retained in its own new state, this information would disseminate through the network in the same manner. Formally we must now assume that $I(s^{t+1} | s^t) \leq I(h^{t+1} | s^t)$ in order for the IDT of the unit to remain unchanged. This assumption is reasonable for our model because a unit has no direct influence on its own next state, i.e., there are no self-loops in the network of interactions.

At subsequent time steps ($t + 2$ and onward), the information about the unit with an amount of I_1^k will disseminate at a constant average rate

$$\hat{I} = \sum_k q(k) \cdot I_1^{k+1} / I_0^{k+1}$$

from its neighbours and other units. This is due to the absence of degree-degree correlations or other structural bias in the network. Here, $q(k) = (k+1)p(k+1)\langle k \rangle^{-1}$ is the probability distribution of the number of additional interactions that a nearest-neighbour unit contains besides the interaction with unit s , called the excess degree distribution (127). As a consequence, the dissemination of information of all nodes occurs at an equal rate except for the initial amount of information I_1^k , which the k neighbour states contain at time $t + 1$, which depends on the degree k of the unit.

Note that, in general, the rate that the information about s_i^t dissipates from $t+2$ and onward equals \hat{I} up to a correction factor that depends on the state-state correlations implied by the conditional transition probabilities $p(s_k^{t+1} | s_j^t)$. For example, if s_A^t dictates 20% of the information stored in its neighbor state s_B^{t+1} , and s_B^{t+1} in turn dictates 10% of the information in s_C^{t+2} , then $I(s_A^t | s_C^{t+2})$ may not necessarily equal $20\% \cdot 10\% = 2\%$ of the information $H(s_C^{t+2})$ stored in s_C^{t+2} . Therefore we assume that our approximation of the absolute dynamical impact of a unit's state is accurate up to a constant factor. Since we are interested in comparing the IDTs of units as function of their degree, a relative measure suffices and a common correction factor does not affect our conclusions.

We are now ready to define the IDT of a unit up to a constant. The number of time steps it takes for the information in the network about unit s with degree k to reach an arbitrarily small constant ε satisfies

$$D(s) \propto \hat{I} \log \frac{\varepsilon}{I_1^k} = \frac{\log \varepsilon - \log I_1^k}{\log \hat{I}}.$$

Note that $D(s)$ is not equivalent to the classical correlation length. The correlation length is a measure of the time it takes for a unit to lose a *certain fraction of its original correlation* with the system state, instead of the time it takes for the unit to reach a certain absolute value of correlation. For our purpose of comparing the dynamical impact of units, the correlation length would not be a suitable measure. For example, if unit A has a large initial correlation with the system state and another unit B has a small initial correlation, but the halftime of their correlation is equal, then in total we consider A to have more impact on the system's state since it dictates more bits of information of the system state.

A note on causation versus correlation

In the general case, the mutual information $I(s_x^t | s_y^{t_0})$ between the state of unit s_x at time t_0 and another unit's state s_y at time t is the sum of two parts: I_{causal} , which is information that is due to a causal relation between the state variables, and I_{corr} , which is information due to 'correlation' that does not overlap with the causal information. Correlation occurs if the units s_x and s_y both causally depend on a third 'external' variable e in a similar manner, i.e., such that $I(e | (s_x^t, s_y^{t_0})^T) < I(e | s_x^t) + I(e | s_y^{t_0})$. This can lead to a non-zero mutual information $I(s_x^t | s_y^{t_0})$ among the two units, even if the two units would not directly depend on each other in a causal manner.

For this reason we do not directly calculate the dependence of $I(S^t | s^{t_0})$ on the time variable t in order to calculate the IDT of a unit s . It would be difficult to tell how much of this information is non-causal at every time point. In order to find this out we would have to understand exactly how each bit of information is passed onward through the system, from one state variable to the next, which we do not yet understand at this time.

To prevent measuring the non-causal information present in the network we use local single-step 'kernels' of information diffusion, namely the I_1^k / I_0^k as discussed previously. The information I_0^k is trivially of causal nature (i.e., non-causal information is zero), since it is fully stored in the state of the unit itself. Although in the general case, I_1^k may consist of a significant non-causal part, in our model we expect this to be zero or at most an insignificant amount. We did not formally derive this property for our model, however, so we must assume it. The rationale is that units do not self-interact (no self-loops) and the network is locally tree-like: if s_x and s_y are direct neighbors, then there is no third s_z with 'short' interaction pathways to both s_x and s_y . The only way that non-causal (i.e., not due to

s_x^t influencing s_y^{t+1}) information can be created between s_x^t and s_y^{t+1} is through the two interaction paths $s_z^{t'} \rightarrow \dots \rightarrow s_y^{t-1} \rightarrow s_x^t$ and $s_z^{t'} \rightarrow \dots \rightarrow s_y^{t+1}$, where $t' < t-1$. That is, one and the same state variable $s_z^{t'}$ must causally influence both s_x^t and s_y^{t+1} , where it can only reach s_x through s_y . We expect any thusly-induced non-causal information in $I(s_y^{t+1} | s_x^t)$ is insignificant compared to the causal information through $s_x^t \rightarrow s_y^{t+1}$, and the reason is three-fold. Firstly, the minimum lengths of the two interaction paths from s_z are two and three interactions respectively, where information is lost through each interaction due to its stochastic nature. Secondly, of the information that remains, not all information $I(s_z^{t'} | s_x^t)$ may overlap with $I(s_z^{t'} | s_y^{t+1})$, but even if it does then the ‘correlation part’ of the mutual information $I(s_y^{t+1} | s_x^t)$ due to this overlap is upper bounded by their minimum: $\min\{I(s_z^{t'} | s_x^t), I(s_z^{t'} | s_y^{t+1})\}$. Thirdly, the mutual information due to correlation may in general overlap with the causal information, i.e., both pieces of information may be partly about the same state variables. That is, the I_{corr} part of $I(s_y^{t+1} | s_x^t)$, which is the error of our assumption, is only that part of the information-due-to-correlation that is not explained by (contained in) I_{causal} . The final step is the observation that I_1^k is the combination of all $I(s_y^{t+1} | s_x^t)$ for all neighbor units $s_y \in h_x$.

Unit dynamics in the local thermodynamic equilibrium

Before we can proceed to show that $D(s)$ diminishes to a constant for highly connected units, we must first define the class of unit dynamics in more detail. That is, we must arrive at an expression for the conditional probabilities $p(s^{t+1} = r | h^t)$.

We assume that every unit in the network can actually be in a wide variety of (unobservable) microstates, which translate many-to-one to the

(observable) macrostates Σ . In fact, each unit may even be a subsystem in its own right in a multi-scale setting, such as a cell in a tissue or a person in a social network. In the general case we have no further knowledge about the statistics of these microstates so we assume that the microstates which yield the same energy in the network are equally probable. In the limit of the size of the full network, we can interpret each unit as the ‘system’ and the rest of the network as the ‘heat bath’ in the canonical ensemble. The whole network itself may be part of a ‘universe’ (heat bath, or surroundings) in turn as well. We focus on Markov networks, so the dynamics of each unit is governed by the same set of conditional probabilities $p(s^{t+1} = r | h^t)$ with the Markov property.

In this LTE description, a unit chooses its next (macro)state depending on the energy of that state, which is induced by the (macro)states of its nearest-neighbors through its interactions. The higher the energy of a state, the less probable the unit chooses the state. Stochasticity can arise if multiple states have an equal energy, and additional stochasticity is introduced by means of the temperature of the heat bath that surrounds the network.

The consequence of the LTE that is relevant to our study is that the state transition probabilities of the highly connected units belong to the exponential family with respect to the energy. That is, in a discrete-time description, s^t chooses $s^{t+1} = r$ as the next state with a probability

$$p(s^{t+1} = r | h^t) \propto \exp \sum_{s_j \in h} E(r | s_j^t) / T,$$

where T is the temperature of the network’s heat bath and $\sum_j E(r | s_j^t)$ is the energy of state r given the states of its interacting neighbors $s_j^t \in h^t$. As a result, the energy landscape of r does not depend on individual states of specific neighbor units; it depends on the distribution of neighbor states.

2.2.2 Diminishing IDT of hubs

As a function of the degree k of unit s ,

$$D(s) \propto \log I_1^k,$$

because \hat{T} and ε are independent of the unit's degree. Since the logarithm preserves order, to show that the dynamical importance diminishes for high-degree units it is sufficient to show that I_1^k decreases to a constant as $k \rightarrow \infty$, which we do next.

The quantity I_1^k is upper bounded by the arithmetic sum

$$\sum_{s_j^{t+1} \in h_i^{t+1}} I(s_j^{t+1} | s_i^t),$$

and lower bounded by zero. In the average case we can write the amount of information in the next system state about the unit's current state as

$$I_1^k = U(k) \cdot k \cdot T(k), \text{ where}$$

$$T(k) = \left\langle I(s_j^{t+1} | s_i^t) \right\rangle_{k_j},$$

where $T(k)$ is the information in a neighbor unit's next state averaged over its degree, and $U(k) \in [k^{-1}, 1]$ is the degree of uniqueness of the next states of the neighbors: it equals unity in case the information of a neighbor does not overlap with that of any other neighbor unit of s_i^t , and it is less than unity to the extent that information does overlap between neighbor units. The operator $\langle \cdot \rangle_{k_j}$ denotes an average over the degree k_j of a neighbor unit s_j , i.e., weighted by the excess degree distribution $q(k_j - 1)$.

Since the factor $U(k) \cdot k$ is at most a linear growing function of k , a sufficient condition for $D(s_i)$ to diminish as $k \rightarrow \infty$ is for $T(k)$ to decrease more strongly than linear in k .

First we write $T(k)$ as an expected energy difference of a high-degree node:

$$\begin{aligned} T(k) &= \left\langle H(s_i^t) - H(s_i^t | s_j^{t+1}) \right\rangle_{k_j} \\ &= \sum_q \sum_{r_j} p(s_j^{t+1} = r_j) \cdot p(s_i^t = q | s_j^{t+1} = r_j) \left\langle \log \frac{p(s_i^t = q | s_j^{t+1} = r_j)}{p(s_j^{t+1} = r_j)} \right\rangle_{k_j} \\ &\propto \sum_q \sum_{r_j} p(s_j^{t+1} = r_j) \cdot p(s_i^t = q | s_j^{t+1} = r_j) \left\langle E(s_i^t = q | s_j^{t+1} = r_j) - E(s_i^t = q) \right\rangle_{k_j}. \end{aligned}$$

We observe that the energy of a unit's state is the sum of the interaction energies of the unit's state with each neighbor state, so

$$\left\langle E(s_i^t = q) \right\rangle = \sum_{j=1}^{k_i} \left\langle e(s_i = q, s_j = r_j) \right\rangle_{k_j, r_j} = k_i \cdot e_q,$$

and

$$\begin{aligned} \left\langle E(s_i^t = q | s_j^{t+1} = r_m) \right\rangle &= \sum_{\substack{j=1 \\ j \neq m}}^{k_i} \left\langle e(s_i = q, s_j = r_j) \right\rangle_{k_j, r_j} + \left\langle e(s_i^t = q | s_m^{t+1} = r_m) \right\rangle_{k_m, r_m} \\ &= k_i \cdot e_q + \left(e_q(s_m^{t+1} = r_m) - e_q \right). \end{aligned}$$

Here, $e(s_i = q, s_j = r_j)$ is the interaction energy between two interacting units with states q and r_j , e_q is the expected energy between two interacting units given that one of the units has state q , and $e_q(s_m^{t+1} = r_m)$ is the expected energy between a unit with state q and another unit whose next state will be r_m .

Averaged over q , r_m and k_m , we first rename the expected energy difference of knowing the next state of one of the interacting units

$$\left(e_q(s_m^{t+1} = r_m) - e_q \right) \equiv \Delta_q^{t+1},$$

for convenience, which is a quantity that is conditionally independent of k_i given q but which is otherwise unknown. Then we obtain

$$T(k) \propto \sum_q \sum_{r_j} p(s_j^{t+1} = r_j) \cdot p(s_i^t = q | s_j^{t+1} = r_j) \cdot \Delta_q^{t+1},$$

simplified as

$$T(k) \propto \sum_q T_q(k), \text{ where}$$

$$T_q(k) \equiv \sum_{r_j} p(s_j^{t+1} = r_j) \cdot p(s_i^t = q | s_j^{t+1} = r_j) \cdot \Delta_q^{t+1}.$$

This quantity converges to a constant since the probability of each Δ_q^{t+1} either decays exponentially to zero, or exponentially decays upward to a positive constant $1/n$ where n is the number of states with equal lowest energy. More specifically, since the average sum of interaction energies of a unit increases linearly with its degree, we find in successive steps of algebra,

$$\begin{aligned}
T(k+1) &\propto \sum_q \Delta_q^{t+1} \cdot \sum_{r_j} p(s_{j^+}^{t+1} = r_j) \cdot p(s_{i^+}^t = q | s_{j^+}^{t+1} = r_j) \\
T(k+1) &\propto \sum_q \Delta_q^{t+1} \cdot p(s_{i^+}^t = q) \\
T(k+1) &\propto \sum_q \Delta_q^{t+1} \cdot \frac{\sum_{r_1, \dots, r_{k+1} \in \Sigma^{k+1}} e^{-\sum_{j=1}^{k+1} e(q, r_j)}}{Z_{k+1}} \\
T(k+1) &\propto \sum_q \Delta_q^{t+1} \cdot \frac{\sum_{r_1, \dots, r_k \in \Sigma^k} e^{-\sum_{j=1}^k e(q, r_j)}}{Z_k} \cdot \frac{\sum_{r_m \in \Sigma} e^{-e(q, r_m)}}{Z_1} \\
T(k+1) &\propto \sum_q p_1(q) \cdot \Delta_q^{t+1} \cdot \sum_{r_j} p(s_j^{t+1} = r_j) \cdot p(s_i^t = q | s_j^{t+1} = r_j) \\
T(k+1) &\propto \sum_q p_1(q) \cdot T_q(k).
\end{aligned}$$

Here, s_i is a unit with degree k with neighbor units $s_j \in h_i$ and s_{i^+} is a unit with degree $k+1$ with neighbor units $s_{j^+} \in h_{i^+}$; further, $e(q, r_j)$ is a shorthand for $e(s_i = q, s_j = r_j)$, defined above. In words, we find that we obtain $T(k+1)$ by multiplying each term of $T(k)$ by a factor $p_1(q)$, which is a shorthand for the state probability $p(s_m = r)$ given that $k_m = 1$. This factor is independent of k and is strictly less than unity (except for the degenerate case where only one unit state is allowed). In formula, we find that

$$T(k+1) \propto a \cdot T(k) \text{ where } a \leq 1.$$

As a result, the expected value of $T(k)$ converges to a constant at an exponential rate as $k \rightarrow \infty$. Since each term is multiplied by a factor $p_1(s_m = q) \leq 1$, this convergence is downward for most systems but never upward even for degenerate system dynamics.

2.2.3 Numerical experiments with networks of Ising spins

For our experimental validation we calculate the dynamical importance $D(s)$ of 6000 ferromagnetic spins with nearest-neighbor interactions in a heavy-tailed network both numerically and in simulations and find that it indeed diminishes for highly connected spins. See Figure 3.

The numerical calculations use the Glauber dynamics (128) to describe locally how each spin updates its state depending on the states of its neighbors. We calculate the upper bound of $D(s)$ by setting $U(k)=1$, that is, all information about a unit's state is assumed to be unique which optimizes its dynamical importance. The network is characterized by a heavy-tailed degree distribution $p(k) \propto k^{-\gamma}$ and is otherwise random, i.e., there are no structural degree-degree correlations or community structures. The results are summarized in Figure 3 for $\gamma=1.6$; see Section A2 for higher values.

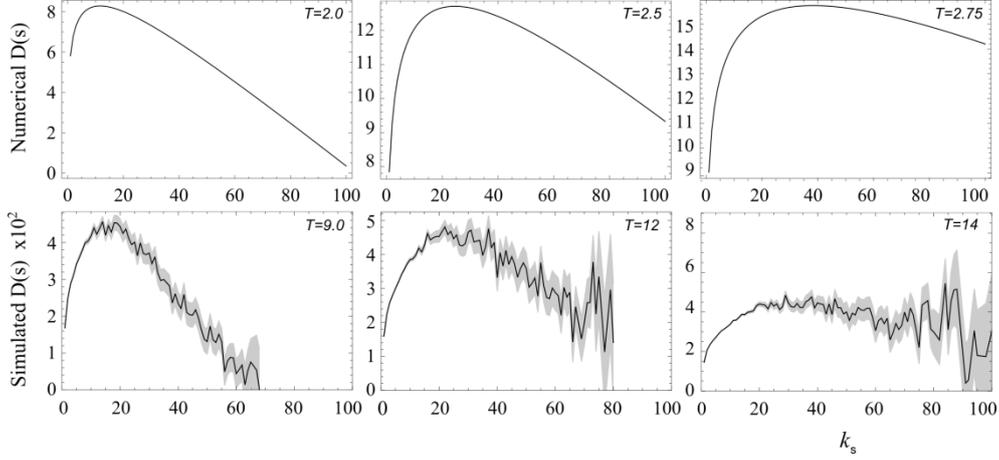


Figure 3: The dynamical impact $D(s)$ of a ferromagnetic Ising spin s as function of its connectivity k_s , from numerical calculation as well as from computer experiments. For the numerical calculations we used Glauber dynamics to describe the behavior of the units; for the computer experiments we used the Metropolis-Hastings algorithm. For the latter we simulate a network of 6000 spins with a power-law degree distribution $p(k) \propto k^{-1.6}$; the plots are the result of six realizations, each of which generated 90,000 time series of unit states that lead up to the same system state, which was chosen randomly after equilibration. The grey area is within two times the standard error of the mean IDT of a unit with a given connectivity.

In addition we perform computer simulations to produce time series of the states of 6000 ferromagnetic Ising spins, and we measure the dynamical importance of each unit by regression. For each temperature value we generate six random networks with $p(k) \propto k^{-\gamma}$ for $\gamma = 1.6$ and record the state of each spin at 90,000 time steps. The state of each unit is updated using the Metropolis-Hastings algorithm. Of the resulting time series of the unit states we computed the time d_i where $I(s_1^{t+d_i}, \dots, s_N^{t+d_i} | s_i^t) = \varepsilon$ of each unit s_i by regression. This is semantically equivalent to $D(s_i)$ but does not assume a locally tree-like structure or a uniform information dissipation rate \hat{I} . See Section A1 for details. The results are presented in the lower panel of Figure 3.

2.3 Empirical evidence

We present empirical measurements from the literature of the impact of units on the behavior of three different systems, namely networks of neurons, social networks, and protein dynamics. These systems are commonly modeled using a Gibbs measure to describe the unit dynamics. In each case, the highly connected units turn out to have a saturating or decreasing impact on the behavior of the system. This provides qualitative evidence that our measure $D(s)$ indeed characterizes the dynamical importance of a unit, and consequently, that highly connected units have a diminishing dynamical importance in a wide variety of complex systems. In each study, it remains an open question which mechanism is responsible for the observed phenomenon. Our work proposes a new candidate explanation for the underlying cause for each case, namely that it is an inherent property of the type of dynamics that govern the units.

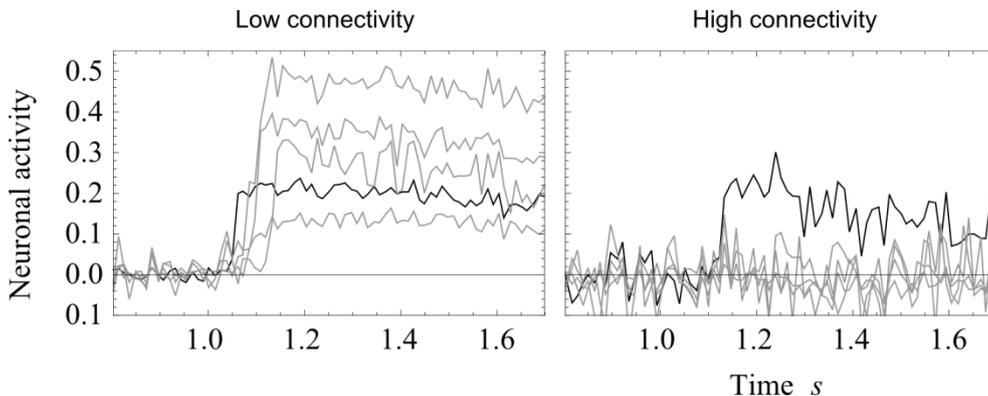


Figure 4: The level of activity of a set of neurons under a microscope as function of time, after seeding one neuron with an electrical potential (black line). The activity was measured by changes in calcium ion concentrations that were detected by imaging fluorescence levels. In the sparse cultures with few synapses per neuron, the stimulated neuron evokes a network burst of activity in all other neurons in the field after a short delay. In contrast, in the dense cultures with many synapses per neuron, only the stimulated neuron has an increased potential. Permission to reproduce this figure as well as its original data was kindly provided by Ivenshitz and Segal (126).

The first evidence is found in the signal processing of *in vitro* networks of neurons (126). The more densely neurons are placed in a specially prepared Petri dish, the more connections (synapses) each neuron creates with other neurons. In their experiments, Ivenshitz and Segal find that sparsely connected neurons are capable of transmitting their electrical potential to neighboring neurons, whereas densely connected neurons are unable to trigger network activity even if they are depolarized in order to discharge several action potentials. Their findings are summarized in Figure 4. In search for the underlying cause the authors exclude some obvious candidates, such as the ratio of excitatory versus inhibitory connections, the presence of compounds that stimulate neuronal excitability, and the size of individual postsynaptic responses. Although the authors do find tell-tale correlations, e.g., between the network density and the structure of the dendritic trees, they conclude that the phenomenon is not yet understood.

Secondly, in a person-to-person recommendation network consisting of four million persons, Leskovec et al. find that the most active recommenders are not necessarily the most successful (124). In the setting of word-of-mouth marketing among friends in social networks, the adoption rate of recommendations saturates or even diminishes for the highly active recommenders, which is shown in Figure 5 for four product categories. This observation is remarkable because in the dataset, the receiver of a recommendation does not know how many other persons receive it as well. As a possible explanation, the authors hypothesize that widely recommended products may not be suitable for viral marketing. Nevertheless the underlying cause remains an open question. We propose an additional hypothesis, namely that highly active recommenders have a diminishing impact on the opinion forming of others in the social network. In fact, the model of Ising spins in our numerical experiments is a widely used model for opinion forming in social networks (75, 77, 110–112). As a consequence, the results in Figure 5 may be interpreted as estimating the dynamical impact of a person's opinion as function of the number of friends that he debates his opinion with.

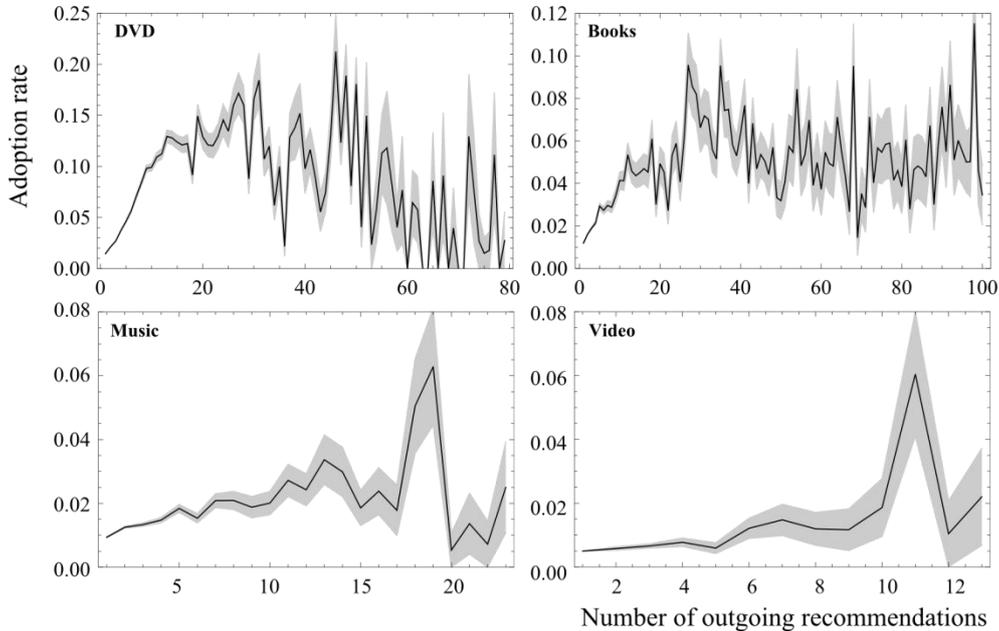


Figure 5: The success of a person’s recommendation of a product as function of the number of recommendations that he sent. A person could only recommend a product to friends after he purchased the product himself. The success is measured is a normalized rate of receivers buying the product upon the recommendation. The normalization counts each product purchase equally in terms of the system’s dynamics, as follows: if a person receives multiple recommendations for the same product from different senders, a ‘successful purchase’ is only accounted to one of the senders. The grey area is within one standard error of the mean. The total recommendation network consists of four million persons who made 16 million recommendations about half a million products. The subnetworks of the books and DVDs categories are by far the largest and most significant, with 73% of the persons recommending books and 52% of the recommendations concerning DVDs. The data in these figures are reproduced with permission from Leskovec (124).

The third empirical evidence is found in the evolutionary conservation of human proteins (125). According to the neutral model of molecular evolution, most successful mutations in proteins are irrelevant to the functioning of the system of protein interactions (129). This means that the evolutionary conservation of a protein is a measure of the intolerance of the organism to a mutation to that protein, i.e., it is a measure of the dynamical

importance of the protein to the reproducibility of the organism (130). Brown and Jurisica measured the conservation of human proteins by mapping the human interactome to mice and rats using orthologs, which is shown in Figure 6. Their analysis reveals that the conservation of highly connected proteins is inversely related with their connectivity. Again, this is consistent with our analytical prediction. The authors conjecture that this effect may be due to the overall high conservation rate, approaching the maximum of 1 and therefore affecting the statistics. We suggest that it may indeed be an inherent property of protein interaction dynamics.

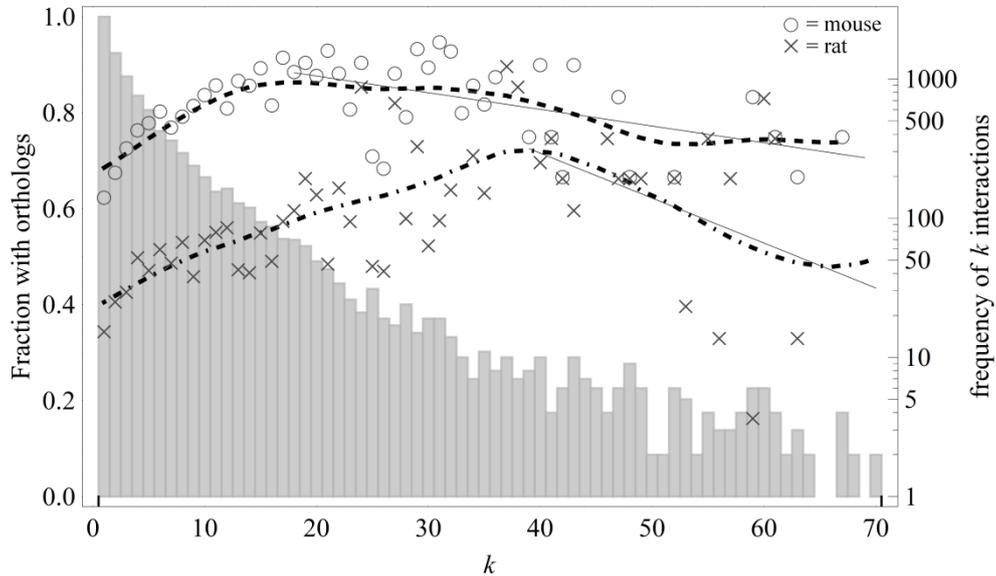


Figure 6: The rate of evolutionary conservation of human proteins as a function of their connectivity k . The rate of conservation is measured as the fraction of proteins that have an orthologous protein in the mouse (circles) and the rat (crosses). The dashed and dot-dashed curves show the trend of the conservation rates compared to mice and rats, respectively. They are calculated using a Gaussian smoothing kernel with a standard deviation of 10 data points. To evaluate the significance of the downward trend of both conservation rates, we performed a least-squares linear regression of the original data points starting from the peaks in the trend lines up to $k = 70$. For the fraction of orthologs with mice, the slope of the regression line is $-0.00347137 \pm 0.0011149$ (mean and standard error); with rats, the slope is $-0.00936715 \pm 0.00594403$. The vertical bars denote the frequency of proteins with k interactions in logarithmic scale. The data in this figure was reproduced with permission from Brown and Jurisica (125) who kindly provided its original data.

2.4 Discussion

We find that various research areas encounter a diminishing dynamical impact of hubs that is unexplained. Our analysis demonstrates that this phenomenon could be caused by the combination of unit dynamics and the topology of their interactions. We show that in large Markov networks, the dynamical behavior of highly connected units have a low impact on the

dynamical behavior of the system as a whole, in the case where units choose their next state depending on the interaction potential induced by their nearest-neighbors.

For highly connected units this type of dynamics enables the local thermodynamic equilibrium (LTE) assumption, originally used for describing radiative transport in a gas or plasma. To illustrate LTE, there is no single temperature value that characterizes an entire star: the outer shell is cooler than the core. Nonetheless, the mean free path of a moving photon inside a star is much smaller than the temperature gradient. Therefore, on a small time scale the photon's movement can be approximated using a local temperature value. A similar effect is found in various systems of coupled units, such as social networks, gene regulatory networks, and brain networks. In such systems, the internal dynamics of a unit is often faster than a change of the local interaction potential. Intuitive examples are the social interactions in blog websites, discussion groups, or product recommendation services. Here, changes that affect a person are relatively slow so that he can assimilate his internal state-of-mind (microstate) to his new local network of friendships and the set of personal messages he received, before he makes the decision to add a new friend or send a reply (macrostate). Indeed, this intuition combined with our analysis is consistent with multiple observations in social networks. Watts and Doods (131) numerically explored the importance of 'influentials', a minority of individuals who influence an exceptional number of their peers. They find counter to intuition that large cascades of influence are usually not driven by influentials, but rather by a critical mass of easily influenced individuals. Granovetter (132) found that even though hubs gather information from different parts of the social network and transmit it, the clustering and centrality of a node provide better characteristics for diffusing innovation (133). Rogers (134) found experimentally that the innovator is usually an individual in the periphery of the network, with few contacts with other individuals.

Our approach can be interpreted in the context of how dynamical systems intrinsically process information (85, 86, 93, 95, 135). That is, the state of each unit can be viewed as a (hidden) storage of information. As one unit interacts with another unit, part of its information is transferred to the state of the other unit (and vice versa). Over time, the information that was stored in the instantaneous state of one unit percolates through the interactions in the system, and at the same time it decays due to thermal noise or randomness. The more information that is stored in the system state originates from a previous state of the unit, the more the unit's state determines the future system state. Integrating this measure over time reveals to what extent the unit dictates the state trajectory of the entire system, i.e., its dynamical importance, which we quantify by $D(s)$.

Our work contributes to the understanding of the behavior of complex systems at a conceptual level. Our results suggest that the concept of information processing can be used as a general framework to infer how dynamical units work together to produce the system's behavior. The inputs to this inference are both the rules of unit dynamics as well as the topology of interactions, which contrasts with most complex systems research. A popular approach to infer the importance of units in general are topology-only measures such as connectedness and betweenness-centrality (7, 8, 116, 118, 136–139), following the intuition that well-connected or centrally located units must be important to the behavior of the system. We demonstrate that this intuition is not necessarily true. A slightly more realistic approach is to consider is to simulate a simple process on the topology, such as the percolation of particles (33), magnetic spin interactions (11, 12, 42, 73–83), or the synchronization of oscillators (8, 24, 42, 65–70, 140). The dynamical importance of a unit in such a model is then translated to that of the complex system under investigation. Among the 'totalistic' approaches that consider the dynamics and interaction topology simultaneously, a common method to infer a unit's dynamical importance is to perform 'knock-out' experiments (117–119). That is, experimentally removing or altering a unit and observing the difference in the system's

behavior. This is a measure of how robust the system is to a perturbation, however, and care must be taken to translate robustness into dynamical importance. In case the perturbation is not part of the natural behavior of the system then the perturbed system is not a representative model of the original system. To illustrate, we find that highly connected ferromagnetic spins hardly explain the observed dynamical behavior of a system, even though removing such a spin would have a large impact on the average magnetization, stability, and critical temperature (141, 142). In summary, our work is an important step towards a unified framework for understanding the interplay of the unit dynamics and network topology from which the system's behavior emerges.