Universality in network dynamics

Supplementary Information

Baruch Barzel[1,2] & Albert-László Barabási[1,2,3]

1. *Center for Complex Network Research and Departments of Physics, Computer Science and Biology, Northeastern University, Boston, Massachusetts 02115, USA.*

2. *Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.*

3. *Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA.*

Contents

I. INTRODUCTION

This supplementary information is organized as follows. In Sec. S.II we introduce the correlation matrix and show how it is obtained from the system's dynamics. We then show how to analyze the local dynamics and obtain the scaling of the stability (S_i) and impact (I_i) in Sec. S.III. The propagation dynamics, $\Gamma(l)$ and β, together with the correlation distribution, $P(G)$ and ν, are derived in Sec. S.IV, followed by the derivation of the cascade sizes, C_i, in Sec. S.V. Next we show how to generalize the theory to dynamics that cannot be factorized in Sec. S.VI. In Sec. S.VII we discuss the numerical simulations and explicitly analyze the four dynamical models that were simulated in the paper. The analysis of the empirical data is discussed in Sec. S.VIII. This analysis raises a general issue regarding the sensitivity of the formalism to the construction process of the network. We address this issue, showing that the formalism is, in fact, robust against different construction schemes as well as partial knowledge of the network topology in Sec. S.IX.

II. THE CORRELATION MATRIX

We characterize the dynamical behavior of a system by the *activity* $x_i(t)$, a time dependent variable assigned to each node i ($i = 1, \ldots, N$). Most generally, the activities follow the dynamical rate equation

$$\frac{dx_i}{dt} = f_i\left(x_1(t), \ldots, x_N(t)\right), \tag{1}$$

in which the temporal dynamics of i is affected by all other x_j. To quantify the impact of a perturbation j we first focus on the *local correlation matrix*

$$R_{ij} = \left|\frac{\partial x_i/x_i}{\partial x_j/x_j}\right|, \tag{2}$$

where x_i, x_j, and the derivative itself are all taken under steady-state condition, namely $f_i(x_1(t), \ldots, x_N(t)) = 0$. The terms R_{ij} capture the impact that a small perturbation in x_j has on the value of x_i, providing a quantitative measure for the influence of j on the activity of i. The partial derivative $(\partial/\partial x)$ in (2) implies that no other node activity has changed, so that the local correlation matrix, R_{ij}, captures the correlation between i and j, disregarding the effect of the rest of the network, not accounting for indirect interactions - hence the term *local*. In fact, $R_{ij} \neq 0$ only if i and j are directly linked, namely only if in the *adjacency matrix* $A_{ij} = 1$.

To account for indirect interactions we introduce the correlation matrix,

$$G_{ij} = \left| \frac{dx_i/x_i}{dx_j/x_j} \right|, \tag{3}$$

in which the full derivative (d/dx) implies that now all nodes are allowed to change in response to j's perturbation, hence indirect effects are also accounted for. Each column in G_{ij} captures the response of the entire network to a permanent perturbation in the state of node j, quantifying the interdependence between every pair of nodes in the network. Clearly, the diagonal terms of this matrix must satisfy

$$\frac{dx_i/x_i}{dx_i/x_i} = 1. \tag{4}$$

For the off-diagonal terms, $i \neq j$, we write

$$\frac{dx_i/x_i}{dx_j/x_j} = \frac{\partial x_i/x_i}{\partial x_j/x_j} + \sum_{\substack{q=1 \\ q \neq j}}^{N} \left(\frac{\partial x_i/x_i}{\partial x_q/x_q} \right) \left(\frac{dx_q/x_q}{dx_j/x_j} \right), \tag{5}$$

where we used the mathematical identity $dx_i/dx_j = \partial x_i/\partial x_j + \sum_{q \neq j} (\partial x_i/\partial x_q)(dx_q/dx_j)$. In matrix form Eq. (5) becomes [1]

$$\begin{cases} G_{ii} = 1 \\ G_{ij} = \sum_{q=1}^{N} R_{iq} G_{qj} \quad (i \neq j) \end{cases}, \tag{6}$$

a set of $N \times N$ coupled linear equations, taking R_{ij} as input and providing G_{ij}.

III. LOCAL DYNAMICS

The local dynamics, namely the stability and the impact functions, are captured by the direct correlations between neighboring nodes, hence we focus on the local correlations, captured by R_{ij} (2). The *impact* of node i on its nearest neighbors is

$$I_i = \sum_{j=1}^{N} A_{ij} R_{ij}^{\mathrm{T}} = k_i \left\langle R_{ij}^{\mathrm{T}} \right\rangle_{j \in K_i(1)} \tag{7}$$

and the *stability* of node i is

$$S_i = \frac{1}{\sum_{j=1}^{N} A_{ij} R_{ij}} = \frac{1}{k_i \left\langle R_{ij} \right\rangle_{j \in K_i(1)}}. \tag{8}$$

Here $K_i(l)$ it the group of all nodes at distance l from i, and $\langle \cdot \rangle_y$ is an average over all terms in y. Note that here we used the local correlation matrix R_{ij} to approximate G_{ij} in the definition of I_i and S_i (see Eqs. (3) and (4) in paper). The meaning of this approximation is that the impact between nearest neighbors is not significantly affected by indirect paths. Indeed, as we later show, correlations decay exponentially with distance, hence only the shortest paths dominate the impact between node pairs. Equations (7) and (8) indicate that the scaling of I_i and S_i on the degree, k_i, is determined by the degree dependence of R_{ij}, prompting us to derive this dependence below.

A. The Scaling of R_{ij}

We start from the dynamical equation

$$\frac{dx_i}{dt} = W(x_i) - \sum_{n=1}^{N} A_{in} Q(x_i, x_n), \tag{9}$$

where $W(x_i)$ describes the dynamics of i in isolation, and $Q(x_i, x_j)$ describes the effect of the pairwise interactions. At the steady state $dx_i/dt = 0$, allowing us to write

$$1 - \sum_{n=1}^{N} A_{in} \tilde{Q}(x_i, x_n) = 0, \tag{10}$$

where $\tilde{Q}(x_i, x_n) = Q(x_i, x_n)/W(x_i)$. To obtain R_{ij} we induce a small perturbation on j (the source) and follow the response of i (the target): $x_j \to x_j + \partial x_j$, resulting in $x_i \to x_i + \partial x_i$.

Note that in R_{ij}, the activities of all other nodes, apart from i and j, remain unchanged ('∂' vs. 'd'). Hence following j's perturbation the system will be driven into a new steady state, in which

$$1 - \sum_{\substack{n=1 \\ n\neq j}}^{N} A_{in}\tilde{Q}(x_i + \partial x_i, x_n) - A_{ij}\tilde{Q}(x_i + \partial x_i, x_j + \partial x_j) = 0. \tag{11}$$

We expand (11) in orders of ∂x_i and ∂x_j using a perturbative approach, providing

$$\begin{aligned} 1 \; - \; & \sum_{\substack{n=1 \\ n\neq j}}^{N} A_{in}[\tilde{Q}(x_i, x_n) + \tilde{Q}'_{x_i}(x_i, x_n)\partial x_i] \\ - \; & A_{ij}\left(\tilde{Q}(x_i, x_j) + \tilde{Q}'_{x_i}(x_i, x_j)\partial x_i + \tilde{Q}'_{x_j}(x_i, x_j)\partial x_j\right) + O(\partial x_i \partial x_j) = 0, \end{aligned} \tag{12}$$

where $\tilde{Q}'_{x_{i(j)}}(x_i, x_j) = \partial\tilde{Q}/\partial x_{i(j)}$, a derivative taken at the steady state. Excluding non-linear terms in ∂x_i and ∂x_j, namely $O(\partial x_i \partial x_j)$, and using the steady state assumption (10), we bring (12) into the form

$$\left(\sum_{n=1}^{N} A_{in}\tilde{Q}'_{x_i}(x_i, x_n)\right)\partial x_i + \left(A_{ij}\tilde{Q}'_{x_j}(x_i, x_j)\right)\partial x_j = 0. \tag{13}$$

The sum on the l.h.s. of (13) can be written as

$$\sum_{n=1}^{N} A_{in}\tilde{Q}'_{x_i}(x_i, x_n) = k_i \left\langle \tilde{Q}'_{x_i}(x_i, x_n)\right\rangle_{n\in K_i(1)}, \tag{14}$$

from which we obtain

$$\frac{\partial x_i}{\partial x_j} = -\frac{A_{ij}\tilde{Q}'_{x_j}(x_i, x_j)}{k_i \left\langle \tilde{Q}'_{x_i}(x_i, x_n)\right\rangle_{n\in K_i(1)}}, \tag{15}$$

and hence the local correlation matrix (2)

$$R_{ij} = \frac{A_{ij}x_j\tilde{Q}'_{x_j}(x_i, x_j)}{k_i x_i \left\langle \tilde{Q}'_{x_i}(x_i, x_n)\right\rangle_{n\in K_i(1)}}. \tag{16}$$

The structure of R_{ij} is a result of two competing dynamical forces. The numerator captures the effect of the network neighbors by quantifying the impact of the perturbation in x_j on the interaction term $\tilde{Q}(x_i, x_j)$. The denominator captures the impact of self-regulation of

node i, given by the dependence of $\tilde{Q}(x_i, x_n)$ on x_i. Thus if the self regulation dominates, the impact of the interactions is small, and if the network regulation dominates, the neighbors play a significant dynamical role, and R_{ij} is large.

To derive the dependence of R_{ij} on k_i and k_j we focus on dynamical models for which $\tilde{Q}(x_i, x_j)$ could be factorized as

$$\tilde{Q}(x_i, x_j) = f(x_i)g(x_j), \tag{17}$$

where $f(x_i)$ describes the self regulating mechanism and $g(x_j)$ describes the interactions. We can now separate the terms of (16) as

$$\begin{aligned} \tilde{Q}'_{x_i}(x_i, x_j) &= f'(x_i)g(x_j) \\ \tilde{Q}'_{x_j}(x_i, x_j) &= f(x_i)g'(x_j) \\ \left\langle \tilde{Q}'_{x_i}(x_i, x_n) \right\rangle_{n\in K_i(1)} &= f'(x_i) \left\langle g(x_n) \right\rangle_{n\in K_i(1)}, \end{aligned} \tag{18}$$

so that R_{ij} takes the form

$$R_{ij} = \frac{A_{ij}}{\langle g(x_n) \rangle_{n\in K_i(1)}} S_g(x_j) S_f(x_i) \tag{19}$$

where

$$\begin{aligned} S_f(x_i) &= \frac{f(x_i)}{k_i x_i f'(x_i)} \\ S_g(x_j) &= x_j g'(x_j). \end{aligned} \tag{20}$$

The dependence of R_{ij} on k_i is determined by $S_f(x_i)$ and its dependence on k_j by $S_g(x_j)$. As all the terms in (20) are calculated under steady state condition, we first derive and analyze the steady state of (9).

B. Analysis of the Steady-state

First we write Eq. (10) as

$$1 - \sum_{n=1}^{N} A_{in} f(x_i) g(x_n) = 0, \tag{21}$$

from which we obtain

$$f(x_i) = \frac{1}{k_i \left\langle g(x_n) \right\rangle_{n \in K_i(1)}} \tag{22}$$

and consequently, if $f(x_i)$ is invertible,

$$x_i = f^{-1}\left(\frac{1}{k_i \left\langle g(x_n) \right\rangle_{n \in K_i(1)}} \right) = f^{-1}\left(\frac{1}{\left\langle g(x_n) \right\rangle_{n \in K_i(1)}} \xi_i \right), \tag{23}$$

where in (22) and (23) we used

$$\left\langle g(x_n) \right\rangle_{n \in K_i(1)} = \frac{1}{k_i} \sum_{n=1}^{N} A_{ij} g(x_n) \tag{24}$$

and $\xi_i = 1/k_i$. Note that $\left\langle g(x_n) \right\rangle_{n \in K_i(1)}$ is an average taken over the neighbors of i, and hence it depends on the degrees k_n of i's nearest neighbors. In the absence of degree correlations, however, k_n is independent of k_i. Hence the average appearing in (24) does not depend on i, but rather is an average over nearest neighbor nodes in general, namely $\langle \cdot \rangle_{K(1)}$. We thus adopt below the simpler notation $\left\langle g(x_n) \right\rangle_{n \in K_i(1)} = \left\langle g(x) \right\rangle_{K(1)}$, emphasizing that this average is independent of i.

We can now expand x_i in powers of ξ_i as

$$x_i = \sum_{n=0}^{\infty} C_n \xi_i^n = C_0 + C_1 \xi_i + C_2 \xi_i^2 + \ldots, \tag{25}$$

where

$$C_n = \frac{1}{n!} \left(\frac{1}{\left\langle g(x) \right\rangle_{K(1)}} \right)^n \left. \frac{d^n f^{-1}(\xi)}{d\xi^n} \right|_{\xi=0}. \tag{26}$$

Equations (25) and (26) show that the dependence of x_i on ξ_i, and hence on k_i, is fully determined by the structure of $f(x_i)$ (through $f^{-1}(\xi)$), namely by the self regulation mechanism.

Next we use (22) and (23) in order to express each of the terms composing $S_f(x_i)$ in (20):

$$\begin{aligned} x_i &\sim f^{-1}(\xi_i) \\ f(x_i) &\sim \xi_i \\ f'(x_i) &\sim \frac{\partial \xi_i}{\partial x_i} \sim \frac{\partial \xi_i}{\partial f^{-1}(\xi_i)}. \end{aligned} \tag{27}$$

This allows us to write (20)

$$S_f(x_i) \sim \xi_i^2 \frac{1}{f^{-1}(\xi_i)} \frac{\partial f^{-1}}{\partial \xi_i}. \tag{28}$$

Following (25), we expand $S_f(x_i)$ in powers of ξ_i, obtaining

$$S_f(x_i) \sim \xi_i^2 \frac{\sum_{n=0}^{\infty} n C_n \xi_i^{n-1}}{\sum_{n=0}^{\infty} C_n \xi_i^n}, \tag{29}$$

where C_n are taken from (26), and for large k_i (small ξ_i) we take only the leading terms.

C. Stability

We can now write S_i (8) using (19) as

$$S_i \sim \frac{1}{k_i \left\langle S_g(x_j) \right\rangle_{j \in K_i(1)} S_f(x_i)} \sim \xi_i^{-1} \frac{\sum_{n=0}^{\infty} C_n \xi_i^n}{\sum_{n=0}^{\infty} n C_n \xi_i^{n-1}}, \tag{30}$$

where we used (29) to substitute for $S_f(x_i)$. We denote by n_0 the leading term in (25) and by n_1 its leading *non-vanishing* term. To be specific if in (25)

$$x_i \sim f^{-1}(\xi_i) \sim C_0 + C_{n_1} \xi_i^{n_1} + \ldots \tag{31}$$

we have $n_0 = 0$ and $n_1 > n_0$, but if in (25) we have $C_0 = 0$ (or $n_1 < 0$) then $n_0 = n_1$, namely the leading term and the leading non-vanishing term coincide. Using this notation we find that in the limit of large k_i (small ξ_i) the leading terms of (30) provide

$$S_i \sim \xi_i^{-\delta} = k_i^{\delta} \tag{32}$$

where

$$\delta = n_1 - n_0. \tag{33}$$

Two universality classes emerge based on the structure of the function $f^{-1}(x)$:

Uniform stability: If in (25) $n_1 = n_0$ Eq. (33) predicts that $\delta = 0$ in (32) and hence *the stability is independent of the node's degree* and consequently the stability distribution $P(S)$ is independent of the degree distribution $P(k)$.

Heterogeneous stability: The only other possibility is that $n_0 = 0$, and $n_1 > 0$, for which $\delta = n_1 > 0$ (33), *the stability is driven by the node's degree* and $P(S)$ is determined by $P(k)$.

A special case is where $C_0 \to \infty$, which occurs when $f^{-1}(\xi)$ diverges for $\xi \to 0$. Under this condition, the series expansion of (25) cannot be carried out around $\xi_i = 0$, and instead we use the Laurent series, which includes negative powers

$$x_i \sim f^{-1}(\xi_i) = \sum_{n=-\infty}^{\infty} C_n \xi_i^n. \tag{34}$$

Repeating the above derivation shows that this case leads to uniform stability as well, so that in effect, Eq. (33) remains valid only now n_0 and n_1 could also take negative values.

D. Impact

We use (19) to write the impact (7) as

$$I_i \sim k_i \left\langle S_f(x_j) \right\rangle_{j \in K_i(1)} S_g(x_i) \sim k_i x_i g'(x_i), \tag{35}$$

where we swapped x_i and x_j to account for the transposed matrix R_{ij}^{T}. To obtain the degree dependence of (35) we focus on $g(x_i)$, which following (23), we write as

$$g(x_i) = g\left[f^{-1}\left(\frac{1}{\langle g(x) \rangle_{K(1)}} \xi_i \right) \right]. \tag{36}$$

Expanding (36) in powers of ξ_i, as before, provides the dependence of $g(x_i)$ on k_i

$$g(x_i) = \sum_{m=0}^{\infty} C_m \xi_i^m = C_0 + C_1 \xi_i + C_2 \xi_i^2 + \ldots, \tag{37}$$

where

$$C_m = \frac{1}{m!} \left(\frac{1}{\langle g(x) \rangle_{K(1)}} \right)^m \left. \frac{d^m g(f^{-1}(\xi))}{d\xi^m} \right|_{\xi=0}. \tag{38}$$

We write the impact (35) as

$$I_i \sim \frac{1}{\xi_i} f^{-1}(\xi_i) g'\left(f^{-1}(\xi_i) \right) \tag{39}$$

where we used (23) to express x_i, giving rise to

$$I_i \sim \frac{1}{\xi_i} f^{-1}(\xi_i) \left(\frac{\partial f^{-1}(\xi_i)}{\partial \xi_i} \right)^{-1} \frac{\partial g\left(f^{-1}(\xi_i)\right)}{\partial \xi_i}. \tag{40}$$

We analyze (40) term by term: according to (28) we have

$$f^{-1}(\xi_i) \left(\frac{\partial f^{-1}(\xi_i)}{\partial \xi_i} \right)^{-1} \sim \xi^2 \frac{1}{S_f(x_i)} \sim \xi_i^{1-\delta}. \tag{41}$$

The contribution of the last term in (40) is determined by the leading *non vanishing* power in the expansion of (37) to be

$$\frac{\partial g\left(f^{-1}(\xi_i)\right)}{\partial \xi_i} \sim \xi_i^{m_1 - 1}. \tag{42}$$

To be specific, whether $g(f^{-1}(\xi_i)) \sim \xi_i^{m_1}$ or $g(f^{-1}(x)) \sim C_0 + C_{m_1}\xi_i^{m_1}$, the derivative in (42) has $\xi_i^{m_1-1}$ as its leading term. Collecting all terms, (41) and (42), we arrive at (40)

$$I_i \sim \xi_i^{-1} \xi_i^{1-\delta} \xi_i^{m_1-1} \sim \xi_i^{-\varphi} \sim k_i^{\varphi} \tag{43}$$

where

$$\varphi = 1 + \delta - m_1. \tag{44}$$

As before, in case the Taylor expansion (37) cannot be carried out at $\xi_i = 0$, namely when $g(f^{-1}(\xi_i))|_{\xi_i \to 0} \to \infty$, we use the Laurent expansion

$$g\left(f^{-1}(\xi_i)\right) = \sum_{m=-\infty}^{\infty} C_m \xi_i^m. \tag{45}$$

Here, if the leading term is associated with a negative power, m_1 in (44) will be negative.

The value of φ determines the patterns of the local impact:

Uniform impact: In case $\varphi = 0$ (namely $m_1 = 1 + \delta$) we have $I_i \sim k_i^0$, describing a system where the impact is independent of the node's degree. As a consequence the impact distribution, $P(I)$, will be independent of the degree distribution, $P(k)$.

Heterogeneous impact: If $\varphi \neq 0$ we have $I_i \sim k_i^{\varphi}$, the impact scales with the node's degree and $P(I)$ is driven by $P(k)$.

Local dynamics - summary:

The degree dependence of S_i is determined by the functional form of $f(x)$ via the leading terms of the Laurent expansion

$$f^{-1}(x) = \sum_{n=\infty}^{\infty} C_n x^n.$$

- If the lowest power in the expansion is $f^{-1}(x) \sim x^{n_1}$ $(n_1 \neq 0)$ we have $\delta = 0$, providing *uniform stability*.
- If the lowest power is $n_0 = 0$ we have $f^{-1}(x) \sim C_0 + C_{n_1} x^{n_1}$ $(n_1 > n_0)$ and $\delta = n_1$, providing *heterogeneous stability*.

The degree dependence of I_i is determined by both $f(x)$ and $g(x)$ through

$$g\left(f^{-1}(x)\right) = \sum_{m=\infty}^{\infty} C_m x^m.$$

Denoting the lowest non vanishing power in the expansion by m_1, we have $\varphi = 1 + \delta - m_1$.

- If $\varphi = 0$, we have *uniform impact*.
- If $\varphi \neq 0$, we have *heterogeneous impact*.

IV. PROPAGATION OF PERTURBATIONS

The correlation function $\Gamma(l)$ follows the perturbations experienced by *all* nodes at distance l from the source, namely

$$\Gamma(l) = \frac{1}{N} \sum_{j=1}^{N} \sum_{i \in K_j(l)} G_{ij}. \tag{46}$$

To derive (46) we first focus on *individual* correlations at distance l, given by

$$G(l) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{|K_j(l)|}\sum_{i\in K_j(l)} G_{ij} = \left\langle \langle G_{ij}\rangle_{i\in K_j(l)}\right\rangle_j , \tag{47}$$

where $|K_j(l)|$ is the number of nodes in the $K_j(l)$ group. Hence the correlation function $\Gamma(l)$ captures the response of all nodes at l, while $G(l)$ captures the response of the *average* node at l. In other words, in $\Gamma(l)$ we sum over all individual correlations $G(l)$ between pairs separated by distance l as

$$\Gamma(l) = K(l)G(l). \tag{48}$$

Using (6) we write (47) as

$$G(l) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{|K_j(l)|}\sum_{i\in K_j(l)}\sum_{q=1}^{N} R_{iq}G_{qj} \tag{49}$$

for all $l > 0$; the $l = 0$ case corresponds to $i = j$, for which $G(0) = 1$. We now focus on the sum at the r.h.s. of (49). As the elements of R_{iq} vanish unless $q \in K_i(1)$, this sum includes only nearest neighbors of i. Moreover, since $i \in K_j(l)$, these neighbors of i must satisfy $q \in K_j(l+n)$, where $n = -1 \ldots l_{\max}$, in case the network is directed, and $n = -1, 0$ or 1 in case it is undirected. Indeed, since i is at distance l from j, i's neighbors q must be at distance $l-1$ or greater from j. Consequently their correlation with j is on average given by

$$G_j(l+n) = \langle G_{qj}\rangle_{q\in K_j(l+n)} , \tag{50}$$

which describes the average impact of the j perturbation at distance $l+n$. Note that

$$G(l) = \frac{1}{N}\sum_{j=1}^{N} G_j(l). \tag{51}$$

We can thus split the sum of Eq. (49) into groups of nodes, the nth group being $K_i(1) \cap K_j(l+n)$, including all nearest neighbors of i, which are at distance $l+n$ from j. Equation (49) now takes the form

$$G(l) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{|K_j(l)|}\sum_{i\in K_j(l)}\sum_{n=-1}^{l_{\max}} \langle R_{iq}\rangle_{q\in K_i(1)} |K_i(1)\cap K_j(l+n)| G_j(l+n), \tag{52}$$

which together with the equation for $l = 0$ (and using (51)) becomes

$$\begin{aligned} G(0) &= 1 \\ G(l) &= \rho \sum_{n=-1}^{l_{\max}} \epsilon_n(l) G(l+n), \end{aligned} \tag{53}$$

where

$$\rho = \frac{1}{N} \sum_{j=1}^{N} \frac{1}{|K_j(l)|} \sum_{i \in K_j(l)} k_i \left\langle R_{iq} \right\rangle_{q \in K_i(1)} = \left\langle \left\langle k_i \left\langle R_{iq} \right\rangle_{q \in K_i(1)} \right\rangle_{i \in K_j(l)} \right\rangle_j \tag{54}$$

and

$$\epsilon_n(l) = \frac{1}{N} \sum_{j=1}^{N} \frac{1}{|K_j(l)|} \sum_{i \in K_j(l)} \frac{|K_i(1) \cap K_j(l+n)|}{|K_i(1)|} \tag{55}$$

is the *expansion factor* of the network. The contribution of the dynamics is accounted for by ρ, which depends on R_{ij}, and the contribution of the network topology is given by $\vec{\epsilon}(l)$, which depends solely on the wiring diagram. Note that

$$\sum_{n-1}^{l_{\max}} \epsilon_n(l) = 1, \tag{56}$$

where for an undirected network one can substitute $l_{\max}$ by $+1$, namely $\epsilon_{-1}(l)+\epsilon_0(l)+\epsilon_{+1}(l) = 1$.

To get a better understanding of the meaning of $\vec{\epsilon}(l)$ consider a pair of nodes in an undirected network, the source j, and the target i at distance l, namely $i \in K_j(l)$. The terms $\vec{\epsilon}(l)$ capture the fraction of i's nearest neighbors that are at $K_j(l-1)$ ($\epsilon_{-1}(l)$), $K_j(l)$ ($\epsilon_0(l)$), and $K_j(l+1)$ ($\epsilon_{+1}(l)$). If the network is directed, then i's neighbors could belong to $K_j(l)$ for $l = -1, \ldots, \infty$, where an infinite distance between a pair of nodes means that there exists no path between them. Thus, in non-technical terms, Eq. (53) can be intuitively illustrated as

$$G(l) = \begin{bmatrix} \text{The impact} \\ \text{of a nearest} \\ \text{neighbor} \end{bmatrix} \times \Bigg(\begin{bmatrix} \text{The fraction} \\ \text{of the target's} \\ \text{neighbors at } l-1 \\ \text{from the source} \end{bmatrix} G(l-1) + \begin{bmatrix} \text{The fraction} \\ \text{of the target's} \\ \text{neighbors at } l \\ \text{from the source} \end{bmatrix} G(l) + \begin{bmatrix} \text{The fraction} \\ \text{of the target's} \\ \text{neighbors at } l+1 \\ \text{from the source} \end{bmatrix} G(l+1) + \dots \Bigg) . \tag{57}$$

The expansion factor (55) can be empirically measured from the topology of the network, as shown in Fig S1, and then used as input in Eq. (53). In Sec. S.IV A we derive the terms of $\vec{\epsilon}(l)$ analytically for large networks.

Note that ρ in (54) is in general a function of l (through the $i \in K_j(l)$ term attributed to the average), namely $\rho = \rho(l)$. This dependence comes from the distribution of R_{iq}, capturing the correlations between nodes i that are at distance l from the source j, and their nearest neighbors q. However, as we have shown in Sec. S.III, R_{iq} is determined by the interacting node's degrees, k_i and k_q. In the absence of degree correlations k_i, k_q and hence ρ, all become independent of l, their degree distribution being the same as any nearest neighbor degree distribution. This allows us to substitute the terms $q \in K_i(1)$ and $i \in K_j(l)$ in (54) by the simpler form $K(1)$, which indicates that the average is carried out over nearest neighbor nodes, but not specific to i or j, as the degrees of all nearest neighbor nodes have the same distribution. Note that above we also substituted l by $l = 1$, as with no degree correlations, the degree distribution of nearest neighbors and that of neighbors at distance l is the same. We can thus simplify (54) as

$$\rho = \left\langle k_i \left\langle R_{iq} \right\rangle_{K(1)} \right\rangle_{K(1)} , \tag{58}$$

in which the l dependence is removed. In the general solution we present below the l dependence of ρ could be accounted for. For simplicity, however, as we will not be using this dependence later on, we take ρ to be l independent throughout the rest of this work.

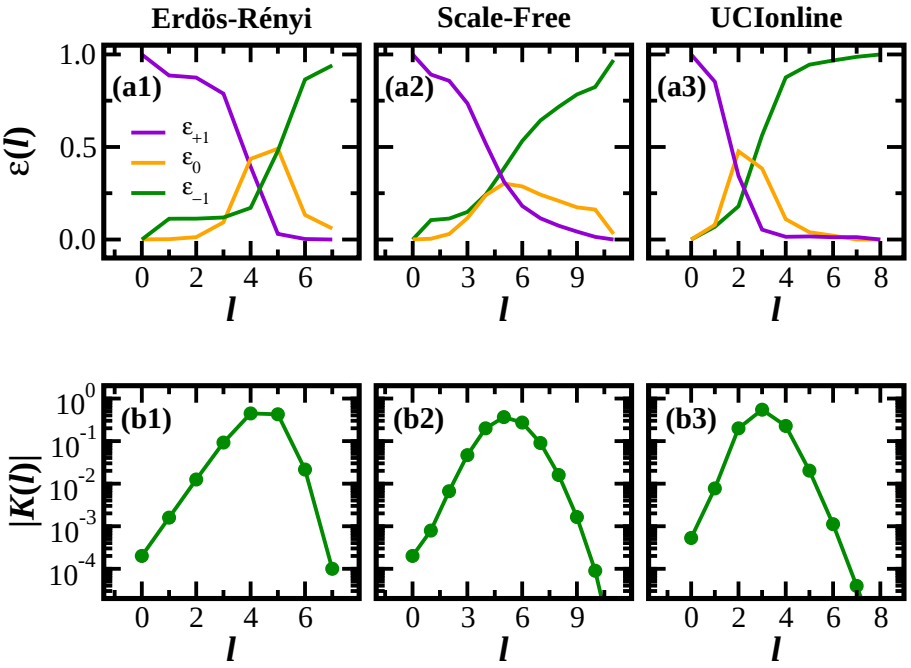

FIG. 1: **The expansion factor and network growth**. The expansion factor $\vec{\epsilon}(l)$ vs. l for (a1) an Erdős-Rényi network; (a2) a scale-free network; (a3) an empirically obtained online social network (UCIonline) [2]. For small l the network is expanding according to (69) and $\epsilon_{+1}(l)$ dominates, indicating that nodes tend to link to outer shells. For large l, as the network expansion saturates $\epsilon_{+1}(l)$ becomes small, indicating that nodes tend to link to inner shells. For $l \sim \langle l \rangle$, $\epsilon_0(l)$ is maximal, indicating that nodes at $\langle l \rangle$ tend to link to equidistant nodes. (b1) - (b3) The probability that a randomly selected pair of nodes is at distance l, $|K(l)|$, grows exponentially for $l < \langle l \rangle$. For $l > \langle l \rangle$ the network growth saturates and the shells begin to contract. The expansion and contraction are reflected by $\vec{\epsilon}(l)$ through the ratio between $\epsilon_{+1}(l)$ and $\epsilon_{-1}(l)$: for small l $\epsilon_{+1}(l)$ dominates and for large l $\epsilon_{-1}(l)$ dominates.

A. General Solution

To solve Eq. (53) we assume that the impact of the source on the target is carried mainly through the nodes closer to the source, allowing us to collapse the terms of $\vec{\epsilon}(l)$ as

$$\sum_{n=1}^{l_{\max}} \epsilon_n(l) \to \epsilon_{+1}(l), \tag{59}$$

having $\epsilon_{+1}(l)$ account for all nodes at distances $l+1$ to $l_{\max}$ from the source. In other words, $\epsilon_{+1}(l)$ denotes the fraction of the target's neighbors that are at any distance greater than l from the source (and not just at $l+1$). This is equivalent to truncating (57) after the $G(l+1)$ term, and substituting the relevant coefficient as

$$\begin{bmatrix} \text{The fraction} \\ \text{of the target's} \\ \text{neighbors at } l+1 \\ \text{from the source} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{The fraction} \\ \text{of the target's} \\ \text{neighbors at } l_{\text{neighbor}} > l \\ \text{from the source} \end{bmatrix} \tag{60}$$

For an undirected network $\epsilon_n(l) = 0$ for $n > 1$, and hence (59) is exact.

Next we use a continuum approximation, referring to l as a continuous variable. This approximation, exact in the limit of large networks ($N \to \infty$), allows us to write Eq. (53) in differential form as

$$\begin{cases} G(0) = 1 \\ G(l) = \rho[\epsilon_{-1}(l)G(l-1) + \epsilon_0(l)G(l) + \epsilon_{+1}(l)G(l+1)] \end{cases}. \tag{61}$$

Taking $G(l) = \exp[f(l)]$ we have

$$1 = \rho\left[\epsilon_{-1}(l)e^{-\frac{df}{dl}} + \epsilon_0(l) + \epsilon_{+1}(l)e^{\frac{df}{dl}}\right], \tag{62}$$

where we used $f(l \pm 1) \approx f(l) \pm df/dl$. From (62) we find

$$\exp\left(\frac{df}{dl}\right) = \frac{\rho^{-1} - \epsilon_0(l) \pm \sqrt{\left(\rho^{-1} - \epsilon_0(l)\right)^2 - 4\epsilon_{-1}(l)\epsilon_{+1}(l)}}{2\epsilon_{+1}(l)}, \tag{63}$$

which, using the fact that $\sum_n \epsilon_n(l) = 1$, provides

$$\begin{aligned} f(l) \;=\; & \int_0^l \ln\left[\frac{1}{\tilde{\epsilon}_+(l) + \tilde{\epsilon}_-(l)}\left(\tilde{\epsilon}_+(l) - 1 + \rho^{-1}\right.\right. \\ & \left.\left. \pm\; \sqrt{(1-\rho^{-1})^2 - 2\tilde{\epsilon}_+(l)(1-\rho^{-1}) + \tilde{\epsilon}_-^2(l)}\right)\right] d\tilde{l}, \end{aligned} \tag{64}$$

where $\tilde{\epsilon}_\pm(l) = \epsilon_{+1}(l) \pm \epsilon_{-1}(l)$. To satisfy the boundary condition in which $G(l)$ approaches zero for large l, we select the solution where the square-root is subtracted in Eq. (64), providing

$$G(l) = \exp\left\{\int_0^l [\ln \Psi(\tilde{l}, \rho)] d\tilde{l}\right\}, \tag{65}$$

where

$$\Psi(l, \rho) = \frac{1}{\tilde{\epsilon}_+(l) + \tilde{\epsilon}_-(l)}\left(\tilde{\epsilon}_+(l) - 1 + \rho^{-1} - \sqrt{(1-\rho^{-1})^2 - 2\tilde{\epsilon}_+(l)(1-\rho^{-1}) + \tilde{\epsilon}_-^2(l)}\right). \tag{66}$$

B. Large networks

Equations (65) and (66) provide the correlation function $G(l)$, receiving $\vec{\epsilon}(l)$ and ρ as input. To obtain an analytical expression for $\vec{\epsilon}(l)$ we focus below on large networks in the limit $N \to \infty$. For simplicity, we focus on undirected networks, for which $\vec{\epsilon}(l) = (\epsilon_{-1}, \epsilon_0, \epsilon_{+1})$ has only three terms.

Consider a node i and its shells $K_i(l)$ of neighbors at distance l, in a large network with an arbitrary degree distribution $P(k)$. While $P(k)$ captures the degree distribution of the whole network, the degree distribution of nodes from $K_i(l)$ is biased towards nodes with a higher degree, as the selection process favors nodes that have more links, giving rise to the distribution [3]

$$q(k) = \frac{kP(k)}{\langle k \rangle}, \tag{67}$$

and hence providing the average degree of nodes in $K_i(l)$ as $\langle k_n \rangle_{n \in K_i(l)} = \langle k^2 \rangle / \langle k \rangle$, and the average residual degree

$$k_{\rm res} = \sum_{j \in K_i(l)} (k_j - 1) q(k_j) = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}. \tag{68}$$

The meaning of (68) is that each node in $K_i(l)$ links to a single node from $K_i(l-1)$ and $k_{\rm res}$ additional nodes from other shells. However, in the $N \to \infty$ limit no loops are present in the network [3], a condition only satisfied if all of the $k_{\rm res}$ remaining edges link to nodes from $K_i(l+1)$. As a result the size of the shells grows according to $|K_i(l+1)| = k_{\rm res}|K_i(l)|$, which leads to an exponential growth of the network shells as

$$|K_i(l)| = \frac{\langle k \rangle^2}{\langle k^2 \rangle - \langle k \rangle} e^{\alpha l} \tag{69}$$

where

$$\alpha = \ln k_{\rm res} \tag{70}$$

characterizes the rate of the network expansion. Using (68) we find that α is related to the average neighbor's degree as

$$e^{\alpha} \approx \left\langle \langle k_n \rangle_{n \in K_i(1)} \right\rangle_i, \tag{71}$$

capturing the average number of next nearest neighbors of a node. Moreover, α is closely related to the small world phenomena, as by integrating both sides of Eq. (69) we arrive at

$$\langle l \rangle \sim \frac{\ln N}{\alpha}, \tag{72}$$

representing the small world property, indicating that the larger is α the more pronounced is the small world nature of the network. For an Erdős-Rényi (ER) network, with $P(k)$ being Poisson, we find that $\alpha = \ln \langle k \rangle$. For a scale-free (SF) network with $P(k) = k^{-\gamma}/\zeta(\gamma)$, we write the mth moment as $\langle k^m \rangle = \zeta(\gamma - m)/\zeta(\gamma)$, where $\zeta(\gamma) = \sum_{k=1}^{\infty} k^{-\gamma}$ is the Riemann zeta function. We thus find that $\alpha = \ln(\zeta(\gamma - 2) - \zeta(\gamma - 1)) - \ln \zeta(\gamma - 1)$, which diverges if $\gamma < 3$, indicating that the number of second neighbors, $|K_i(2)|$, approaches infinity. In general, heterogeneity in the degrees, characterized by a large $\langle k^2 \rangle$, leads to a larger value for α, so that for a fat tailed degree distribution $\alpha > \ln \langle k \rangle$, and as the heterogeneity increases, α increases as well. While (70) is exact in the $N \to \infty$ limit, it tends to overestimate the inflation rate of $|K(l)|$ for finite networks. A discussion on the evaluation of α for real networks appears in Fig. S2.

Consider the number of links $E_{+1}(l)$ drawn from nodes in the shell $K_i(l)$, to neighboring nodes in the shell $K_i(l+1)$. It is proportional to the number of nodes in the shell $|K_i(l)|$, to the average degree of nodes in $K_i(l)$ and to $\epsilon_{+1}(l)$, the fraction of these links that reach a node in $K_i(l+1)$. This has to be consistent with the number of links in the opposite direction, emerging from $K_i(l+1)$, connecting to nodes from $K_i(l)$. Thus we have

$$|K_i(l)| \langle k \rangle_{K_i(l)} \epsilon_{+1}(l) = |K_i(l+1)| \langle k \rangle_{K_i(l+1)} \epsilon_{-1}(l+1). \tag{73}$$

The absence of loops in the $N \to \infty$ limit prohibits links between nodes within the same shell, hence $\epsilon_0(l) = 0$, which using (56) provides $\epsilon_{-1}(l+1) = 1 - \epsilon_{+1}(l+1)$. This enables us to write (73) as

$$|K_i(l)|\epsilon_{+1}(l) = |K_i(l+1)|(1 - \epsilon_{+1}(l+1)), \tag{74}$$

a recursive formulation of $\vec{\epsilon}(l)$. To complete the formulation consider $\vec{\epsilon}(l)$ at $l = 0$. It simply accounts for the nearest neighbors of a node, who are clearly all at the $K_i(1)$ shell, so that $\vec{\epsilon}(0) = (0, 0, 1)$ by definition. Together with Eqs. (69) and (74) this provides

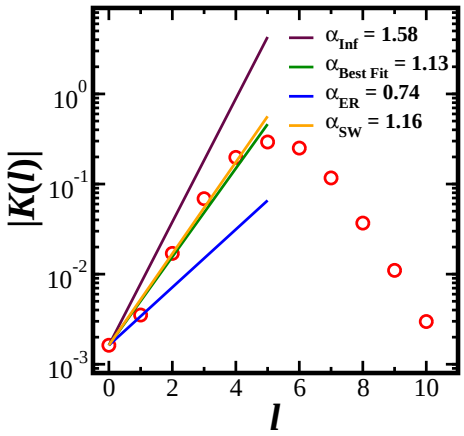

FIG. 2: **Evaluating the expansion rate of a network**. In a small world network the fraction of node pairs at distance l increases exponentially with l as $|K(l)| \sim e^{\alpha l}$. For an infinite network α is determined by the average number of next nearest neighbors of a node following (70). In case the network is finite, however, α becomes smaller, because of the saturation observed as l approaches $\langle l \rangle$. Here we demonstrate the evaluation of α for a scale-free network obtained from a series of genetic perturbation experiments [4] (S.VIII B). As the figure indicates for $l < \langle l \rangle$ the network indeed features an exponential expansion, which saturates for $l > \langle l \rangle$ (circles). We present four approaches for the evaluation of α: Equation (70), exact for $N \to \infty$ is shown to overestimate α (purple); taking $\alpha_{\rm ER} = \ln \langle k \rangle$, exact for an Erdős-Rényi random network underestimates the actual expansion rate (blue); using linear regression (LR) to obtain the best fit for the data points (for $l \leq \langle l \rangle$) provides the most reliable result (green). In this work we used Eq. (72) to evaluate α directly from the network topology as $\alpha_{\rm SW} = \ln N / \langle l \rangle$ (orange), which as the figure indicates, provides a good estimate for the network expansion.

$$\begin{cases} \epsilon_{+1}(0) = 1 \\ \epsilon_{+1}(l+1) = 1 - e^{-\alpha} \epsilon_{+1}(l) \end{cases}, \tag{75}$$

from which $\vec{\epsilon}(l)$ can be obtained for all l. The solution of (75) is

$$\epsilon_{+1}(l) = 1 - e^{-\alpha} + O(e^{-2\alpha}) \tag{76}$$

providing

$$\vec{\epsilon}(l) = \left(e^{-\alpha}, 0, 1 - e^{-\alpha}\right) = \left(\frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}, 0, \frac{\langle k^2 \rangle - 2\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}\right), \tag{77}$$

which is independent of l. For an ER network we have

$$\vec{\epsilon}(l) = \left(\frac{1}{\langle k \rangle}, 0, \frac{\langle k \rangle - 1}{\langle k \rangle}\right), \tag{78}$$

and for a SF network

$$\vec{\epsilon}(l) = \begin{cases} (0,0,1) & \gamma \leq 3 \\ \left(\frac{\zeta(\gamma-1)}{\zeta(\gamma-2)-\zeta(\gamma-1)}, 0, \frac{\zeta(\gamma-2)-2\zeta(\gamma-1}{\zeta(\gamma-2)-\zeta(\gamma-1)}\right) & \gamma > 3 \end{cases}. \tag{79}$$

To summarize, Eq. (55) provides the most general description of $\vec{\epsilon}(l)$, which can be obtained directly from the network topology; Eq. (77) provides $\vec{\epsilon}(l)$ in the limit $N \to \infty$, which is independent of l, and fully determined by the first and second moments of the degree distribution; Eqs. (78) and (79) are valid for ER and SF networks, respectively.

The expansion factor is closely related to the small world property, as it captures the rate of the network expansion. If $\epsilon_{+1}(l)$ is large, the meaning is that nodes in $K_i(l)$ tend to draw more links to nodes in $K_i(l+1)$ than to nodes in $K_i(l)$ or $K_i(l-1)$. This results in a rapid exponential growth of the shells as appears in Eq. (69). Using Eq. (77) together with (70) we find that

$$\ln\left(\frac{\epsilon_{+1}(l)}{\epsilon_{-1}(l)}\right) \approx \alpha, \tag{80}$$

indicating that like α, $\vec{\epsilon}(l)$ also characterizes the degree of *small worldness* of the network, by capturing the rate by which the shells $K(l)$ expand with l.

C. The Correlation Function for Large Networks

Above we have shown that in the $N \to \infty$ limit the expansion factor, $\vec{\epsilon}(l)$, becomes independent of l (77). Under these conditions $\Psi(l,\rho)$ (66) also becomes independent of l, allowing us to perform the integral (65), providing

$$G(l) = e^{-\frac{l}{\lambda}}, \tag{81}$$

where

$$\lambda = -\frac{1}{\ln\left(\Psi(\alpha,\rho)\right)} \tag{82}$$

is the *correlation length* of the system. The parameter λ provides the radius of impact of a node, expressing the average penetration depth of a perturbation. Equation (66) takes the form

$$\Psi(\alpha,\rho)=\frac{\rho^{-1}}{2(1-e^{-\alpha})}\left[1-\sqrt{1-4\frac{e^{-\alpha}(1-e^{-\alpha})}{\rho^{-2}}}\right], \tag{83}$$

which in the small world limit (large α) can be approximated by

$$\Psi(\alpha,\rho)=\rho e^{-\alpha}+O(\rho^2 e^{-2\alpha)}). \tag{84}$$

Consequently we find that the correlation length (82) is

$$\lambda=\frac{1}{\alpha-\ln\rho}, \tag{85}$$

where α characterizes the topology of the underlying network and ρ (58) characterizes the dynamical mechanism of the interactions.

For $l>\langle l\rangle$ the exponentially growing shells have exhausted most of the nodes, hence the network no longer features the exponential expansion (69), and the shells $K_i(l)$ begin to contract (Fig. S1(b1 - b3)). The terms of the expansion factor now satisfy

$$\epsilon_{-1}(l)\gg\epsilon_{+1}(l), \tag{86}$$

as nodes tend to link mainly to inner shells, at lower l. We follow the same derivation as in (81) - (85), only this time we take the limit (86), finding that now (82) predicts

$$\tilde{\lambda}=-\frac{1}{\ln\rho}. \tag{87}$$

We thus arrive at the prediction that $G(l)$ is characterized by two correlation lengths, λ (85) and $\tilde{\lambda}$ (87), as

$$G(l)=\begin{cases} e^{-\frac{l}{\lambda}} & l<\langle l\rangle \\ e^{-\left(\frac{\langle l\rangle}{\lambda}+\frac{l-\langle l\rangle}{\tilde{\lambda}}\right)} & l>\langle l\rangle \end{cases}. \tag{88}$$

Equation (88) describes the propagation of an individual perturbation, providing the correlation lengths, λ and $\tilde{\lambda}$, which characterize the rate of decay of perturbations. The specific value of λ and $\tilde{\lambda}$ depends on the system's dynamics through R_{ij} in ρ (58), hence our next step is to use the degree dependence of R_{ij} in order to evaluate ρ. We analyze R_{ij} starting with Eq. (19)

$$R_{ij} = \frac{A_{ij}}{\langle g(x_n)\rangle_{n\in K_i(1)}} S_g(x_j) S_f(x_i), \tag{89}$$

which we break down term by term. From (30) and (35)we have

$$S_f(x_i) \sim \frac{1}{k_i S_i} \sim k_i^{-\delta-1} \tag{90}$$

and

$$S_g(x_j) \sim \frac{I_j}{k_j} \sim k_j^{\varphi-1}. \tag{91}$$

Using (45) we write the denominator in R_{ij} (89) as

$$\langle g(x_n)\rangle_{n\in K_i(1)} = \left\langle \sum_{m=-\infty}^{\infty} C_m \xi_n^m \right\rangle_{n\in K_i(1)} \sim \langle \xi_n^{m_0}\rangle_{n\in K_i(1)}, \sim \left\langle k_n^{-m_0}\right\rangle_{n\in K_i(1)}, \tag{92}$$

where in the last step we took only the leading term m_0 in the Laurent expansion.

Collecting all three terms (90) - (92) we find that the local correlation matrix (89) follows

$$R_{ij} = C \frac{A_{ij}}{\left\langle k_n^{-m_0}\right\rangle_{n\in K_i(1)}} k_i^{-\delta-1} k_j^{\varphi-1}. \tag{93}$$

where the constant C depends on the microscopic details of the pairwise dynamics, such as the rate constants for the different dynamical processes in (9). To obtain ρ we must average over R_{ij} as appears in (54)

$$\rho = \left\langle \left\langle k_i \left\langle R_{iq}\right\rangle_{q\in K_i(1)} \right\rangle_{i\in K_j(l)} \right\rangle_j \tag{94}$$

First we use (93) to write the average of R_{ij} over i's nearest neighbors as

$$\langle R_{ij}\rangle_{j\in K_i(1)} = C \left\langle \frac{k_i^{-\delta-1} k_j^{\varphi-1}}{\left\langle k_n^{-m_0}\right\rangle_{n\in K_i(1)}} \right\rangle_{j\in K_i(1)} = C \frac{\left\langle k_j^{\varphi-1}\right\rangle_{j\in K_i(1)}}{\left\langle k_n^{-m_0}\right\rangle_{n\in K_i(1)}} k_i^{-\delta-1}, \tag{95}$$

after which we can write (94) as

$$\rho = C \left\langle \left\langle \frac{\left\langle k_j^{\varphi-1}\right\rangle_{j\in K_i(1)}}{\left\langle k_n^{-m_1}\right\rangle_{n\in K_i(1)}} k_i^{-\delta} \right\rangle_{i\in K_q(l)} \right\rangle_q \approx \frac{1}{\left\langle k_n^{-m_1}\right\rangle_{K(1)}} \left\langle k_j^{\varphi-1}\right\rangle_{K(1)} \left\langle k_i^{-\delta}\right\rangle_{K(1)}. \tag{96}$$

The last step, factorizing the inner product average into a product of separate averages, is exact in the absence of degree correlations. This also enables us to substitute the i and

q dependent terms, $n \in K_i(1)$ and $i \in K_q(l)$, by $K(1)$, denoting an average over nearest neighbor nodes, not specific to i, q or l (see disscusion surrounding Eq. (58)). Using (71) we have $e^{\alpha} = \langle k \rangle_{K(1)}$, from which it follows that

$$\rho \approx C e^{(m_1 - \delta + \varphi - 1)\alpha}. \tag{97}$$

Extracting the logarithm and taking φ from (44), we arrive at

$$\ln \rho = (m_0 - m_1)\alpha + \ln C, \tag{98}$$

where m_0 is the leading term in (45) and m_1 is the leading *non-vanishing* term in the expansion. In the small world limit, where α is large, we write (98) as $\ln \rho \approx (m_0 - m_1)\alpha$, accurate up to a logarithmic correction which depends on the specific rate constants, but is not inherent to the dynamical model. The correlation lengths in (88), given by (85) and (87), become

$$\begin{aligned} \lambda &= \frac{1}{(\beta+1)\alpha} \\ \tilde{\lambda} &= \frac{1}{\beta\alpha}, \end{aligned} \tag{99}$$

where

$$\beta = m_1 - m_0. \tag{100}$$

Finally, the correlation function $\Gamma(l)$ is obtained for $l < \langle l \rangle$ using (48) and (69), providing

$$\Gamma(l) = e^{\alpha l} e^{-l/\lambda} = e^{-\beta\alpha l}. \tag{101}$$

The value of β is determined by the structure of $g(f^{-1}(x))$, leading to two distinct universality classes:

Conservative dynamics: If the leading term in the expansion (45) is $g(f^{-1}(x)) \sim x^{m_1}$ ($m_1 \neq 0$) then the *leading term*, m_0, and the *leading non vanishing term*, m_1, coincide. As a result $\beta = 0$ (100), and $\Gamma(l) = 1$ (101). This describes a *conservative* process, in which perturbations propagate without loss. Note that the individual correlations do decay as $G(l) = e^{-\alpha l}$, but this decay is due to the topological expansion of the network, in which

the original perturbation is distributed over $K(l) \sim e^{\alpha l}$ nodes. Hence the total effect is conserved.

Dissipative dynamics: The only other possibility is that $g(f^{-1}(x)) \sim C_0 + C_{m_1} x^{m_1}$ ($m_1 > 0$), in which the *leading term* is $m_0 = 0$ and hence $\beta = m_1 > 0$ (100). Now $\Gamma(l)$ decays exponentially with l, describing a *dissipative* process, in which perturbations remain localized in the vicinity of the perturbed node. The decay of individual correlations is $G(l) = e^{-\alpha(\beta+1)l}$, which has two sources: the topological expansion ($e^{-\alpha l}$) and the dissipation ($e^{-\beta\alpha l}$).

D. The Correlation Distribution

Consider the probability density that a randomly selected term in G_{ij} is between G and $G+dG$. Following (81) we can translate this to the probability $P(l)$ that a randomly selected node pair is between l and $l + dl$ where

$$l = -\lambda \ln G. \tag{102}$$

In a small world network, for which the topological expansion follows (69) we have $P(l) \sim e^{\alpha l}$, and hence

$$P(G)dg = P(l)\frac{dl}{dG}dG \sim \frac{e^{\alpha l}}{G}dG. \tag{103}$$

Substituting (102) for l we arrive at

$$P(G) \sim G^{-\nu} \tag{104}$$

where $\nu = \alpha\lambda + 1$ or

$$\nu = \frac{\beta + 2}{\beta + 1} \tag{105}$$

in which we used $\lambda = 1/(\beta + 1)\alpha$ (99).

Propagation - summary:

The propagation of perturbations is captured by the correlation function

$$\Gamma(l) = e^{-\beta\alpha l}$$

and the correlation distribution

$$P(G) \sim G^{-\nu}.$$

The dissipation rate, β, is determined by the Laurent expansion

$$g\left(f^{-1}(x)\right) = \sum_{m=-\infty}^{\infty} C_m x^m.$$

and the exponent ν follows

$$\nu = \frac{\beta+2}{\beta+1}.$$

- If the lowest power in the expansion is $g(f^{-1}(x))|_{x\to 0} \sim x^{m_0}$ $(m_0 \neq 0)$ we have $\beta = 0$ and $\nu = 2$, describing *conservative dynamics*.
- If the lowest power is $m_0 = 0$ we have $g(f^{-1}(x))|_{x\to 0} \sim C_0 + C_{m_1} x^{m_1}$ $(m_1 > 0)$, providing $\beta = m_1$ and $\nu = (\beta+2)/(\beta+1) < 2$, describing *dissipative dynamics*.

E. The Role of the Small World Property

The derivation above relies on the small world property, which is present in all random networks [3], and expressed here by the exponential expansion of Eq. (69). Hence while the δ and φ exponents are fully independent of the underlying topology, the universality of β and ν, and hence $\Gamma(l)$ and $P(G)$, depends on the presence of the small world property. To gain a deeper understanding of the role of the small world property in generating the dynamical universality, we consider again the logarithmic correction of (98) which in (101) and (81) leads to

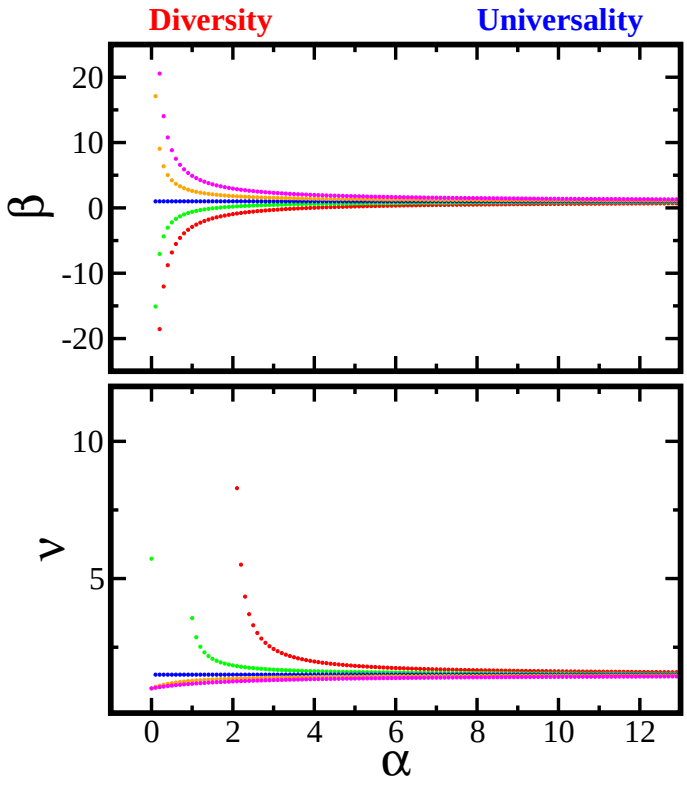

FIG. 3: **The small world property and dynamical universality**. The parameters β and ν are universally determined by the leading terms of (45) up to a logarithmic correction. In the small world limit (large α) this correction becomes negligible and universality emerges. Here we present β and ν obtained for epidemic dynamics ($\mathcal{E}$) with varying rate constants. For non-small world networks (small α) β and ν express strong diversity, but as we enter the small world regime they converge to their universal value, dependent only on the dynamical class of $\mathcal{E}$ ($\beta = 1$, $\nu = 3/2$).

$$\begin{aligned} \Gamma(l) &= e^{-\left(\beta+\frac{\ln C}{\alpha}\right)\alpha l} \\ G(l) &= e^{-\left(1+\beta+\frac{\ln C}{\alpha}\right)\alpha l} \end{aligned} \tag{106}$$

and in (105) provides

$$\nu = \frac{\beta + 2 + \frac{\ln C}{\alpha}}{\beta + 1 + \frac{\ln C}{\alpha}}, \tag{107}$$

namely

$$\beta \to \beta_{\text{Eff}} = \beta + \ln C/\alpha. \tag{108}$$

In this logarithmic term are encapsulated all the *non-universal* details, *e.g.* rate constants used in (9). When α is small, this non-universal correction dominates the system's dynamics,

determining the decay of correlations in $G(l)$ and $\Gamma(l)$ and their distribution in $P(G)$. For instance in a lattice, where the small world property is absent, we have $\alpha \to 0$ and (106) predicts $\Gamma(l) = e^{-l \ln C}$, which is fully governed by the microscopic details of (9). Hence in non small world networks the diversity prevails, as the microscopic details of (9) impact the propagation dynamics, while in the small world limit, these details become marginal and universality emerges (Fig. S3).

V. GLOBAL DYNAMICS: CASCADES

While the impact I_i captures the local effect of a perturbation on the close neighbors and the correlation function, $\Gamma(l)$, captures its propagation to more distant nodes, the full effect of a perturbation is captured by the cascade size C_i, which describes the global response of the system to an individual perturbation. The cascade includes all nodes whose activity changes beyond a threshold q following a perturbation. To derive the cascade size distribution, $P(C)$, consider a perturbation induced on node i with degree k_i. First the perturbation impacts each of i's k_i nearest neighbors, whose average response is given by (7)

$$\langle R_{ij} \rangle_{j \in K_i(1)} \sim \frac{I_i}{k_i} \sim k_i^{\varphi - 1}. \tag{109}$$

Following the response of each of these nearest neighbors the impact is propagated to the rest of the network following $G(l) \sim e^{-(\beta+1)\alpha l}$ ((88) and (99)). Here we use the assumption that there are no degree correlations in our network, allowing us to describe the propagation from i's nearest neighbors using the *average* correlation function, independently of k_i. Indeed, in the absence of degree correlations, while the node degrees k_i may be highly heterogeneous, the degrees of the neighbor's, next neighbor's etc. are quite uniform. Hence we follow the propagation of individual perturbations from i as

$$G_i(l) \sim k_i^{\varphi - 1} e^{-(\beta+1)\alpha l}. \tag{110}$$

For a node to be part of i's cascade its response to i's perturbation must be greater than the threshold, namely $G_i(l) > q$, which allows us to obtain the *cascade radius* from (110) as

$$l_{\rm C} \sim -\frac{1}{(\beta+1)\alpha} \ln\left(q k_i^{1-\varphi}\right). \tag{111}$$

On average, all nodes within a distance of $l_{\rm C}$ from i will be included in C_i. To obtain the number of nodes within this radius we follow $|K_i(l)|$, the number of nodes at distance l from i. Clearly, $K_i(1) = k_i$, after which the expansion continues as $K_i(l) \sim e^{\alpha l}$, using, once again, the absence of degree correlations. Hence the expansion from i follows

$$|K_i(l)| \sim k_i e^{\alpha l} \tag{112}$$

and the cascade size becomes

$$C_i \sim |K_i(l_{\rm C})| \sim k_i \left(k_i^{1-\varphi}\right)^{-\frac{1}{\beta+1}}. \tag{113}$$

Gathering all the terms we arrive at

$$C_i \sim k_i^{\omega} \tag{114}$$

where

$$\omega = \frac{\varphi+\beta}{\beta+1} \tag{115}$$

depends only on the dynamical model ($f(x)$ and $g(x)$) through the behavior of the local impact (φ) and the propagation dynamics (β). See also Fig. 4 where we show a geometric derivation of ω.

The precise value of ω leads to four classes of dynamical behavior, based on the values of β and φ:

(i) **Uniform cascades** ($\beta = \varphi = 0$): For a conservative system with uniform local impact (115) predicts $\omega = 0$. Hence all nodes generate comparable cascades, independent of their degree, providing a uniform cascade size distribution $P(C)$. Remarkably, in such systems even if $P(k)$ is fat-tailed, the cascade size distribution $P(C)$ will be bounded, so that the dynamical behavior is independent of the topological heterogeneity.

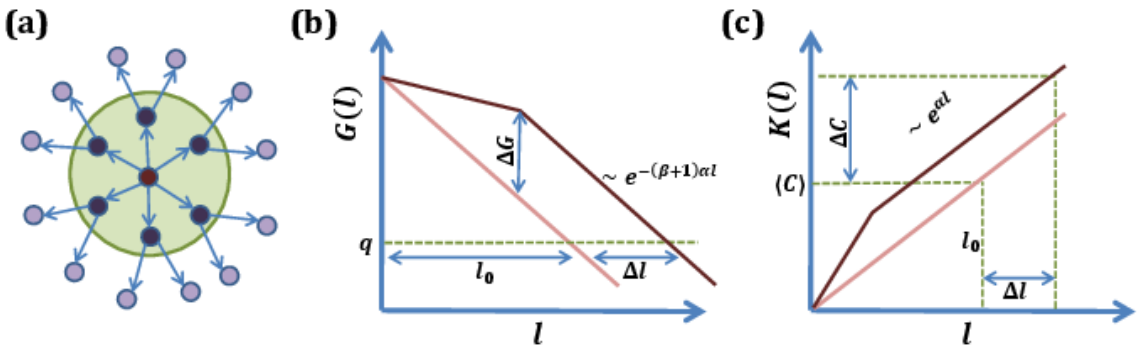

FIG. 4: **Evaluating the cascade size**. (a) The cascade includes all nodes whose activity changes above a threshold q following the perturbation of the central node (source). (b) The cascade radius defines the distance from the source for which the correlations are above q. For the average node we write $G(l_0) = q$, providing the average cascade radius l_0. For a node with degree k_i we first evaluate the impact on its nearest neighbors. Taking advantage of the *geometry* of the semi-logarithmic plot we write $\Delta G = (\varphi - 1) \ln k_i$. From the nearest neighbors of the source correlations keep propagating as (81), which is geometrically described by the straight line $G(l) \sim -(\beta+1)\alpha l$. The cascade radius is thus enlarged by $\Delta l = \Delta G/(\beta + 1)\alpha = [(\varphi - 1)/(\beta + 1)\alpha] \ln k_i$. (c) To evaluate the number of nodes within the cascade radius we present $|K_i(l)|$ (69) and $|K(l)|$ on the semi-logarithmic plane. The average cascade is $\langle C \rangle = |K_i(l_0)|$, and the increase observed in C_i is geometrically described by ΔC. It has two sources: the more rapid expansion $|K_i(l)| \sim k_i |K(l)|$ and the larger cascade radius $l_i = l_0 + \Delta l$. Geometrically we obtain $\Delta C = \ln k_i + \alpha \Delta l = (\varphi + \beta)/(\beta + 1) \ln k_i$. The translation back to the linear scale provides $C_i \sim k_i^{\omega}$ where ω follows (115).

(ii) **Locally heterogeneous cascades** ($\beta = 0$, $\varphi \neq 0$): For a conservative system with heterogeneous local impact (115) predicts $\omega = \varphi$. Hence C_i scales with a nodes degree, k_i, and consequently $P(C)$ is driven by $P(k)$, becoming fat-tailed if $P(k)$ is fat-tailed. The cascade heterogeneity is driven by the local dynamics through the heterogeneous local impact that nodes have on their nearest neighbors (I_i, φ). Hence $\omega = \varphi$ and $P(C) \sim P(I)$.

(iii) **Propagation generated heterogeneous cascades** ($\beta > 0$, $\varphi = 0$): For dissipative dynamics with uniform local impact (115) predicts $\omega = \beta/(\beta + 1) > 0$, so that $P(C)$ is heterogeneous. The source of the cascade heterogeneity is the dissipative nature of the spreading dynamics $(\Gamma(l), \beta)$, rather than the local impact between neighbors. Hence, remarkably, while all nodes have comparable local impact, their global impact on the network can become highly heterogeneous.

(iv) **Heterogeneous cascades** ($\beta > 0$, $\varphi \neq 0$): In these systems $\omega > 0$ and the global

dynamics is characterized by heterogeneous cascades. The heterogeneity originates in both the local dynamics (heterogeneous local impact) and the spreading dynamics (dissipative), hence both $P(C)$ and $P(I)$ are fat-tailed, but as opposed to (ii) $P(C)$ may follow a different form than $P(I)$, as it is not only a consequence of the local dynamics, but also of the spreading patterns.

For a directed network we must distinguish between the role of k_i in Eq. (111) and that in Eq. (112). The degree appearing in the derivation of (111) is the *in-degree*, $k_i^{\rm In}$, which charaterizes the number of neighbors acting on i. To understand this consider Eq. (22), from which the scaling of R_{ij} was derived. There, k_i is the number of neighbors apearing in the interaction term of the dynamical equation (9), namely the number of nodes acting on i, or $k_i^{\rm In}$. However, the network expansion, desctribed by (112) depends on the number of outgoing links $k_i^{\rm Out}$. Hence, repeating the above calculation we find that for directed networks

$$C_i \sim \left(k_i^{\rm In}\right)^{\omega_{\rm In}} k_i^{\rm Out} \tag{116}$$

where $\omega_{\rm In} = (\varphi - 1)/(\beta + 1)$.

VI. GENERALIZING THE DYNAMICS

The derivations of Secs. III - V can be generalized to account for all dynamics following

$$\frac{dx_i}{dt} = W(x_i) + \sum_{j=1}^{N} A_{ij} Q(x_i, x_j) \tag{117}$$

with a general $Q(x_i, x_j)$, including even for functions that cannot be factorized as in (17). We start from the general result derived in Sec. III for R_{ij}, (see Eq. (16))

$$R_{ij} = \frac{A_{ij} x_j \tilde{Q}'_{x_j}(x_i, x_j)}{k_i x_i \left\langle \tilde{Q}'_{x_i}(x_i, x_n) \right\rangle_{n \in K_i(1)}}. \tag{118}$$

The stability S_i is obtained by summing over i, while the impact I_j by summing over j, which following a derivation analogous to that of Sec. III A, can be shown to depend on the three functions

$$f(x_i) \;=\; \left\langle \tilde{Q}(x_i, x_j) \right\rangle_{j \in K_i(1)} \tag{119}$$

$$h(x_i) \;=\; \left\langle \frac{\partial \tilde{Q}(x_i, x_j)}{\partial x_j} \right\rangle_{j \in K_i(1)} \tag{120}$$

$$g(x_j) \;=\; \left\langle \tilde{Q}(x_i, x_j) \right\rangle_{i \in K_j(1)}. \tag{121}$$

The pertinent Laurent expansions are thus

$$f^{-1}(x) \;=\; \sum_{n=-\infty}^{\infty} a_n x^n \tag{122}$$

$$h\left(f^{-1}(x)\right) \;=\; \sum_{w=-\infty}^{\infty} b_w x^w \tag{123}$$

$$g\left(f^{-1}(x)\right) \;=\; \sum_{m=-\infty}^{\infty} c_m x^m, \tag{124}$$

whose leading terms uniquely determine the dynamical exponents of the system.

We follow all the same steps to derive I_i and S_i (Sec. III), $\Gamma(l)$ and $P(G)$ (Sec. IV) and C_i (Sec. V). This results in

$$\delta \;=\; n_1 - n_0 + w_0 - 1 \tag{125}$$

$$\varphi \;=\; \delta - m_1 + 1 \tag{126}$$

$$\beta \;=\; m_1 - m_0 + w_0 - 1. \tag{127}$$

As before $\nu = (\beta + 2)/(\beta + 1)$ and $\omega = (\beta + \varphi)/(\beta + 1)$. In case $\tilde{Q}(x_i, x_j)$ can be factorized as in (17) we have $h(f^{-1}(x)) \sim x$, hence $w_0 = 1$ and the results of (125) - (127) converge to those documented in the main paper.

Note, however, that with no separation of variables the leading terms of (122) - (124) may depend on the steady-state values x_i. Indeed that fact that x_j cannot be factored out in (119) and (120) (or x_i in (121)) means that the relevant Laurent expansion may have powers that depend on the specific values of x_j (or x_i). As a result the dynamical exponents will also depend on the steady state activities, no longer having the intrinsic discrete values observed for the factorized models (17). Instead, the same dynamical model may be characterized by different exponents, depending on the detailed steady state of the system. The formalism outlined in Secs. III - V can be applied to this case as well, leading to potentially novel behavior. At this point, lacking systems to motivate further work in this direction, we have not followed this path.

VII. NUMERICAL SUPPORT

To test the predictions of the dynamical theory, we performed extensive numerical simulations, incorporating a set of widely used dynamical models (see Table I in paper) on both model and real network topologies. In each of these numerical tests we ran Eq. (9) using a fourth-order Runge-Kutta stepper, having it reach steady-sate for all node activities. We then obtained the full correlation matrix, G_{ij} (3) by perturbing every x_j as

$$x_j \to (1+\chi)x_j \tag{128}$$

with $\chi = 0.1$, and running the Runge-Kutta stepper again until the system reaches the perturbed steady state, in which,

$$x_i \to (1+\chi_{ij})x_i. \tag{129}$$

From this we constructed the correlation matrix according to

$$G_{ij} = \frac{\chi_{ij}}{\chi}. \tag{130}$$

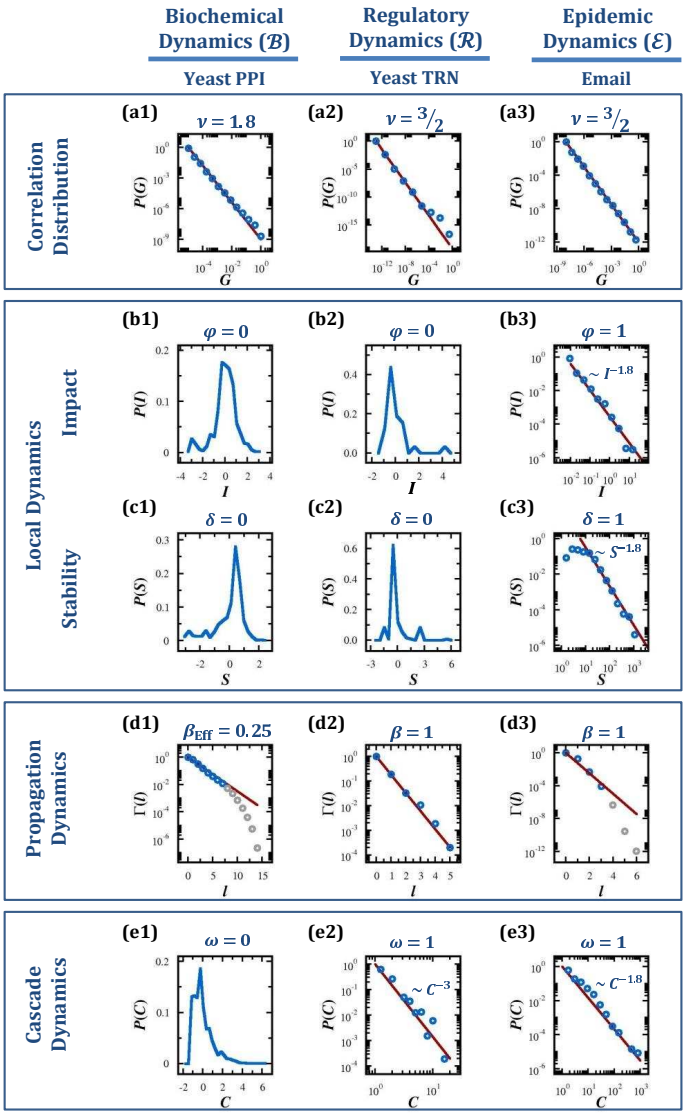

FIG. 5: **Universality classes in empirical networks**. To test our predictions on real networks, we measured the dynamical functions on three real topologies, each with an appropriate dynamics. We applied biochemical dynamics ($\mathcal{B}$) to the yeast protein-protein interaction (PPI) network, consisting of $1,647$ nodes and $5,036$ links [5]; regulatory dynamics ($\mathcal{R}$) to the yeast transcriptional regulatory network (TRN), consisting of 915 nodes and $1,063$ directional links [6] and epidemic dynamics ($\mathcal{E}$) to the email dataset analyzed in Sec. S.VIII.A, generating a network of $2,688$ nodes and $47,578$ links [7].

In the simulations described below, we set all rate constants to one, unless explicitly mentioned otherwise.

A. Epidemic Dynamics - $\mathcal{E}$

In the susceptible-infected-susceptible (SIS) model each node may be in one of two potential states: infected (I) and susceptible (S). The dynamics is given by the two processes

$$I + S \to 2I, \tag{131}$$

where a susceptible node is infected by one of its nearest neighbors, and

$$I \to S, \tag{132}$$

where an infected node is recovered, becoming susceptible again. The activity of a node, $0 \leq x_i \leq 1$ denotes the probability that the node is in the infected state. The dynamics of the system is governed by [8]

$$\frac{dx_i}{dt} = -Bx_i + \sum_{j=1}^{N} A_{ij} R(1 - x_i) x_j. \tag{133}$$

The first term on the r.h.s. accounts for the process of recovery and the second term accounts for the process of infection, a node could only be infected if its in the susceptible state $(1-x_i)$ and one of its neighbors is infected (x_j).

We first define the dynamical functions $f(x)$ and $g(x)$ which here are

$$\begin{aligned} f(x) &= \frac{1-x}{x} \\ g(x) &= \frac{R}{B}x. \end{aligned} \tag{134}$$

We obtain δ by expanding $f^{-1}(x)$ as in (34)

$$f^{-1}(x) = \frac{1}{1+x} \sim 1 - x + \frac{1}{2}x^2 + \dots. \tag{135}$$

Since the leading term has $n_0 = 0$ and the leading non-vanishing term is $n_1 = 1$, we have *heterogeneous stability* with $\delta = n_1 = 1$ (33).

Next we focus on the structure of $g(f^{-1}(x))$, which, since $g(x) \sim x$ (134), is the same as (135), namely $g(f^{-1}(x)) \sim 1 - x + O(x^2)$. Hence the leading term is $m_0 = 0$ and the leading non-vanishing term is $m_1 = 1$. This provides *heterogeneous impact* with $\phi = \delta + 1 - m_1 = 1$

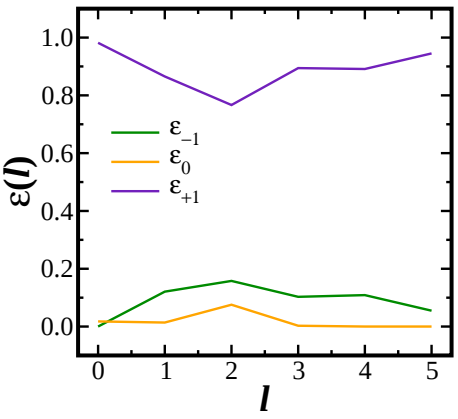

FIG. 6: **The expansion factor for the yeast transcriptional regulatory network**. The tree-like structure of the transcriptional regulatory network, in which loops are scarce, is expressed in the expansion factor by an almost vanishing $\epsilon_0(l)$ and $\epsilon_{-1}(l)$, and a dominant $\epsilon_{+1}(l)$, persistent for all l.

(44), and *dissipative* propagation with $\beta = m_1 - m_0 = 1$ (100). Finally, for $P(G)$ we have $\nu = (\beta + 2)/(\beta + 1) = 3/2$ (105), and for the cascades we have $\omega = (\phi + \beta)/(\beta + 1) = 1$ (115), resulting in *heterogeneous cascades*.

As an empirical network for the SIS model we used an email dataset [7], which records all 3×10^5 emails sent between $3,188$ individuals over the course of $T = 161$ days. We constructed the network by linking every pair of nodes where i sent at least one email to j over the sampled period, resulting in a giant component with $2,688$ nodes and $47,578$ links (Fig. S5a3 - e3).

B. Regulatory Dynamics - $\mathcal{R}$

To model regulatory interactions we referred to the commonly uses Michaelis-Menten dynamics which take the form [9, 10],

$$\frac{dx_i}{dt} = -Bx_i + \sum_{j=1}^{N} R\mathcal{H}(x_j), \tag{136}$$

where $\mathcal{H}(x_j)$ is the Hill function characterizing the activation/inhibition of x_i by x_j. As the regulation of x_i depends on the presence or absence of x_j, having little sensitivity to j's specific abundance, the Hill function is designed to be a *switch*-like function, satisfying $\mathcal{H}(x_j) \to 1$ ($\mathcal{H}(x_j) \to 0$) for large (small) x_j in case x_j activates x_i, or $\mathcal{H}(x_j) \to 1$ ($\mathcal{H}(x_j) \to 0$) for (small) large x_j in the case of inhibition. More specifically, $\mathcal{H}(x_j)$ has the form [9]

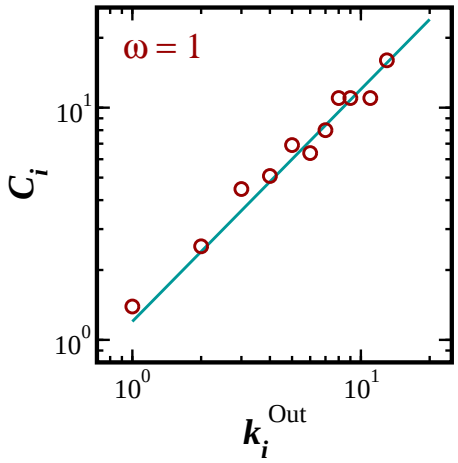

FIG. 7: **Cascades in a directed network**. For the directed yeast transcriptional regulation network we calculate C_i using (116), predicting $C_i \sim k_i^{\rm Out}$.

$$\mathcal{H}(x_j) = \begin{cases} \frac{x_j^h}{1+x_j^h} & \text{for activation} \\ \frac{1}{1+x_j^h} & \text{for inhibition} \end{cases}, \tag{137}$$

where h is the Hill coefficient [10]. Equation (136) can be mapped on to (9) by taking

$$\begin{aligned} f(x) &\sim \frac{1}{x} \\ g(x) &\sim \mathcal{H}(x). \end{aligned} \tag{138}$$

Here, as $f^{-1}(x) \sim x^{-1}$, the leading terms are $n_0 = n_1 = -1$, and hence the stability is *uniform*, namely $\delta = 0$ (33).

For the second Laurent expansion (45) we write

$$g(f^{-1}(x)) \sim \mathcal{H}(x^{-1}) \sim \begin{cases} 1 - x^h + O(x^{2h}) & \text{for activation} \\ x^h + O(x^{2h}) & \text{for inhibition} \end{cases}. \tag{139}$$

Hence for activation we have $m_0 = 0$ and $m_1 = h$, predicting: $\phi = 1 - h$ (44), $\beta = h$ (100) and consequently $\nu = (h+2)/(h+1)$ (105) and $\omega = 1/(1+h)$ (115). For inhibition, following the same steps we find $m_0 = m_1 = h$, hence $\phi = 1 - h$ as for activation, but as opposed to that $\beta = 0$, $\nu = 2$ and $\omega = 1 - h$.

We tested our predictions on the yeast transcriptional regulatory network (TRN) [6]), which is dominated by activation, accounting for over 80% of the links. Hence from the two predictions above, activation versus inhibition, we focus on the first, namely the one for

activation. Choosing $h = 1$ in our simulations, we predict $\delta = 0$ (uniform stability), $\phi = 0$ (uniform impact), $\beta = 1$ (dissipative dynamics), $\nu = 3/2$ and $\omega = 1/2$ (spread generated heterogeneous cascades). In Fig. S5a2 - e2 we show the results obtained for the yeast TRN, all in line with our predictions. As Fig. S6 indicates, the yeast TRN has $\epsilon_{+1}(l) > \epsilon_{-1}(l)$ for all l, describing a unique tree-like structure, which does not saturate. This is because this directed network has almost no loops and thus continues on branching, forming a hierarchical structure for all $l_{\max} = 5$ layers. As a result, most pairs of nodes are at distance $l \to \infty$, namely there is no directed path connecting them, and the average path length, $\langle l \rangle \approx 1.4$ calculated only among the finite paths, is greatly underestimated. Hence we cannot use Eq. (72) to obtain a reliable estimate for α in this case. As an alternative we use the empirically measured $\vec{\epsilon}(l)$, which as shown in Fig. S6, is independent of l. We thus extracted α using (77), providing $\alpha = -\ln(1 - \epsilon_{+1}(l)) \approx 1.7$. As this network is directed we used (116) to obtain the scaling of C_i as $C_i \sim k_i^{\text{Out}}$ (Fig. S7).

C. Biochemical Dynamics - $\mathcal{B}$

As a biochemical example we consider protein-protein interactions (PPI), which include the processes $\varnothing \to X_i$ describing the synthesis of an protein i at rate F; $X_i \to \varnothing$ describing protein degradation at rate B; $X_i + X_j \leftrightharpoons X_i X_j$ describing the binding (unbinding) of a pair of interacting proteins at rate R (U). The hetero-dimer $X_i X_j$ undergoes degradation $X_i X_j \to \varnothing$ at rate Q. The dynamical equations for this system are [11]

$$\begin{aligned} \frac{dx_i}{dt} &= F - Bx_i + \sum_{j=1}^{N} U x_{ij} - \sum_{j=1}^{N} A_{ij} R x_i x_j \\ \frac{dx_{ij}}{dt} &= A_{ij} R x_i x_j - (U+Q) x_{ij}, \end{aligned} \tag{140}$$

where $x_i(t)$ is the concentration of i and $x_{ij}(t)$ is the concentration of the hetero-dimer $X_i X_j$. Assuming steady state for the hetero-dimer concentration, we set $dx_{ij}/dt = 0$, obtaining

$$\frac{dx_i}{dt} = F - Bx_i - \sum_{j=1}^{N} A_{ij} \tilde{R} x_i x_j \tag{141}$$

in the form (9), where the effective binding rate $\tilde{R} = QR/(U+Q)$. The dynamical functions are thus

$$\begin{aligned} f(x) &= \frac{x}{F - Bx} \\ g(x) &= \tilde{R}x, \end{aligned} \tag{142}$$

from which we obtain (34) and (45)

$$\begin{aligned} f^{-1}(x) &= \frac{Fx}{1 + Bx} \sim x + O(x^2) \\ g\left(f^{-1}(x)\right) &= \frac{\tilde{R}Fx}{1 + Bx} \sim x + O(x^2). \end{aligned} \tag{143}$$

The first expansion in (143) provides $n_0 = n_1 = 1$, predicting $\delta = 0$ (33) (uniform stability). From the second expansion we write $m_0 = m_1 = 1$, and hence $\phi = 0$ (44) (uniform impact) and $\beta = 0$ (100) (conservative propagation). Finally, as for all conservative propagation dynamics, $\nu = 2$ (105) and following (115) we have $\omega = 0$ (uniform cascades).

Note that the conservative nature of the dynamics is a consequence of only the leading terms in the expansion of (142). This implied a conservative propagation even if, at the microscopic level, the pairwise dynamics is non-conservative. Indeed, in (141) processes such as influx (F) and degradation (B, Q) violate the conservation of mass as they draw proteins in and out of the system. Still, we predict that perturbations propagate without loss ($\beta = 0$). These non-conservative processes are expressed in the logarithmic correction of (46), and are expected to have little effect in the small world limit (see Sec S.IV E).

In Fig. S5a1 - e1 we display the results of numerical simulations obtained for the yeast PPI network [5]. A slight discrepancy is observed in the value of ν (1.8 vs. 2). To understand this deviation consider the logarithmic correction in the value of β (98), which we rendered negligible in the small world limit where α is large. As the yeast PPI network has a relatively low average degree of $\langle k \rangle \approx 3$ it also has a small $\alpha = 1.3$. Consequently the logarithmic correction has a detectable impact on ν (107). This is especially significant in the case of conservative dynamics, where $\beta_{\text{Eff}} = 0 + \ln C/\alpha$ (108), so that the role of the logarithmic correction is more pronounced (In the SIS or MM models , $\beta_{\text{Eff}} = 1 + \ln C/\alpha$, so that the logarithmic correction has a less significant impact). We can now use the measured value of ν to evaluate the effective dissipation rate β_{Eff} using (105), obtaining $\beta_{\text{Eff}} = 0.25$, in prefect agreement with the simulation results (Fig. S5d1). Choosing networks deeper in the small

world limit (such as the model networks featured in the paper) eliminates the effect of the logarithmic correction, and generates the expected universal dynamical behavior.

D. Birth-Death Processes - $\mathcal{BD}$

Birth-death processes have many applications in population dynamics [12], queuing theory [13] or biology [12]. We consider a network in which the nodes represent sites, each site i having a population x_i, where population flow is enabled between neighboring sites. This process can be described by a dynamical equation of the form (9) as

$$\frac{dx_i}{dt} = -Bx_i^b + \sum_{j=1}^{N} A_{ij} x_j^a. \tag{144}$$

The first term on the r.h.s. represents the internal dynamics of site i, characterized by the exponent b. In queuing dynamics, choosing $b = 0$ represents a constant influx (outflux) into (out of) site i; in population dynamics mortality can be represented by setting $b = 1$, indicating that the number of mortality instances per unit time is proportional to the current population at i; below we set $b = 2$ to represent pairwise depletion, as frequently used in ecology to account for competition within a population over limited resources [14], or in biochemistry to model dimerization [11]. The second term describes the flow from i's neighboring sites j into i, which is typically linear in x_j, namely $a = 1$.

Equation (144) can be cast into (10) with

$$\begin{aligned} f(x) &= -\frac{1}{Bx^b} \\ g(x) &= x^a, \end{aligned} \tag{145}$$

providing

$$\begin{aligned} f^{-1}(x) &\sim x^{-1/b} \\ g\left(f^{-1}(x)\right) &\sim x^{-a/b}. \end{aligned} \tag{146}$$

Hence $\mathcal{BD}$ has $n_0 = n_1 = -1/b \neq 0$ and $m_0 = m_1 = -a/b$, predicting $\delta = 0$ (uniform stability), $\phi = 1 + a/b = 3/2 > 0$ (heterogeneous impact) and $\beta = 0$ (conservative dynamics).

Consequently we also predict $\nu = 2$ and $\omega = \phi = 3/2$ (locally generated heterogeneous cascades).

E. Scale-free Networks

Our formalism predicts five parameters, δ, ϕ, β, ν and ω, whose values are determined only by the dynamics of the system, independent of topology. These parameters, in turn, predict the behavior of the pertinent distributions, $P(S)$, $P(I)$ and $P(C)$ and their dependence on the network topology. To be specific the uniform/heterogeneous distinction *is an intrinsic property of the dynamics, independent of the underlying topology*, however the specific form of these distributions may depend on the topology through the degree distribution, $P(k)$. Consider a random variable X_i (representing the stability, impact or cascades above), which scales as

$$X_i \sim k_i^{\psi}. \tag{147}$$

For $\psi \neq 0$ this variable displays heterogeneous dynamics in which $P(X)$ is driven by $P(k)$ via

$$P(X) \sim \int_0^{\infty} P(k)\delta(k^{\psi} - X)dk. \tag{148}$$

For a scale-free network with $P(k) \sim k^{-\gamma}$ we solve this integral by substituting $x = k^{\psi}$, obtaining

$$P(X) \sim \int_0^{\infty} x^{-\Psi}\delta(x - X)dx \sim X^{-\Psi}, \tag{149}$$

where

$$\Psi = \frac{\gamma + \psi - 1}{\psi}. \tag{150}$$

This allows us the obtain the precise form of $P(S)$, $P(I)$ and $P(C)$ for a scale-free network substituting δ, ϕ or ω for ψ in (150). For the scale-free model network used in the numerical simulations we had $\gamma = 3$, for the real networks, we measured γ separately for each network, finding $\gamma = 2.2$ (yeast PPI), $\gamma = 3.0$ (yeast TRN) and $\gamma = 1.9$ (Email).

VIII. EMPIRICAL SUPPORT

To test the predictive power of the theory on empirical data we focused on two systems pertaining to human and cellular dynamics.

A. Human Dynamics

The dataset [7] records all 3×10^5 emails sent between $3,188$ individuals over the course of $T = 161$ days. We constructed the network by linking every pair of nodes where i sent at least one email to j over the sampled period. While this allows for the construction of a directed network, in practice, we found that almost all links are reciprocal, to the extent that no significant difference was detected between the directed and the undirected versions of the network. The dynamics of node i is given by $x_i(t)$, denoting the number of emails sent by a user over a period of $\Delta t = 6, 10$ or 24 hours (the results shown in the paper were obtained for $\Delta t = 6$ hours). To evaluate the correlation between the usage patterns we measured

$$G_{ij} = \frac{\frac{1}{T}\int_0^T x_i(t)x_j(t)\,dt}{\frac{1}{T}\int_0^T x_i^2(t)\,dt} = \frac{\langle x_i x_j\rangle}{\langle x_i^2\rangle}, \tag{151}$$

providing the degree to which the usage pattern of j is correlated with that of i. To evaluate the propagation we first used the non-normalized $G_{ij} = \langle x_i x_j\rangle$, which is symmetric, to obtain $\Gamma_{\rm Raw}(l)$. We then set $\Gamma(l) = \Gamma_{\rm Raw}(l)/\Gamma_{\rm Raw}(0)$ to ensure $\Gamma(0) = 1$.

Note that the only arbitrary parameter in this analysis is Δt, which is chosen to reflect the typical time scales of email activity (between several hours to one day). To test if this parameter affects the results we measured $x_i(t)$ and (151) using the three specified values of Δt. The results, shown in Fig. S8a - d indicate that our analysis is not sensitive to the selection of Δt.

B. Cellular Dynamics

The microarray data obtained from [4] includes $J = 110$ experiments in which 55 yeast genes where perturbed (twice for each gene), measuring the resulting change in the expression of the remaining $N = 6,222$ genes. This results is an $N \times J$ matrix in which the elements

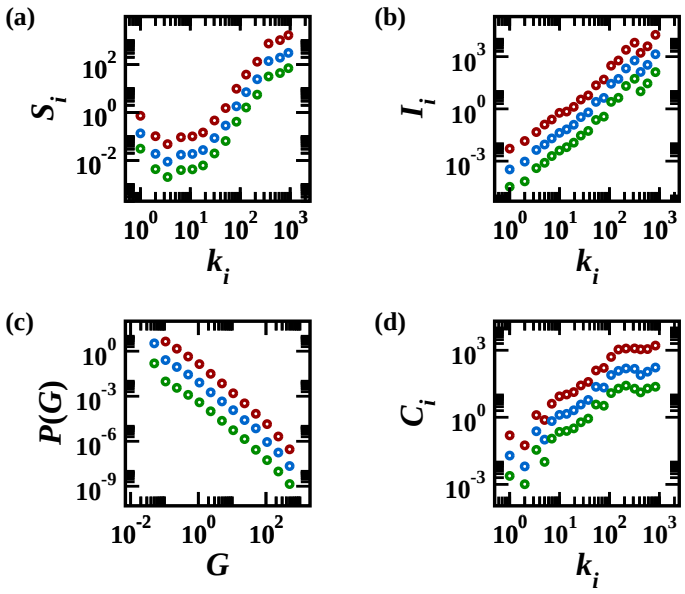

FIG. 8: **Analysis of the Email dataset**. The analysis of the human activity requires the arbitrary selection of the time resolution Δt. Here we repeated the analysis for three different choices of Δt: $\Delta t = 24$ (red), $\Delta t = 10$ (blue) and $\Delta t = 6$ (green) hours, showing that our results are insensitive to the specific choice of Δt.

are given by

$$G_{ij} = \frac{x_i(j)}{x_i(0)}, \tag{152}$$

where $x_i(0)$ is the expression level of gene i in the control, and $x_i(j)$ is the expression level of the i gene following the perturbation of j. Lacking the underlying topology, we approximated the stability and impact by

$$S_i = \frac{1}{J}\sum_{j=1}^{J} G_{ij} \tag{153}$$

and

$$I_i = \frac{1}{N}\sum_{j=1}^{N} G^{\mathrm{T}}_{ij}. \tag{154}$$

While we could not directly measure δ and ϕ, as they require us to know the degrees k_i, we could indirectly infer the pertinent class from $P(S)$ and $P(I)$, whose measurement requires no knowledge of the underlying network. Similarly, we could not directly measure β and ω, but the associated distributions $P(G)$ and $P(C)$ could be obtained as they too do require us to know the topology.

IX. ROBUSTNESS OF THE THEORY

A. Topological Uncertainty

In many cases, the challenge in analyzing network dynamics begins at the stage of defining the network, namely selecting the criteria by which to draw the links. A *liberal* strategy, in which these criteria are loose, will generate a relatively dense network, while a *conservative* strategy, following strict criteria, raises the bar for linking between nodes, resulting in a relatively sparse network, with the nodes having typically low degrees. For instance, in the email network analyzed above (Sec. S.VIII A), the structure of the underlying network strongly depends on the definition of a link: does a single email exchange between a pair of nodes constitute a significant social tie, or perhaps several emails (q) are required. This arbitrary choice impacts the degrees of all nodes, the distance between them, and consequently the value of α - a liberal choice (q low) leads to a dense network with a large α, and a conservative choice (q high) results in a sparse network with a small α. However, we argue that the dynamical functions presented in this work, are all robust against such variability in the network construction strategies, and can thus provide dynamical predictions even in the face of topological uncertainty.

This robustness is clear in the case of $P(S)$, $P(I)$, $P(G)$ and $P(C)$, which can be measured directly from the dynamical data (*e.g.* the microarray experiments), completely independent of the network structure. Indeed, to obtain these distributions one does not need to construct the network topology at all, so that it is fully independent of any structural knowledge about the network (see Sec. S.VIII B, where we explicitly achieve this).

To understand the impact of the threshold q on the scaling of S_i and I_i, consider its effect on the degrees of all nodes. Adopting a liberal (conservative) strategy will shift the degrees upwards (downwards), effectively rescaling them as

$$k_i(q') \sim C(q,q')k_i(q), \tag{155}$$

where $C(q,q') > 1$ ($C(q,q') < 1$) if $q' < q$ ($q' > q$). Such a shift in the degrees of all nodes has no effect on the scaling of S_i and I_i so that δ, and φ remain unaffected. Of course, one

can define criteria for the network construction, where the degrees of all nodes change in a non-monotonic fashion, such that nodes with a low degree in one construction scheme end up having a high degree in the other, affecting the scaling of S_i and I_i. However, such an inconsistency between the construction schemes is unlikely, if both are to capture the true patterns of interactions in the system. Indeed, one expects that a highly connected node under one network construction scheme will remain highly connected in the other. Thus as long as the degree ranking is preserved, the scaling exponents are not likely to significantly change.

The measure which is most sensitive to the topology of the network is $\Gamma(l)$, which depends on α, and consequently on the distance between all pairs of nodes. Still, as we next show, the self-consistency of the theory ensures that even $\Gamma(l)$ maintains its validity under different network construction strategies. To understand this consider a system, *e.g.* the email dataset, for which the network was constructed using two thresholds, $q > q'$, giving rise to the topologies T sparser than T', with $\alpha < \alpha'$. We now focus on a pair of nodes, i and j, whose dynamical correlation is G_{ij}. Using the sparser topology T we find that these two nodes are more distant than with the denser topology T', namely $l > l'$. However, from (99), as λ is inversely dependent on α, we also find that $\lambda > \lambda'$, so that while the nodes are more distant in T than in T', $G(l)$ and $\Gamma(l)$ decay more slowly. The predicted correlation between this pair will be approximately the same in both topologies, namely

$$e^{-\frac{l}{\lambda}} \approx e^{-\frac{l'}{\lambda'}}. \tag{156}$$

A similar argument applies also for other sources of topological uncertainty. For instance, consider a system for which only a fraction f of the links are known. If this fraction represents a random selection of all links, then on average all degrees will be rescaled as $k_{i_{\text{Known}}} = f k_{i_{\text{Real}}}$, a similar rescaling to that of (155).

B. Empirical Realization

The theory presented here is exact in the limit of small perturbations, where the linear response of G_{ij} is a valid approximation. In practice, however, measuring G_{ij} can take different forms: Sometimes statistical correlation measures are used as a proxy for G_{ij}, such as in our analysis of the human dynamics (email). In other cases, the perturbations are

uncontrolled, so that they may be rather large, such as in the microarray data, where large perturbations have been applied, at times up to the complete knockout of an entire gene. As our empirical results clearly indicate, the theory's predictions are robust against such deviations from the small perturbation limit. Of course, any prediction regarding specific terms of G_{ij} is expected to strongly depend on the size and form of the induced perturbations. Yet the exponents we predict capture the relationship between these terms, which are largely independent of the specific empirical realization of G_{ij}. For instance, if node i has high stability in response to small perturbations, it is natural to expect that it will also have a high stability if the perturbations are large. While the specific value of S_i might change, its relationship with all other S_j is likely to remain the same, hence δ, which quantifies this relationship, will not be sensitive to such deviations in the empirical realization. Similarly, if the system features dissipative dynamics, one expects this property to be expressed for all types of perturbations, since the dissipation is an intrinsic characteristic of the flow of perturbations in the system. A similar argument holds for all other parameters predicted by our framework. To test this, in Fig. 9 we present results obtained for regulatory dynamics ($\mathcal{R}$), in which G_{ij} is obtained from both small perturbations (blue) and extremely large perturbation, in which the node is completely removed from the system, *i.e.* knockout (red). Clearly, the scaling exponents, which are intrinsic to $\mathcal{R}$, are identical in both realizations, indicating that our theoretical predictions are insensitive to the size of the induced perturbations. This robustness is key to the theory's empirical relevance, as in actual empirical settings, the precise mathematical conditions upon which the theory builds, cannot always be realized.

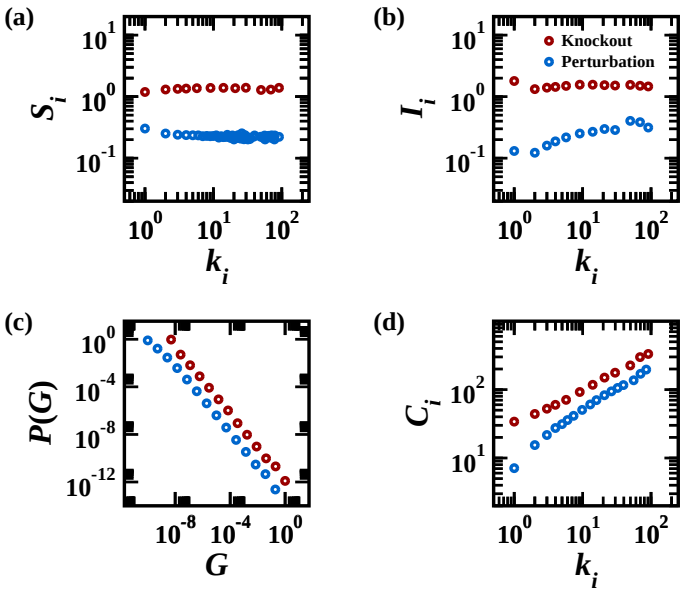

FIG. 9: **Testing the theory in the limit of large perturbations**. To test the theory's applicability in the limit of large perturbations we obtained G_{ij} for $\mathcal{R}$ using node removal, namely setting $x_j = 0$, and measuring the response of x_i (red). The results show that all the dynamical exponents remain unchanged compared to the ones obtained from small perturbations (blue). Hence our theoretical predictions, derived under the assumption of small perturbations, are insensitive to the specific realization of G_{ij}.

[1] B. Barzel and O. Biham. Quantifying the connectivity of a network: The network correlation function method. *Phys. Rev. E* **80**, 046104–15 (2009).

[2] T. Opsahl and P. Panzarasa. Clustering in weighted networks. *Social Networks* **31**, 155–163 (2009).

[3] M.E.J. Newman. *Networks - an introduction*. Oxford University Press, New York, (2010).

[4] G. Chua *et al.* Identifying transcription factor functions and targets by phenotypic activation. *Proc. Natl. Acad. Sci. US* **103**, 12045–50 (2006).

[5] H. Yu *et al.* High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104–110 (2008).

[6] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii and U. Alon. Network motifs: Simple building blocks of complex networks. *Science* **298**, 824–827 (2002).

[7] J.-P. Eckmann, E. Moses and D. Sergi. Entropy of dialogues creates coherent structures in e-mail traffic. *Proc. Natl. Acad. Sci. US* **101**, 14333–7 (2004).

[8] P.S. Dodds and D.J. Watts. A generalized model of social and biological contagion. *Journal of Theoretical Biology* **232**, 587–604 (2005).

[9] U. Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall, London, U.K., (2006).

[10] G. Karlebach and R. Shamir. Modelling and analysis of gene regulatory networks. *Nature Reviews* **9**, 770–780 (2008).

[11] E.O. Voit. *Computational Analysis of Biochemical Systems*. Cambridge University Press, New York, NY, (2000).

[12] A.S. Novozhilov, G.P. Karev and E.V. Koonin. Biological applications of the theory of birth-and-death processes. *Briefings in Bioinformatics* **7**, 70–85 (2006).

[13] J.F. Hayes and T.V.J. Ganesh Babu. *Modeling and Analysis of Telecommunications Networks*. John Wiley & Sons, Inc., Hoboken, NJ, USA, (2004).

[14] M.G. Waldon. Competition models. *The American Naturalist* **109**, 487–489 (1975).

[15] K.M. Harris, C.T. Halpern, E. Whitsel, J. Hussey, J. Tabor, P. Entzel and J.R. Udry. The National Longitudinal Study of Adolescent Health: Research Design [WWW document], (2009). URL: http://www.cpc.unc.edu/projects/addhealth/design.