nature physics

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41567-018-0409-0

In the format provided by the authors and unedited.

Spatiotemporal signal propagation in complex networks

Chittaranjan Hens[1,2,3], Uzi Harush[1,3], Simi Haber[1], Reuven Cohen[1] and Baruch Barzel[1]*

[1]Department of Mathematics, Bar-Ilan University, Ramat-Gan, Israel. [2]Physics and Applied Mathematics Unit, Indian Statistical Institute, Kolkata, India. [3]These authors contributed equally: Chittaranjan Hens, Uzi Harush. *e-mail: baruchbarzel@gmail.com

Spatio-temporal signal propagation in complex networks

Supplementary Material

Chittaranjan Hens, Uzi Harush, Simi Haber, Reuven Cohen & Baruch Barzel

November 30, 2018

Contents

1 Analytical derivations - from M to τ_i

To construct the propagation times $T(j \to i)$ we must first obtain the individual response times τ_i, capturing the transient response of each node i to direct incoming perturbations from its interacting neighbors. Indeed, as we show in the main paper (Eq. (7)), the propagation times $T(j \to i)$ can be constructed from the sequence of local responses τ_i along each path. Hence we use a perturbative approach to derive the response time of a node to a directly neighboring perturbation Δx_m.

Starting from the dynamic equation

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = M_0(x_i(t)) + \sum_{j=1}^{\mathrm{N}} A_{ij} M_1(x_i(t)) M_2(x_j(t)), \tag{1.1}$$

we obtain the steady state x_i by setting the derivative on the l.h.s. to zero, and then introduce a time-independent perturbation $x_m(t) = x_m + \Delta x_m$ on the activity of node m, one of i's nearest neighbors. Throughout the derivation we denote time-independent variables, *i.e.* steady-states, by $x_i, \Delta x_i$, omitting the argument t, and time-dependent variables by $x_i(t), \Delta x_i(t)$. Node i's response to the permanent Δx_m signal will follow

$$x_i(t) = x_i + \Delta x_i(t), \tag{1.2}$$

with τ_i representing the relaxation time of $\Delta x_i(t)$. Below, we show in detail how to calculate τ_i (for a brief outline of this derivation see Methods).

Our theoretical framework. Our derivation predicts the scaling relationship between τ_i and each node's weighted degree S_i, directly from the system's dynamics $\mathbf{M} = (M_0(x), M_1(x), M_2(x))$ in (1.1). Throughout this derivation we rely on two main approximate assumptions:

- *Perturbative limit.* We assume that the signal Δx_m is small, allowing us to employ linear response theoretic tools to treat (1.1) analytically.
- *Configuration model.* We allow A_{ij} to feature any arbitrary degree/(positive) weight distribution, including scale-free or other fat-tailed density functions, but assume that it is otherwise random [1]. Such approximation may overlook certain aspects of the network's fine-structure, such as degree-degree correlations [2], or clustering, which, in the limit of sparse networks ($\langle k \rangle \ll N \to \infty$) become negligible due to the random connectivity.

In Sec. 7 we systematically test the robustness of our predictions against these approximations. We examine the impact of large perturbtaions, including the system's response to complete node knockout, an unambiguously *large* perturbation. We also observe our theory's performance under increasing levels of degree-correlations, clustering, hidden nodes and parametric noise in $\mathbf{M}$. We find, that our predictions are generally insensitive to these approximations, successfully withstanding empirically relevant levels of clustering and degree-correlations as well as large signals, all of which have but a marginal - and in fact non-visible - effect on our predicted scaling and its consequent dynamic regimes. The origins of this robustness are also discussed in Sec. 7. We also test the level of uncertainty, *i.e.* hidden nodes and noise, beyond which our predictions break down.

1.1 Configuration model

Throughout our analysis below we use the configuration model framework to analyze A_{ij} [1]. Within this framework A_{ij} represents a general weighted network with arbitrary degree and weight distributions, but otherwise random structure. Hence we assume negligible correlation between the number of neighbors of a node k_i, and its link weights A_{ij}, namely $P(A_{ij} = a|k_i) = P(A_{ij} = a)$. Another significant implication of the configuration model assumption is that we neglect minor structural correlations between nodes and their immediate surrounding. For instance, while two nodes, i and j, may have extremely different topological characteristics, say i is a hub and j is a low degree node, their neighborhoods are assumed to share similar statistical properties, namely i's (many) neighbors are extracted from the same statistical pool as j's (few) neighbors. Specifically, let us denote by $\mathbf{G}(S)$ the group of all nodes whose weighted degrees are between S and $S + \mathrm{d}S$, *i.e.* $\mathbf{G}(S) = \{i|S_i \in (S, S + \mathrm{d}S)\}$. This group can be characterized by one or more random variables Q_i, capturing, for instance the activity $x_i(t)$ or the response time τ_i associated with a randomly selected node $i \in \mathbf{G}(S)$. The corresponding distribution

$$P_S(Q_i = q) \equiv P\big(Q_i = q\big|i \in \mathbf{G}(S)\big) \tag{1.3}$$

is unique to $\mathbf{G}(S)$, since nodes in $\mathbf{G}(S)$ are distinct from nodes in $\mathbf{G}(S')$, hence, in general $P_S(Q_i = q) \neq P_{S'}(Q_i = q)$. This distinction translates also to statistics extracted from $\mathbf{G}(S)$. For instance, the mean value of Q_i over nodes in $\mathbf{G}(S)$, denoted by

$$Q(S) = \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} Q_i, \tag{1.4}$$

may, in general, differ from $Q(S')$; in (1.4) $|\mathbf{G}(S)|$ represents the number of nodes in $\mathbf{G}(S)$. Taking Q_i to represent, *e.g.*, response time, the implication is that the typical response time of nodes in S is potentially different than that of nodes in S', *i.e.* $\tau(S) \neq \tau(S')$.

Next we consider the random variable

$$Q_{i,\odot} = \frac{1}{S_i} \sum_{n=1}^{N} A_{in} Q_n, \tag{1.5}$$

a weighted average over i's nearest neighbors, whose probability distribution is given by $P(Q_{i,\odot} = q)$. Averaging over nodes in $\mathbf{G}(S)$ we obtain

$$Q_\odot(S) = \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} Q_{i,\odot} = \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} \frac{1}{S} \sum_{n=1}^{N} A_{in} Q_n, \tag{1.6}$$

analogous to $Q(S)$ in (1.4). According to the configuration model the nearest neighbors of $i \in \mathbf{G}(S)$ and those of $j \in \mathbf{G}(S')$ follow similar statistics, hence we have

$$P\big(Q_{i,\odot} = q\big|i \in \mathbf{G}(S)\big) = P\big(Q_{i,\odot} = q\big|i \in \mathbf{G}(S')\big), \tag{1.7}$$

or more generally

$$P\big(Q_{i,\odot} = q\big|i \in \mathbf{G}(S)\big) = P(Q_{i,\odot} = q), \tag{1.8}$$

capturing the independence of the neighborhood ($\odot$) statistics on its central node. Therefore, Eq. (1.8) allows us to substitute the specific $Q_{i,\odot}$-distribution extracted from nodes in $\mathbf{G}(S)$ by the general $Q_{i,\odot}$-distribution taken over *all* nodes in the network. The meaning is that while the statistical properties of Q_i may, generally, depend on S, with $Q(S) \neq Q(S')$, those of $Q_{i,\odot}$ are independent of S, *i.e.* $Q_\odot(S) = Q_\odot(S')$, ultimately providing $Q_\odot(S) = \langle Q \rangle_\odot$, an average over all nodes in the network. This translates to

$$Q_\odot(S) \equiv \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} \frac{1}{S} \sum_{n=1}^{N} A_{in} Q_n = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{S_i} \sum_{n=1}^{N} A_{in} Q_n \equiv \langle Q \rangle_\odot \tag{1.9}$$

where the l.h.s. represents a nearest neighbor average over nodes within $\mathbf{G}(S)$ and the r.h.s. represents a nearest neighbor average over all nodes, a characteristic of the *network*, independent of S.

1.2 Steady state analysis

We consider systems of the form (1.1) that exhibit at least one fully positive steady state x_i $(i = 1, \ldots, N)$. We focus on the dependence of this steady-state, x_i, on a node's weighted (incoming) degree $S_i = \sum_{j=1}^{N} A_{ij}$. Therefore, we seek the average (time-dependent) activity $x(S,t)$ characterizing all nodes $i \in \mathbf{G}(S)$. Substituting $x_i(t)$ for the random variable Q_i in (1.4), this activity takes the form

$$x(S,t) = \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} x_i(t). \tag{1.10}$$

We now use (1.1) to write

$$\frac{\mathrm{d}x(S,t)}{\mathrm{d}t} = \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} \Big[M_0\big(x_i(t)\big) + \sum_{n=1}^{N} A_{in} M_1\big(x_i(t)\big) M_2\big(x_n(t)\big) \Big], \tag{1.11}$$

which we approximate by

$$\frac{\mathrm{d}x(S,t)}{\mathrm{d}t} = M_0\big(x(S,t)\big) + M_1\big(x(S,t)\big) \frac{1}{|G(S)|} \sum_{i \in \mathbf{G}(S)} \sum_{n=1}^{N} A_{in} M_2\big(x_n(t)\big). \tag{1.12}$$

Equation (1.12) approximates

$$\frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} M_q(x_i) \approx M_q \left(\frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} x_i \right), \tag{1.13}$$

$(q = 0, 1)$, namely it *introduces the average into the functions* (see Box). This is exact if $M_q(x)$ is linear, or in the limit where $x_i(t)$ are narrowly distributed within $\mathbf{G}(S)$. Indeed, if all *environments* are uniform, as per the configuration model, one expects nodes of equal degree S to exhibit similar characteristics, expressed by a narrow $P_S(x_i(t))$ distribution.

The mean-field approximation in Eq. (1.13). To make analytical advances, we use the mean-field approximation to write (shorthand) $\langle M_q(x)\rangle \approx M_q(\langle x\rangle)$. This approximation assumes three limits, relevant in the context of the present discussion:

- *Linear functions.* In case $M_q(x)$ are linear the mean-field approximation is exact. Indeed, while in general the dynamics (1.1) are nonlinear, in some cases (*e.g.*, $\mathbb{R}_1, \mathbb{E}$) the individual functions $M_q(x)$ are linear. Under these conditions Eq. (1.13) becomes exact.
- *Uniform activities.* While $x_i(t)$ can, in general, be highly diverse, driven by the extreme levels of heterogeneity of most networks (*e.g.*, scale-free), our averaging in Eq. (1.13) is limited to nodes in $\mathbf{G}(S)$, namely nodes of equal weighted degree. Within this group we expect the activities $\{x_i(t) \mid i \in \mathbf{G}(S)\}$ to be rather homogeneous, following a bounded distribution $P_S(x_i)$. This is especially relevant under the configuration model, in which all nodes of equal degree witness a similar surrounding neighborhood. Such homogeneity enables the *degree-based* mean-field offered in (1.13).
- *Sub-linearity or saturation.* The homogeneity of $x_i(t) \in \mathbf{G}(S)$ is further strengthened in case $M_q(x)$ is sub-linear, *e.g.*, a saturating function $(\mathbb{R}_1, \mathbb{R}_2, \mathbb{M}, \mathbb{N})$ or a fractional exponent $(\mathbb{R}_2, \mathbb{P})$. Indeed, subjecting a distributed x_i to a sub-linear or saturating function $f(x_i)$ reduces the spread of the resulting random variable. For example if x_i is spread across, say, two orders of magnitude, then the sub-linear $\sqrt{x_i}$ spreads only across one order of magnitude, and the saturating $\mathcal{H}(x_i)$ is bounded between zero and unity. While not guaranteed, such functional forms are often encountered in many of the most frequently used dynamic models.

Therefore, despite the approximate nature of our derivation, it is no surprise that the numerical validation fully supports its predicted outcomes.

We can now use (1.9) to express the sum on the r.h.s. of (1.12) as

$$\frac{1}{|\mathbf{G}(S)|}\sum_{i\in\mathbf{G}(S)}\sum_{n=1}^{N}A_{in}M_2\big(x_n(t)\big) = S\left\langle M_2\big(x(t)\big)\right\rangle_{\odot}, \tag{1.14}$$

where $\langle M_2(x(t))\rangle_{\odot}$, an average over all nearest neighbor nodes in the network, is independent of S. Equation (1.12) then takes the form

$$\frac{\mathrm{d}x(S,t)}{\mathrm{d}t} = M_0\big(x(S,t)\big) + SM_1\big(x(S,t)\big)\left\langle M_2(x(t))\right\rangle_{\odot}. \tag{1.15}$$

To obtain the steady state $x(S)$ (no t) we set the l.h.s. of (1.15) to zero, providing

$$R\big(x(S)\big) = \frac{1}{\langle M_2(x)\rangle_{\odot}\, S}, \tag{1.16}$$

where

$$R(x) = -\frac{M_1(x)}{M_0(x)}. \tag{1.17}$$

Extracting $x(S)$ from (1.16) we write

$$x(S) \sim R^{-1}(\lambda), \tag{1.18}$$

where $R^{-1}(x)$ is the inverse function of $R(x)$ and

$$\lambda = \frac{1}{\langle M_2(x)\rangle_{\odot} S} \propto S^{-1} \tag{1.19}$$

is (proportional to) the inverse weighted degree. Equation (1.18) expresses the average steady-state activity over all nodes with in-degree S ($i \in \mathbf{G}(S)$) in function of their inverted degree $\lambda \propto S^{-1}$.

1.3 The scaling of τ_i

We now calculate the response time τ_i of a node to a neighboring perturbation. Hence, we induce a small permanent perturbation Δx_m on the steady state activity of node m, a nearest neighbor of i, setting

$$x_m(t) = x_m + \Delta x_m. \tag{1.20}$$

The dynamic equation (1.1) then becomes

$$\begin{aligned} \frac{\mathrm{d}(x_i + \Delta x_i)}{\mathrm{d}t} &= M_0\big(x_i + \Delta x_i(t)\big) + \sum_{\substack{j=1 \\ j\neq m}}^{N} A_{ij} M_1\big(x_i + \Delta x_i(t)\big) M_2\big(x_j + \Delta x_j(t)\big) \\ &+ A_{im} M_1\big(x_i + \Delta x_i(t)\big) M_2\big(x_m + \Delta x_m\big), \end{aligned} \tag{1.21}$$

where $\Delta x_i(t)$ and $\Delta x_j(t)$ ($j = 1, \ldots, N, j \neq m$) are all time dependent, while Δx_m is constant; the terms x_i, x_j and x_m represent time-independent steady-state activities. Linearizing around the steady state we obtain

$$\begin{aligned} \frac{\mathrm{d}\Delta x_i}{\mathrm{d}t} &= \left(M_0'(x_i) + M_1'(x_i) \sum_{j=1}^{N} A_{ij} M_2(x_j) \right) \Delta x_i(t) \\ &+ M_1(x_i) \sum_{j=1}^{N} A_{ij} M_2'(x_j) \Delta x_j(t) + O(\Delta x^2), \end{aligned} \tag{1.22}$$

where $M_q'(x)$ ($q = 0, 1, 2$) represents the derivative $\mathrm{d}M_q/\,\mathrm{d}x$ with x taken at the steady-state, which according to (1.18) can be expressed by $x = R^{-1}(\lambda)$. Next, following a similar derivation as the one leading to (1.15), we average of over all nodes in $\mathbf{G}(S)$ to obtain a direct equation for the response of nodes with weighted degree $S_i \in (S, S + \mathrm{d}S)$, namely we seek the dynamic equation for

$$\Delta x(S, t) = \frac{1}{|\mathbf{G}(S)|} \sum_{i \in \mathbf{G}(S)} \Delta x_i(t). \tag{1.23}$$

Using (1.22) to express the time derivative of $\Delta x_i(t)$ in (1.23) and neglecting super-linear terms $O(\Delta x^2)$, we obtain

$$\begin{aligned}\frac{\mathrm{d}\Delta x(S,t)}{\mathrm{d}t} &= \left(M_0'\big(x(S)\big) + M_1'\big(x(S)\big)\frac{1}{|\mathbf{G}(S)|}\sum_{i\in\mathbf{G}(S)}\sum_{j=1}^{N}A_{ij}M_2(x_j)\right)\Delta x(S,t)\\ &+ M_1\big(x(S)\big)\frac{1}{|\mathbf{G}(S)|}\sum_{i\in\mathbf{G}(S)}\sum_{j=1}^{N}A_{ij}M_2'(x_j)\Delta x_j(t),\end{aligned} \tag{1.24}$$

where $x(S)$ is the steady state activity of nodes in $\mathbf{G}(S)$, as expressed in (1.18). Finally, the configuration model assumption, allows us to simplify the first sum on the r.h.s. using (1.9), providing us with

$$\frac{\mathrm{d}\Delta x(S,t)}{\mathrm{d}t} = \Big(M_0'\big(x(S)\big) + SM_1'\big(x(S)\big)\left\langle M_2(x)\right\rangle_{\odot}\Big)\Delta x(S,t) + f(S,t), \tag{1.25}$$

where

$$f(S,t) = M_1\big(x(S)\big)\frac{1}{|\mathbf{G}(S)|}\sum_{i\in\mathbf{G}(S)}\sum_{j=1}^{N}A_{ij}M_2'(x_j)\Delta x_j(t). \tag{1.26}$$

Equation (1.25) can be written in the form

$$\frac{\mathrm{d}\Delta x(S,t)}{\mathrm{d}t} = -\frac{1}{\tau(S)}\Delta x + f(S,t), \tag{1.27}$$

in which the average relaxation time $\tau(S)$ follows

$$\frac{1}{\tau(S)} = -M_0'\big(x(S)\big) - SM_1'\big(x(S)\big)\left\langle M_2(x)\right\rangle_{\odot}. \tag{1.28}$$

Equation (1.27) is a non-homogeneous linear differential equation, describing the average time dependent response $\Delta x(S,t)$ of nodes in $\mathbf{G}(S)$ to a neighboring permanent perturbation Δx_m. Its solution takes the form

$$\Delta x(S,t) = Ce^{-\frac{t}{\tau(S)}} + e^{-\frac{t}{\tau(S)}}\int_0^t f(S,t')e^{\frac{t'}{\tau(S)}}\Delta t', \tag{1.29}$$

where the constant C is set to zero to satisfy the initial condition $\Delta x(S,t=0)=0$. The relaxation of $\Delta x(S,t)$ (1.29) to its final, perturbed, state is governed by $\tau(S)$ (1.28), which depends on the weighted degree S, both explicitly, and implicitly through $x(S)$ in (1.18). To observe this we focus on each of the two terms on the r.h.s. of (1.28) independently. First we write

$$M_0'\big(x(S)\big) = \left.\frac{\mathrm{d}M_0}{\mathrm{d}x}\right|_{x=R^{-1}(\lambda)}, \tag{1.30}$$

a derivative around the steady state $x(S)$, which we expressed using (1.18). Using the

definition of $R(x)$ (1.17) we further develop (1.30), writing it as

$$\begin{aligned} M_0'\big(x(S)\big) &= \left(-\frac{M_1'(x)}{R(x)}+\frac{M_1(x)}{R^2(x)}R'(x)\right)\Bigg|_{x=R^{-1}(\lambda)} \\ &= -\frac{M_1'\big(R^{-1}(\lambda)\big)}{\lambda}+\frac{M_1\big(R^{-1}(\lambda)\big)R'\big(R^{-1}(\lambda)\big)}{\lambda^2}, \end{aligned} \tag{1.31}$$

where in the last step we used the fact that $R(R^{-1}(\lambda))=\lambda$. In a similar fashion we express the second term of (1.28) as

$$SM_1'\big(x(S)\big)\left\langle M_2(x)\right\rangle_{\odot}=\left\langle M_2(x)\right\rangle_{\odot}\frac{M_1'\big(R^{-1}(\lambda)\big)}{\lambda}. \tag{1.32}$$

Collecting all the terms we arrive at

$$\frac{1}{\tau(S)}=c_1\frac{M_1'\big(R^{-1}(\lambda)\big)}{\lambda}+c_2\frac{M_1\big(R^{-1}(\lambda)\big)R'\big(R^{-1}(\lambda)\big)}{\lambda^2}, \tag{1.33}$$

where the coefficients are

$$\begin{aligned} c_1 &= 1-\left\langle M_2(x)\right\rangle_{\odot} \\ c_2 &= -1. \end{aligned} \tag{1.34}$$

As we are only interested in the asymptotic scaling of $\tau(S)$ with S (or λ) in the limit of large S (small λ), we can rewrite (1.33) without the coefficients. Indeed, for sufficiently large S, only the leading terms where S is raised to the highest power dominate the equation, providing $1/\tau(S)\sim c_1S^a+c_2S^b\sim S^{\max(a,b)}$, independent of c_1 and c_2. Hence, preserving only the terms relevant to the scaling, Eq. (1.33) becomes

$$\frac{1}{\tau(S)}\sim\frac{1}{\lambda^2}\Big[R\left(R^{-1}(\lambda)\right)M_1'\left(R^{-1}(\lambda)\right)+M_1\left(R^{-1}(\lambda)\right)R'\left(R^{-1}(\lambda)\right)\Big], \tag{1.35}$$

where, once again, we used the fact that $\lambda=R(R^{-1}(\lambda))$, allowing us to extract the pre-factor of λ^{-2}, preceding the square brackets. Interestingly, the terms within the square brackets can be written in derivative form as $\mathrm{d}[M_1(x)R(x)]/\,\mathrm{d}x$, taken at $x=R^{-1}(\lambda)$. Therefore we can further simplify (1.35) to the form

$$\frac{1}{\tau(S)}\sim\frac{1}{\lambda^2}\frac{\mathrm{d}\big(M_1(x)R(x)\big)}{\mathrm{d}x}\Bigg|_{x=R^{-1}(\lambda)}, \tag{1.36}$$

providing us with

$$\tau(S)\sim\lambda^2Y\left(R^{-1}(\lambda)\right) \tag{1.37}$$

where

$$Y(x)=\left(\frac{\mathrm{d}(M_1R)}{\mathrm{d}x}\right)^{-1}. \tag{1.38}$$

Equation (1.37) expresses $\tau(S)$ as a function of λ (1.19), from which its dependence on S can be obtained. It indicates that the scaling of $\tau(S)$ with S is determined directly by the dynamical functions $M_1(x)$ and $R(x)$, or, using (1.17), $M_1(x)$ and $M_0(x)$. Next we express $Y(R^{-1}(\lambda))$ as a Hahn series [3] around $\lambda = 0$

$$Y\big(R^{-1}(\lambda)\big) = \sum_{n=0}^{\infty} C_n \lambda^{\Gamma(n)}, \tag{1.39}$$

allowing us to systematically consider the asymptotic behavior at $S \to \infty$, equivalent to $\lambda \to 0$. The Hahn series is a generalization of the Taylor expansion to allow for both negative and real powers, as represented by $\Gamma(n)$, a countable set of real numbers, ordered such that $\Gamma(n-1) < \Gamma(n) < \Gamma(n+1)$. Hence the leading power of (1.39) is $\Gamma(0)$, the next leading power is $\Gamma(1)$, etc. For large S we only keep the leading order term, namely $\lambda^{\Gamma(0)}$. Using (1.37) this provides us with

$$\tau(\lambda) \sim \lambda^2 \lambda^{\Gamma(0)}, \tag{1.40}$$

or, substituting S^{-1} for λ,

$$\tau(S) \sim S^{\theta}, \tag{1.41}$$

where

$$\theta = -2 - \Gamma(0), \tag{1.42}$$

as presented in Eqs. (4) - (6) in the main paper text.

2 Classification of the dynamic models

We analyzed the propagation patterns in six different frequently used dynamic models, for each obtaining θ (1.42), and hence their class as distance limited ($\theta = 0$), degree limited ($\theta > 0$) or composite ($\theta < 0$). The detailed derivations appear below.

2.1 Regulatory dynamics - $\mathbb{R}_1$ and $\mathbb{R}_2$

Gene regulation is often modeled using Michaelis-Menten dynamics, in which the activity, *i.e.* expression, of all genes follows [4, 5],

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -Bx_i^a(t) + \sum_{j=1}^{N} A_{ij}\mathcal{H}\big(x_j(t)\big), \tag{2.1}$$

where $\mathcal{H}(x)$ is the Hill function describing the activation/inhibition of $x_i(t)$ by its neighbor's activity $x_j(t)$. Since regulation depends primarily on the presence or absence of the regulator j, with little sensitivity to j's specific abundance, the Hill function is designed to be a *switch-like* function satisfying $\mathcal{H}(x) \to 1$ ($\mathcal{H}(x) \to 0$) for large (small) x in case of activation, or $\mathcal{H}(x) \to 1$ ($\mathcal{H}(x) \to 0$) for small (large) x in the case of inhibition. A

most common choice is [4, 5]

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -Bx_i^a(t) + \sum_{j=1}^{N} A_{ij} \frac{x_j^h(t)}{1 + x_j^h(t)}, \tag{2.2}$$

where the Hill coefficient h governs the rate of saturation of $\mathcal{H}(x)$. Equation (2.2) can be cast in the form (1.1) with $M_0(x) = -Bx^a$, $M_1(x) = 1$ and $M_2(x) = x^h/(1 + x^h)$. Hence Eq. (1.17) provides $R(x) = -B^{-1}x^{-a}$, and its inverse follows

$$R^{-1}(x) = B^{-\frac{1}{a}} x^{-\frac{1}{a}} \propto x^{-\frac{1}{a}}. \tag{2.3}$$

Next we use (1.38) to write

$$Y(x) = \left(\frac{\mathrm{d}M_1 R}{\mathrm{d}x} \right)^{-1}, \tag{2.4}$$

which taking $M_1(x) = 1$ and the above $R(x)$ becomes

$$Y(x) = \left(\frac{\mathrm{d}x^{-a}}{\mathrm{d}x} \right)^{-1} \sim x^{a+1}. \tag{2.5}$$

Using (2.3) in (2.5) we arrive at the Hahn expansion of (1.39)

$$Y\left(R^{-1}(\lambda)\right) \sim Y\left(\lambda^{-\frac{1}{a}}\right) \sim \lambda^{-\frac{a+1}{a}}, \tag{2.6}$$

whose leading (indeed, only) power is $\Gamma(0) = -(a + 1)/a$. Finally, we predict θ from (1.42) to be

$$\theta = -2 - \Gamma(0) = -2 + \frac{a+1}{a} = \frac{1-a}{a}. \tag{2.7}$$

For $\mathbb{R}_1$ we set $a = 1$ and $h = 1$, predicting $\theta = 0$, a distance limited dynamics; for $\mathbb{R}_2$ we set $a = 0.4$ and $h = 0.2$, predicting $\theta = 3/2$, a degree limited system. Both predictions are perfectly confirmed on both model and real networks in Figs. 3 and 4 of the main text.

2.2 Population dynamics - $\mathbb{P}$

Birth-death processes have many applications in population dynamics [6], queuing theory [7] or biology [6]. We consider a network in which the nodes represent sites, each site i having a population $x_i(t)$, with population flow enabled between neighboring sites. This process can be described by

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -Bx_i^b(t) + \sum_{j=1}^{N} A_{ij} x_j^a(t), \tag{2.8}$$

where the first term on the r.h.s. represents the internal dynamics of site i, characterized by the exponent b, which distinguishes between processes such as [8] in/out flux ($b = 0$),

mortality ($b = 1$), pairwise annihilation ($b = 2$), etc.. The second term describes the nonlinear flow from i's neighboring site j into i. Here we have $M_0(x) = -Bx^b$, $M_1(x) = 1$ and $M_2(x) = x^a$, therefore $R(x) = -B^{-1}x^{-b}$. Following the same steps leading from (2.3) to (2.6) we find

$$Y\big(R^{-1}(\lambda)\big) \sim \lambda^{-\frac{b+1}{b}}, \tag{2.9}$$

predicting

$$\theta = -2 - \Gamma(0) = -2 + \frac{b+1}{b}. \tag{2.10}$$

In $\mathbb{P}$ we set $b = 0.5$ and $a = 0.2$, predicting $\theta = 1$, a degree limited system, as confirmed in Figs. 3 and 4 of the main text.

The dynamic exponents a, b. In both $\mathbb{R}$ and $\mathbb{P}$ we observe equation terms of the form x^α, where α takes either integer values, *e.g.*, 1 in $\mathbb{R}_1$, or fractional values, *i.e.* $0.2, 0.4$ and 0.5 in $\mathbb{R}_2$ and $\mathbb{P}$. These exponents, as opposed to the rate constant B or the weights A_{ij}, are intrinsic to the dynamic mechanisms driving the system components. For example $\alpha = 1$ represents a first-order reaction, such as mortality/degradation or duplication, in which a single copy is depleted/added to the system; $\alpha = 2$ captures a second order process, such competition or dimerization, in which the interaction rate is driven by the number of pairs of reacting individuals/molecules. Such distinctions are intrinsic to the *physics* of the interacting components, and are thus not subject to perturbations or sensitive to environmental conditions. Therefore the fact the θ is determined by these exponents indicates that it is a robust observable, ingrained in the system's intrinsic dynamics.

A non-integer α is often encountered under fractional-order reactions, which arise when mechanisms of different order combine to form chain-reactions. A classic example is given by the catalytic chain reaction

$$\begin{aligned} A + x &\xrightarrow{k_1} 2A + x^\star \\ \beta A &\xrightarrow{k_2} \emptyset, \end{aligned} \tag{2.11}$$

under $k_2 \gg k_1$. The kinetic equations follow

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= -k_1 A x \\ \frac{\mathrm{d}A}{\mathrm{d}t} &= k_1 A x - k_2 A, \end{aligned} \tag{2.12}$$

where, due to the time-scale separation of k_2 vs. k_1 we take A to be at quasi-steady-state, setting $\mathrm{d}A/\,\mathrm{d}t = 0$. This provides $A \propto -x^{1/(\beta-1)}$, which leads, in the first equation to

$$\frac{\mathrm{d}x}{\mathrm{d}t} \propto -x^\alpha, 0 \tag{2.13}$$

where $\alpha = \beta/(\beta-1)$, an effective non-integer order reaction. More complex chain reactions can lead to equation terms with any desired order. For most real systems the effective order of a reaction is determined experimentally, and terms of the form $\dot{x} = -x^{\alpha} + \cdots$, with $0 < \alpha < 1$, are quite common in biochemical settings.

2.3 Epidemics - $\mathbb{E}$

In the susceptible-infected-susceptible (SIS) model, each node may be in one of two potential states: infected (I) or susceptible (S). The spreading dynamics is driven by the two process

$$I + S \to 2I, \tag{2.14}$$

where a susceptible node becomes infected through contact with one of its infected neighbors, and

$$I \to S, \tag{2.15}$$

an infected node recovering and becoming susceptible again. The activity $x_i(t)$ denotes the probability that i is in the infected state. The infection and recovery processes above are captured by [9]

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -Bx_i + \sum_{j=1}^{N} A_{ij}\big(1 - x_i(t)\big)x_j(t). \tag{2.16}$$

The first term on the r.h.s. accounts for the process of recovery, at a rate B, and the second term accounts for the process of infection, where a node could only become infected if it is in the susceptible state, with probability $1 - x_i(t)$, and its neighbor is in the infected state, with probability $x_j(t)$. We have $M_0(x) = -Bx$, $M_1(x) = 1 - x$ and $M_2(x) = x$, providing (1.17)

$$R(x) = \frac{1-x}{Bx}, \tag{2.17}$$

and therefore

$$R^{-1}(x) = \frac{1}{1+Bx}. \tag{2.18}$$

Equation (1.38) takes the form

$$Y(x) = \left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{(1-x)^2}{x}\right)\right)^{-1} \propto \frac{1}{1-x^{-2}}, \tag{2.19}$$

allowing us to obtain the Hahn expansion (1.39) as

$$Y\big(R^{-1}(\lambda)\big) = Y\left(\frac{1}{1+B\lambda}\right) = \frac{1}{1-(1+B\lambda)^2} = \frac{1}{2B}\lambda^{-1} + \frac{1}{4}\lambda^0 + \frac{1}{8}B\lambda^1 + O(\lambda^2), \tag{2.20}$$

whose leading power is $\Gamma(0) = -1$. Using (1.42) this predicts

$$\theta = -2 - \Gamma(0) = -2 - (-1) = -1, \tag{2.21}$$

a composite dynamics, in which hubs respond most rapidly (Figs. 3 and 4 in main paper).

2.4 Mutualistic dynamics in ecology - $\mathbb{M}$

We consider symbiotic eco-systems, such as plant-pollinator networks, in which the interacting species exhibit mutualistic relationships. The species populations follow the dynamic equation

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = Bx_i(t)\left(1 - \frac{x_i^a(t)}{C}\right) + \sum_{j=1}^{N} A_{ij}x_i(t)F\big(x_j(t)\big). \tag{2.22}$$

The self dynamics

$$M_0(x) = Bx\left(1 - \frac{x^a}{C}\right) \tag{2.23}$$

is a generalization of the frequently used logistic growth: when the population is small, the species reproduce at a rate B, yet, as x_i approaches the carrying capacity of the system C, growth is hindered by competition over limited resources [10], captured by the nonlinear $-x_i^{a+1}$ term. For $a = 1$ we arrive at the classic quadratic growth deficiency term, in which competition scales with the number of competing *pairs*. In case $a > 1$ growth is hindered through higher order internal competition within a species.

The mutualistic inter-species interactions are captured by

$$\begin{aligned} M_1(x) &= x \\ M_2(x) &= F(x), \end{aligned} \tag{2.24}$$

where $F(x)$ represents the *functional response*, describing the positive impact that species j has on species i. This functional response can take one of several forms [11]: **Type I:** Linear impact

$$F(x) = \alpha x. \tag{2.25}$$

Type II: Saturating impact

$$F(x) = \frac{\alpha x}{1 + \alpha x}. \tag{2.26}$$

Type III: A generalization of Type II, where

$$F(x) = \frac{\alpha x^h}{1 + \alpha x^h}. \tag{2.27}$$

In our simulations we used Type II mutualistic interactions and set the competition coefficient to $a = 2$, providing

$$\begin{aligned} M_0(x) &= Bx\left(1 - \frac{x^2}{C}\right) \\ M_1(x) &= x \\ M_2(x) &= \frac{\alpha x}{1 + \alpha x}, \end{aligned} \tag{2.28}$$

where, for simplicity, we set $B = C = \alpha = 1$. Hence we have (1.17)

$$R(x) = \frac{1}{x^2 - 1}, \tag{2.29}$$

and therefore

$$R^{-1}(x) = \left(\frac{x+1}{x}\right)^{\frac{1}{2}}. \tag{2.30}$$

Next we use (1.38) to write

$$Y(x) = \left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x}{x^2-1}\right)\right)^{-1} = -\frac{\left(x^2-1\right)^2}{1+x^2}. \tag{2.31}$$

Consequently, the Hahn expansion (1.39) takes the form

$$Y\left(R^{-1}(\lambda)\right) = Y\left(\left(\frac{\lambda+1}{\lambda}\right)^{\frac{1}{2}}\right) = \frac{1}{\lambda + 2\lambda^2} = \lambda^{-1} - 2\lambda^0 + 4\lambda^1 + O(\lambda^2), \tag{2.32}$$

for which the leading power $\Gamma(0) = -1$. As a result we predict (1.42)

$$\theta = -2 - \Gamma(0) = -1, \tag{2.33}$$

classifying $\mathbb{M}$ in the composite dynamics class, as fully confirmed by the results presented in Figs. 3 and 4 of the main paper.

2.5 Excitation dynamics between brain regions - $\mathbb{N}$

Brain regions, comprising a macroscopic ensemble of neurons each, interact through neuronal bundels in the connectome to activate each other. The dynamics of such excitation interactions have been offered as [12]

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -Bx_i(t) + C\tanh x_i(t) + \sum_{j=1}^{N} A_{ij}\tanh x_j(t), \tag{2.34}$$

where the first term captures the relaxation of node i, at a rate B, the second term captures the self-activation between neurons within region i, characterized by the rate C, and the interaction term sums over the impact of i's network neighbors. Here we set $B = 2$ and $C = 2.5$.

In (2.34) we have $M_0(x) = -Bx + C\tanh x$, $M_1(x) = 1$ and $M_2(x) = \tanh x$, providing via (1.17)

$$R(x) = \frac{1}{Bx - C\tanh x}. \tag{2.35}$$

We cannot explicitly invert (2.35), however we can approximate its inverse in the $\lambda \to 0$ limit, which is all we need to evaluate θ. In the limit of large x, we have $R(x) \to 0$,

therefore the inverse function $R^{-1}(\lambda)$, taken at $\lambda \to 0$, is governed mainly by the large x limit. Indeed inverting $R(x)$ is equivalent to switching the R-axis and the x-axis. Hence the inverse function around zero, $R^{-1}(x \to 0)$ is the mirror image of the segment in the original function where $R(x) \to 0$. The latter only occurs when $x \to \infty$. In this limit we can approximate $\tanh x \approx 1$, enabling us to invert (2.35) as

$$R^{-1}(\lambda) = \frac{C}{B} + \frac{1}{B}\lambda^{-1}. \tag{2.36}$$

Equation (1.38) takes the form

$$Y(x) = \left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{Bx - C}\right)\right)^{-1} = \frac{(Bx - C)^2}{B}. \tag{2.37}$$

Substituting $R^{-1}(\lambda)$ in (2.36) in place of x, we obtain the Hahn expansion (1.39) as

$$Y\big(R^{-1}(\lambda)\big) = \frac{1}{B}\lambda^{-2}, \tag{2.38}$$

with the leading power of $\Gamma(0) = -2$. Hence (1.42) predicts

$$\theta = -2 - \Gamma(0) = -2 - (-2) = 0, \tag{2.39}$$

a distance limited propagation, as observed in Figs. 3 and 4 in main text.

Model	Dynamic equation	θ	Class
Regulatory Dynamics ($\mathbb{R}_1$)	$\frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} A_{ij} \frac{x_j}{1+x_j}$	0	Distance Limited
Neuronal Dynamics ($\mathbb{N}$)	$\frac{dx_i}{dt} = -Bx_i + C\tanh x_i + \sum_{j=1}^{N} A_{ij} \tanh x_j$	0	Distance Limited
Regulatory Dynamics ($\mathbb{R}_2$)	$\frac{dx_i}{dt} = -x_i^{0.4} + \sum_{j=1}^{N} A_{ij} \frac{x_j^{0.2}}{1+x_j^{0.2}}$	$\frac{3}{2}$	Degree Limited
Population Dynamics ($\mathbb{P}$)	$\frac{dx_i}{dt} = -x_i^{0.5} + \sum_{j=1}^{N} A_{ij}\, x_j^{0.2}$	1	Degree Limited
Epidemic Dynamics ($\mathbb{E}$)	$\frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} A_{ij}\,(1-x_i)\,x_j$	-1	Composite
Mutualistic Dynamics ($\mathbb{M}$)	$\frac{dx_i}{dt} = x_i\,(1-x_i^2) + \sum_{j=1}^{N} A_{ij} \frac{x_i x_j}{1+x_j}$	-1	Composite

Table 1: **Dynamic models.** Summary and classification of all dynamic models.

3 Methods and data analysis

3.1 Numerical integration

To numerically test our predictions we constructed Eq. (1.1) for each of the systems in Table 1, using the appropriate A_{ij} (Scale-free, Erdős-Rényi, empirical, etc.). We then used a fourth-order Runge-Kutta stepper (Matlab's ode45) to numerically solve the resulting equations. Starting from an arbitrary initial condition $x_i(t=0)$, $i = 1, \ldots, N$ we allowed the system to reach steady-state by waiting for $\dot{x}_i \to 0$. To numerically realize this limit we implemented the termination condition

$$\max_{i=1}^{N} \left| \frac{x_i(t_n) - x_i(t_{n-1})}{x_i(t_n)\Delta t_n} \right| < \varepsilon, \tag{3.1}$$

where t_n is the time stamp of the nth Runge-Kutta step and $\Delta t_n = t_n - t_{n-1}$. As the system approaches a steady-state, the activities $x_i(t_n)$ become almost independent of time, and the numerical derivative $\dot{x}_i = x_i(t_n) - x_i(t_{n-1})/\Delta t_n$ becomes small compared to $x_i(t_n)$. The condition (3.1) guarantees that the maximum of $\dot{x}_i/x_i$ over all activities $x_i(t_n)$ is smaller than the pre-defined termination variable ε. In our simulations, across the different dynamics we tested, we set $\varepsilon \leq 10^{-12}$, a rather strict condition, to ensure that our system is sufficiently close to the *true* steady-state.

3.2 Measuring $T(j \to i)$ and τ_i

To observed the spatio-temporal propagation of a perturbation we set the initial condition of the system to its numerically obtained steady-state above. We then introduce a boundary condition on the source node j, as

$$x_j(t) = x_j + \Delta x_j, \tag{3.2}$$

a signal in the form of a permanent perturbation to j's steady state activity x_j. In our simulated results we used $\Delta x_j = \alpha x_j$, setting $\alpha = 0.1$, a 10% activity perturbation. The remaining $N-1$ nodes continue to follow the original dynamics (1.1), responding to the signal Δx_j. To be explicit, we simulate this propagating perturbation by numerically solving the *perturbed* Eq. (1.1), which now takes the form

$$\begin{cases} \dfrac{\mathrm{d}x_j}{\mathrm{d}t} &= 0 \\ \dfrac{\mathrm{d}x_i}{\mathrm{d}t} &= M_0\big(x_i(t)\big) + \displaystyle\sum_{n=1}^{N} A_{in} M_1\big(x_i(t)\big) M_2\big(x_n(t)\big) \quad i \neq j \end{cases}, \tag{3.3}$$

in which the perturbation on j is held constant in time, and the remaining $N-1$ nodes respond via the system's intrinsic dynamics. The system's response is then obtained as

$$x_i(t) = x_i + \Delta x_{ij}(t), \tag{3.4}$$

in which $\Delta x_{ij}(t)$ represents i's temporal response to the signal Δx_j. We continue running (3.3) until the termination condition (3.1) is realized again, and the system reaches its new perturbed state with $\Delta x_{ij}(t \to \infty) = \Delta x_{ij}$, i's final response to j's signal. To focus on the response time of each node, we define i's normalized response as

$$f_{ij}(t) = \frac{\Delta x_{ij}(t)}{\Delta x_{ij}}, \tag{3.5}$$

which transitions smoothly between $f_{ij}(t) = 0$ at $t = 0$ to $f_{ij}(t) = 1$ at $t \to \infty$, as i approaches its final perturbed state. The function $f_{ij}(t)$ captures the spatio-temporal response of the system in the discrete network space, namely the level of response obtained at time t in location i, as a result of perturbation j. When $f_{ij}(t) = \eta$, we say that i has reached an η-fraction of its final response to the traveling signal Δx_j. For instance, setting $\eta = 1/2$ allows us to evaluate the *half-life* of i's response. We then evaluate the propagation time $T(j \to i)$ as the time when

$$f_{ij}\big(t = T(j \to i)\big) = \eta \tag{3.6}$$

or alternatively

$$T(j \to i) = f_{ij}^{-1}(\eta). \tag{3.7}$$

The parameter η can be set to any value between zero and unity, $\eta \in (0, 1)$, with the typical choice being of order $\eta \sim 1/2$. All results presented in the main paper were obtained for $\eta = 0.7$, however, as we show in Fig. 1 changing the value of η has no detectable effect on the observed behavior of τ_i and hence of $T(j \to i)$.

Local response. To obtain the local response τ_i we must measure the response time to a signal in the direct vicinity of i, namely $T(j \to i)$ where j is directly linked to i ($A_{ij} > 0$). Hence we denote by K_i the group of incoming neighbors of i

$$K_i = \{j = 1, \ldots, N | A_{ij} \neq 0\} \tag{3.8}$$

and average i's response time to these neighbors as

$$\tau_i = \frac{1}{|K_i|} \sum_{j \in K_i} T(j \to i), \tag{3.9}$$

where $|K_i|$ is the number of nodes in K_i.

Visualizing the spatio-temporal spread. To construct the visualizations of Fig. 3a - f in the main paper we used Gephi [13] to layout the weighted scale-free network SF, placing the source node j at the center. As the propagation unfolds we measured the response of all nodes $i = 1, \ldots, N$, setting the size and color depth of each node to be linearly proportional to $f_{ij}(t)$ (3.5). Later, in Fig. 6 of the main text we present the exact same data only this time we laid out the nodes according to our universal metric $\mathcal{L}(j \to i)$, as described in Eq. (7) of the main text. Hence we located j at the center as before ($\mathcal{L}(j \to j) = 0$), and then placed all target nodes i at a radial distance r proportional to $\mathcal{L}(j \to i)$, while randomly selecting the azimuth $\varphi \in [0, 2\pi]$. In the case of degree-driven propagation ($\theta > 0$, red) since $\mathcal{L}(j \to i)$ (and $T(j \to i)$) span several orders of magnitude

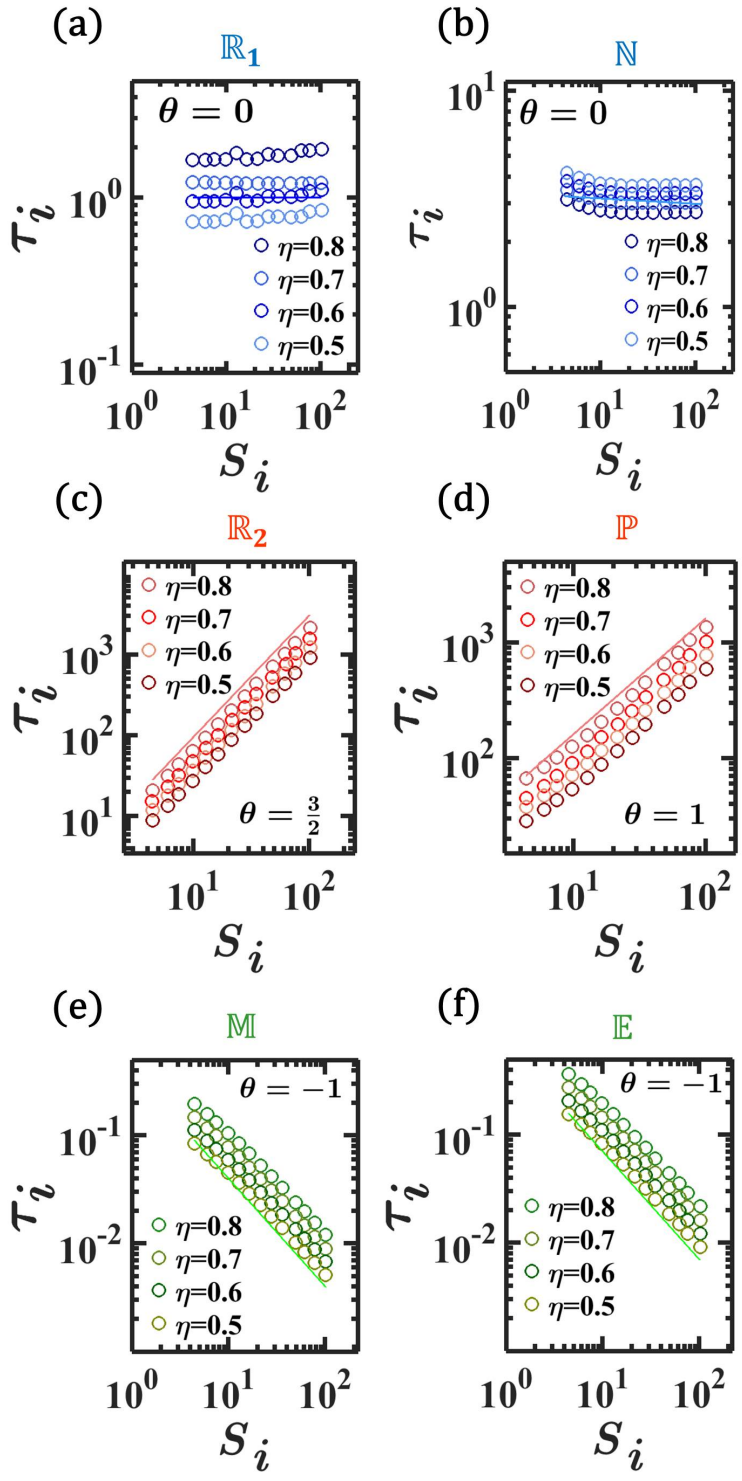

Figure 1: **The impact of the arbitrary parameter** η. Measuring the response times requires to select the value of η in (3.7), an arbitrary parameter between zero and unity. To test the impact of this parameter we tested our results for the scaling θ (1.41) under different selected values of η. As expected, we find that our results are not affected by the selected η.

we set $r \propto \ln \mathcal{L}(j \to i)$.

3.3 Logarithmic binning

The scaling $\tau_i \sim S_i^{\theta}$ is shown in log-scale in Fig. 4 of the main paper, with θ captured by the linear slope of τ_i vs. S_i. To construct these plots we employed logarithmic binning [14]. First we divide all nodes into W bins

$$\mathcal{B}(w) = \left\{ i = 1, \ldots, N \middle| c^{w-1} < S_i \leq c^{w} \right\}, \tag{3.10}$$

where $w = 1, ..., W$ and c is a constant. In (3.10) the wth bin includes all nodes i whose weighted degrees S_i are between c^{w-1} and c^{w}. The parameter c is selected such that the unity of all bins $\cup_{w=1}^{W} \mathcal{B}(w)$ includes all nodes, hence we set $c^{W} = \max S_i$. We then plot the average degree of the nodes in each bin

$$S_w = \langle S_i \rangle_{i \in \mathcal{B}(w)} = \frac{1}{|\mathcal{B}(w)|} \sum_{i \in \mathcal{B}(w)} S_i \tag{3.11}$$

versus the average response time of nodes in that bin

$$\tau_w = \langle \tau_i \rangle_{i \in \mathcal{B}(w)} = \frac{1}{|\mathcal{B}(w)|} \sum_{i \in \mathcal{B}(w)} \tau_i. \tag{3.12}$$

To evaluate the measurement error for each bin we first calculated the variance in the observed τ_i across all nodes in the bin $\sigma_w^2 = \langle \tau_i^2 \rangle_{i \in \mathcal{B}(w)} - \langle \tau_i \rangle_{i \in \mathcal{B}(w)}^2$. We then set the error-bar to represent the 95% confidence interval as [15]

$$E_w = \frac{1.96 \sigma_w}{\sqrt{|\mathcal{B}(w)|}}. \tag{3.13}$$

A similar scheme was used to present $\langle T(j \to i) \rangle$ vs. N in Fig. 4g - i and $T(j \to i)$ vs. $\mathcal{L}(j \to i)$ in Fig. 6h - j of the main paper. In most cases the error bars were tiny, smaller than the size of the plot markers, and hence, despite the fact that they are plotted, they do not show in the display.

3.4 Model and empirical networks

To test our predictions we used model and real networks, as summarized below:
ER. An Erdős-Rényi random network with $N = 6,000$ nodes and an average degree of $\langle S \rangle = 4$.
SF. A binary scale-free network with $N = 6,000$ nodes, $\langle S \rangle = 4$ and degree distribution following $P(S) \sim S^{-\gamma}$ with $\gamma = 3$, constructed using the Barabási-Albert model [16].
SF1. Using the underlying topology of SF we added uniformly distributed weights W_{ij}, extracted from the uniform distribution $W_{ij} \sim \mathcal{U}(0.1 - 0.9)$.
SF2. Using the underlying topology of SF we extracted the weights W_{ij} from a scale-free probability density function $P(w) \sim w^{-\nu}$ with $\nu = 3$, resulting in an extremely heterogeneous network, featuring both scale-free topology and scale-free weights.

Network	N	$\langle S \rangle$	$P(S)$	$P(w)$	Used Models
ER	6,000	4.71	Poisson	Binary	$\mathbb{R}_1, \mathbb{N}, \mathbb{R}_2, \mathbb{P}, \mathbb{M}, \mathbb{E}$
SF	6,000	3.99	SF, $\gamma = 3$	Binary	$\mathbb{R}_1, \mathbb{N}, \mathbb{R}_2, \mathbb{P}, \mathbb{M}, \mathbb{E}$
SF1	6,000	2.2	SF, $\gamma = 3$	Uniform	$\mathbb{R}_1, \mathbb{N}, \mathbb{R}_2, \mathbb{P}, \mathbb{M}, \mathbb{E}$
SF2	6,000	7.99	SF, $\gamma = 3$	SF, $\upsilon = 3$	$\mathbb{R}_1, \mathbb{N}, \mathbb{R}_2, \mathbb{P}, \mathbb{M}, \mathbb{E}$

Table 2: **Model networks.** Summary of model networks used to exemplify our formalism.

UCIonline. An instant messaging network from the University of California Irvine [17], capturing 61,040 transactions between 1,893 users during a $T = 218$ day period. Connecting all individuals who exchanged messages throughout the period, we obtain a network of 1,893 nodes with 27,670 links, exhibiting a fat-tailed degree distribution.
Email Epoch. This dataset monitors $\sim 3 \times 10^5$ emails exchanged between 3,185 individuals over the course of ~ 6 months [18], giving rise to a scale-free social network with 31,885 binary links.
Epinions. A binary online social network linking 467 individuals via 6,538 social ties [19].
ATN. The Advogato trust network, a symmetric social network constructed from connections within the community of open source developers, comprising 6,539 nodes and 23,540 links [20].
PPI1. The yeast scale-free protein-protein interaction network, consisting of 1,647 nodes (proteins) and 5,036 undirected links, representing chemical interactions between proteins [21].
PPI2. The human protein-protein interaction network, a scale-free network, consisting of $N = 2,035$ nodes (protein) and $L = 13,806$ protein-protein interaction links [22].
PPI3. Binary protein-protein interaction network of *Arabidopsis thaliana*, whose giant connected component comprises 2,938 nodes and 7,720 links [23].
PPI4. Multiplex genetic and protein interactions network of *Rattus norvegicus*, composed of 2,350 nodes and 3,484 links [24].
Brain. Mapping the physical fibre bundle connections between 998 brain regions, as measured using diffusion tensor imaging techniques. The weights represent empirically observed bundle densities [25].
ECO1 and ECO2. To construct the mutualistic ecological networks we collected data on symbiotic interactions of plants and pollinators in Carlinville Illinois from [26]. The resulting $456 \times 1,429$ network M_{ik} is a bipartite graph linking the 456 plants with their 1,429 pollinators. When a pair of plants is visited by the same pollinator they mutually benefit each other indirectly, by increasing the pollinator populations. Similarly pollinators sharing the same plants also posses an indirect mutualistic interaction. Hence we can collapse M_{ik} to construct two mutualistic networks: The $1,429 \times 1,429$ pollinator

Network	N	$\langle S \rangle$	$P(S)$	$P(w)$	Models
PPI1	1,647	3.05	SF, $\gamma = 2.7$	Binary	$\mathbb{R}_1, \mathbb{R}_2$
PPI2	2,035	6.78	SF, $\gamma = 2.3$	Binary	$\mathbb{R}_1, \mathbb{R}_2$
PPI3	2,938	5.25	SF, $\gamma = 2.5$	Binary	$\mathbb{R}_1, \mathbb{R}_2$
PPI4	2,350	2.96	SF, $\gamma = 2.9$	Binary	$\mathbb{R}_1, \mathbb{R}_2$
Brain	998	0.44	Bounded	Bounded	$\mathbb{N}$
ECO1	1,044	14.17	SF, $\gamma = 2.0$	SF, $\upsilon = 2.9$	$\mathbb{M}$,$\mathbb{P}$
ECO2	456	62.32	Bounded	SF, $\upsilon = 3.2$	$\mathbb{M}$,$\mathbb{P}$
UCIonline	1,893	14.61	SF, $\gamma = 2.1$	Binary	$\mathbb{E}$
Epoch	3,185	20.00	SF, $\gamma = 2.4$	Binary	$\mathbb{E}$
Epinions	467	28.54	Bounded	Binary	$\mathbb{E}$
ATN	6,539	12.01	SF, $\gamma = 2.2$	Binary	$\mathbb{E}$

Table 3: **Real networks.** We implemented our theory on a set of highly diverse empirical networks, including social, biological and ecological networks. For each network we characterize the weighted degree distribution $P(S)$, bounded versus scale-free (SF), and show the empirically extracted scaling exponent γ, where relevant. For the weighted networks ECO1/2 we also present the scaling exponent ν of $P(w)$. In BRAIN the weights represent empirically measured link densities from diffusion MRI. On each network we ran the relevant models, *e.g.*, epidemic spreading ($\mathbb{E}$) on the social networks.

network ECO1 and the 456×456 plant network ECO2. The resulting networks are

$$B_{kl} = \sum_{i=1}^{456} \frac{M_{ik} M_{il}}{\sum_{s=1}^{n} M_{is}}, \tag{3.14}$$

for the pollinator network (ECO1), and

$$A_{ij} = \sum_{k=1}^{1,429} \frac{M_{ik} M_{jk}}{\sum_{s=1}^{n} M_{sk}}, \tag{3.15}$$

for the plant network (ECO2). In both networks the numerator equals to the number of mutual plants (B_{kl}) or pollinators (A_{ij}). For each mutual plant i (pollinator k) we divide by the overall number of plants (pollinators) that share i (k). Hence, the weight of

the mutualistic interaction in, *e.g.*, A_{ij} is determined by the density of mutual symbiotic relationships between all plants, where: (i) the more mutual pollinators k that plants i and j share the stronger the mutualistic interaction between them; (ii) on the other hand the more plants pollinated by k the smaller is its contribution to each plant. A similar logic applies also for the pollinator network B_{ij}. This process potentially allows us to have isolated components, *e.g.*, single disconnected nodes. The state of these isolated nodes is decoupled from the state of the rest of the network, and hence in our analysis we only focused on the giant connected component of A_{ij} and B_{ij}, comprising all 456 plants, rendering A_{ij} to be a fully connected component, but only $1,044$ pollinators, eliminating 385 isolated pollinators in B_{ij}.

4 Complementary results from empirical networks

To complement the results presented in the main paper we include here observations extracted from our set of empirical networks (Sec. 3.4), comprising 17 combinations of networks and dynamics, as appear in Table 3. The scaling relationship $\tau_i \sim S_i^{\theta}$ has already been tested in the main text (Fig. 4) on all 41 systems, model and empirical, hence we focus below on the structure of $P(T)$ and on the layouts predicted by our universal metric $\mathcal{L}(j \to i)$ (Eq. (7) in main text).

4.1 Distance limited propagation

In Fig. 2a - e we show $P(T)$ vs. T (blue) as obtained for the distance limited $\mathbb{R}_1$ and $\mathbb{N}$ implemented on PPI1 - PPI4 ($\mathbb{R}_1$) and on Brain ($\mathbb{N}$). As predicted for these distance limited dynamics, $P(T)$ exhibits separated sharp peaks, corresponding to the discrete lengths of all paths L_{ij}. A slight overlap is observed for $\mathbb{N}$ (Fig. 2e), where the peaks, while clearly separated, show a smoother pattern than in ER, SF1 - 3 and PPI1 - 4. This is due to Brain being a weighted network, allowing some level of heterogeneity within each L_{ij}-shell as a consequence of the distributed weights across all links. We further tested our universal dynamic metric $\mathcal{L}(j \to i)$, as predicted in Eq. (7) of the main text. The results, presented in Fig. 4 confirm that indeed, these five systems all exhibit distance limited propagation, expressed through the discrete shells characterizing the traveling signals.

4.2 Degree limited propagation

Our testing ground includes six degree limited systems: $\mathbb{R}_2$ combined with PPI1 - PPI4 and $\mathbb{P}$ combined with ECO1 and ECO2. The density $P(T)$ for these systems is presented in Fig. 2f - k (red), following precisely the anticipated form, as predicted and observed on the model networks of Fig. 3 in the main text. The spatio-temporal layout, $\mathcal{L}(j \to i)$,

for these six systems appears in Fig. 5. While the results for the four protein interaction networks and ECO1 follow our predictions with high accuracy, we find that for ECO2, the $\mathcal{L}(j \to i)$ prediction exhibits rather high levels of noise (Fig. 5f). Still, the average propagation is well approximated by $\mathcal{L}(j \to i)$ (Fig. 5h). Indeed, ECO2, a small ($N = 456$) and relatively dense ($\langle S \rangle = 62$) network, is characterized by many loops ($C \approx 0.1$), and extremely short paths ($\max(L_{ij}) = 2$), and hence does not adhere to our model assumptions (large configuration model network). This has little effect on our macroscopic predictions, θ, $P(T)$, average $\mathcal{L}(j \to i)$ vs. $T(j \to i)$, but does impact the quality of the microscopic, node-specific, layouts of Fig. 5f.

4.3 Composite propagation

Our six empirical systems in the composite class include $\mathbb{M}$, applied to ECO1 and ECO2, and $\mathbb{E}$, applied to UCIonline, Epoch, Epinion and ATN. In Fig. 2l - q we show $P(T)$ (green), as obtained from these six systems. As predicted, we find multiple overlapping peaks - the fingerprint of the composite dynamic regime. Interestingly, in these empirical settings the composite interplay between L_{ij} and $P(S)$ is more complex that that observed on the *clean* model networks. For instance, in ECO1, Epoch and UCIonline, while the inner functions, representing $P(T \cap L)$ (shades of green), indeed show the anticipated effect of network distance, they also feature secondary peaks within the same shell. Specifically, we find that $P(T \cap L = 2)$ (dark green) is bi-modal, showing that within the same distance, we observe two typical response times. This is a direct consequence of the composite dynamics, in which $P(T)$ depends both on L_{ij} and on S_i. To observe this we focus on these three systems in Fig. 3, this time showing $P(T|S_i)$, the $T(j \to i)$ density conditional on the target nodes having degree S_i. We find, indeed, that the secondary peaks are driven by the low degree nodes within each shell, whose response time is large, due to their low weighted degree (as expected under $\theta < 0$). This illustrates the essence of the composite regime, where $T(j \to i)$ is determined both by the L_{ij}-shells, but also by the distribution of S_i within each shell, leading, in the case of these empirical networks, to the observed non-trivial structure of $P(T)$. The spatio-temporal layouts for these systems appear in Fig. 6.

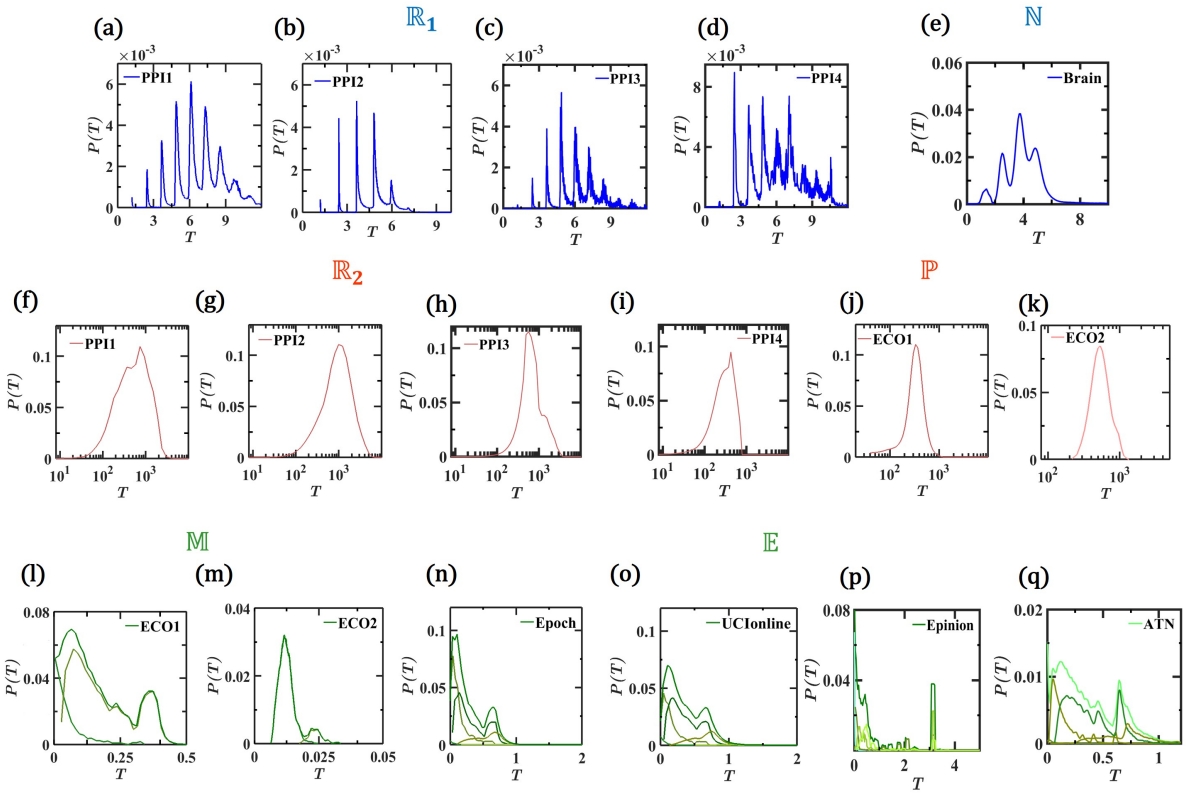

Figure 2: **Response time density function in empirical networks.** (a) - (d) $P(T)$ vs. T as obtained from $\mathbb{R}_1$ on the protein interaction networks PPI1 - PPI4. (e) $P(T)$ vs. T as obtained from $\mathbb{N}$ on Brain. The slight overlap between peaks is a consequence of the weight heterogeneity in Brain, in which equidistant nodes may have slightly different response times, owing to the different weights along their paths. (f) - (i) $P(T)$ for the degree limited $\mathbb{R}_2$ on PPI1 - PPI4. (j) - (k) $P(T)$ vs. T as obtained from $\mathbb{P}$ on the ecological networks ECO1 and ECO2. (l) - (q) In the composite dynamics $\mathbb{M}$ and $\mathbb{E}$, $P(T)$ is characterized by multiple overlapping peaks. The density $P(T|L_{ij})$, capturing the response times within each L_{ij}-shell, is also shown (shades of green). ECO2, whose diameter equals 2 shows only two peaks, as expected. ECO1, Epoch and UCIonline exhibit each two or three peaks within each shell - a consequence of the composite dynamics, in which low degree nodes respond at later times.

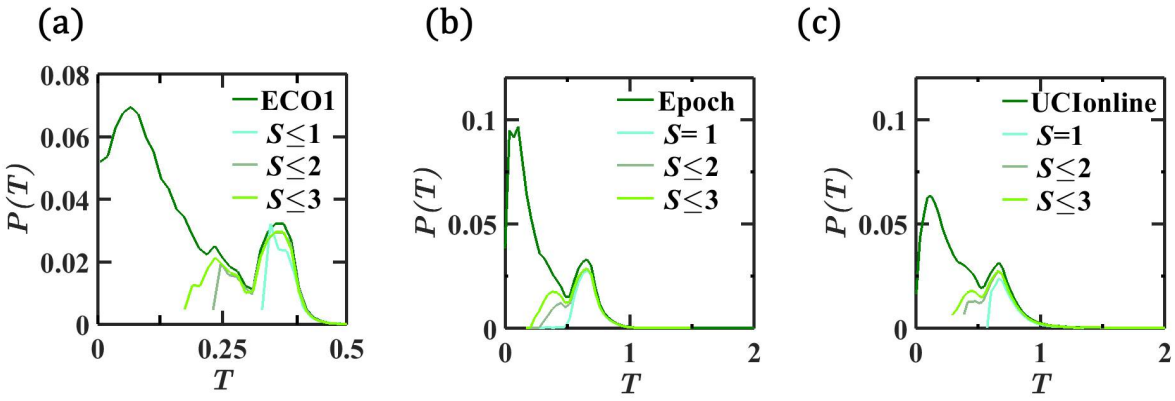

Figure 3: **A close up on the composite structure of $P(T)$.** We focus on $P(T)$ vs. T on (a) $\mathbb{M}$ on ECO1; (b) $\mathbb{E}$ on Epoch and (c) $\mathbb{E}$ on UCIonline. The different peaks, in these systems, correspond to the low degree target nodes, which respond late (large T) under composite dynamics. This structure of $P(T)$ exposes the interplay of distance and degrees, characterizing the composite universality class.

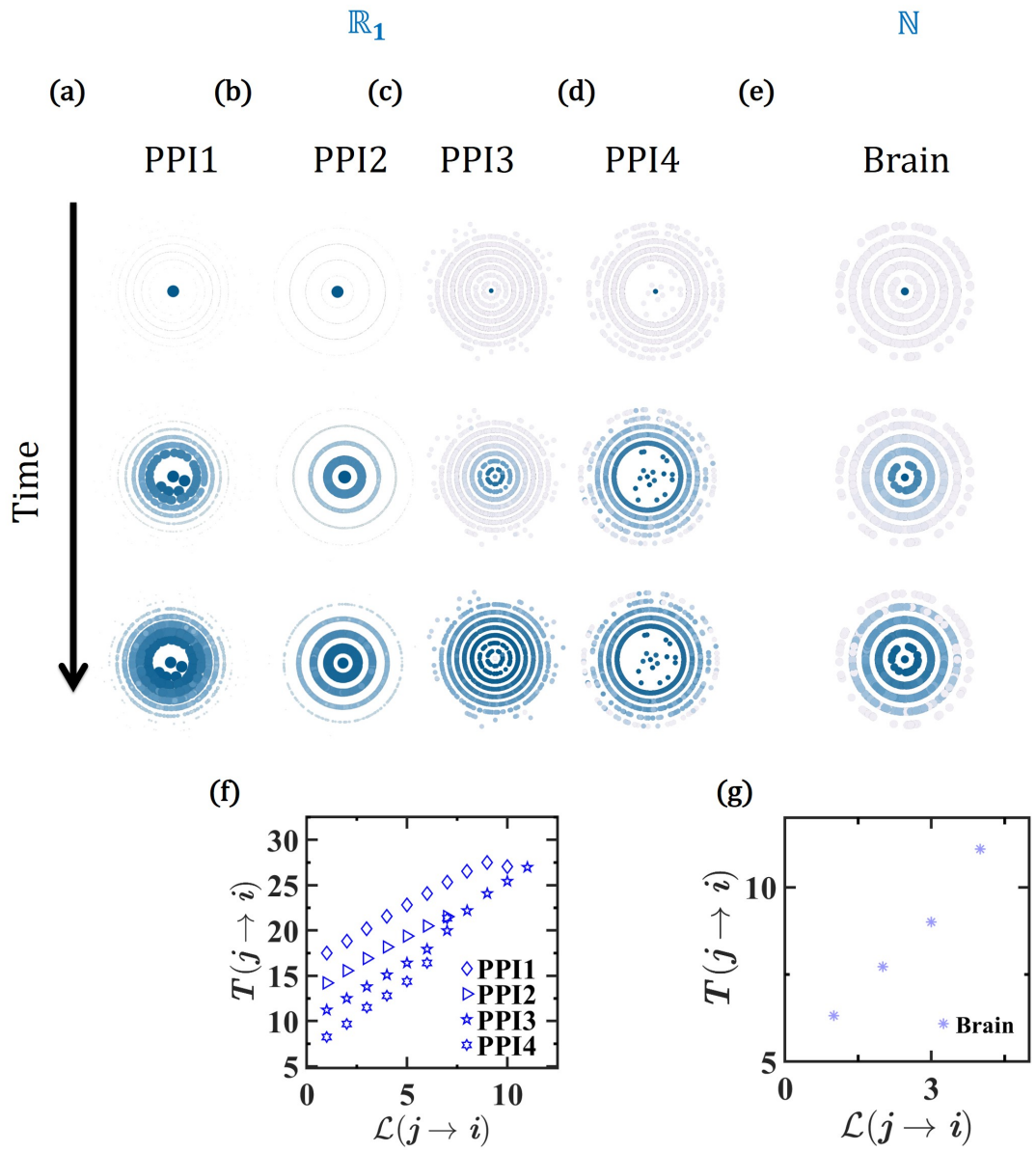

Figure 4: **The universal distance $\mathcal{L}(j \to i)$ in empirical networks.** Results obtained from our five distance limited systems.

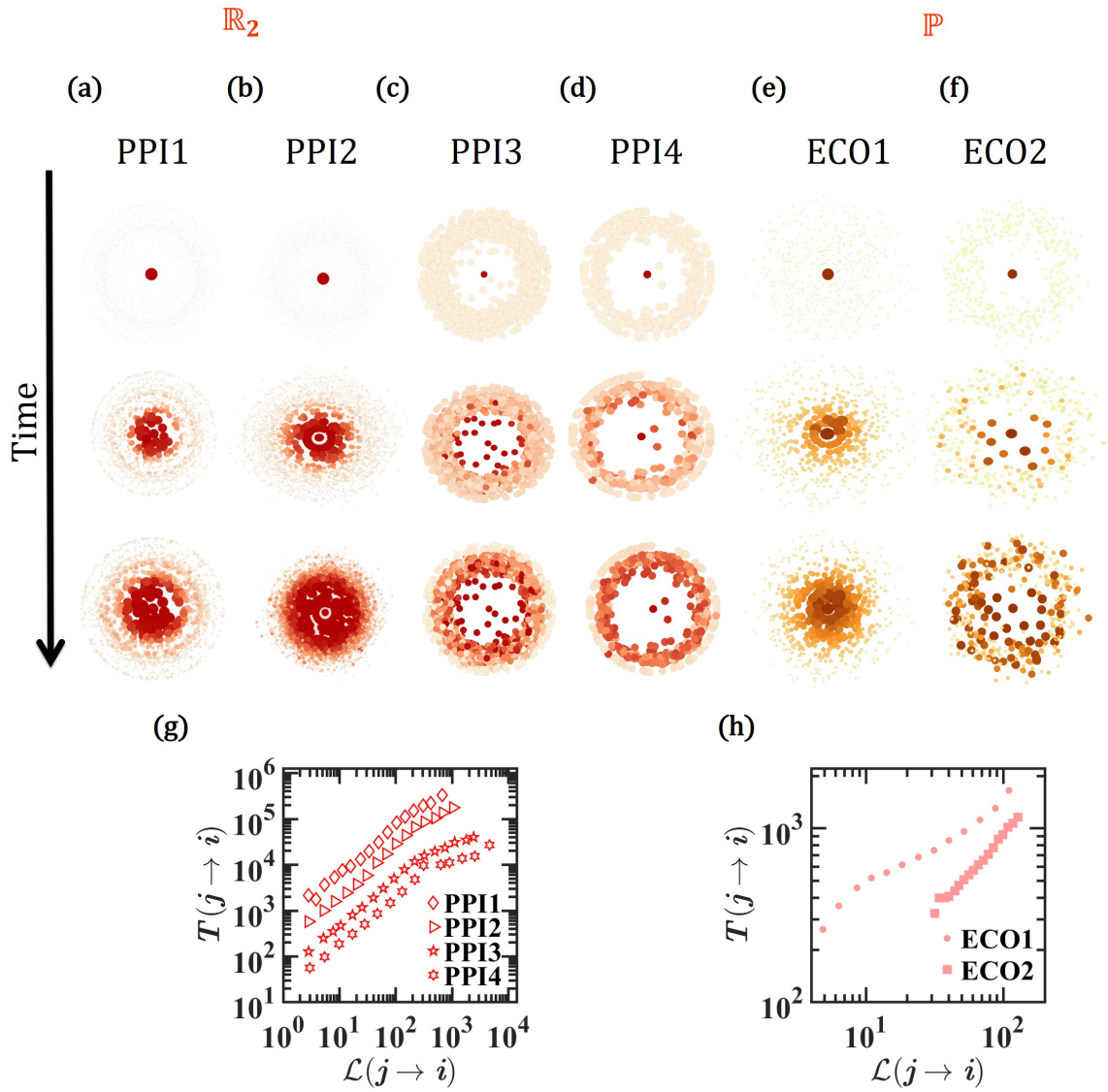

Figure 5: **The universal distance $\mathcal{L}(j \to i)$ in empirical networks.** Results obtained from our six degree limited systems.

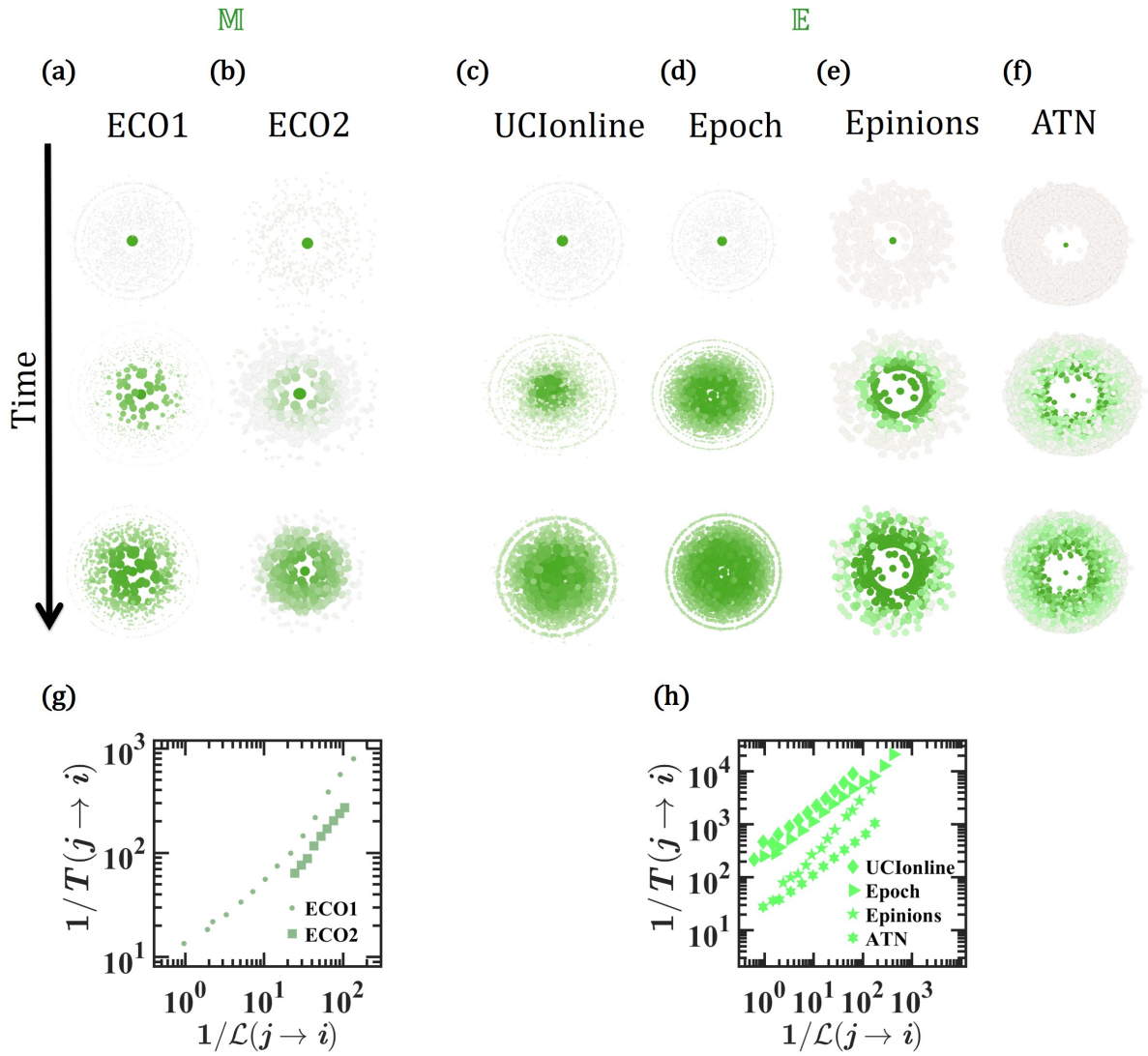

Figure 6: **The universal distance $\mathcal{L}(j \to i)$ in empirical networks.** Results obtained from our six composite systems.

Epidemics and power failure - propagation in empirical systems. Below and in Sec. 6 we examine two realistic systems that extend beyond our analytical framework. We hypothesize that our proposed *approach* - to seek the scaling of individual response times (τ_i), then combine these response times to generate global propagation patterns ($P(T), \mathcal{L}(j \to i)$), will indeed apply under more general conditions than those underlying our analytical derivations. Hence in the following two sections we examine this extended applicability and discuss its merits and limitations.

5 Propagation of a global epidemic

5.1 Overview

We used the international air-traffic network to model the spread of a global epidemic. Here A_{ij} is constructed from human mobility data, capturing the flux of travelers between $N = 1,292$ international airports. The local dynamics, within each location (node), is driven by the susceptible-infected-recovered (SIR) model [27], the links represent the volume of human travel [28], and the *signal* is captured by a local outbreak. This system challenges our analytical framework in two ways: the dynamics are outside the restrictions of Eq. (1.1) and the propagation represents a macroscopic state-transition from healthy to pandemic, beyond the linear regime of small Δx. Therefore we cannot analytically predict θ. However, we can still *measure* θ, finding that despite the theoretical discrepancies, the system does follow Eq. (1.41), exhibiting a (weak) degree limited propagation with $\theta \approx 0.15$ and $P(T)$, aptly, having a single peak (Fig. 7b,c). Using this observed θ in $\mathcal{L}(j \to i)$, we can now predict $T(j \to i)$, providing an accurate account of the actual disease propagation (Fig. 7e). The observed scaling is outside the scope of our analytical framework, as this dynamics cannot be cast within the form (1.1). However, it illustrates the merits of our strategic approach: first, examining the rules governing the local response times, *i.e.* θ, then piecing together the individual components to project the global propagation patterns, *e.g.*, $\mathcal{L}(j \to i)$. Below we provide the detailed analysis of this system.

5.2 SIR Model

To model epidemic spreading via air travel we used a local susceptible-infected-recovered (SIR) model with diffusive coupling, following Ref. [28]. In this framework each node n ($n = 1, \ldots, N$) represents a local population of M_n individuals, of which $S_n(t)$ are in the susceptible state $\mathcal{S}$, $I_n(t)$ are infected ($\mathcal{I}$) and $R_n(t)$ are recovered ($\mathcal{R}$), hence $S_n(t)+I_n(t)+R_n(t) = M_n$ for all t. Within each node, we assume a well-mixed population that locally follows SIR dynamics, namely

$$\begin{aligned} \mathcal{I} + \mathcal{S} &\xrightarrow{\alpha} 2\mathcal{I} \\ \mathcal{I} &\xrightarrow{\beta} \mathcal{R}, \end{aligned} \tag{5.1}$$

where α and β are the infection and recovery rates, respectively. The coupling between two meta-populations n and m is mediated by the flux of incoming/outgoing travelers

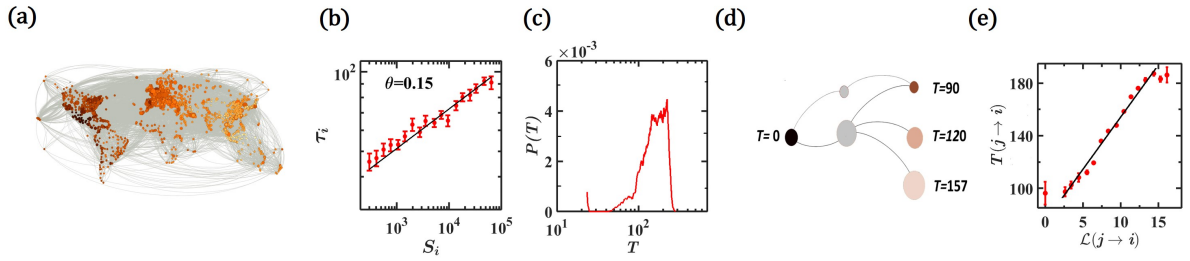

Figure 7: **Propagation in epidemic spreading**. (a) The spread of a global epidemic, driven by SIR dynamics, trough the weighted air-traffic network. This system cannot be cast within (1.1), yet its propagation can still be classified using our the three regimes, distance/degree limited or composite. (b) τ_i vs. S_i indicates a weak degree driven dynamics with $\theta \approx 0.15$. The roots of the hubs' delayed response to neighboring outbreaks is their typically higher populations, which require more time for the disease to penetrate. (c) $P(T)$ vs. T as obtained from 100 random disease outbreaks. As expected, $P(T)$ features the characteristic shape of degree limited dynamics. (d) In this regime we predict that hubs delay the propagation. We exemplify this here on a sub-network of (a), showing three nodes at distance $L_{ij} = 2$ from the disease source. Despite their identical distance, they respond at different times. As predicted, the higher degree nodes (node size) respond at later times. (e) $T(j \to i)$ vs. $\mathcal{L}(j \to i)$ with $\theta = 0.15$ successfully predicts the observed spread.

via air-traffic F_{nm}, quantifying the number of individuals flying from m to n per day. In this network the weighted in-degree of all nodes

$$S_n = \sum_{m=1}^{N} F_{nm} \tag{5.2}$$

captures the number of incoming passengers into n on an average day. Its out-degree $S_n^{\text{Out}} = \sum_{m=1}^{N} F_{mn}$ is the number of passengers departing daily from n.

This form of epidemic spreading gives rise to the following dynamic equations

$$\frac{\mathrm{d}S_n}{\mathrm{d}t} = -\alpha \frac{S_n(t) I_n(t)}{M_n} + \sum_{m=1}^{N} \Big(w_{nm} S_m(t) - w_{mn} S_n(t) \Big) \tag{5.3}$$

$$\frac{\mathrm{d}I_n}{\mathrm{d}t} = \alpha \frac{S_n(t) I_n(t)}{M_n} - \beta I_n(t) + \sum_{m=1}^{N} \Big(w_{nm} I_m(t) - w_{mn} I_n(t) \Big) \tag{5.4}$$

$$\frac{\mathrm{d}R_n}{\mathrm{d}t} = \beta \mathcal{I}_n(t) + \sum_{m=1}^{N} \Big(w_{nm} R_m(t) - w_{mn} R_n(t) \Big). \tag{5.5}$$

The first term/s on the r.h.s. capture the processes of infection, proportional to the product of susceptible and infected individuals, and recovery, proportional to $I_n(t)$. The summation terms describe the diffusion of $\mathcal{S}, \mathcal{I}$ or $\mathcal{R}$ individuals between local populations, where

$$w_{nm} = \frac{F_{nm}}{M_m} \tag{5.6}$$

is the per-capita flux from m to n, hence, *e.g.*, $w_{nm}S_m(t)$ is the volume of susceptible passengers leaving m and entering n per day.

Finally, we introduce an invasion threshold ε, which *activates* the local SIR dynamics only if the infected population rises above an ε fraction of the local population. We apply this by adding a sigmoidal function [28]

$$\sigma(x) = \frac{(x/\varepsilon)^h}{1 + (x/\varepsilon)^h} \tag{5.7}$$

to the equations, providing

$$\frac{\mathrm{d}S_n}{\mathrm{d}t} = -\alpha \frac{S_n(t)I_n(t)}{M_n}\sigma\left(\frac{I_n(t)}{M_n}\right) + \sum_{m=1}^{N}\Big(w_{nm}S_m(t) - w_{mn}S_n(t)\Big) \tag{5.8}$$

$$\frac{\mathrm{d}I_n}{\mathrm{d}t} = \alpha \frac{S_n(t)I_n(t)}{M_n}\sigma\left(\frac{I_n(t)}{M_n}\right) - \beta I_n(t) + \sum_{m=1}^{N}\Big(w_{nm}I_m(t) - w_{mn}I_n(t)\Big) \tag{5.9}$$

$$\frac{\mathrm{d}R_n}{\mathrm{d}t} = \beta I_n(t) + \sum_{m=1}^{N}\Big(w_{nm}R_m(t) - w_{mn}R_n(t)\Big), \tag{5.10}$$

hence *infection* is locally initiated only when $I_n(t)/M_n$ exceeds ε, an *invasion* of n.

5.3 Normalized equations

Next we rewrite Eqs. (5.3) - (5.5) for the normalized populations $s_n(t) = S_n(t)/M_n$, $j_n(t) = I_n(t)/M_n$ and $r_n(t) = R_n(t)/M_n = 1 - s_n(t) - j_n(t)$, obtaining

$$\frac{\mathrm{d}s_n}{\mathrm{d}t} = -\alpha s_n(t)j_n(t)\sigma(j_n) + \sum_{m=1}^{N}\left(\frac{M_m}{M_n}w_{nm}s_m(t) - w_{mn}s_n(t)\right) \tag{5.11}$$

$$\frac{\mathrm{d}j_n}{\mathrm{d}t} = \alpha s_n(t)j_n(t) - \beta j_n(t)\sigma(j_n) + \sum_{m=1}^{N}\left(\frac{M_m}{M_n}w_{nm}j_m(t) - w_{mn}j_n(t)\right) \tag{5.12}$$

$$\frac{\mathrm{d}r_n}{\mathrm{d}t} = \beta j_n(t) + \sum_{m=1}^{N}\left(\frac{M_m}{M_n}w_{nm}r_m(t) - w_{mn}r_n(t)\right) \tag{5.13}$$

Assuming a negligible fraction of one-directional trips, *i.e.* that immigration accounts for a marginal part of the overall international mobility, we can write $F_{nm} = F_{mn}$, stating that, on average, the number of passengers flying daily from m to n is the same as those flying from n to m. This enables us, using (5.6) to write

$$F_{nm} = F_{mn} \Rightarrow w_{nm}M_m = w_{mn}M_n, \tag{5.14}$$

which allows us to further simplify Eqs. (5.11) - (5.13), bringing them to their final normalized form

$$\frac{\mathrm{d}s_n}{\mathrm{d}t} = -\alpha s_n(t) j_n(t) \sigma(j_n) + \sum_{m=1}^{N} A_{nm}\big(s_m(t) - s_n(t)\big) \tag{5.15}$$

$$\frac{\mathrm{d}j_n}{\mathrm{d}t} = \alpha s_n(t) j_n(t) \sigma(j_n) - \beta j_n(t) + \sum_{m=1}^{N} A_{nm}\big(j_m(t) - j_n(t)\big) \tag{5.16}$$

$$\frac{\mathrm{d}r_n}{\mathrm{d}t} = \beta j_n(t) + \sum_{m=1}^{N} A_{nm}\big(r_m(t) - r_n(t)\big), \tag{5.17}$$

where

$$A_{nm} = \frac{M_m}{M_n} w_{nm} = w_{mn} \tag{5.18}$$

is the normalized human flux matrix.

To construct A_{nm} from data we denote the mobility rate by [28]

$$\mu = \frac{\sum_{m,n=1}^{N} F_{nm}}{\sum_{n=1}^{N} M_n}, \tag{5.19}$$

in which the numerator quantifies the number of individuals flying per day and the denominator equals to the total (global) population, therefore providing the daily fraction of individuals who seek air travel. This allows us to estimate a local population M_n by

$$M_n = \frac{1}{\mu} \sum_{k=1}^{N} F_{kn}, \tag{5.20}$$

assuming that all people departing from n, *i.e.* $\sum_{k=1}^{N} F_{kn}$, represent a μ fraction of n's total population, M_n. We can now write

$$w_{nm} = \frac{F_{nm}}{M_m} = \mu \frac{F_{nm}}{\sum_{k=1}^{N} F_{km}} \tag{5.21}$$

and

$$\frac{M_m}{M_n} = \frac{\sum_{k=1}^{N} F_{km}}{\sum_{k=1}^{N} F_{kn}}, \tag{5.22}$$

providing

$$A_{nm} = \mu \frac{F_{nm}}{\sum_{k=1}^{N} F_{kn}}, \tag{5.23}$$

which allows us to extract A_{nm} directly from the mobility data (F_{nm}, μ).

The resulting dynamics, a highly realistic account of global epidemic spreading, goes beyond the limits of our universal dynamic equation (1.1) in several aspects: (i) the

equation is multi-dimensional, describing the state of each node with more than a single variable $x_i(t)$; (ii) the diffusive interaction term cannot be factorized as $M_1(x_i)M_2(x_j)$; (iii) for the appropriate selection of α and β the healthy state ($s_n = 1$) becomes unstable, hence the initial perturbation causes a macroscopic state-transition, rather than a small linear response.

5.4 The origins of the observed scaling

The SIR model in (5.15) - (5.17) cannot be covered within our analytical framework, and hence the observed scaling of $\theta \approx 0.15$, while compatible with our classification, is not rooted in the same theoretical arguments as those outlines in Sec. 1. Still the existence of an analogous scaling relationship demonstrates the conceptual utility of our approach, that by characterizing the intrinsic time scales associated with each individual component (θ), it helps piece together the puzzle of the system's global propagation patterns (*e.g.*, $\mathcal{L}(j \to i)$).

To understand the observed scaling, we consider the two steps leading to a node's infection. First, the source node m is infected, reaching its peak infection through the local SIR dynamics. As infected individuals accumulate in m they begin to diffuse to its neighboring node n, through the link A_{nm}. The local SIR process in n, however, is only initiated once it becomes *penetrated*, *i.e.* entered by enough individuals from m such that $j_n > \epsilon$. Up to this point, the *penetration time* τ_1 of n, the SIR dynamics remains effectively dormant. Hence, the first stage in spreading the epidemic to n, lasting for time τ_1, is governed by diffusion only, almost independent of the disease parameters α and β. Once n is penetrated ($j_n > \epsilon$) local SIR dynamics takes over, reaching peak infection within the *infection time* τ_2. The total response time τ_n is, therefore, the sum of these two dynamic processes (Fig. 9a)

$$\tau_n = \tau_1 + \tau_2, \tag{5.24}$$

the former driven by diffusion (A_{nm}) and the latter by the SIR dynamics (α, β). The infection time τ_2 is independent of the network, driven by the disease's infection/recovery rates, and thus unaffected by n's degree S_n. This component of τ_n is, in effect, distance limited, characterized by $\theta_2 = 0$. The penetration time τ_1, on the other hand, depends on the diffusion rate A_{nm} from m to n, which, according to (5.18) is inversely dependent on n's populations size M_n. Since $M_n \propto S_n$ (5.20), we find that τ_1 depends on n's degree $S_n = \sum_{k=1}^{N} F_{kn}$. Therefore, τ_1 is degree limited, having $\theta_1 > 0$, expressing the fact that it required more time to penetrate nodes with a larger population (Fig. 9b).

In Fig. 9d we show τ_1 (squares) and τ_2 (triangles) vs. S. As expected τ_1 increases with S, while τ_2 is uniform for all nodes. Their sum provides the total response time τ (circles), resulting in an effective (weak) scaling with degree, as observed in Fig. 7.

Implementation. We used the Global Mobility Network dataset [29] to extract the empirical fluxes F_{nm}. The dataset lists the daily travel of $\sim 8.91 \times 10^6$ passengers per day over the course of 365 days, *i.e.* $i = 1, \ldots, 365$, between $N = 1,292$ airports, linked through $38,377$ directional air-routes. The mean flux from m to n is thus captured by

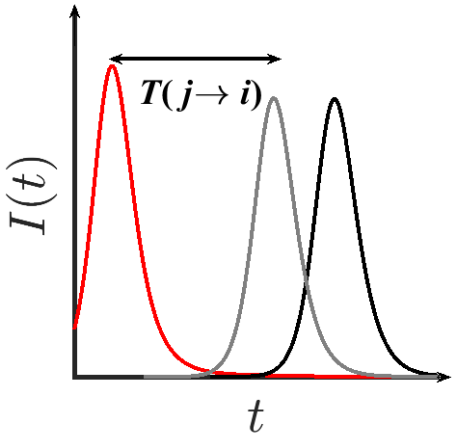

Figure 8: **Measuring epidemic propagation times**. To evaluate $T(j \to i)$ we measured the time from the peak infection of the source j (red) to that of the target i (grey, black). Here we show two targets: the grey represents a low degree node at distance $L_{ij} = 3$ from the source (red); the black represents a hub at distance $L_{ij} = 2$. Due to the degree limited dynamics the closer hub exhibits an increased response time.

$$F_{n \leftarrow m} = \frac{1}{365} \sum_{i=1}^{365} F_{n \leftarrow m}^{\{i\}} \tag{5.25}$$

where $F_{n \leftarrow m}^{\{i\}}$ is the number of passengers flying from m to n in day i. In our analysis we assume an almost symmetric flux matrix, as expressed in Eq. (5.14). Indeed we find that practically all links are bi-directional, and that in the vast majority of existing links we have $F_{n \leftarrow m} \approx F_{n \to m}$, a similar passenger flux in both directions. Therefore, to construct F_{nm} we used the symmetrized

$$F_{nm} = \frac{F_{n \leftarrow m} + F_{n \to m}}{2}, \tag{5.26}$$

a network comprising $19,614$ bi-directional links, that represents a symmetric approximation of the empirical fluxes. Evaluating the local populations M_n as in (5.20) we obtained A_{nm} via (5.23). Following Ref. [28], we used $\mu = 0.0028$ day^{-1} for the mobility rate. For the disease parameters we set $\beta = 0.285$ day^{-1}, capturing a typical $3 - 4$ day recovery time, and $\alpha = 1.5\beta$, to secure the system above the epidemic threshold ($\alpha/\beta > 1$). For the penetration threshold we set $\epsilon = 10^{-5}$ and $h = 2$ in (5.7). In each realization we induced at $t = 0$ an outbreak on a randomly selected node q and measured the response of all other nodes. The response times τ_i and $T(q \to i)$ were defined as the time from the peak infection at the source q, j_q^{Max}, to the time of the peak infection at the target, j_i^{Max} (Fig. 8).

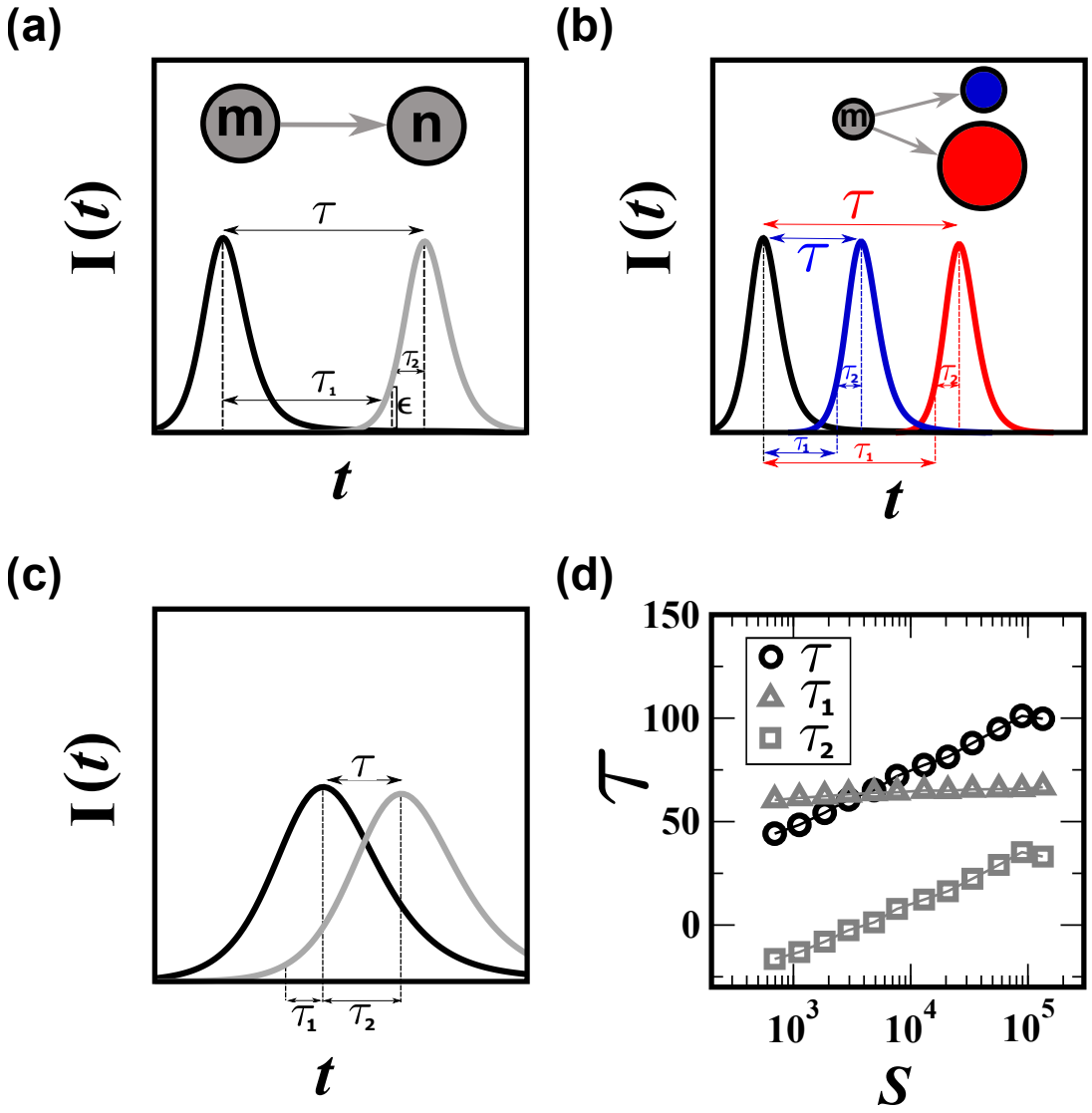

Figure 9: **The two time scales driving epidemic spread**. (a) The propagation time from source (m) to target (n), τ, as measured in Fig. 8, comprises two approximately separate stages: τ_1, the penetration time, defined as the time for the target population to accumulate an ϵ fraction of infected individuals; followed by τ_2, the infection time, capturing the time required for the local SIR dynamics to reach peak infection. (b) τ_1 is small for low degree nodes (blue) and large for hubs (red), while τ_2 is independent of node degree. (c) For extremely small nodes, penetration, *i.e.* $j_n(t) = \epsilon$, might occur prior to the source's peak infection. In such cases we observe a negative τ_1. (d) The target response time τ vs. its weighted degree S (black circles), is obtained as an aggregation of the two time-scales: the degree limited τ_1 (squares) added to the degree independent τ_2 (triangles).

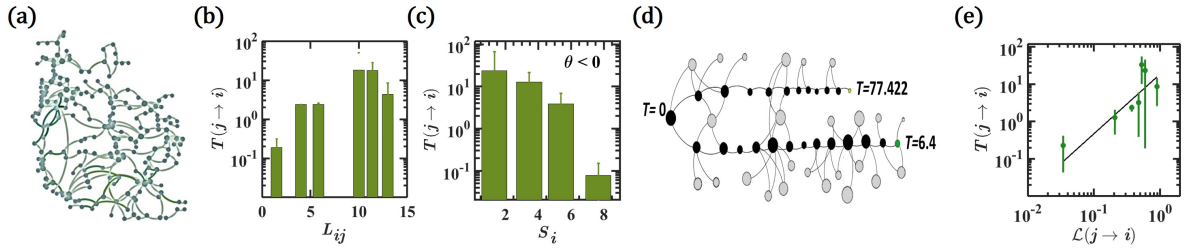

Figure 10: **Propagation of a power component failure**. (a) We tracked the propagation of a cascading failure in the a power supply network, impacting a total of 49 nodes. (b) The mean failure time $T(j \to i)$ of power components vs. their distance from the source L_{ij}. We find that L_{ij} cannot always predict the failure propagation. (c) $T(j \to i)$ vs. component degree S_i shows that higher degree components respond at earlier times, indicating a composite dynamics in which $\theta < 0$. Having access only to a single spontaneously occurring perturbation, we cannot measure τ_i for this system, as we lack data on the local response times of nodes to neighboring perturbations. Consequently we cannot directly measure the precise value of θ. (d) The composite dynamics predict that the failure propagates faster on hub-enriched pathways. We exemplify this on a sub-network of (a), where the more distant node (bottom) responds earlier due to the many high degree nodes along its path. (j) $T(j \to i)$ vs. $\mathcal{L}(j \to i)$ as obtained with $\theta = -1$, helping us approximate the observed propagation.

6 Failure propagation in power-supply

6.1 Overview

We collected data from a cascading failure in a US power supply network [31], starting from a local malfunction that lead to the failure of 49 nodes. Unable to induce controlled perturbations, we cannot directly measure τ_i for this system, and hence cannot evaluate the precise value of θ. Still, we find that $T(j \to i)$ is, on average, inversely dependent on S_i, indicating that this system is in the composite regime, having $\theta < 0$ (Fig. 10c). Indeed, setting $\theta = -1$ in $\mathcal{L}(j \to i)$ we obtain an approximate account of the empirical spatio-temporal spread (Fig. 10e). The composite classification indicates that long paths may be potentially faster if they are enriched with hubs. This is demonstrated in Fig. 10d, where the failure traverses the longer bottom path in less time than it does the shorter top one.

6.2 Implementation

To track the propagation of a local failure in a power supply network we collected data from a 1996 (August 10) disturbance in the Western Interconnection grid, in which a transmission line trip resulted in a sequence of 49 component failures [30, 31]. The cascade, which unfolded over a $\sim$ 3 hour period, documents the failure time of each component (Table 4), ultimately resulting in a breakdown, which disintegrated the 212 node network into four isolated component.

Data analysis. As indicated in the official report [30], the first 5 events preceded the

actual spread, hence the propagation times are tracked from event 6, marking the failure of transmission line $j = 249$ at time point $t_j = 76.588$ minutes. Measuring the propagation times relative to this initial disturbance we define the failure time of all other components as $T(j \to i) = t_i - t_j$, taking t_i from the right column of Table 4.

Table 4 further indicates that the majority of failures affected transmission lines (Type 2) rather than generating units (Type 1), *i.e.* most failures are link failures, as opposed to node failures. Therefore, to track the propagation we mapped the original power supply network A_{ij} into its adjoint graph $B_{\mu\nu}$, in which the role of nodes/links is reversed. In this mapping each of the 316 transmission lines is mapped into a node in $B_{\mu\nu}$, $\mu, \nu = 1, \ldots, 316$. Nodes μ and ν are linked if, in the original network, their corresponding links share a node. For instance, consider a simple linear graph with four nodes, $1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4$, and three links A_{12}, A_{23} and A_{34}. Its adjoint graph will have three nodes $\mu = 1$, corresponding to link A_{12}, $\mu = 2$, corresponding to link A_{23} and $\mu = 3$, corresponding to link A_{34}. The links in the adjoint network $B_{\mu\nu}$ with thus be B_{12}, owing to the shared node 2 between A_{12} and A_{23}, and B_{23}, owing to the shared node 3 between A_{23} and A_{34}. The $1, 3$ link, however, will not exist in $B_{\mu\nu}$, *i.e.* $B_{13} = 0$, as the corresponding links in the original graph, A_{12} and A_{34}, have no shared node. Focusing solely on links failures, and disregarding events $1 - 5$, we are left with a sequence of 34 events that propagate over the course of $150 - t_j \approx 70 - 80$ minutes.

In Fig. 10 we measure L_{ij} and S_i from the adjoint graph. In Fig. 10e we calculated $\mathcal{L}(j \to i)$ using Eq. (5) of the main text, taking S_p from the adjoint graph and setting $\theta = -1$. Having a small number of data points (34) we used uniform binning, grouping 6 data points into each bin, apart from the last bin, which consisted of the remaining points. The error bars were calculated following Eq. (3.13).

Events	Type of faliure	Node/edge	Time
0	2	246	
1	2	243	
2	2	40	
3	2	45	
4	2	43	
5	2	129	
6	2	249	76.588
7	2	251	76.613
8	2	261	76.618
9	2	252	76.632
10	2	262	76.641
11	2	245	76.682
12	2	254	76.74
13	2	253	76.778
14	2	255	76.782
15	2	248	76.878
16	2	256	77
17	2	260	77.255
18	1	75	77.622
19	2	314	77.708
20	2	177	77.93
21	2	309	78.415
22	2	149	78.576
23	2	152	78.622
24	2	202	78.7
25	2	294	78.703
26	2	204	78.765
27	2	163	78.825
28	2	116	78.846
29	2	293	78.936
30	1	200	78.938
31	2	208	79
31	2	102	79
31	2	207	79
32	1	129	79.141
33	1	129	79.405
34	2	224	82.496
35	2	161	82.803
36	2	295	83
37	1	176	83.341
38	1	176	83.492
39	1	175	84
40	1	176	91
41	1	143	97.315
41	1	143	97.405
42	1	174	98
43	1	174	104.83
44	2	158	129.271
45	2	159	132
46	1	148	143.723
47	1	128	144
48	2	211	144
49	2	30	150

Table 4: **Failure propagation in power supply**. The first column represents the event index. The second column describes the the type of failure: 1 - node failure, *i.e.* generating units; 2 - link failure, *i.e.* transmission lines. Column three notes the index of the failed node/link, and column four is its failure time (minutes). The propagation begins at event 6, marking the failure of transmission line $j = 249$ at $t_j = 76.588$ minutes.

7 Extended testing

Our analytical derivations, outlined in Secs. 1 and 2 are exact under two main assumptions: small signals Δx, which allow us to use linear response theoretic tools; the configuration model [1] framework, according to which node i's nearest neighbor statistics are independent of i. In real scenarios, however, we are often confronted by large perturbations, or by empirical networks, which may violate, to some extent, the *clean* picture of the configuration model. Moreover, realistic scenarios are compounded by uncertainties, such as partially mapped networks or noisy dynamics. Therefore we tested the robustness of our analytically predicted scaling, (1.41), against different types of discrepancies. To test the impact of large signals, we introduce perturbations ranging from 10% to 100%, culminating in full node knockout. To examine our formalsim's performance against topological disrepancies we introduce two frequently encountered network features that violate the configuration model framework: degree-degree correlations [2] and clustering. Finally, to examine the limits of predictability of our method we test the effects of hidden links and of parametric noise on our ability to predict and observe θ.

7.1 The effect of large perturbations

Measuring $T(j \to i)$ entails introducing a signal, Δx_j, to the steady state activity x_j of the source node j, and observing the flow of information as it propagates from j to i. In our derivations we resort to the perturbative limit, where $\alpha = \Delta x_j/x_j \ll 1$, a small perturbation, that allows us to use linearization to achieve analytical advances. Specifically, in our numerical experiments we set the magnitude of our signals to 10% of the source's steady-state activity, namely $\alpha = 0.1$. In Fig. 11 we examine the impact of larger perturbations, setting $\alpha = 0.4$ (squares), a 40% perturbation, $\alpha = 0.7$ (down-triangles), a larger perturbation of 70%, and even $\alpha = 1$ (up-triangles) a signal of the same size as the node's unperturbed activity. We find that the predicted scaling θ is extremely robust, with the size of the perturbation having no visible effect. We further tested the propagation following the full knockout of the source node, namely removing node j and observing the spatio-temporal system response (diamonds). Such node removal represents a common procedure to observe sub-cellular dynamics via controlled genetic knockouts [32]. It also arises in naturally occurring settings, such as in spontaneous component failure in *e.g.*, the power grid. We find that even under these extreme conditions our predicted scaling remains valid, indicating that our predictions are highly robust against perturbation size.

This lack of sensitivity is rooted in the well-established robustness of scaling relationships, which are often unaffected by small deviations and discrepancies [33]. This is especially relevant in a network environment, where local perturbations rapidly decay (exponentially) as they penetrate the network [34]. Under these conditions even a large local perturbations will have only a small effect on all individual nodes in its vicinity. Therefore, the consequent responses of the signal's direct neighbors, next neighbors and so on, can be well-approximated by the perturbative limit, even if the original j-signal violated this limit. Hence we find that the linear response framework remains valid even under unambiguously *large* perturbations.

7.2 The effect of clustering

Next we consider the impact of clustering C, representing the network's tendency to from triads, in which there is an increased probability for an n, m link, if n and m share a mutual neighbor i. Under the configuration model assumption, clustering tends to zero if the network is sparse and $N \to \infty$ [1]. Most empirical networks, however, feature non vanishing C, in some cases reaching an order of $C \sim 10^{-1}$ [35], significantly higher than that expected in a random connectivity. To measure node i's clustering we write

$$C_i = \frac{\sum_{m,n=1}^{N} A_{im}A_{in}A_{nm}}{\binom{k_i}{2}}, \tag{7.1}$$

in which the numerator counts the number of *actual* triads involving nearest neighbors of i, and the denominator equals to the number of *possible* triads around i, *i.e.* the number of potential pairs among i's k_i nearest neighbors. Hence $0 \leq C_i \leq 1$ is the fraction of potential triads that are actually present among i's neighbors. The clustering of the *network* is then obtained by averaging over all nodes as

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{7.2}$$

In Table 5 we show the clustering C as obtained from our set of empirical networks. We find that for some of these networks C is rather high, in some cases reaching as much as $C = 0.2567$ (Epoch) or $C = 0.4637$ (Brain). Still, as demonstrated in the main text, our analytical predictions performed well, even under these challenging conditions of extreme clustering. This indicates that our predictions are robust against empirically observed levels of clustering. To further examine the effects of clustering in a controlled fashion, we used the scale-free network SF, and gradually rewired it to increase its clustering to $C_1 = 0.05, C_2 = 0.1$ and $C_3 = 0.15$ (Table 5), generating three model networks, SFC1 - SFC3, with controlled levels of clustering. We then measured τ_i vs. S_i on each of these networks. We find again that even extreme levels of clustering ($C_3 = 0.15$ is two orders of magnitude higher than the expected value from the configuration model), our theoretical predictions are consistently sustained (Fig. 12).

7.3 The effect of degree-degree correlations

Next, we examine the effect of degree correlations Q, as defined in Ref. [2]. As before, we first observe the correlation levels exhibited by our set of empirical networks, finding that they feature rather high levels of degree correlations (Table 5). The fact that our predictions cover these networks is, as before, an indication of our theory's robustness against empirically observed correlation levels. To complement this finding we rewired SF, once again, this time to exhibit increasing levels of positive and negative degree correlations, producing SFQ1 - SFQ4, as detailed in Table 5. As in the case of clustering, the results, presented in Fig. 13, show that our predictions are largely unharmed by Q, indicating their low sensitivity to the configuration model assumption of Sec. 1.

Robustness of predicted scaling and dynamic regimes. Our theory provides both quantitative as well as qualitative predictions. At the quantitative level, we predict the precise value of θ, allowing us to provide the precise response times of all nodes (Fig. 4 in main text). No less important are, however, our qualitative predictions, that allow us to translate θ into direct insights on the macroscopic propagation patterns of a networked system. This is observed by the distinct structures of $P(T)$ (Fig. 3g - l in main text), the different roles of network paths L_{ij} (Fig. 4d - f in main text), and the class-specific contribution of $P(S)$ (Fig. 4g - l and Fig. 5 in main text). All of these observations represent macro-level dynamic patterns that determine how the *system* (as opposed to specific nodes) manifests information propagation. Such intrinsic characteristics are seldom sensitive to microscopic discrepancies.

We further argue that even if the precise value of θ deviates due to some specific departures from our model assumptions - deviations that we have not observed in our extensive numerical tests - still, the implications on the macro-scale behavior of the system, indeed, the *qualitative* insight that our theory aims to provide, will ultimately be marginal. For instance, consider a deviation in one of our dynamics, say the degree limited $\mathbb{R}_2$, which under some hypothetical conditions features, *e.g.*, a decrease in its observed θ from the theoretically predicted $\theta = 3/2$ to, say, $\theta \approx 1$. This may constitute a significant discrepancy in terms of our quantitatively predicted scaling, but will not significantly impact the observed propagation patterns, which will remain within the degree-driven class. Indeed, micro, or even meso-scopic discrepancies from our model assumptions are extremely unlikely to cause a qualitative shift to a different class, turning the system, for instance from degree-driven to distance-driven or composite dynamics. Such transition can only be done by altering the system's internal mechanisms, such as shifting its dynamics from $\mathbb{R}_2$ ($\theta = 3/2$) to $\mathbb{E}$ ($\theta = -1$), a change in the *physics* of the node interactions, which requires a fundamental intervention, unattainable by minor discrepancies.

7.4 The effect of parametric noise

As it is written, the dynamic Eq. (1.1), is subject to intrinsic noise, introduced through the random (weighted) topology A_{ij}, which is, in most cases, extremely heterogeneous and diverse. This source of *topological* noise seems to have little effect on the accuracy of our predictions, as indicated by our results. We now examine the performance of our predicted scaling (1.41) in the presence of *dynamic* noise, in which, in addition to the random A_{ij}, the dynamic functions $\mathbf{M}$ also feature varying levels of randomness across all nodes. Hence we consider the $\mathbb{R}$-family of dynamics

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -Bx_i^{a_i}(t) + \sum_{j=1}^{N} A_{ij} \frac{x_j^h(t)}{1 + x_j^h(t)}, \tag{7.3}$$

in which the exponent a_i is now a random variable extracted from a skewed distribution

$$a_i \sim \mathcal{N}_+(\mu, \sigma^2), \tag{7.4}$$

a normal distribution truncated such that $P(a_i \leq 0) = 0$, namely a_i is a non-negative

random variable. This introduces a source of parametric noise into (1.1), in which each node follows a different dynamics. The level of this noise is quantified by the variance σ^2 in (7.4). The randomness is introduced exclusively on a_i, as this is the only parameter affecting θ in this dynamics; see Eq. (2.7). To predict θ we use the mean value of a_i, given by $\langle a \rangle = N^{-1} \sum_{i=1}^{N} a_i$, and substitute it for a in Eq. (2.7), yielding

$$\theta = \frac{1 - \langle a \rangle}{\langle a \rangle}. \tag{7.5}$$

We expect the observed scaling to fluctuate around the prediction θ due to the noise, with the fluctuation magnitude growing as σ^2 is increased. To quantify these deviations we fit the observed τ_i with the predicted scaling $\tau_i \sim S_i^{\theta}$, and measure the mean-square-error (MSE) between all data points and the theoretical scaling. To be specific, we use a logarithmic transformation, plotting $\log \tau_i$ vs. $\log S_i$. We then fit a linear plot to the data

$$\log \tau_i^{\mathrm{Th}} = \theta \log S_i + C, \tag{7.6}$$

a perfect linear function, in which we constrain the slope to equal θ, as calculated from (7.5), then seek the best fit for the arbitrary constant C. Finally, we evaluate the error using

$$\mathrm{MSE} = \frac{1}{N} \sum_{i=1}^{N} \left(\tau_i - \tau_i^{\mathrm{Th}} \right)^2, \tag{7.7}$$

quantifying how noisy the data is in relation to the *perfect* scaling of (7.6).

In Fig. 14 we show the results obtained for three different dynamics, implemented on the network SF: $\mathbb{R}_1$, where we set $\mu = 1$, $\mathbb{R}_2$, where we set $\mu = 0.4$ and $\mathbb{R}_3$, where we set $\mu = 2$, representing our three classes - blue, red and green - respectively. In each panel we show τ_i vs. S_i for all individual nodes (grey), as well as the logarithmically binned averages (circles). The calculated θ from Eq. (7.5) is also shown, representing the slope of the fitted theoretical line (7.6) plotted in each panel (black solid lines). As we gradually increase the noise level σ, we find that the individual data points become more scattered, exhibiting a growing MSE, and, consequently leading to observed discrepancies between the theory (solid black lines) and the data (grey dots). This is especially expressed in the distance limited $\mathbb{R}_1$, where the expected θ vanishes, and hence it is rather sensitive even to minor discrepancies. Note that in $\mathbb{R}_2$, the predicted θ decreases with σ, a consequence of the skewed $\mathcal{N}_+(\mu, \sigma^2)$, which, under large variance, drives $\langle a \rangle$ to be greater than 0.4, and hence θ (7.5) to become smaller than 3/2. Indeed, for $\mu = 0.4$, as σ is increased, the truncation at zero, removes a significant portion of the a_i density from $a_i < 0$ to $a_i > 0$, a bias that leads to an increase in $\langle a \rangle$.

In Fig. 15 we collect the MSE vs. σ for all three dynamics $\mathbb{R}_1 - \mathbb{R}_3$. Note that MSE$> 0$ even for $\sigma = 0$. This is the baseline error associated with the system's intrinsic *topological* noise. The increase above this baseline captures the impact of the *dynamic* noise we introduce here. As MSE increases, an unavoidable consequence of the growing noise levels, it may potentially impact our *quantitative* prediction, affecting the measurement error in our evaluation of θ. Still, even for large MSE, such discrepancies have no bearing on the *qualitative* classification of our systems as distance limited, degree limited or composite (see grey Box in Sec. 7.3).

7.5 The effect of hidden links

Often the challenge in observing network dynamics is that the underlying network is only partially mapped, including many unknown interactions. For example, in sub-cellular networks it is estimated that we have only uncovered $\sim 20\%$ of all links. In social networks, relying on proxies, such as online connections, leaves us with a fraction of unknown off-line social ties that leave no trace in the digital arena, and are hence hidden from the observer. Such topological uncertainties have no impact on some of our predictions, *e.g.*, $P(T)$, which can be measured absent any knowledge of A_{ij}. Our scaling prediction $\tau_i \sim S_i^{\theta}$, however, requires us to know the degrees of all nodes, and hence relies on a reliable mapping of A_{ij}. Therefore we now examine our ability to observe θ in the presence of hidden links.

An advantage of scaling relationships is that they are, potentially, robust against hidden links. To understand this consider the case where there is a random fraction f of un-mapped links in A_{ij}. On average, such uncertainty will reduce the observed degrees of all nodes as

$$S_i^{\text{Obs}} \sim (1-f)S_i, \tag{7.8}$$

as indeed, lacking a random fraction f of all links, a typical node is expected to *miss* a fraction f of its actual neighbors. Fortunately, as long as f is sufficiently smaller than unity, such re-scaling of all degrees has little impact on Eq. (1.41), which is insensitive to the multiplicative constant $1-f$, namely

$$\tau_i \sim S_i^{\theta} \Longrightarrow \tau_i \sim \left((1-f)S_i\right)^{\theta} \sim \left(S_i^{\text{Obs}}\right)^{\theta}. \tag{7.9}$$

Such robustness is, of course, dependent on the hidden links being randomly scattered across all nodes.

We tested this by implementing $\mathbb{R}_1, \mathbb{R}_2$ and $\mathbb{R}_3$ on SF, and systematically *hiding* links at random, *i.e.* simulating the complete network, but assuming that a fraction f of the links are empirically inaccessible. We find that even when 30% of the links are hidden there is no visible effect on the observed scaling (Fig. 17b). A significant deviation is observed when f approaches $0.6-0.8$, depending on the dynamics (Fig. 17c). As expected, when f continues to grow, we reach a state where we completely fail to observe the scaling relationship, here shown for $f=0.9$, *i.e.* our network is only 10% known (Fig. 17d).

7.6 The effect of multiple fixed-points

Nonlinear systems often feature multiple steady-states, which may be characterized by different propagation patterns. Indeed, extracting the exponent θ depends on the asymptotic behavior of $R^{-1}(x)$ in (1.39), which is potentially different around different steady-states. We demonstrate this here in Fig. 16 on the regulatory model $\mathbb{R}_1$ and on the SIS model $\mathbb{E}$.

For $\mathbb{R}_1$ the system exhibits two fixed-points (Fig. 16a) the inactive dead state where all $x_i = 0$, which is always stable, and the active state, which is only stable if A_{ij} is dense enough to support sufficiently strong activation of all nodes [36]. We set the system at each of these fixed-points and tested its response to local signals. We find that the

propagation patterns are distance limited around both states, with $\theta = 0$, $T(j \to i) \propto L_{ij}$ and $P(T)$ featuring the characteristic discrete peaks (Fig. 16c - e).

For $\mathbb{E}$ the system transitions between an exclusively stable healthy state for low network density and an exclusively stable pandemic state as the network crosses the epidemic threshold (Fig. 16b). Here we find that the propagation patterns are, in fact, different across the states of the system: while the pandemic state features composite dynamics (as also shown in the paper), the healthy state is degree limited (Fig. 16f - h). Indeed, linearizing Eq. (2.16) around these two different steady-states predicts a different behavior in function of S_i.

State transitions are a fundamental aspect of nonlinear dynamics, capturing different *phases* that can be exhibited by the system. Our observation, that the propagation class is also state-dependent uncovers an additional layer associated with these phases - it shows that they are not only characterized by different activities x_i, but also by different propagation patterns. Hence, as the system transitions from, *e.g.*, healthy to pandemic, it also transitions from distance limited to composite.

Network		Degree Correlation (Q)	Clustering Coefficient (C)
PPI1		-0.1059	0.1908
PPI2		-0.2192	0.0437
PPI3		-0.1929	0.1256
PPI4		-0.1858	0.0853
Brain		0.5108	0.4637
ECO1		-0.1740	0.0376
ECO2		-0.0216	0.1082
UCIonline		-0.1880	0.1097
Epoch		-0.2583	0.2567
Epinions		0.6021	0.8085
ATN		-0.0951	0.1953
SF	SF	-0.0530	0.0082
	SFC1	-0.0348	0.0506
	SFC2	0.0022	0.1012
	SFC3	0.0918	0.1503
	SFQ1	-0.1855	0.0002
	SFQ2	-0.1065	0.0070
	SFQ3	0.1182	0.0044
	SFQ4	0.2258	0.0054

Table 5: **Degree correlations (Q) and clustering (C) of our model and empirical networks.** We measured Q and C from our set of empirical networks. Many of our networks features rather high levels of Q, C (compared to a random network), yet this had but a negligible effect on our results, as displayed in the main text. We also rewired our model scale-free network SF to increase its clustering (SFC1 - SFC3) and degree-correlations (SFQ1 - SFQ4) in a controlled fashion, allowing us to systematically examine the effect of Q and C on the performance of our formalism. Note that rewiring a random network to increase clustering also impacts its degree correlations and vice versa.

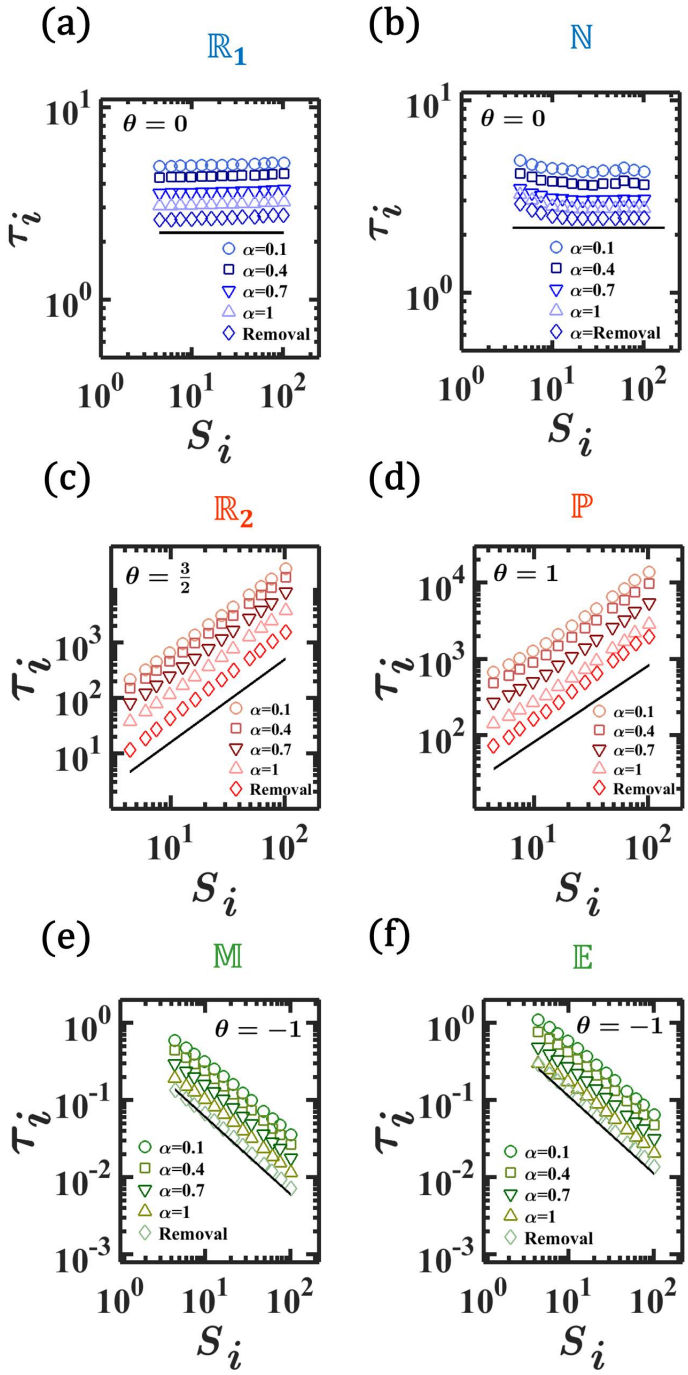

Figure 11: **The effect of perturbation size**. To test the limits of our linear response framework we measured τ_i vs. S_i, as obtained for large signals, representing an $\alpha = 10\%$ (circles), 40% (squares), 70% (down-triangles) and 100% (up-triangles) perturbation. We also tested the scaling under complete node knockout (Removal, diamonds). We find that perturbation size has no visible effect on the macroscopic patterns of flow, with θ consistently adhering to its theoretically predicted value (solid lines).

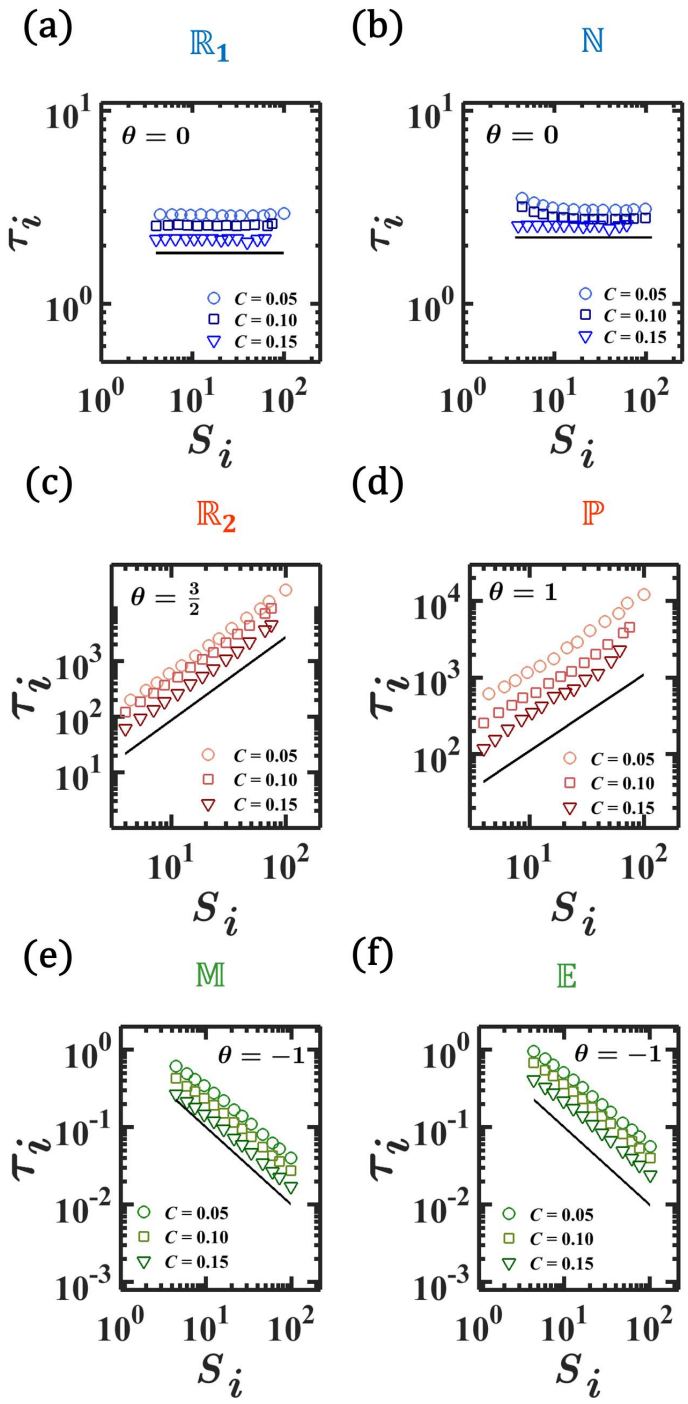

Figure 12: **The impact of clustering** C. τ_i vs. S_i as obtained from SFC1 - SFC3, featuring increasing levels of clustering $C = 0.05$ to 0.15. Despite the clustering, the predicted scaling in each dynamics (θ, solid lines) remains valid. Rewiring SF to increase C beyond these levels was numerically prohibitive without fragmenting the network. Still, some of our empirical networks exceeded these levels, *e.g.*, PPI1, Epoch or Brain.

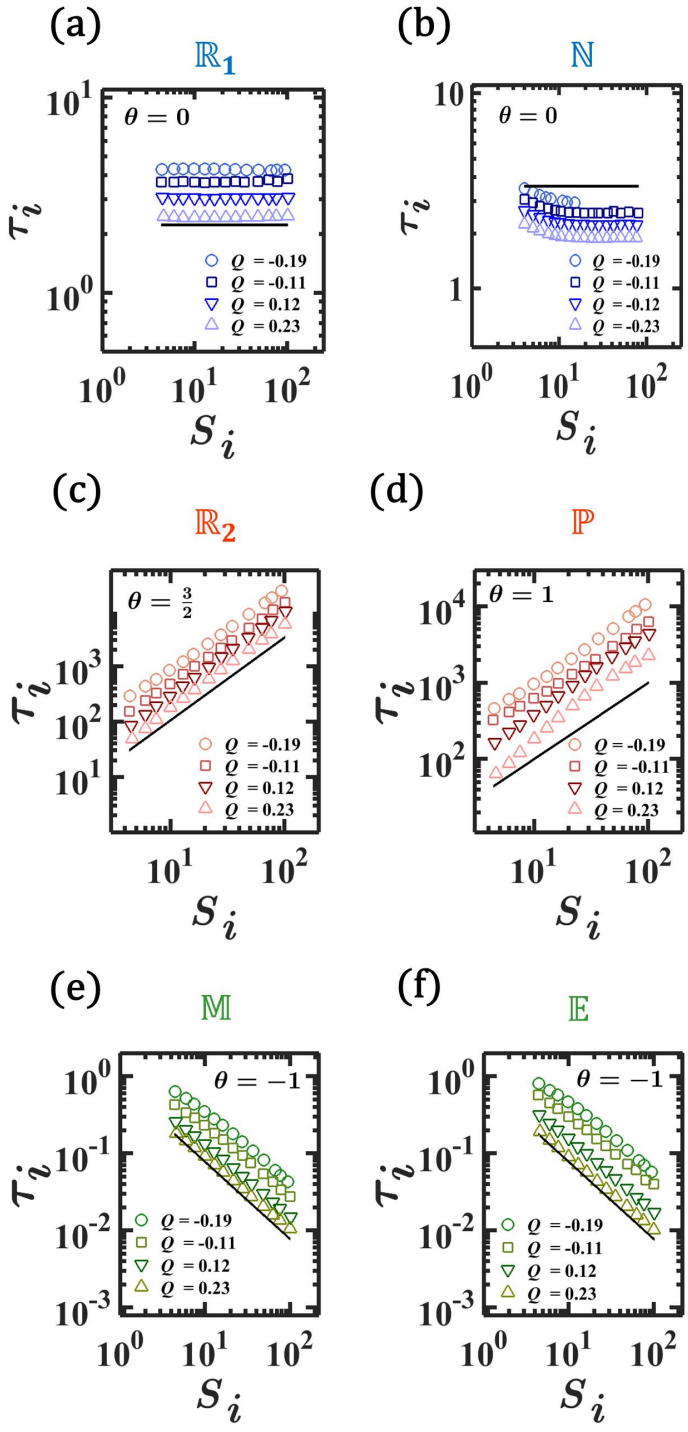

Figure 13: **The impact of degree correlations** Q. τ_i vs. S_i as obtained from SFQ1 - SFQ4, featuring negative and positive degree correlations. Despite these correlation levels the predicted scaling in each dynamics (θ, solid lines) remains valid.

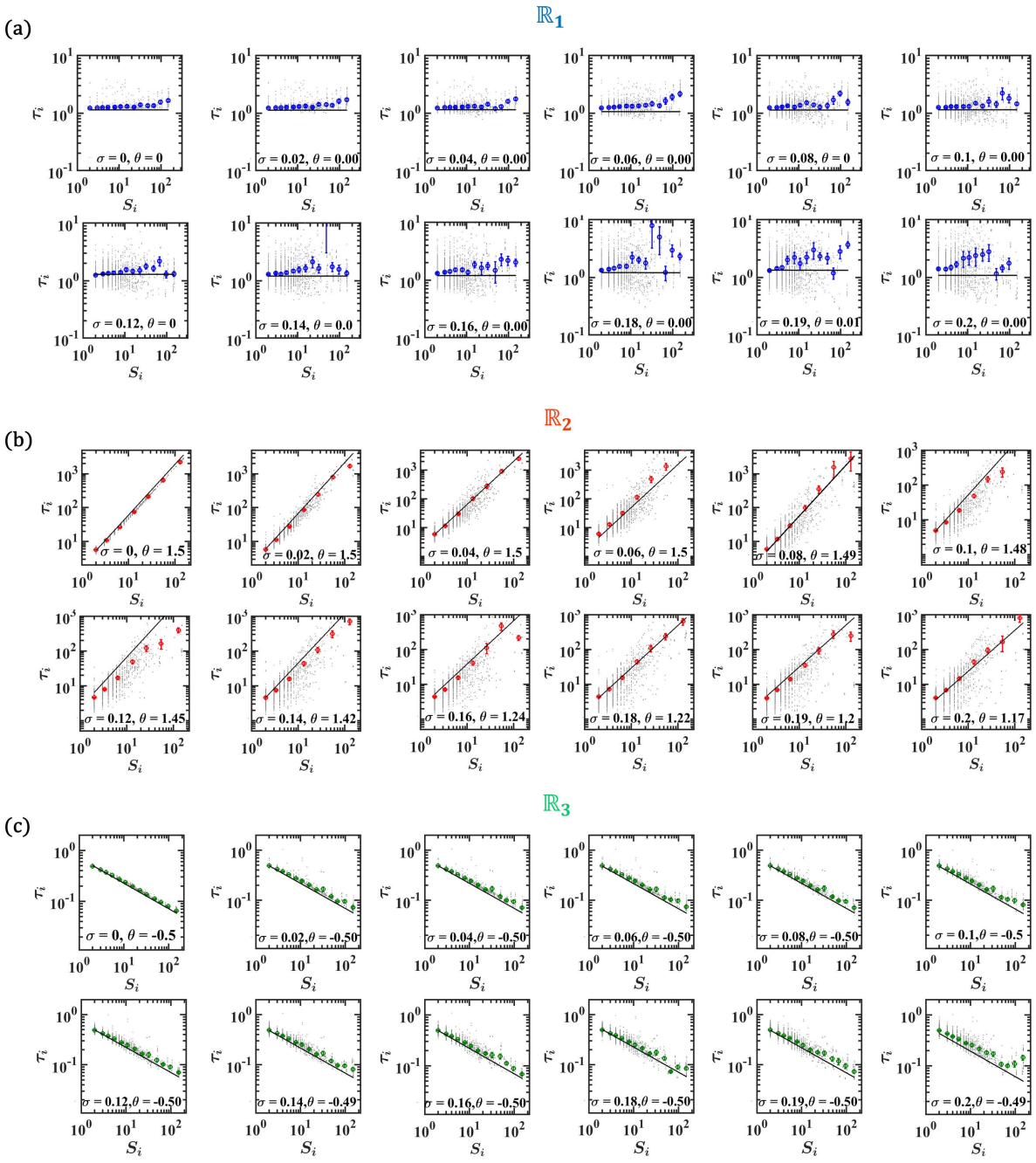

Figure 14: **The impact of parametric noise**. τ_i vs. S_i as obtained from Eq. (7.3), setting μ in (7.4) to $\mu = 1$ ($\mathbb{R}_1$, blue), $\mu = 0.4$ ($\mathbb{R}_2$, red) and $\mu = 2$ ($\mathbb{R}_3$, green). In each panel we show the raw data obtained for all individual nodes (grey dots), the logarithmically-binned data, capturing the average behavior (circles), with error bars calculated as in Sec. 3.3, and the fitted linear plot from (7.6). The noise level σ and the expected θ from (7.5) are also shown.

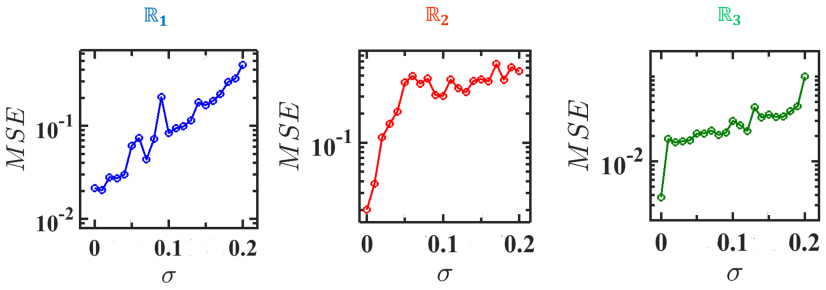

Figure 15: **Error induced by parametric noise**. The mean-squared-error (MSE) as obtained from Eq. (7.7) vs. σ for all simulations presented in Fig. 14, on (a) $\mathbb{R}_1$, (b) $\mathbb{R}_2$ and (c) $\mathbb{R}_3$ dynamics.

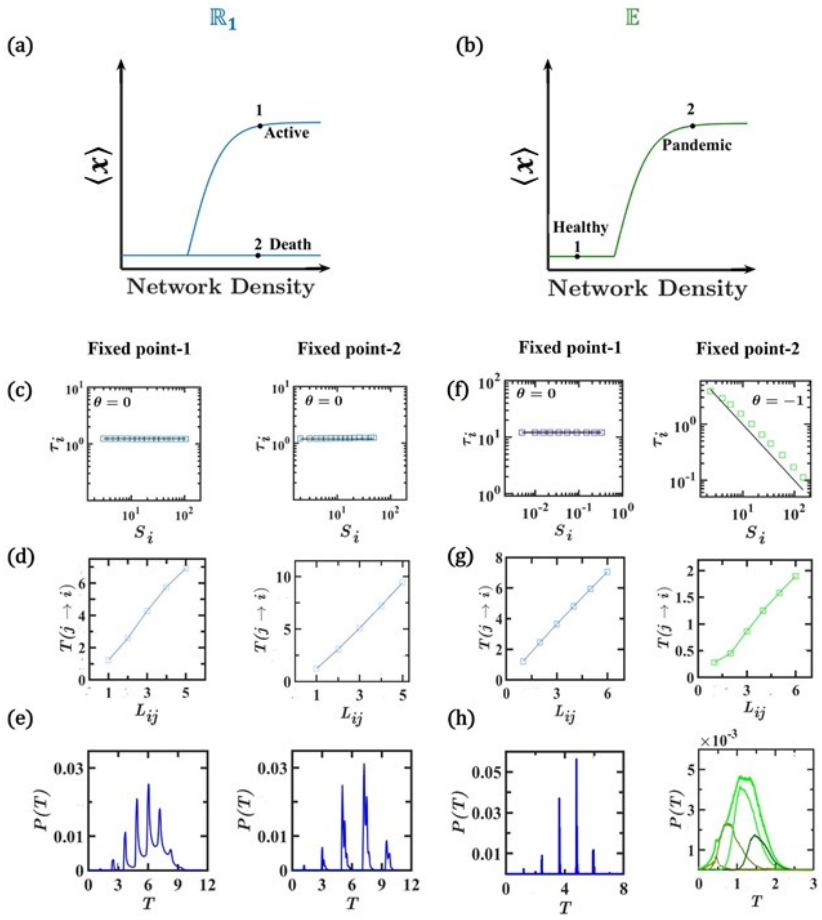

Figure 16: **State-dependent propagation patterns**. (a) $\mathbb{R}_1$ features to simulataneously stable fixed-points - active vs. dead. (b) $\mathbb{E}$ features a second order transition from a stable healthy state to a pandemic state. (c) - (e) In both states of $\mathbb{R}_1$ the propagation is in the $\theta = 0$ universality class, featuring distance limited propagation. (f) - (g) $\mathbb{E}$ is distance limited ($\theta = 0$) around the healthy fixed point (left), but transitions to composite ($\theta = -1$) around the pandemic state. Hence, interestingly, the state transition is also a propagation class transition.

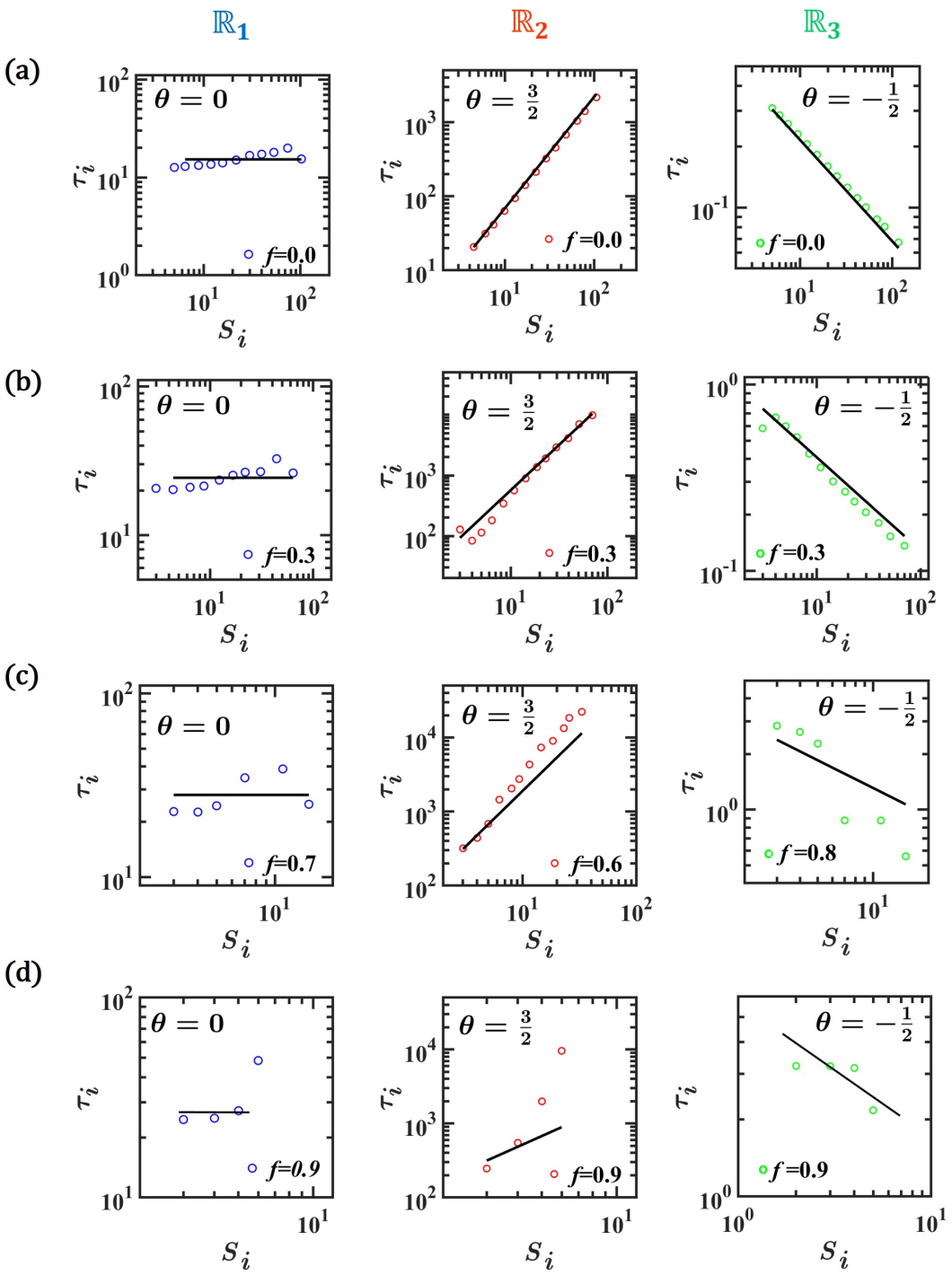

Figure 17: **The effect of hidden links**. (a) τ_i vs. S_i as obtained from a fully mapped network ($f = 0$) for $\mathbb{R}_1 - \mathbb{R}_3$. (b) The scaling continues to be observable even for a 70% mapped network ($f = 0.3$). (c) Hiding a large fraction of the links begins to impact the observed scaling at around $f = 0.7$ for $\mathbb{R}_1$, $f = 0.6$ for $\mathbb{R}_2$ and $f = 0.8$ for $\mathbb{R}_3$. (d) Hiding 90% of all links, we are no longer able to extract θ.

8 Empirical accessibility - complementary discussion

Despite our continuously growing access to data on complex networked systems, we continue to lack complete empirical freedom in some of the relevant systems, limiting our ability to observe the detailed signal propagation patterns. Below we consider several scenarios, from best to worst case, and discuss, for each scenario, the limits and the possibilities of empirically observing our predicted signal propagation.

Known network and controlled perturbations. With access to the complete (weighted) network and with the ability to apply controlled perturbations to selected nodes, we can empirically measure the entire set of analytical predictions. Indeed, perturbing each node sequentially, we can measure its direct neighbor response times τ_i, as well as the propagation times $T(j \to i)$ to distant nodes. Knowing the network we can now empirically observe the scaling $\tau_i \sim S_i^{\theta}$, and calculate the temporal distances $\mathcal{L}(j \to i)$ • Such best case scenario is difficult to attain in a fully empirical setting, however our simulated global epidemic spreading (Sec. 5) illustrates its potential merits. Here all weighted links can be empirically calculated from data on human travel and the disease parameters can be selected within a well-documented realistic range, allowing us to capture the propagation in a highly realistic setting.

Partially known network and controlled perturbations. In many cases, we are confronted with hidden links/unknown weights, confounding our ability to fully access the propagation. As discussed and demonstrated in Sec. 7.5 this has no bearing on observing $P(T)$ or $\langle T \rangle$, and to a limited extent may also allow us to observe the universal exponent θ (Fig. 17).

Unknown network. If we lack all access to A_{ij} we cannot identify direct neighbors, and hence we are unable to directly measure τ_i. Under these conditions we can only observe macroscopic properties, such as $P(T)$ and $\langle T \rangle$, which do not require knowledge of the network. Fortunately $P(T)$ provides a distinctive fingerprint to distinguish between positive (fat-tailed), zero (discrete peaks) or negative (overlapping peaks) θ. It cannot, however provide the specific value of *theta*, just the relevant dynamic regime, *i.e.* θ greater/smaller than or equal to zero.

Uncontrolled perturbations. Often we are unable to induce controlled perturbations to the system, such as in social or infrastructure networks. Under these conditions we must rely on data collected following spontaneously occurring outbreaks or failures. This limits our ability to measure τ_i, which requires us to systematically perturb each node's nearest neighbors, but still allows us to observe $P(T)$, $\langle T \rangle$ and $T(j \to i)$ This scenario is demonstrated in Sec. 6, where we examine the propagation of a local failure in a power network - a challenging scenario, in which we have access to just a single uncontrolled perturbation, whose spatio-temporal propagation affected only a small sample of nodes, providing limited statistics.

References

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

[2] M.E.J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701 – 4, 2002.

[3] L. Schmetterer and K. Sigmund (Eds.). *Hans Hahn Gesammelte Abhandlungen Band 1/Hans Hahn Collected Works Volume 1*. Springer, Vienna, Austria, 1995.

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

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

[6] 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.

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

[8] B. Barzel and O. Biham. Binomial moment equations for stochastic reaction systems. *Phys. Rev. Lett.*, 106:150602–5, 2011.

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

[10] R.M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.

[11] C.S. Holling. Some characteristics of simple types of predation and parasitism. *The Canadian Entomologist*, 91:385–398, 1970.

[12] M. Stern, H. Sompolinsky and L. F. Abbott. Dynamics of random neural networks with bistable units. *Phys. Rev. E*, 90:062710, 2014.

[13] M. Bastian, S. Heymann and M. Jacomy. Gephi: an open source software for exploring and manioulating networks. In *International AAAI Conference on Weblogs and Social Media. Associaltion for Advancement of Artificial Intelligence*, 2009.

[14] S. Milojević. Power-law distributions in information science: making the case for logarithmic binning. *Journal of the American Society for Information Science and Technology*, 61:2417–2425, 2010.

[15] D.R. Cox and D.V. Hinkley. *Theoretical Statistics*. Chapman & Hall, London, 1974.

[16] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

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

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

[19] J. Tang, T. Lou, J. Kleinberg, and S.Wu. Transfer link prediction across heterogeneous social networks. *ACM Trans. Inf. Syst*, 9(4):39:1–43, 2010.

[20] P. Massa, M. Salvetti, and D. Tomasoni. Bowling alone and trust decline in social network sites. *2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, page 658.

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

[22] J.F. Rual *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, 437:1173–1178, 2005.

[23] K. Ikehara and Aaron Clauset. Characterizing the structural diversity of complex networks across domains. *arXiv:1710.11304v1*, pages 1–23.

[24] M. De Domenico, V.Nicosia, A. Arenas and V. Latora. Structural reducibility of multilayer networks. *Nature Communications*, 6:6864, 2015.

[25] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10:186 – 198, 2009.

[26] Interaction web database. http://www.nceas.ucsb.edu/interactionweb/resources.

[27] R. Pastor-Satorras, C. Castellano, P. Van Mieghem and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–958, 2015.

[28] D. Brockmann and D. Helbing. The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342:1337–1342, 2013.

[29] OAG Ltd. 2018-06-11. *https://www.oag.com/.*

[30] North American Electric Reliability Council. 1996 System Disturbances: Review of Selected 1996 Electric System Disturbances in North America. Princeton, New Jersey, 2002.

[31] L. Daqing, J.Yinan, K.Rui and S. Havlin. Spatial correlation analysis of cascading failures: Congestions and blackouts. *Nature Communications*, 4:5381, 2014.

[32] S. Kauffman. The ensemble approach to understand genetic regulatory networks. *Physica A*, 340:733–740, 2004.

[33] K.G. Wilson. The renormalization group: Critical phenomena and the Kondo problem. *Rev. Mod. Phys.*, 47:773, 1975.

[34] B. Barzel and A.-L. Barabási. Universality in network dynamics. *Nature Physics*, 9:673 – 681, 2013.

[35] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47, 2002.

[36] J. Gao, B. Barzel and A.-L. Barabási. Universal resilience patterns in complex networks. *Nature*, 530:307–312, 2016.