NETWORK SCIENCE

Network structural origin of instabilities in large complex systems

Chao Duan[1,2]†, Takashi Nishikawa[2,3]*†, Deniz Eroglu[2,4], Adilson E. Motter[2,3]

A central issue in the study of large complex network systems, such as power grids, financial networks, and ecological systems, is to understand their response to dynamical perturbations. Recent studies recognize that many real networks show nonnormality and that nonnormality can give rise to reactivity—the capacity of a linearly stable system to amplify its response to perturbations, oftentimes exciting nonlinear instabilities. Here, we identify network structural properties underlying the pervasiveness of nonnormality and reactivity in real directed networks, which we establish using the most extensive dataset of such networks studied in this context to date. The identified properties are imbalances between incoming and outgoing network links and paths at each node. On the basis of this characterization, we develop a theory that quantitatively predicts nonnormality and reactivity and explains the observed pervasiveness. We suggest that these results can be used to design, upgrade, control, and manage networks to avoid or promote network instabilities.

INTRODUCTION

The dynamical stability of large complex network systems is an intriguing problem. The basic question of what properties of such systems govern their stability has attracted much interest, which was initially sparked by a 1972 article by Robert May predicting that sufficiently large systems should be linearly unstable despite the observed stability of large real ecological systems (*1*). While the literature on this problem and discrepancies between theory and observation has focused mostly on May's original context [ecological networks (*2*–*7*), including microbiome communities (*8*–*10*)], the problem is relevant for large network systems in general, including financial networks (*11*, *12*), power networks (*13*), and immune system networks (*13*). The problem acquires a new dimension when the Jacobian matrix M determining the linear stability is nonnormal (*14*) (i.e., $MM^T \neq M^TM$, where M^T denotes the transpose of M). This is because a small perturbation in such a system can cause the resulting state deviation to initially grow and become large enough to excite nonlinear instabilities, even when the system is linearly stable (*14*–*17*). The system's capacity to exhibit initial growth of deviations in the linear regime is termed reactivity (*15*), which is known to relate to a spectral property of the matrix M. Although the initial interest in nonnormality and reactivity emerged in hydrodynamics (*18*–*20*), these properties have recently gained attention in the study of network systems, including ecological (*5*, *21*), neuronal (*22*, *23*), chemical reaction (*17*), and communication networks (*24*), as well as in the study of pattern formation in networks (*25*, *26*) and control of networks (*27*). The literature has begun to reveal how prevalent nonnormality and reactivity are in real-world networks (*16*), but the fundamental question of which network structural mechanisms underlie the apparent prevalence of nonnormality and reactivity has not yet been addressed.

Here, we address this question by deriving rigorous conditions for nonnormality and reactivity that can be applied to any directed network [rather than to the average over an ensemble of networks (*5*)] and interpreted in terms of the structure of the given network. For nonnormality, the condition is that there is an imbalance between incoming and outgoing links at a node or a pair of nodes in terms of their numbers and/or weights. For reactivity, the condition is that there is an imbalance between the eigenvector centrality of a node associated with incoming network paths (including their weights) and the eigenvector centrality associated with outgoing paths. We use these conditions to show that, in a broad class of directed networks, the probability that the coupling matrices are both nonnormal and reactive approaches 1 quickly as the network size increases. We prove our results for large networks using a general network model that permits arbitrary distributions of possibly correlated in- and out-degrees (the number of incoming and outgoing links at a node, respectively) and arbitrary distributions of link weights. We also validate the prevalence of nonnormality and reactivity using a dataset of 251 real networks. This set is the largest and most diverse collection of directed networks, which also includes the largest networks (with up to nearly 8 million nodes), ever considered in this context.

These findings indicate that no additional global organization of connectivity is necessary to generically observe nonnormality and reactivity. This is important given that large-scale structures can have dynamical consequences, such as the stability promoted by trophic coherence in food-web networks (*28*) and the nonmonotonicity supported by linear chain structures in chemical reaction networks (*17*). In addition to establishing the prevalence of nonnormality and reactivity, we develop a quantitative theory that directly relates the extent of the degree and centrality imbalances in a given network to the extent of nonnormality and reactivity, respectively. Thus, our results reveal the network structural features responsible for nonnormality and reactivity, contributing to the much needed fundamental understanding of the relationship between dynamical and structural properties of directed networks (*29*–*39*).

RESULTS

Nonnormality and reactivity of network systems

Given that many real network systems operate near an equilibrium, we consider the class of (nonlinear) systems whose linearization around a given reference equilibrium state is described by

[1]School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, 710049, China. [2]Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA. [3]Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208, USA. [4]Department of Molecular Biology and Genetics, Kadir Has University, 34083 Istanbul, Turkey.
*Corresponding author. Email: tnishik21@gmail.com
†These authors contributed equally to this work.

$$\dot{x}_i = -\alpha_i x_i + \sum_{j=1}^{n} A_{ij} x_j \quad (1)$$

for $i = 1, \ldots, n$, where x_i is the (scalar) deviation from the reference state for the ith node. The matrix $A = (A_{ij})$ can be regarded as the weighted adjacency matrix of the system's directed interaction network: $A_{ij} \neq 0$ if node j is connected to node i, and $A_{ij} = 0$ otherwise. The parameter α_i and the diagonal element A_{ii} represent the node dynamics and any self-link at node i, respectively. We assume that the time scales of the dynamics in Eq. 1 are much shorter than those of the evolution of the interaction network structure, so that A_{ij} can be regarded as constant. For concreteness, we also assume $\alpha_i = \alpha$ for all i in the following unless otherwise indicated. We note, however, that our results on random networks are valid for heterogeneous α_i and that the presence of the heterogeneity is generally expected to increase both the nonnormality and the reactivity of the Jacobian matrix of the system (see Materials and Methods for details). For $\alpha_i = \alpha$, the Jacobian matrix M and the adjacency matrix A are related as $M = A - \alpha I_n$, with I_n denoting the $n \times n$ identity matrix. Thus, the reference state is asymptotically stable if and only if $\mathrm{Re}\,\lambda_1(M) < 0$, or $\mathrm{Re}\,\lambda_1(A) < \alpha$, where $\lambda_1(X)$ denotes the eigenvalue with the largest real part for any matrix X. Assuming $A_{ij} \geq 0$ for $i \neq j$, as observed in many real networks, $\lambda_1(A)$ is guaranteed to be real by the Perron-Frobenius theorem for nonnegative matrices (*40*), regardless of whether the network is strongly connected (i.e., any two nodes are connected by directed paths in both directions). While we have implicitly assumed one-dimensional node dynamics in Eq. 1 for clarity, we also establish an exact relation between the Jacobian and adjacency matrices for a broader class of systems. In particular, this relation shows how the nonnormality and reactivity of the adjacency matrix generically imply the same properties for the Jacobian matrix (see Supplementary Materials, section S1, for details).

We first show how the nonnormality and reactivity of the system in Eq. 1 can be expressed as properties of the network structure described by A under the uniform α_i assumption. Noting that matrix $D := MM^T - M^TM$ represents the deviation of the Jacobian matrix M from being normal, the nonnormality of the system can be quantified by the Frobenius norm $\|D\|_F := \sqrt{\sum_i \sum_j |D_{ij}|^2}$ (*41*). Since $D = MM^T - M^TM = AA^T - A^TA$ (and hence does not depend on α), the nonnormality of M is reduced to the nonnormality of A. To characterize the reactivity of the system in Eq. 1 mathematically, we consider the maximum exponential rate of initial growth of the state deviation vector that can result from a perturbation of the initial state. This rate is given by $\lambda_1((M + M^T)/2) = \lambda_1(H) - \alpha$, where $H := (A + A^T)/2$ is the symmetric part of A. Thus, the state deviation can grow if $\lambda_1(H) > \alpha$. Combining this with the stability requirement $\lambda_1(A) < \alpha$, we define reactivity as a property of the interaction network structure A (including its dependence on the reference state): A is said to be reactive if there are α values for which $\lambda_1(H) > \alpha > \lambda_1(A)$. That is, A is reactive if, within the linear regime, the system can be both stable and capable of exhibiting initial growth of state deviations. Such growth can push the system state out of the region in which the linearization in Eq. 1 is valid, potentially inducing nonlinear instabilities. The reactivity condition can be expressed as

$$\lambda_\Delta(A) := \lambda_1(H) - \lambda_1(A) > 0 \quad (2)$$

and thus we use $\lambda_\Delta(A)$ as a measure of the reactivity of the interaction network structure A [this definition can be extended to allow for negative weights by replacing $\lambda_1(A)$ with $\mathrm{Re}\,\lambda_1(A)$]. This measure is independent of α, in contrast to the reactivity for a specific α, which could be defined as $\lambda_1(H) - \alpha$ following previous studies (*14*, *15*, *19*, *20*). In the general case of heterogeneous α_i, both nonnormality and reactivity can be defined in the same way after absorbing the heterogeneity into the diagonal elements of the matrix A (see Materials and Methods for details).

We note that $\lambda_\Delta(A) \geq 0$ always holds (see Eq. 7 in Materials and Methods) and $\lambda_\Delta(A) > 0$ implies that the initial growth rate of state vector deviation $\frac{d\|x\|}{dt}\Big|_{t=0} \Big/ \|x_0\|$ in Eq. 1 for $\alpha = 0$ can be strictly larger than $\lambda_1(A)$ when the initial state vector x_0 is chosen to be the eigenvector corresponding to $\lambda_1(H)$, where we use $\|\cdot\|$ to denote the 2-norm and $x := (x_1, \ldots, x_n)^T$ to denote the state vector. This is the case because $\frac{d\|x\|}{dt}\Big|_{t=0} \Big/ \|x_0\| = x_0^T H x_0 \Big/ \left(x_0^T x_0\right) = \lambda_1(H) > \lambda_1(A)$. We also note that $\lambda_\Delta(A)$ is invariant under any coordinate transformation if we concurrently apply the transformation to the observable for the system in Eq. 1, since $\lambda_\Delta(A)$ is based on the value of the observable rather than its coordinate-specific representation. For the reactivity measure defined in Eq. 2, the observable is the 2-norm of the system state vector. This choice is standard in the literature [see, e.g., (*5*, *14*–*16*, *19*, *21*, *25*, *26*)], as it permits a convenient characterization through eigenvalues, but for certain network processes, different measures of deviations may be more natural and would lead to different definitions of reactivity [see, e.g., the 1-norm used in (*42*–*44*) and the absolute value of a one-dimensional projection used in (*17*)].

We used the conditions and measures just defined to study nonnormality and reactivity in a large dataset of 251 real directed networks, which consists of 63 biological, 51 informational, 79 social, 39 technological, and 19 economic/game networks and avoids repetition of similar networks. The dataset used here substantially expands on an earlier study (*16*) with respect to the number of networks, the largest network size, and the diversity of network types. We verified that A is nonnormal (i.e., $\|D\|_F > 0$) for all 251 networks and is reactive [i.e., $\lambda_\Delta(A) > 0$] for all but one network (see Materials and Methods for details on the network data and our findings). Quantitatively, Fig. 1A reveals a strong positive correlation between the level of nonnormality $\|D\|_F$ and the overall level of imbalance between the nodes' in- and out-degrees, measured by the average $\langle (d^{\mathrm{in}} - d^{\mathrm{out}})^2 \rangle := \sum_i \left(d_i^{\mathrm{in}} - d_i^{\mathrm{out}}\right)^2 / n$, where d_i^{in} and d_i^{out} are the in- and out-degrees of node i, respectively (i.e., the numbers of incoming and outgoing links from/to other nodes). Similarly, Fig. 1B shows a monotonic relation between the reactivity $\lambda_\Delta(A)$ and a measure of imbalance, in this case between the eigenvector centrality associated with the incoming and outgoing paths, quantified by the angle θ_1 between the left and right eigenvectors corresponding to $\lambda_1(A)$ (to be precisely defined below). Both relations are observed for each type of networks (fig. S1) and will be theoretically derived and computationally validated in a later section.

Imbalance conditions for nonnormality and reactivity

To understand the mechanisms underlying the correlations observed above, we first examine the network structural features that are responsible for the nonnormality and reactivity of A. The measure of nonnormality introduced above can be expressed as

$$\|D\|_F^2 = \sum_{i=1}^{n} \left(\delta_i^{\mathrm{in}} - \delta_i^{\mathrm{out}}\right)^2 + 2 \sum_{i<j} \sum \left(d_{ij}^{\mathrm{in}} - d_{ij}^{\mathrm{out}}\right)^2 \quad (3)$$

where $\delta_i^{\mathrm{in}} := \sum_{k \neq i} A_{ik}^2$, $\delta_i^{\mathrm{out}} := \sum_{k \neq i} A_{ki}^2$, $d_{ij}^{\mathrm{in}} := \sum_k A_{ik} A_{jk}$, and $d_{ij}^{\mathrm{out}} := \sum_k A_{ki} A_{kj}$. The variables δ_i^{in} and δ_i^{out} can be regarded as

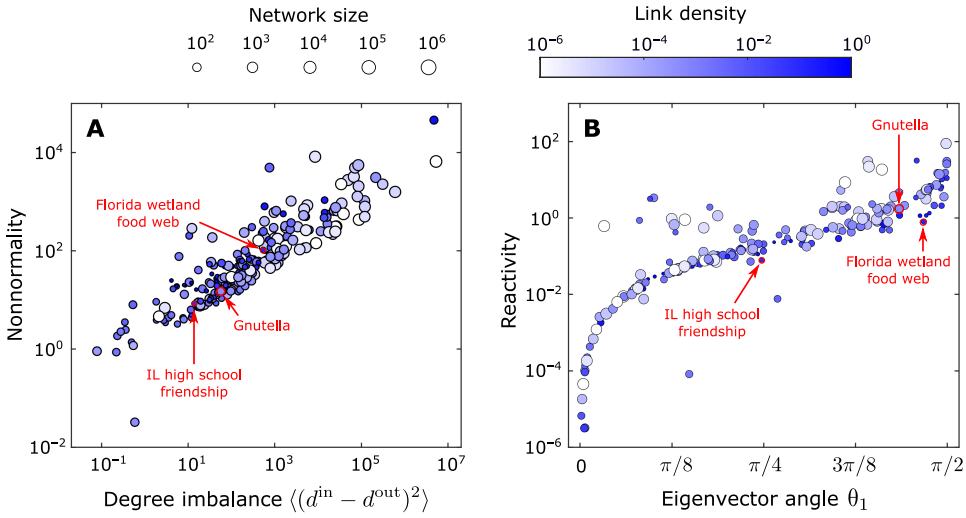

Fig. 1. Nonnormality and reactivity of real networks. (**A**) Nonnormality versus the in- and out-degree imbalance quantified by the average $\langle (d^{in} - d^{out})^2 \rangle$ of the squared difference between d_i^{in} and d_i^{out} over all nodes i for the 251 real networks considered (see Materials and Methods for details on the data). The measure of nonnormality $\|D\|_F$ is normalized by $\sqrt{n}\,\langle w^2 \rangle$ based on the scaling predicted by our theory for random networks, where $\langle w^2 \rangle := \sum_i \sum_j A_{ij}^2/(n\bar{d})$ is the mean squared weight over all links in the network and $\bar{d}$ is the average degree; see Eq. 22. The size and color of a circle indicate the network size and link density, respectively. (**B**) Reactivity versus the in- and out-centrality imbalance quantified by the angle θ_1 (in radian) between the left and right eigenvectors associated with the largest eigenvalue of A. The measure of reactivity $\lambda_\Delta(A)$ is normalized by $\lambda_1(A)$ based on our theory in Eq. 5. The plot shows only the 212 nondegenerate reactive networks [i.e., those having nondegenerate $\lambda_1(A)$ and satisfying $\lambda_\Delta(A) > 0$; see Materials and Methods for details]. We observe appreciable levels of nonnormality and reactivity in a vast majority of the networks: $\|D\|_F/(\sqrt{n}\,\langle w^2 \rangle) > 1$ for 248 networks (≈99%) and $\lambda_\Delta(A)/\lambda_1(A) > 0.01$ for 188 nondegenerate networks (≈88%). IL, Illinois.

generalized in- and out-degrees, respectively, since they reduce to d_i^{in} and d_i^{out} for unweighted networks (i.e., if $A_{ij} \in \{0,1\}$). For distinct nodes i and j, the variable d_{ij}^{in} (d_{ij}^{out}) can be interpreted as a further generalization of the in-degree (out-degree) to a pair of nodes, and it reduces to the number of common in-neighbors (out-neighbors) shared by the two nodes in the case of unweighted networks. Equation 3, despite being immediate from the definition of D and its Frobenius norm, provides an insightful decomposition of nonnormality into two types of imbalances between incoming and outgoing links: The first term is the square sum of the generalized in- and out-degree differences at individual nodes, while the second is an analogous square sum for node pairs. From Eq. 3, we see that A is normal if and only if the generalized in- and out-degrees are equal for each node and for each pair of nodes (see Fig. 2A for an illustrative example). It is thus sufficient to have just a single node whose in- and out-degrees differ in order to make A nonnormal. However, even without such a node, A can still be nonnormal if there is a node pair with an imbalance between their common weighted in-neighbors and the common weighted out-neighbors (making the second sum nonzero in Eq. 3), as illustrated in Fig. 2B. Furthermore, Eq. 3 shows that a larger total imbalance between incoming and outgoing links implies a more nonnormal A.

For reactivity, a stronger condition is needed, since nonnormal A does not have to be reactive, as illustrated by the network in Fig. 2B. In particular, for a d-regular network, defined as an unweighted network in which the in- and out-degrees are all equal to a constant integer $d \geq 0$, the adjacency matrix A can be nonnormal but can never be reactive. To see this, we first note that $\lambda_1(A)$ is bounded between the minimum and maximum row sum of A because this eigenvalue is the spectral radius of A [see, e.g., theorem 8.1.22 in (*40*)] and thus is equal to d for any d-regular network. Since the same argument applies to the symmetric part of A, we have $\lambda_1(H) = d$, implying that $\lambda_\Delta(A) = d - d = 0$, i.e., A is nonreactive. Nonetheless, for many d-regular networks, the second term in Eq. 3 is strictly positive, rendering A nonnormal.

Given that nonnormality is only a necessary condition for reactivity, we now present a condition guaranteeing reactivity: A is reactive if the left and right eigenspaces associated with its largest eigenvalue $\lambda_1(A)$ are distinct (see Materials and Methods for a proof). In the generic case in which these eigenspaces are one dimensional, the reactivity condition is equivalent to having a strictly positive angle θ_1 between the right eigenvector v_1 and left eigenvector u_1. We always choose the acute angle so that $\theta_1 \leq \pi/2$, where the quantity $1/|\cos\theta_1|$ is known as the eigenvalue condition number in the literature (*14*). This condition is illustrated by the example network in Fig. 2C. While the nonorthogonality of right eigenvectors defining nonnormality has been shown to often lead to reactivity [e.g., in (*45*, *46*)], our condition $\theta_1 > 0$ captures reactivity more precisely, as it is equivalent to the specific nonorthogonality between the right eigenvector v_1 and some other right eigenvector. The condition $\theta_1 > 0$ also translates to the existence of at least one node i for which $v_{1i} \neq u_{1i}$, where v_{1i} is the eigenvector in-centrality, defined to be the ith component of the right eigenvector v_1 (and hence associated with incoming paths to the node), and u_{1i} is the eigenvector out-centrality, defined similarly through the left eigenvector u_1 (and the outgoing paths). The inequality $v_{1i} \neq u_{1i}$ can thus be interpreted as an imbalance between the incoming and outgoing "flow" of centrality. In the example of d-regular networks above, if $\lambda_1(A)$ is nondegenerate (so that v_1 and u_1 are unique and θ_1 is well defined), we find that $v_{1i} = u_{1i}$ for all i and thus $\theta_1 = 0$, as expected from the result above that A is nonreactive for all d-regular networks. In general, the imbalances $\delta_i^{in} - \delta_i^{out}$ and $v_{1i} - u_{1i}$ for individual nodes and $d_{ij}^{in} - d_{ij}^{out}$ for pairs of nodes can be either positive or negative and tend to be distributed heterogeneously across a given network, as illustrated by the three example real networks in fig. S2.

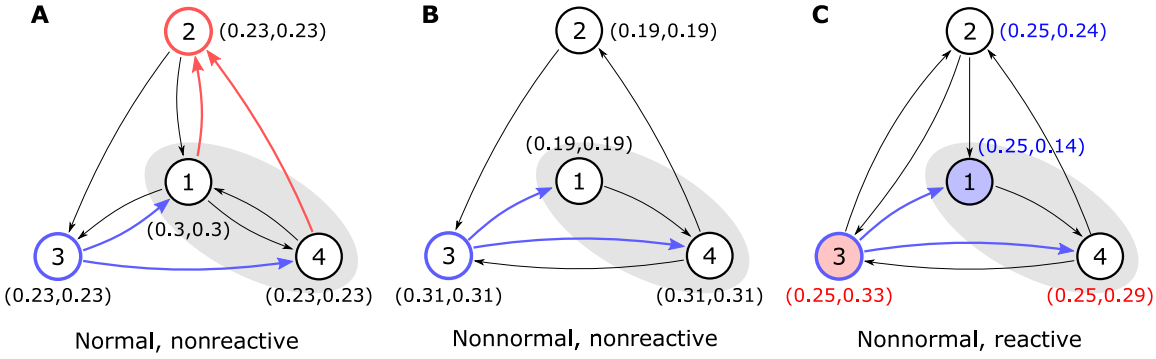

Fig. 2. Topological and spectral features inducing nonnormality and reactivity. The four-node unweighted example networks shown illustrate all three possible cases (the normal-reactive case is not possible). The eigenvector in-centrality v_{1i} and the eigenvector out-centrality u_{1i}, normalized so that $\|v_1\| = \|u_1\| = 1$, are indicated near each node i as (v_{1i}, u_{1i}). Note that we have $\delta_i^{\text{in}} = d_i^{\text{in}}$ and $\delta_i^{\text{out}} = d_i^{\text{out}}$ as the networks are unweighted. (**A**) Network for which A is normal (and thus nonreactive) because the in- and out-degrees match ($d_i^{\text{in}} = d_i^{\text{out}} = 2$) for all nodes i and the number of shared in- and out-neighbors match ($d_{ij}^{\text{in}} = d_{ij}^{\text{out}} = 1$) for all pairs of nodes i and $j \neq i$, as illustrated for $i = 1, j = 4$ (blue, shared in-neighbor; red, shared out-neighbor). (**B**) Network for which A is nonnormal because the second sum in Eq. 3 is nonzero due to $d_{ij}^{\text{in}} \neq d_{ij}^{\text{out}}$ for some $i \neq j$, e.g., $d_{14}^{\text{in}} = 1 \neq d_{14}^{\text{out}} = 0$. However, A is nonreactive because the in- and out-centralities are balanced at each node, i.e., $v_{1i} = u_{1i}$ for all i. (**C**) Network for which A is nonnormal and reactive. Nonnormality in this case is guaranteed not only by $d_{ij}^{\text{in}} \neq d_{ij}^{\text{out}}$ for some $i \neq j$ (e.g., for $i = 1, j = 4$) but also by $d_i^{\text{in}} \neq d_i^{\text{out}}$ for some i (blue node, $d_i^{\text{in}} > d_i^{\text{out}}$; red node, $d_i^{\text{in}} < d_i^{\text{out}}$). Reactivity is guaranteed by $v_{1i} \neq u_{1i}$ for all nodes (blue numbers, $v_{1i} > u_{1i}$; red numbers, $v_{1i} < u_{1i}$).

While we focus mainly on the adjacency matrix A, different interaction matrices can also be considered in Eq. 1. This includes the Laplacian matrix, which we show can be nonnormal even if A is not and vice versa (a fact that has been generally overlooked; see fig. S3 for illustrative examples).

Prevalence of nonnormality and reactivity

We now address the question of how often nonnormality and reactivity are expected to be observed by considering a model of random directed weighted networks with a given number of nodes n, a given joint probability distribution for the in- and out-degrees, and a given distribution of link weights. By allowing for arbitrary distributions for the in- and out-degrees, an arbitrary correlation between them, and an arbitrary distribution of weights, the model is capable of capturing the essential elements of real networks involved in the conditions established above for nonnormality and reactivity. This model is a generalization of the Chung-Lu-Vu model studied in (*47*) and allows us to specify a general joint distribution of in- and out-degrees (rather than the expected node degrees) and a general weight distribution. A network realization under this model is generated as follows. First, the (possibly non-integer) expected in- and out-degrees of each node are randomly drawn from the given joint distribution. Then, for each i and j, a directed link is created from node j to node i with a probability proportional to the product of the expected in-degree of node i and the expected out-degree of node j. Last, a random weight A_{ij} is drawn from the given distribution for each link $j \to i$. We note that a special case of this model produces unweighted networks, and self-links can be excluded if desired. For further details on this generalized Chung-Lu-Vu (GCLV) model, see Materials and Methods.

For weighted networks generated by the GCLV model, we show that the adjacency matrix A is almost always nonnormal in the limit of large network size n (see Supplementary Materials, section S2, for a proof for unweighted networks and its extensions to weighted and Laplacian-coupled networks). More precisely, we show that the probability of having at least one node whose generalized in-degree and out-degree are different (implying nonnormality of A, as discussed above) tends to one as $n \to \infty$. In practice, the probability that A is nonnormal grows quickly and is very close to one even for networks with less than 10 nodes, as shown numerically for both weighted and unweighted networks in fig. S4A and table S1 for two classes of in-/out-degree distributions: (i) the gamma distribution, whose probability density function is $p(x) \sim x^{a-1}e^{-bx}$, $x > 0$, where a and b are parameters, and (ii) the Dirac delta distribution centered at d, which renders the model equivalent to the Erdős-Réyi (ER) networks with fixed mean degree d (and thus with n-dependent connection probability $p = d/n$). The same appears to hold true for other random network models, as verified in fig. S4A for the unweighted ER networks with fixed p and random d-regular networks and also verified in table S1 for the weighted versions of these random networks.

Nonnormality does not necessarily imply reactivity. For large networks, however, we can show that, if A is nonnormal, then it is also reactive in almost all cases. This is because the conditional probability $\mathbb{P}(A \text{ is reactive} \mid A \text{ is nonnormal}) \geq \mathbb{P}(A \text{ is reactive})$, as reactivity implies nonnormality, and because we show that $\mathbb{P}(A \text{ is reactive})$ approaches one as $n \to \infty$ for the GCLV model (see Supplementary Materials, section S2, for a proof, including the case of Laplacian-coupled networks). More precisely, we prove that the probability of having distinct left and right eigenspaces associated with $\lambda_1(A)$ (sufficient for reactivity, as noted earlier) converges to one. When estimated numerically for finite n, the actual probability that A is reactive and the conditional probability that A is reactive given that it is nonnormal are again very close to one even for small n, as shown in fig. S4 (B and C), respectively (which also shows similar results for two other random network models). All these results support the observation from Fig. 1 that the reactivity of A is prevalent among real networks.

Quantitative characterization of nonnormality and reactivity

To understand the correlations observed for the real networks in Fig. 1, we now derive theoretical estimates of the nonnormality $\|D\|_{\text{F}}$ and the reactivity $\lambda_\Delta(A)$ for random networks. For the nonnormality, we first consider networks generated by the unweighted GCLV model with no self-links (i.e., $A_{ij} \in \{0,1\}$ and $A_{ii} = 0$ for all i and j) given a fixed set of in-degrees d_i^{in} and out-degrees d_i^{out} for all

nodes. In this case, the first sum in Eq. 3 is fixed and proportional to the (constant) average single-node degree imbalance $\langle (d^{\text{in}} - d^{\text{out}})^2 \rangle$, while the second sum representing the imbalances at the node pair level is a random variable. Assuming $n \gg 1$ and approximating the second sum with its expected value (see Materials and Methods for details), we have

$$\begin{aligned} \|D\|_{\text{F}}^2 &\approx n\left\langle (d^{\text{in}} - d^{\text{out}})^2 \right\rangle + n\left[\left\langle (d^{\text{in}} - d^{\text{out}})^2 \right\rangle + 2\langle d^{\text{in}} d^{\text{out}} \rangle - 2\bar{d}\right] \\ &= 2n\left[\left\langle (d^{\text{in}} - d^{\text{out}})^2 \right\rangle + \langle d^{\text{in}} d^{\text{out}} \rangle - \bar{d}\right] \end{aligned} \tag{4}$$

where $\langle z \rangle \coloneqq \sum_i z_i/n$ denotes the average of z_i over nodes i and $\bar{d} \coloneqq \langle d^{\text{in}} \rangle = \langle d^{\text{out}} \rangle$ is the average degree. Thus, in the limit of large networks, nonnormality increases with the network size n as $\|D\|_{\text{F}} \sim n^{1/2}$, with an n-independent prefactor expressed as a simple function of the in- and out-degree imbalance $\langle (d^{\text{in}} - d^{\text{out}})^2 \rangle$, the in- and out-degree correlation $\langle d^{\text{in}} d^{\text{out}} \rangle$, and the average degree $\bar{d}$. We also derive extensions of this formula to the weighted GCLV model (allowing self-links) and to Laplacian-coupled networks (see Materials and Methods), which have additional terms and factors involving the diagonal elements A_{ii} and the statistics of link weights (see Eqs. 22, 32, and 36). For the reactivity $\lambda_\Delta(A)$, we make use of the observation that the spectral gap between the leading eigenvalue $\lambda_1(A)$ and the remaining eigenvalues $\lambda_2(A)$, …, $\lambda_n(A)$ is often large for the adjacency matrix A of large random networks [which is the case, e.g., for the ER networks (*48*) and for the GCLV model when the mean degree is large (*49*)]. For such networks, we expect $\lambda_1(H)$ to be well-approximated by $\lambda_1(H_1)$, where $H_1 \coloneqq (A_1 + A_1^T)/2$ is the symmetric part of the leading component $A_1 \coloneqq \lambda_1(A)\, v_1 u_1^T$ and u_1 and v_1 are respectively the left and right (column) eigenvectors associated with the leading eigenvalue $\lambda_1(A)$, normalized so that $\|v_1\| = 1$ and $u_1^T v_1 = 1$. Among the 212 real networks in Fig. 1B [which have nondegenerate $\lambda_1(A)$, ensuring that H_1 is well defined], we indeed observe $\lambda_1(H) \approx \lambda_1(H_1)$ for most, as shown in fig. S5A. For any (possibly weighted) network satisfying $\lambda_1(H) \approx \lambda_1(H_1)$ and having nondegenerate $\lambda_1(A)$ (even when it is not strongly connected), we derive a simple expression:

$$\lambda_\Delta(A) \approx \frac{1 - \cos\theta_1}{2\cos\theta_1} \cdot \lambda_1(A) \tag{5}$$

with a prefactor that depends monotonically on the angle θ_1 between the leading left and right eigenvectors u_1 and v_1 (see Materials and Methods for a derivation).

We find that Eqs. 4 and 5 provide good approximations, as validated in Fig. 3 for several classes of random networks generated by the GCLV model (green, orange, and red dots) as well as for the real networks used in Fig. 1 (blue dots). In addition, they capture the general tendency observed in Fig. 1 for nonnormality and reactivity to increase with the degree imbalance and the eigenvector angle in real networks.

Our results show that A is more nonnormal and reactive when the network is weighted. In particular, Eqs. 3 and 22 indicate that allowing for weights in random networks can only increase the probability of observing imbalances that induce nonnormality (i.e., the probability that $\|D\|_{\text{F}} \neq 0$) and that a larger variance for the weights leads to a larger extent of nonnormality. Likewise, the eigenvector condition and Eq. 5 indicate that the presence of weights, whose randomness would be reflected in the eigenvectors, is expected to increase in-/out-centrality imbalances, leading to a higher probability that A is reactive and to a larger extent of reactivity.

The tendency for A to be dominated by the leading eigenvalue $\lambda_1(A)$ and the corresponding eigen-component A_1 also explains why large directed networks are almost always both nonnormal and reactive. To see this, we first note that A_1 satisfies the following exact relation linking its nonnormality and reactivity to each other and to the eigenvector angle:

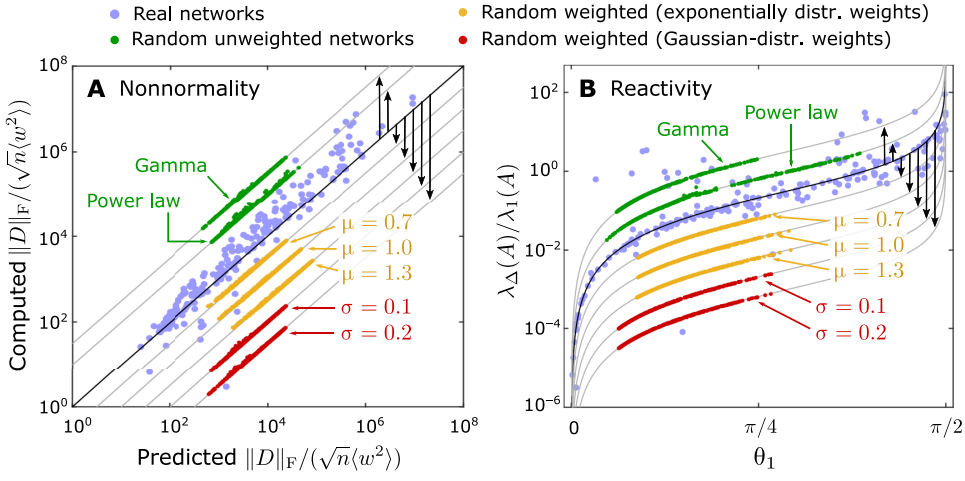

Fig. 3. Validating theoretical predictions for nonnormality and reactivity. (**A**) Computed nonnormality versus the corresponding prediction from Eq. 4 for the real networks in Fig. 1A (blue). Plots are also shown for random networks generated by the GCLV model: unweighted networks with the gamma or power-law distribution for the in- and out-degrees (green) and weighted networks with gamma-distributed degree distributions and exponentially distributed weights (orange; mean $\mu = 0.7$, 1.0, or 1.3) or Gaussian-distributed weights (red; mean one and standard deviation (SD) $\sigma = 0.1$ or 0.2), keeping only positive weights. For each class of networks, we used 20 equally spaced values of n between 10^3 and 10^5 on the logarithmic scale. For each n, we show 20 realizations of the model, with the parameters of the distribution drawn randomly for each network realization (see Materials and Methods for details). The continuous lines correspond to Eq. 22. The plots for the random networks and the corresponding lines are shifted vertically to avoid overlapping. (**B**) Reactivity versus the eigenvector angle θ_1. The continuous curves correspond to Eq. 5. The plot shows the same random networks as in (A) and the real networks in Fig. 1B, to which Eq. 5 is applicable.

$$\lambda_1^2 + \sqrt{\lambda_1^4 + 2\|D_1\|_F^2} = 2(\lambda_1 + 2\lambda_\Delta(A_1))^2 = \frac{2\lambda_1^2}{\cos^2\theta_1} \quad (6)$$

where we denote $\lambda_1 = \lambda_1(A_1) = \lambda_1(A)$ for brevity. Here, we note that A_1 and A have the same eigenvector angle θ_1 associated with λ_1, and we introduce $D_1 := A_1 A_1^T - A_1^T A_1$ to quantify the nonnormality of A_1. From Eq. 6, it immediately follows that, for the leading component A_1, nonnormality ($\|D_1\|_F > 0$), reactivity [$\lambda_\Delta(A_1) > 0$], and having a strictly positive eigenvector angle ($\theta_1 > 0$) are all mathematically equivalent to each other. Moreover, Eq. 6 shows that an increase in any one of these three measures implies an increase in all the other measures. Building on these observations, we find that random networks whose A is dominated by A_1 generally satisfy Eq. 6 approximately with $\|D_1\|_F$ and $\lambda_\Delta(A_1)$ replaced by $\|D\|_F$ and $\lambda_\Delta(A)$, respectively, which implies that nonnormality and reactivity are approximately equivalent for such networks. This finite-n approximate equivalence complements the rigorous results we established above in the limit of large network size. For the real networks in Fig. 3B, Eq. 6 approximately holds true when $\|D_1\|_F$ and $\lambda_\Delta(A_1)$ are replaced by $\|D\|_F$ and $\lambda_\Delta(A)$, respectively (fig. S5, B to D).

For networks satisfying Eqs. 4 and 6 approximately, we see that, if $\lambda_1 = \lambda_1(A)$ is bounded as n increases, nonnormality and reactivity would scale with n as $\|D\|_F \sim n^{1/2}$ and $\lambda_\Delta(A) \sim n^{1/4}$, respectively. For the real networks, $\lambda_\Delta(A)$ tends to increase with n (even when normalized by $\langle w^2 \rangle$, the root mean square of the link weights), which is a trend observed even more strongly when the networks are randomized while holding the in- and out-degrees fixed (Fig. 4A). For random networks with a power-law degree distribution $p(x) \sim x^{-\beta}$ and Gaussian-distributed link weights, the reactivity $\lambda_\Delta(A)$ scales with n with an exponent that depends on the power-law parameter β (Fig. 4B). The increase of $\lambda_\Delta(A)$ with n indicates that, as the system becomes larger, there will be a wider range of α for which the system can simultaneously exhibit a more pronounced transient response to a small perturbation and stronger linear stability. Specifically, Fig. 4C establishes that, if the node stability parameter α in Eq. 1 has an n-dependence $\alpha(n) \sim n^\ell$ with a constant ℓ, then there is a region in the β versus ℓ parameter space (shaded red) for which the maximum initial growth rate of perturbations given by $\lambda_1(H) - \alpha(n)$ increases with n even though the first-order dynamics become more stable [i.e., $\lambda_1(A) - \alpha(n)$ decreases].

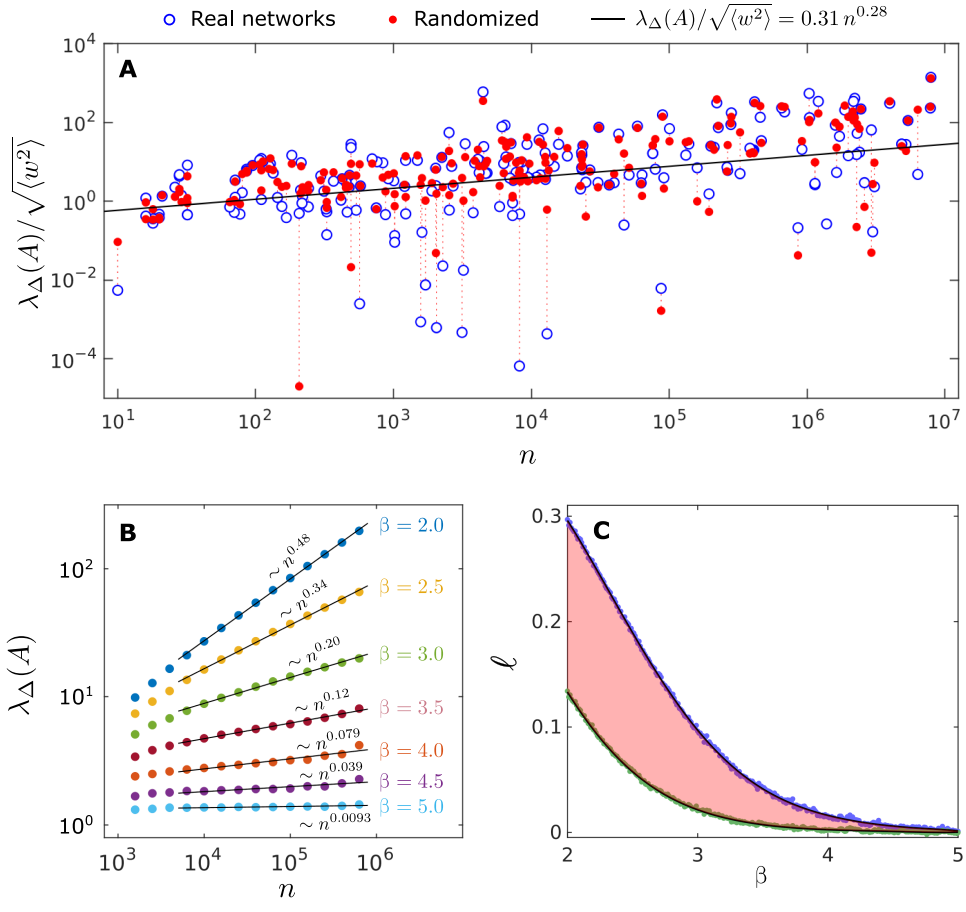

Fig. 4. Increase of reactivity with network size. (**A**) Reactivity versus network size n for the real networks in Fig. 3B (blue circles). We normalize $\lambda_\Delta(A)$ by $\sqrt{\langle w^2 \rangle}$ to facilitate comparison between networks with different scales for link weights. The black line indicates the least-squares fit (on the logarithmic scale), reflecting an increasing trend for the reactivity. Red dots indicate randomized versions of the real networks (see Materials and Methods for details), which more closely follow the trend line. (**B**) Scaling of $\lambda_\Delta(A)$ with n for the GCLV model with the same discrete power-law degree distribution $p(x) \sim x^{-\beta}$ used in Fig. 3, but for a given value of β and uncorrelated $\tilde{d}_i^{\text{in}}$ and $\tilde{d}_i^{\text{out}}$, with link weights drawn from the positive domain of the Gaussian distribution with mean one and SD 0.1. For each β, the plot shows the average over 200 network realizations. We observe that the scaling exponent for $\lambda_\Delta(A)$ depends on the parameter β of the degree distribution. (**C**) Reactivity–stability phase diagram for the system in Eq. 1 with an n-dependent parameter $\alpha = \alpha(n) \sim n^\ell$, where ℓ is a constant, and for the same power-law distribution with parameter β as in (B). Shaded in red is the region in which the maximum growth rate of state deviations $\lambda_1(H) - \alpha(n)$ increases with n despite the concurrent increase of linear stability [i.e., decrease of $\lambda_1(A) - \alpha(n)$]. See Materials and Methods for details on the procedure to identify the boundaries of this region.

Below this region the growth rate increases while linear stability decreases with n, which is doubly destabilizing, and above this region the opposite is observed.

DISCUSSION

Our demonstration that nonnormality and reactivity are stronger and more prevalent for larger networks implies that conventional modal stability analysis alone is not sufficiently informative: The system can be less stable against perturbations than the modal analysis indicates, and this gap can grow with the system size. Even if such a system is linearly stable and the perturbation is well within the linear regime, the transient response may bring the system sufficiently far from the equilibrium that nonlinear instabilities can be induced, possibly in the form of a cascade (*50*). Our findings thus suggest that, following a perturbation, there is a parameter region for which larger networks can exhibit larger oscillations in the linearized system and thus the possibility of a rare but substantial instability in the nonlinear system. Does such an instability cause a permanent transition to a different state? If so, to which state does the system transition? The answers to these questions depend on the size and direction of the perturbation, the extent of reactivity in the system, and the global structure of the state space. The latter, in particular, requires specific knowledge of the nonlinearity of the system under consideration.

Given our validation with extensive network data and general mathematical analysis, these conclusions are applicable to a wide range of real network systems and are thus not limited to the networks previously considered in addressing the complexity–stability problem. In particular, knowing that nonnormality and reactivity can be caused by the imbalance of incoming/outgoing connections or of the node's in-/out-centrality will likely be useful in designing large complex technological (and possibly synthetic biological) systems. Importantly, our results show that nonuniform link weights and self-links, which are aspects of complexity often neglected in the systematic study of networks, tend to increase nonnormality. While here we focused on reactivity as a widely used measure of transient response, we anticipate similar results for other measures, such as pseudospectra (*14*), singular values (*19*), and the perturbation-averaged evolution of the state vector (*51*). We also suggest that the structural characterization of nonnormality and reactivity can be relevant for the study of generalized networks of current interest, such as those accounting for temporal, multilayer, and higher-order (nonpairwise) interactions. Ultimately, the demonstration that nonnormality and reactivity tend to be more prevalent and more consequential for stability as the system grows in size opens up new vistas of network dynamics in complex systems.

MATERIALS AND METHODS

Impact of α_i heterogeneity on nonnormality and reactivity

Any heterogeneity in α_i can be subtracted from the first term and absorbed into the second in Eq. 1 by replacing α_i and A_{ii} with $\alpha = \max_i \alpha_i$ and $A_{ii} + (\alpha - \alpha_i)$, respectively. In Eq. 31, this would contribute to the heterogeneity of β_i, extending the validity of the nonnormality approximation in that equation to the general case of heterogeneous α_i. It then follows from Eqs. 23 and 31 that the expected nonnormality $\|D\|_F$ for the GCLV model generally increases with the heterogeneity of α_i. Combining this with the approximate relation between nonnormality and reactivity in Eq. 6, we see that the reactivity also tends to increase with the heterogeneity of α_i in random networks. Moreover, the modification of α_i and A_{ii} in Eq. 1 used here can also be applied to Eq. 5, which would extend the reactivity approximation formula to heterogeneous α_i. The conclusion above on the impact of α_i heterogeneity in random networks holds true in particular when α_i and A_{ii} are uncorrelated, while specific correlations can, in principle, lead to a reduction in the heterogeneity of β_i and thus in $\mathbb{E}\left(\|D\|_F^2\right)$.

After absorbing α_i heterogeneity into A_{ii} in Eq. 1 as in the previous paragraph, the nonnormality and reactivity of the system can be defined in the same fashion, i.e., as $\|D\|_F = \|AA^T - A^TA\|_F$ and $\lambda_\Delta(A) = \lambda_1(H) - \lambda_1(A)$, respectively, but using the modified A. While this definition allows us to derive results for random networks (as described in the previous paragraph), the real networks were analyzed under the uniform α_i assumption since the values of α_i were not available in the dataset.

Sufficient condition for reactivity

To establish that having distinct (real) left and right eigenspaces associated with $\lambda_1(A)$ implies reactivity [i.e., $\lambda_\Delta(A) > 0$], we will prove the contrapositive: $\lambda_\Delta(A) = 0$ implies that these eigenspaces coincide. Suppose that $\lambda_\Delta(A) = 0$, which implies $\lambda_1(H) = \lambda_1(A)$. We will seek to show that any right eigenvector is also a left eigenvector and vice versa. Let v be a (real) right eigenvector of A associated with the eigenvalue $\lambda_1(A)$. Without loss of generality, we can assume v to be normalized so that $v^Tv = 1$. Then, we have

$$\lambda_1(H) = \max_{x \neq 0} \frac{x^T H x}{x^T x} \geq v^T H v = \lambda_1(A) \tag{7}$$

where we note that $v^THv = v^T(A + A^T)v/2 = \lambda_1(A)v^Tv = \lambda_1(A)$. Since $\lambda_1(H) = \lambda_1(A)$, the inequality in Eq. 7 becomes an equality, implying that v is a solution of the maximization problem and thus a (right) eigenvector of H associated with $\lambda_1(H)$, i.e., $Hv = \lambda_1(H)v$. This, together with $2H = A + A^T$ and $Av = \lambda_1(A)v$, yields $v^TA = (2\lambda_1(H) - \lambda_1(A))v^T = \lambda_1(A)v^T$. We thus conclude that v is a left eigenvector of A associated with $\lambda_1(A)$, in addition to being a right eigenvector. To show the opposite direction, we now let v be any left eigenvector of A associated with the eigenvalue $\lambda_1(A)$. This again implies $v^THv = \lambda_1(A)$ and turns Eq. 7 into an equality, implying that v is a (left) eigenvector of H associated with $\lambda_1(H)$, i.e., $v^TH = v^T\lambda_1(H)$. From this, an argument similar to the one above shows that v is also a right eigenvector of A associated with the eigenvalue $\lambda_1(A)$. Therefore, we conclude that the left and right eigenspaces of A corresponding to $\lambda_1(A)$ must coincide, which completes the proof.

Real network data

We retrieved raw data for 534 distinct directed networks available from Koblenz Network Collection (KONECT) (*52*, *53*), the Netzschleuder Network Catalogue and Repository (*54*), the Colorado Index of Complex Networks (ICON) (*55*), and the Matrix Market Repository (*56*). From KONECT, Netzschleuder, and ICON, we obtained all distinct directed non-bipartite networks that were available in the form of adjacency lists. From the Matrix Market Repository, we retrieved a total of 17 directed nonbipartite networks among the largest available in each technological or economic application domain to compensate for the relatively few networks of those types available from the other three data sources.

To eliminate redundancy among these networks, we identified and excluded duplicates by comparing the description, number of

nodes, number of links, and other network statistics. Since the resulting set still contained a disproportionately large number of wiki-talk and wiki-link networks (28 and 167, respectively), we removed this bias by excluding all but the largest network from each of these two groups. This led to our final selection of 251 networks, consisting of 62 biological networks, 51 informational networks, 79 social networks, 40 technological networks, and 19 economic/game networks. The network size ranges from $n = 10$ to $n = 7.9 \times 10^6$, and the link density ranges from 4.0×10^{-7} to 1.0. The adjacency matrix A of each network was constructed as follows. For each link, the corresponding A_{ij} was set to the weight from the data if available and set to one otherwise. Multiple links between the same pair of nodes were combined into a single link whose weight equals the sum of the weights of the original links. For a self-link, the corresponding A_{ii} was set based on the information from the data if available and $A_{ii} = 0$ otherwise. The constructed adjacency matrices were then used to compute the nonnormality measure $\|D\|_F$, reactivity measure $\lambda_\Delta(A)$, degree imbalance $\langle (d^{\text{in}} - d^{\text{out}})^2 \rangle$, eigenvector angle θ_1, and other properties of the networks. Of these 251 networks, 38 had degenerate or nearly degenerate $\lambda_1(A)$, which we numerically identified using the criterion $\theta_1 > 1.57$ [recalling that $\theta_1 \leq \pi/2$ by definition and that $\theta_1 = \pi/2$ when $\lambda_1(A)$ is degenerate].

The minimum level of nonnormality observed among the 251 networks in Fig. 1A was for a 8192-node network of interactions in a square dielectric waveguide, with $\|D\|_F/(\sqrt{n}\langle w^2 \rangle) \approx 0.032$. The only nonreactive network in the dataset was one representing child-parent relationships from an online genealogical website called WikiTree, for which $\lambda_\Delta(A)$ and θ_1 are both estimated to be zero to machine precision. Among the reactive networks that are nondegenerate, shown in Fig. 1B, the minimum value of $\lambda_\Delta(A)/\lambda_1(A) \approx 3.1 \times 10^{-6}$ was observed for the 3140-node network of intercounty migration in the United States.

Directed random network model

To sample a network realization of size n from the GCLV model, we first draw the expected in-degree $\tilde{d}_i^{\text{in}} \geq 1$ and out-degree $\tilde{d}_i^{\text{out}} \geq 1$ of different nodes $i = 1, \ldots, n$ independently from a common joint probability distribution that does not depend on the network size n and has finite second moments. For each i, the random variables $\tilde{d}_i^{\text{in}}$ and $\tilde{d}_i^{\text{out}}$ are not restricted to be integers and need not to be independent (and thus can be correlated). Following (*47*), the actual in- and out-degrees d_i^{in} and d_i^{out} of the network are then determined as a result of randomly creating a directed link from node j to node i with the probability

$$\mathbb{P}\left(\text{link } j \to i \text{ exists} \mid \tilde{d}_1^{\text{in}}, \ldots, \tilde{d}_n^{\text{in}}, \tilde{d}_1^{\text{out}}, \ldots, \tilde{d}_n^{\text{out}}\right) = \rho_{ij} := \frac{\tilde{d}_i^{\text{in}} \tilde{d}_j^{\text{out}}}{\sum_{k=1}^n \tilde{d}_k^{\text{in}}} \tag{8}$$

for each i and j. We then independently draw random weights $A_{ij} > 0$ for the resulting links from a given distribution, while we set $A_{ij} = 0$ if node j is not connected to node i. Note that this process can create self-links with random weights, but we could choose to prohibit them by setting $A_{ii} = 0$ for all i. The unweighted version of this model is obtained if we instead set $A_{ij} = 1$ for all links.

For the (marginal) distributions of $\tilde{d}_i^{\text{in}}$ and $\tilde{d}_i^{\text{out}}$, we assume that $\rho_{ij} \leq c$ for some constant $\frac{1}{2} \leq c < 1$ and that the extreme values $\tilde{d}_{\max}^{\text{in}} := \max_{1 \leq i \leq n} \tilde{d}_i^{\text{in}}$ and $\tilde{d}_{\max}^{\text{out}} := \max_{1 \leq i \leq n} \tilde{d}_i^{\text{out}}$ asymptotically follow the so-called generalized extreme value (GEV) distributions (*57*) after appropriate normalization (see below for more details). We note that the first assumption is only a slight addition to the second, as the second implies that $\rho_{ij} \to 0$ in probability as $n \to \infty$ (see Supplementary Materials, section S2.1, for a proof). We also assume that $\mathbb{E}\left(\tilde{d}_i^{\text{in}}\right) = \mathbb{E}\left(\tilde{d}_i^{\text{out}}\right)$, so that we have $\lim_{n \to \infty} \left(\sum_{i=1}^n \tilde{d}_i^{\text{in}} - \sum_{i=1}^n \tilde{d}_i^{\text{out}}\right)/n = 0$ almost surely (which can be shown using the strong law of large numbers). Denoting the mean of the expected in- or out-degree by $\tilde{d} := \mathbb{E}\left(\tilde{d}_i^{\text{in}}\right) = \mathbb{E}\left(\tilde{d}_i^{\text{out}}\right)$ (which does not depend on i), we assume $\tilde{d} > 1$. This holds true because $\tilde{d}_i^{\text{in}}, \tilde{d}_i^{\text{out}} \geq 1$ unless the joint distribution of $\tilde{d}_i^{\text{in}}$ and $\tilde{d}_i^{\text{out}}$ is singular with all probability density concentrated at $\tilde{d}_i^{\text{in}} = \tilde{d}_i^{\text{out}} = 1$. These assumptions guarantee that, with probability one, the means of the actual in-degrees d_i^{in} and out-degrees d_i^{out} over the Bernoulli distribution of A_{ij} indeed match $\tilde{d}_i^{\text{in}}$ and $\tilde{d}_i^{\text{out}}$, respectively, in the limit of $n \to \infty$. Thus, the given joint distribution of $\tilde{d}_i^{\text{in}}$ and $\tilde{d}_i^{\text{out}}$ can be interpreted as the expected degree distribution for this random network model.

For $\tilde{d}_{\max}^{\text{in}}$, the GEV assumption mentioned above is more precisely described as follows: there exist sequences of normalization constants $a_n > 0$ and b_n such that

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\tilde{d}_{\max}^{\text{in}} - b_n}{a_n} \leq x\right) = G_\gamma(x) \tag{9}$$

and $\lim_{n \to \infty} a_n/n = \lim_{n \to \infty} b_n/n = 0$. Here, the GEV cumulative distribution function $G_\gamma(x)$ is given by

$$\begin{aligned} &\text{For } \gamma < 0 : G_\gamma(x) = \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right), & x \leq -1/\gamma \\ 1, & x > -1/\gamma \end{cases} \\ &\text{For } \gamma = 0 : G_\gamma(x) = \exp\left(-e^{-x}\right) \text{ for all } x \\ &\text{For } \gamma > 0 : G_\gamma(x) = \begin{cases} 0, & x \leq -1/\gamma \\ \exp\left(-(1+\gamma x)^{-1/\gamma}\right), & x > -1/\gamma \end{cases} \end{aligned} \tag{10}$$

The parameter γ is called the extreme value index and is determined solely by the distribution of $\tilde{d}_i^{\text{in}}$. For $\tilde{d}_{\max}^{\text{out}}$, we assume the same but possibly with different a_n, b_n, and γ. The GEV distributions are the only ones that can arise as the limit distribution of the maximum of independent and identically distributed random variables if the limit in Eq. 9 exists, and they include the well-known Gumbel, Fréchet, and Weibull distribution as special cases.

The GEV assumption is mild and is satisfied for almost all commonly encountered distributions. It can be verified using the explicit conditions for Eq. 9 given in (*57*). For example, one of these conditions (theorem 1.1.8) shows that it is satisfied by any power-law distribution with exponent $\beta > 2$, minimum expected degree $\tilde{d}_{\min} \geq 1$, and the density function given by $p(x) = Cx^{-\beta}$ if $x \geq \tilde{d}_{\min}$, and $p(x) = 0$ otherwise, where $C = (\beta - 1)\tilde{d}_{\min}^{\beta - 1}$ is the normalization constant. The extreme value index in this case is $\gamma = 1/(\beta - 1)$, and the sequences of constants in Eq. 9 are $a_n = n^\gamma \gamma \tilde{d}_{\min}$ and $b_n = n^\gamma \tilde{d}_{\min}$. Since $\beta > 2$ (which is also the condition for this distribution to have a finite mean), we have $\gamma < 1$, implying $\lim_{n \to \infty} a_n/n = \lim_{n \to \infty} \tilde{d}_{\min} \gamma / n^{1-\gamma} = 0$ and $\lim_{n \to \infty} b_n/n = \lim_{n \to \infty} \tilde{d}_{\min}/n^{1-\gamma} = 0$.

Nonnormality approximation for unweighted networks

Our derivation of Eq. 4 is for large random networks generated by the unweighted version of the GCLV model. Since Eq. 4

expresses $\|D\|_F^2$ as a function of the in-degrees d_i^{in} and out-degrees d_i^{out}, the derivation is based on calculating the conditional expected value $\mathbb{E}(\|D\|_F^2 | d_1^{in},\ldots,d_n^{in}, d_1^{out},\ldots,d_n^{out})$, which we denote simply by $\mathbb{E}(\|D\|_F^2)$ in this section (in which we also use similar simplified notations for other expected values, variances, and probabilities). Taking the (conditional) expected value of Eq. 3, we obtain

$$\begin{aligned}\mathbb{E}(\|D\|_F^2) &= \sum_{i=1}^{n}\left(d_i^{in}-d_i^{out}\right)^2 + \mathbb{E}\left(\sum_{i\neq j}\sum\left(d_{ij}^{in}-d_{ij}^{out}\right)^2\right)\\ &= n\left\langle\left(d^{in}-d^{out}\right)^2\right\rangle + \sum_{i\neq j}\sum\mathbb{E}(Y_{ij}^2)\end{aligned} \tag{11}$$

where we define $Y_{ij} \coloneqq d_{ij}^{in} - d_{ij}^{out} = \sum_{k\neq i,j}(A_{ik}A_{jk} - A_{ki}A_{kj})$ (recalling the assumption $A_{ii}=0$, which will be relaxed in the next section). We note that, while d_i^{in} and d_i^{out} are held fixed, d_{ij}^{in} and d_{ij}^{out} are random.

We now seek to estimate $\mathbb{E}\left(Y_{ij}^2\right)$ in Eq. 11 by calculating $\mathbb{E}(Y_{ij})$ and $\mathrm{Var}(Y_{ij})$, which can be broken down to individual terms $\mathbb{E}(A_{ik}A_{jk})$ and $\mathbb{E}(A_{ki}A_{kj})$. We first note that $\mathbb{E}(A_{ik}A_{jk}) = \mathbb{P}(A_{ik}A_{jk}=1) = \mathbb{P}(A_{ik}=1 \text{ and } A_{jk}=1)$. To compute this probability, we define

$$Q_{i,n}^{in}(d) \coloneqq \mathbb{P}\left(\sum_{k=1}^{n}A_{ik} = d\right) \text{ and } Q_{j,n}^{out}(d) \coloneqq \mathbb{P}\left(\sum_{k=1}^{n}A_{kj} = d\right) \tag{12}$$

to represent the probability that the in-degree of node i is d and the probability that the out-degree of node j is d, respectively. Assuming the random variable $\tilde{d}_i^{in}$ fixed for the moment, we have

$$\begin{aligned}Q_{i,n}^{in}(d) &= \sum_{\alpha}\prod_{k=1}^{n}\rho_{ik}^{\alpha_k}(1-\rho_{ik})^{1-\alpha_k}\\ &= \sum_{\alpha}\left(\prod_{k:\alpha_k=1}\rho_{ik}\right)\left(\prod_{k:\alpha_k=0}(1-\rho_{ik})\right)\\ &\approx \sum_{\alpha}\left(\prod_{k:\alpha_k=1}\frac{\tilde{d}_i^{in}\tilde{d}_k^{out}}{n\bar{d}}\right)\exp\left(-\sum_{k:\alpha_k=0}\frac{\tilde{d}_i^{in}\tilde{d}_k^{out}}{n\bar{d}}\right)\\ &\approx \binom{n}{d}\frac{\left(\tilde{d}_i^{in}\right)^d}{n^d\bar{d}^d}\cdot\frac{1}{\binom{n}{d}}\sum_{\alpha}\left(\prod_{k:\alpha_k=1}\tilde{d}_k^{out}\right)\exp\left(-\frac{\tilde{d}_i^{in}}{\bar{d}}\cdot\frac{1}{n}\sum_{k:\alpha_k=0}\tilde{d}_k^{out}\right)\\ &\approx \binom{n}{d}\cdot\frac{\left(\tilde{d}_i^{in}\right)^d}{n^d}\cdot\exp\left(-\tilde{d}_i^{in}\right)\end{aligned} \tag{13}$$

where $\sum_\alpha$ denotes the summation over all $\alpha=\{\alpha_1,\ldots,\alpha_n\}\in\{0,1\}^n$ in which exactly d elements are equal to one, and k: $\alpha_k=0,1$ denotes all k such that $\alpha_k=0,1$, respectively. Equation 13 also used the following approximations for large network size n (with d fixed): $1-\rho_{ik}\approx e^{-\rho_{ik}}$ valid when ρ_{ik} is small (which is the case because $\rho_{ik}\to 0$ in probability as $n\to\infty$; see Supplementary Materials, section S2.1); $\frac{1}{n}\sum_k\tilde{d}_k^{in}\approx\bar{d}$ and $\frac{1}{n}\sum_{k:\alpha_k=0}\tilde{d}_k^{out}\approx\frac{1}{n}\sum_k\tilde{d}_k^{out}\approx\bar{d}$ (based on the strong law of large numbers); and $\sum_\alpha\prod_{k:\alpha_k=1}\tilde{d}_k^{out}/\binom{n}{d}\approx\bar{d}^d$ [a special case of the theorem proved in (58)]. Analogously, holding $\tilde{d}_j^{out}$ fixed temporarily, we have

$$Q_{j,n}^{out}(d) \approx \binom{n}{d}\cdot\frac{\left(\tilde{d}_j^{out}\right)^d}{n^d}\cdot\exp\left(-\tilde{d}_j^{out}\right) \tag{14}$$

and similar arguments also show that

$$\begin{aligned}&\mathbb{P}\left(\sum_{k\neq j}A_{ik}=d\right)=Q_{i,n-1}^{in}(d),\ \mathbb{P}\left(\sum_{k\neq j_1,j_2}A_{ik}=d\right)=Q_{i,n-2}^{in}(d),\\ &\mathbb{P}\left(\sum_{k\neq i}A_{kj}=d\right)=Q_{j,n-1}^{out}(d),\ \mathbb{P}\left(\sum_{k\neq i_1,i_2}A_{kj}=d\right)=Q_{j,n-2}^{out}(d)\end{aligned} \tag{15}$$

We can now calculate $\mathbb{E}(A_{ik}A_{jk})$ and $\mathbb{E}(A_{ki}A_{kj})$ from the definition of the conditional expected values as

$$\begin{aligned}\mathbb{E}(A_{ik}A_{jk}) &= \mathbb{P}\left(A_{ik}=1 \text{ and } A_{jk}=1 \mid d_i^{in},d_j^{in},d_k^{out}\right)\\ &= \frac{\rho_{ik}\rho_{jk}Q_{i,n-1}^{in}\left(d_i^{in}-1\right)Q_{j,n-1}^{in}\left(d_j^{in}-1\right)Q_{k,n-2}^{out}\left(d_k^{out}-2\right)}{Q_{i,n}^{in}\left(d_i^{in}\right)Q_{j,n}^{in}\left(d_j^{in}\right)Q_{k,n}^{out}\left(d_k^{out}\right)}\\ &\approx \frac{d_i^{in}d_j^{in}d_k^{out}\left(d_k^{out}-1\right)}{(\bar{d}n)^2},\\ \mathbb{E}(A_{ki}A_{kj}) &\approx \frac{d_i^{out}d_j^{out}d_k^{in}\left(d_k^{in}-1\right)}{(\bar{d}n)^2}\end{aligned} \tag{16}$$

We note that these expressions do not depend on $\tilde{d}_i^{in}$ nor $\tilde{d}_i^{out}$, and hence, the estimates are valid even when these variables are allowed to be random. We thus have

$$\begin{aligned}\mathbb{E}(A_{ik}A_{jk}-A_{ki}A_{kj}) &= \mathbb{E}(A_{ik}A_{jk})-\mathbb{E}(A_{ki}A_{kj})\\ &\approx \frac{d_i^{in}d_j^{in}d_k^{out}\left(d_k^{out}-1\right)}{(\bar{d}n)^2}-\frac{d_i^{out}d_j^{out}d_k^{in}\left(d_k^{in}-1\right)}{(\bar{d}n)^2},\\ \mathrm{Var}(A_{ik}A_{jk}-A_{ki}A_{kj}) &= \mathrm{Var}(A_{ik}A_{jk})+\mathrm{Var}(A_{ki}A_{kj})\\ &= \mathbb{E}(A_{ik}A_{jk})-\mathbb{E}(A_{ik}A_{jk})^2+\mathbb{E}(A_{ki}A_{kj})-\mathbb{E}(A_{ki}A_{kj})^2\\ &\approx \mathbb{E}(A_{ik}A_{jk})+\mathbb{E}(A_{ki}A_{kj})\\ &\approx \frac{d_i^{in}d_j^{in}d_k^{out}\left(d_k^{out}-1\right)}{(\bar{d}n)^2}+\frac{d_i^{out}d_j^{out}d_k^{in}\left(d_k^{in}-1\right)}{(\bar{d}n)^2}\end{aligned} \tag{17}$$

where we used that the high-order terms $\mathbb{E}(A_{ik}A_{jk})^2$ and $\mathbb{E}(A_{ki}A_{kj})^2$ are negligible for large n. Summing these over k (including the terms $k=i,j$, which are small compared to other terms and hence do not affect the sums for large n), we have

$$\mathbb{E}(Y_{ij}) \approx \frac{1}{n}\left[\frac{d_i^{in}d_j^{in}(\langle(d^{out})^2\rangle-\bar{d})}{\bar{d}^2}-\frac{d_i^{out}d_j^{out}(\langle(d^{in})^2\rangle-\bar{d})}{\bar{d}^2}\right] \tag{18}$$

$$\mathrm{Var}(Y_{ij}) \approx \frac{1}{n}\left[\frac{d_i^{in}d_j^{in}(\langle(d^{out})^2\rangle-\bar{d})}{\bar{d}^2}+\frac{d_i^{out}d_j^{out}(\langle(d^{in})^2\rangle-\bar{d})}{\bar{d}^2}\right] \tag{19}$$

The term $\mathbb{E}\left(Y_{ij}^2\right)$ in Eq. 11 can now be estimated for large n as

$$\mathbb{E}\left(Y_{ij}^2\right) = \mathrm{Var}(Y_{ij})+\mathbb{E}(Y_{ij})^2 \approx \mathrm{Var}(Y_{ij}) \tag{20}$$

since $\mathbb{E}(Y_{ij})^2$ is negligible compared to $\mathrm{Var}(Y_{ij})$, and we have

$$\sum_{i\neq j}\sum \mathbb{E}\left(Y_{ij}^2\right) \approx \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{d_i^{\text{in}}\, d_j^{\text{in}}(\langle (d^{\text{out}})^2\rangle - \bar{d}\,)}{\bar{d}^2} + \frac{d_i^{\text{out}}\, d_j^{\text{out}}(\langle (d^{\text{in}})^2\rangle - \bar{d}\,)}{\bar{d}^2}\right]$$
$$= n\left[\langle (d^{\text{in}})^2\rangle + \langle (d^{\text{out}})^2\rangle - 2\bar{d}\right] \quad (21)$$

(retaining the $i = j$ term, which is small compared to the sum of the other terms). Substituting this into Eq. 11 yields Eq. 4.

Nonnormality approximation for weighted networks

Here, we consider the weighted version of the GCLV model, assuming that the distribution of link weights squared has a finite mean $\mathbb{E}(w^2)$ and a finite variance $\mathrm{Var}(w^2)$. Assuming first that there are no self-links, i.e., $A_{ii} = 0$, we will show that the nonnormality $\|D\|_{\text{F}}$ can be approximated for large n as

$$\|D\|_{\text{F}}^2 \approx 2n\left[\langle (d^{\text{in}} - d^{\text{out}})^2\rangle + \langle d^{\text{in}}\, d^{\text{out}}\rangle - \bar{d}\,\right]\mathbb{E}^2(w^2) + 2n\bar{d}\,\mathrm{Var}(w^2) \quad (22)$$

This extends Eq. 4 to weighted networks. When the link weight variance $\mathrm{Var}(w^2)$ is small, this equation yields the scaling $\|D\|_{\text{F}} \sim \sqrt{n}\,\mathbb{E}(w^2) \approx \sqrt{n}\,\langle w^2\rangle$. As in the case of unweighted networks, the derivation is based on taking the conditional expectation given fixed values of d_i^{in} and d_i^{out}. The weighted version of Eq. 11 reads:

$$\mathbb{E}\left(\|D\|_{\text{F}}^2\right) = \sum_{i=1}^{n}\mathbb{E}\left(Y_{ii}^2\right) + \sum_{i\neq j}\sum \mathbb{E}\left(Y_{ij}^2\right) \quad (23)$$

where $Y_{ij} = \sum_{k\neq i,j}(A_{ik}A_{jk} - A_{ki}A_{kj})$ and thus $Y_{ii} = \sum_{k\neq i}\left(A_{ik}^2 - A_{ki}^2\right)$. Following the derivation of Eq. 17 while accounting for the random weights, we have

$$\mathbb{E}(A_{ik}A_{jk} - A_{ki}A_{kj}) \approx \left(\frac{d_i^{\text{in}}\, d_j^{\text{in}}\, d_k^{\text{out}}\left(d_k^{\text{out}} - 1\right)}{(\bar{d}n)^2} - \frac{d_i^{\text{out}}\, d_j^{\text{out}}\, d_k^{\text{in}}\left(d_k^{\text{in}} - 1\right)}{(\bar{d}n)^2}\right)\mathbb{E}(w^2) \quad (24)$$

$$\mathrm{Var}(A_{ik}A_{jk} - A_{ki}A_{kj}) \approx \mathbb{E}\left(A_{ik}^2A_{jk}^2\right) + \mathbb{E}\left(A_{ki}^2A_{kj}^2\right) \approx \left(\frac{d_i^{\text{in}}\, d_j^{\text{in}}\, d_k^{\text{out}}\left(d_k^{\text{out}} - 1\right)}{(\bar{d}n)^2} + \frac{d_i^{\text{out}}\, d_j^{\text{out}}\, d_k^{\text{in}}\left(d_k^{\text{in}} - 1\right)}{(\bar{d}n)^2}\right)\left(\mathbb{E}(w^2)\right)^2 \quad (25)$$

This leads to the weighted version of Eq. 21:

$$\sum_{i\neq j}\sum \mathbb{E}\left(Y_{ij}^2\right) \approx n\left[\langle (d^{\text{in}})^2\rangle + \langle (d^{\text{out}})^2\rangle - 2\bar{d}\right]\left(\mathbb{E}(w^2)\right)^2 \quad (26)$$

To approximate the first term on the right side of Eq. 23, we note that

$$\mathbb{E}\left(A_{ik}^2\right) \approx \frac{d_i^{\text{in}}\, d_k^{\text{out}}}{\bar{d}n}\mathbb{E}(w^2),\quad \mathbb{E}\left(A_{ki}^2\right) \approx \frac{d_i^{\text{out}}\, d_k^{\text{in}}}{\bar{d}n}\mathbb{E}(w^2) \quad (27)$$

and that there are exactly d_i^{in} and d_i^{out} nonzero terms in $\sum_{k\neq i}A_{ik}^2$ and $\sum_{k\neq i}A_{ki}^2$, respectively. Thus, given fixed values of d_i^{in} and d_i^{out}, the expectation and variance of $Y_{ii} = \sum_{k\neq i}\left(A_{ik}^2 - A_{ki}^2\right)$ can be estimated as

$$\mathbb{E}(Y_{ii}) \approx \left(d_i^{\text{in}} - d_i^{\text{out}}\right)\mathbb{E}(w^2) \quad (28)$$

$$\mathrm{Var}(Y_{ii}) \approx \left(d_i^{\text{in}} + d_i^{\text{out}}\right)\mathrm{Var}(w^2) \quad (29)$$

and hence we have

$$\mathbb{E}\left(Y_{ii}^2\right) = \mathbb{E}(Y_{ii})^2 + \mathrm{Var}(Y_{ii}) \approx \left(d_i^{\text{in}} - d_i^{\text{out}}\right)^2\left(\mathbb{E}(w^2)\right)^2 + \left(d_i^{\text{in}} + d_i^{\text{out}}\right)\mathrm{Var}(w^2) \quad (30)$$

Combining Eqs. 23, 26, and 30 and approximating $\|D\|_{\text{F}}^2$ by its conditional expectation $\mathbb{E}\left(\|D\|_{\text{F}}^2\right)$ yield Eq. 22.

The above results can be further extended to allow for any given assignment of self-links. Assume that $\widehat{A} = A + \mathrm{diag}(\beta)$ with diagonal elements β_i, $i = 1,2,\cdots,n$. Then, $Y_{ij} = \sum_{k\neq i,j}(A_{ik}A_{jk} - A_{ki}A_{kj}) - (A_{ij} - A_{ji})(\beta_i - \beta_j)$ and Y_{ii} remains unchanged. Note that

$$\mathbb{E}\left(Y_{ij}^2\right) \approx \mathrm{Var}(Y_{ij}) = \mathrm{Var}\left(\sum_{k\neq i,j}(A_{ik}A_{jk} - A_{ki}A_{kj})\right) + \mathrm{Var}((A_{ij} - A_{\text{ji}})(\beta_i - \beta_j))$$
$$\approx \left[\frac{d_i^{\text{in}}\, d_j^{\text{in}}(\langle (d^{\text{out}})^2\rangle - \bar{d})}{\bar{d}^2} + \frac{d_i^{\text{out}}\, d_j^{\text{out}}(\langle (d^{in})^2\rangle - \bar{d})}{\bar{d}^2}\right]\cdot\frac{\mathbb{E}^2(w^2)}{n} + (\beta_i - \beta_j)^2\cdot\frac{d_i^{\text{in}}\, d_j^{\text{out}} + d_j^{\text{in}}\, d_i^{\text{out}}}{\bar{d}}\cdot\frac{\mathbb{E}(w^2)}{n} \quad (31)$$

Therefore, we have

$$\|D(\widehat{A})\|_{\text{F}}^2 \approx 2n\left[\langle (d^{\text{in}} - d^{\text{out}})^2\rangle + \langle d^{\text{in}}\, d^{\text{out}}\rangle - \bar{d}\right]\mathbb{E}^2(w^2) + 2n\bar{d}\,\mathrm{Var}(w^2) + 2n\mathbb{E}(w^2)\left[\langle \beta^2 d^{\text{in}}\rangle + \langle \beta^2 d^{\text{out}}\rangle - \frac{2}{\bar{d}}\langle \beta\, d^{\text{in}}\rangle\langle \beta\, d^{\text{out}}\rangle\right]$$
$$= \|D(A)\|_{\text{F}}^2 + 2n\mathbb{E}(w^2)\left[\langle \beta^2 d^{\text{in}}\rangle + \langle \beta^2 d^{\text{out}}\rangle - \frac{2}{\bar{d}}\langle \beta\, d^{\text{in}}\rangle\langle \beta\, d^{\text{out}}\rangle\right] \quad (32)$$

Extension of the nonnormality approximation to Laplacian-coupled networks

Here, we consider the nonnormality of the Laplacian matrices of networks generated by the weighted GCLV model under the same conditions as in the derivation of Eq. 22. The Laplacian matrix of a network with adjacency matrix A is defined by $L := K - A$, where K denotes the diagonal matrix with $K_{ii} = \sum_k A_{ik}$. A straightforward calculation yields

$$\|D(L)\|_{\text{F}}^2 = \|L^TL - LL^T\|_{\text{F}}^2 = \sum_{i=1}^{n}\mathbb{E}\left(\widetilde{Y}_{ii}^2\right) + \sum_{i\neq j}\sum \mathbb{E}\left(\widetilde{Y}_{ij}^2\right) \quad (33)$$

where $\widetilde{Y}_{ij} = \sum_{k\neq i,j}(A_{ik}A_{jk} - A_{ki}A_{kj}) + (A_{ij} - A_{\text{ji}})\left(d_i^{\text{in}} - d_j^{\text{in}}\right)$ and, in particular, $\widetilde{Y}_{ii} = \sum_{k\neq i}\left(A_{ik}^2 - A_{ki}^2\right)$. Following Eq. 30, we have

$$\mathbb{E}\left(\widetilde{Y}_{ii}^2\right) \approx \left(d_i^{\text{in}} - d_i^{\text{out}}\right)^2\left(\mathbb{E}(w^2)\right)^2 + \left(d_i^{\text{in}} + d_i^{\text{out}}\right)\mathrm{Var}(w^2) \quad (34)$$

In addition, following Eqs. 20 and 26, we obtain

$$\mathbb{E}\left(\widetilde{Y}_{ij}^2\right) \approx \mathrm{Var}\left(\widetilde{Y}_{ij}\right) = \mathrm{Var}\left(\sum_{k\neq i,j}(A_{ik}A_{jk} - A_{ki}A_{kj})\right) + \mathrm{Var}((A_{ij} - A_{\text{ji}})(d_i^{\text{in}} - d_j^{\text{in}}))$$
$$\approx \frac{\mathbb{E}^2(w^2)}{n}\left[\frac{d_i^{\text{in}}\, d_j^{\text{in}}(-\langle (d^{\text{out}})^2\rangle - \bar{d})}{\bar{d}^2} + \frac{d_i^{\text{out}}\, d_j^{\text{out}}(\langle (d^{\text{in}})^2\rangle - \bar{d})}{\bar{d}^2}\right] + (d_i^{\text{in}} - d_j^{\text{in}})^2\cdot\frac{d_i^{\text{in}}\, d_j^{\text{out}} + d_j^{\text{in}}\, d_i^{\text{out}}}{n\bar{d}}\cdot\mathbb{E}(w^2) \quad (35)$$

Therefore, we have

$$\|D(L)\|_F^2 \approx 2n[\langle (d^{in}-d^{out})^2\rangle + \langle d^{in}d^{out}\rangle - \bar{d}]\mathbb{E}^2(w^2) + 2n\bar{d}\mathrm{Var}(w^2) + 2n\mathbb{E}(w^2)\left[\langle (d^{in})^3\rangle + \langle (d^{in})^2 d^{out}\rangle - \frac{2}{\bar{d}}\langle (d^{in})^2\rangle\langle d^{in}d^{out}\rangle\right] = \|D(A)\|_F^2 + 2n\mathbb{E}(w^2)\left[\langle (d^{in})^3\rangle + \langle (d^{in})^2 d^{out}\rangle - \frac{2}{\bar{d}}\langle (d^{in})^2\rangle\langle d^{in}d^{out}\rangle\right] \quad (36)$$

Reactivity approximation for networks with dominant largest eigenvalue

Here, we derive Eq. 5 assuming $\lambda_1(H) \approx \lambda_1(H_1)$ and the nondegeneracy of $\lambda_1(A)$. We normalize the left and right eigenvectors associated with $\lambda_1(A)$, so that we have $v_1^T v_1 = 1$ and $u_1^T v_1 = v_1^T u_1 = 1$, and thus, the angle θ_1 between the left and right eigenvectors is given by $\cos\theta_1 = 1/\sqrt{u_1^T u_1}$. Noting that the symmetric part of A_1 can be written as $H_1 = \frac{1}{2}\lambda_1(A)\left(u_1 v_1^T + v_1 u_1^T\right)$ and approximating the eigenvector associated with its largest eigenvalue by a linear combination $\alpha u_1 + \beta v_1$, we have

$$\begin{aligned}\lambda_1(H) \approx \lambda_1(H_1) &= \max_{x\in\mathbb{R}^n, x\neq 0}\frac{x^T H_1 x}{x^T x} \\ &= \frac{1}{2}\lambda_1(A)\max_{\alpha^2+\beta^2\neq 0}\frac{(\alpha u_1+\beta v_1)^T(u_1 v_1^T + v_1 u_1^T)(\alpha u_1+\beta v_1)}{(\alpha u_1+\beta v_1)^T(\alpha u_1+\beta v_1)} \\ &= \lambda_1(A)\max_{y\in\mathbb{R}^2, y\neq 0}\frac{y^T\Phi y}{y^T\Psi y}\end{aligned} \quad (37)$$

where we defined

$$\Phi := \Phi(\theta_1) = \begin{bmatrix} 1 & \frac{1}{2}u_1^T u_1 + \frac{1}{2} \\ \frac{1}{2}u_1^T u_1 + \frac{1}{2} & u_1^T u_1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2\cos^2\theta_1} + \frac{1}{2} \\ \frac{1}{2\cos^2\theta_1} + \frac{1}{2} & \frac{1}{\cos^2\theta_1} \end{bmatrix} \quad (38)$$

$$\Psi := \Psi(\theta_1) = \begin{bmatrix} 1 & 1 \\ 1 & u_1^T u_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & \frac{1}{\cos^2\theta_1} \end{bmatrix} \quad (39)$$

The last maximum in Eq. 37 can be computed as the largest generalized eigenvalue μ of the matrix pencil $(\Phi(\theta_1), \Psi(\theta_1))$ and satisfies the equation

$$\det(\Phi(\theta_1) - \mu\Psi(\theta_1)) = 0 \quad (40)$$

This equation can be explicitly solved to yield $\mu = \frac{1+\cos\theta_1}{2\cos\theta_1}$. Substituting this into Eq. 37 and using the definition of λ_Δ, we obtain Eq. 5 as

$$\lambda_\Delta = \lambda_1(H) - \lambda_1(A) \approx \lambda_1(A)\cdot\frac{1+\cos\theta_1}{2\cos\theta_1} - \lambda_1(A) = \lambda_1(A)\cdot\frac{1-\cos\theta_1}{2\cos\theta_1} \quad (41)$$

Random networks used in Fig. 3

To generate these networks, we used the GCLV model with a given correlated identical distributions of the expected in-degree $\tilde{d}_i^{in}$ and the expected out-degree $\tilde{d}_i^{out}$ for each node i. To realize such a joint distribution, we first drew $\tilde{d}_i^{in}$ and $\tilde{d}_i^{out}$ (independently for each node i) from a common distribution, which was either (i) the gamma distribution, with probability density function $p(x) \sim (x - d_{min})^{a-1}e^{-b(x-d_{min})}$, $x > d_{min}$, where d_{min} is the minimum degree, $a > 0$ is the shape parameter, and $b > 0$ is the rate parameter; or (ii) a discrete power-law distribution with the probability mass function $p(x) \sim x^{-\beta}$ for integers $x = d_{min}, d_{min}+1, \ldots, d_{max}$, where $\beta > 0$ is the scaling exponent, d_{min} is the minimum degree, and $d_{max} = \sqrt{cn\,d_{min}}$ is the maximum degree imposed to ensure that the model assumption $\rho_{ij} \le c$ is satisfied (and here, we set $c = 0.99$). For the gamma distribution, the mean degree d and the parameter $1/b$ were drawn randomly from the intervals [10,50] and [2,10], respectively, and the parameter a was then set to be $a = b(d - d_{min})$ (to ensure that the mean degree equals d). For the power-law distribution, the scaling exponent β was drawn randomly from the interval [2,4]. For both distributions, we used the minimum degree $d_{min} = 10$.

After generating $\tilde{d}_i^{in}$ and $\tilde{d}_i^{out}$, correlation was added between them using a parameter ρ randomly chosen from the interval [−1,1] to increase the range of $\|D\|_F$ and $\lambda_\Delta(A)$ observed. The correlation was created by sorting $\tilde{d}_i^{in}$ and $\tilde{d}_i^{out}$ for a randomly chosen subset of nodes (with mean fraction $|\rho|$) in the same order for $\tilde{d}_i^{in}$ and $\tilde{d}_i^{out}$ if $\rho > 0$ (leading to positive correlation) and in the opposite order if $\rho < 0$ (leading to negative correlation).

For the weighted networks, the random weights were drawn either from the exponential distribution with mean μ and variance μ^2 (for $\mu = 0.7$, 1.0, or 1.3) or from the positive domain of the Gaussian distribution with mean one and variance σ^2 (for $\sigma = 0.1$ or 0.2). For this figure, we allowed self-links in the GCLV model (with random weights drawn from the same distribution as the other links).

Randomization of real networks in Fig. 4A

For a given network from the dataset, we generated its randomization using the GCLV model. The in- and out-degrees of node i in the network were used as the expected degrees $\tilde{d}_i^{in}$ and $\tilde{d}_i^{out}$, respectively, to generate a random network topology. Note that, for nodes i and j with the right side of Eq. 8 exceeding one, we set the connection probability $\rho_{ij} = 1$. The link weights of the network were resampled (with replacement) to generate random weights.

Determination of the shaded region in Fig. 4C

First, for each β, we computed the averages of $\lambda_1(H)$ and $\lambda_1(A)$ over 10 realizations of the GCLV model for 15 values of n, equally spaced on the logarithmic scale between $n = 10^2$ and $n = 10^5$. We then performed a least-squares fit of these averages as functions of n on the logarithmic scale for $n \ge 300$ and used the resulting slopes as the scaling exponents for $\lambda_1(H)$ and $\lambda_1(A)$ for the given β (shown in the background by blue and green dots, respectively). The boundary curves were obtained by fitting these scaling exponents with seventh-order polynomials. The shaded area between these two curves thus represents the parameter region in which the scaling exponent ℓ for $\alpha(n)$ is smaller than that for $\lambda_1(H)$ but larger than that for $\lambda_1(A)$, implying that, as n increases, $\lambda_1(H) - \alpha(n)$ increases while $\lambda_1(A) - \alpha(n)$ decreases.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abm8310

REFERENCES AND NOTES

1. R. M. May, Will a large complex system be stable? *Nature* **238**, 413–414 (1972).
2. R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, 1973).
3. K. S. McCann, The diversity–stability debate. *Nature* **405**, 228–233 (2000).
4. S. Allesina, S. Tang, Stability criteria for complex ecosystems. *Nature* **483**, 205–208 (2012).
5. S. Tang, S. Allesina, Reactivity and stability of large ecosystems. *Front. Ecol. Evol.* **2**, 21 (2014).
6. D. Gravel, F. Massol, M. A. Leibold, Stability and complexity in model meta-ecosystems. *Nat. Commun.* **7**, 12457 (2016).
7. C. Jacquet, C. Moritz, L. Morissette, P. Legagneux, F. Massol, P. Archambault, D. Gravel, No complexity–stability relationship in empirical ecosystems. *Nat. Commun.* **7**, 12573 (2016).
8. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350**, 663–666 (2015).
9. C. L. Murall, J. L. Abbate, M. P. Touzel, E. Allen-Vercoe, S. Alizon, R. Froissart, K. McCann, "Invasions of host-associated microbiome networks" in *Networks of Invasion: Empirical Evidence and Case Studies* (Advances in Ecological Research, Academic Press, Oxford, U.K., 2017), vol. 57, pp. 201–281.
10. S. Butler, J. P. O'Dwyer, Stability criteria for complex microbial communities. *Nat. Commun.* **9**, 2970 (2018).
11. A. G. Haldane, R. M. May, Systemic risk in banking ecosystems. *Nature* **469**, 351–355 (2011).
12. J. Moran, J.-P. Bouchaud, May's instability in large economies. *Phys. Rev. E* **100**, 032307 (2019).
13. H. M. Hastings, Stability of large systems. *Biosystems* **17**, 171–177 (1984).
14. L. N. Trefethen, M. Embree, Spectra and Pseudospectra: *The Behavior of Nonnormal Matrices and Operators* (Princeton University Press, 2005).
15. M. G. Neubert, H. Caswell, Alternatives to resilience for measuring the responses of ecological systems to perturbations. *Ecology* **78**, 653–665 (1997).
16. M. Asllani, R. Lambiotte, T. Carletti, Structure and dynamical behavior of non-normal networks. *Sci. Adv.* **4**, eaau9403 (2018).
17. Z. G. Nicolaou, T. Nishikawa, S. B. Nicholson, J. R. Green, A. E. Motter, Non-normality and non-monotonic dynamics in complex reaction networks. *Phys. Rev. Res.* **2**, 043059 (2020).
18. L. N. Trefethen, A. E. Trefethen, S. C. Reddy, T. A. Driscoll, Hydrodynamic stability without eigenvalues. *Science* **261**, 578–584 (1993).
19. B. F. Farrell, P. J. Ioannou, Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.* **53**, 2025–2040 (1996).
20. P. J. Schmid, Nonmodal stability theory. *Annu. Rev. Fluid Mech.* **39**, 129–162 (2007).
21. M. Asllani, T. Carletti, Topological resilience in non-normal networked systems. *Phys. Rev. E* **97**, 042302 (2018).
22. G. Hennequin, T. P. Vogels, W. Gerstner, Non-normal amplification in random balanced neuronal networks. *Phys. Rev. E* **86**, 011909 (2012).
23. E. Gudowska-Nowak, M. A. Nowak, D. R. Chialvo, J. K. Ochab, W. Tarnowski, From synaptic interactions to collective dynamics in random neuronal networks models: Critical role of eigenvectors and transient behavior. *Neural Comput.* **32**, 395–423 (2020).
24. G. Baggio, V. Rutten, G. Hennequin, S. Zampieri, Efficient communication over complex dynamical networks: The role of matrix non-normality. *Sci. Adv.* **6**, eaba2282 (2020).
25. T. Biancalani, F. Jafarpour, N. Goldenfeld, Giant amplification of noise in fluctuation-induced pattern formation. *Phys. Rev. Lett.* **118**, 018101 (2017).
26. R. Muolo, M. Asllani, D. Fanelli, P. K. Maini, T. Carletti, Patterns of non-normality in networked systems. *J. Theor. Biol.* **480**, 81–91 (2019).
27. G. Lindmark, C. Altafini, Centrality measures and the role of non-normality for network control energy reduction. *IEEE Control Syst. Lett.* **5**, 1013–1018 (2021).
28. S. Johnson, V. Domínguez-García, L. Donetti, M. A. Muñoz, Trophic coherence determines food-web stability. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17923–17928 (2014).
29. S. N. Dorogovtsev, J. F. F. Mendes, A. N. Samukhin, Giant strongly connected component of directed networks. *Phys. Rev. E* **64**, 025101 (2001).
30. N. Schwartz, R. Cohen, D. Ben-Avraham, A.-L. Barabási, S. Havlin, Percolation in directed scale-free networks. *Phys. Rev. E* **66**, 015104 (2002).
31. D. Garlaschelli, M. I. Loffredo, Patterns of link reciprocity in directed networks. *Phys. Rev. Lett.* **93**, 268701 (2004).
32. G. Bianconi, N. Gulbahce, A. E. Motter, Local structure of directed networks. *Phys. Rev. Lett.* **100**, 118701 (2008).
33. E. A. Leicht, M. E. J. Newman, Community structure in directed networks. *Phys. Rev. Lett.* **100**, 118703 (2008).
34. N. Masuda, Y. Kawamura, H. Kori, Impact of hierarchical modular structure on ranking of individual nodes in directed networks. *New J. Phys.* **11**, 113002 (2009).
35. F. D. Malliaros, M. Vazirgiannis, Clustering and community detection in directed networks: A survey. *Phys. Rep.* **533**, 95–142 (2013).
36. L. Ermann, K. M. Frahm, D. L. Shepelyansky, Google matrix analysis of directed networks. *Rev. Mod. Phys.* **87**, 1261–1310 (2015).
37. X. Liu, H. E. Stanley, J. Gao, Breakdown of interdependent directed networks. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1138–1143 (2016).
38. G. Timár, A. V. Goltsev, S. N. Dorogovtsev, J. F. F. Mendes, Mapping the structure of directed networks: Beyond the bow-tie diagram. *Phys. Rev. Lett.* **118**, 078301 (2017).
39. M. Fruchart, R. Hanai, P. B. Littlewood, V. Vitelli, Non-reciprocal phase transitions. *Nature* **592**, 363–369 (2021).
40. R. A. Horn, C. R. Johnson, *Matrix Analysis* (Cambridge University Press, ed. 2, 2012).
41. L. Elsner, M. H. C. Paardekooper, On measures of nonnormality of matrices. *Linear Algebra Appl.* **92**, 107–123 (1987).
42. S. Townley, D. Carslake, O. Kellie-Smith, D. McCarthy, D. Hodgson, Predicting transient amplification in perturbed ecological systems. *J. Appl. Ecol.* **44**, 1243–1251 (2007).
43. I. Stott, S. Townley, D. J. Hodgson, A framework for studying transient dynamics of population projection matrix models. *Ecol. Lett.* **14**, 959–970 (2011).
44. Q. Huang, M. A. Lewis, Homing fidelity and reproductive rate for migratory populations. *Theor. Ecol.* **8**, 187–205 (2015).
45. J. T. Chalker, B. Mehlig, Eigenvector statistics in non-Hermitian random matrix ensembles. *Phys. Rev. Lett.* **81**, 3367–3370 (1998).
46. B. K. Murphy, K. D. Miller, Balanced amplification: A new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).
47. F. Chung, L. Lu, V. Vu, Spectra of random graphs with given expected degrees. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6313–6318 (2003).
48. F. Juhász, On the asymptotic behaviour of the spectra of non-symmetric random (0,1) matrices. *Discrete Math.* **41**, 161–165 (1982).
49. I. Neri, F. L. Metz, Linear stability analysis of large dynamical systems on random directed graphs. *Phys. Rev. Res.* **2**, 033313 (2020).
50. A. E. Motter, Y. Yang, The unfolding and control of network cascades. *Phys. Today* **70**, 32–39 (2017).
51. W. Tarnowski, I. Neri, P. Vivo, Universal transient behavior in large dynamical systems on networks. *Phys. Rev. Res.* **2**, 023333 (2020).
52. Koblenz Network Collection (KONECT), http://konect.cc/.
53. Handbook of Network Analysis: The KONECT Project, http://github.com/kunegis/konect-handbook/raw/master/konect-handbook.pdf.
54. The Netzschleuder Network Catalogue and Repository, http://networks.skewed.de/.
55. The Colorado Index of Complex Networks, http://icon.colorado.edu/.
56. The Matrix Market Repository, http://math.nist.gov/MatrixMarket/.
57. L. de Haan, A. Ferreira, *Extreme Value Theory: An Introduction* (Springer Science and Business Media, 2006).
58. G. Halász, G. J. Székely, On the elementary symmetric polynomials of independent random variables. *Acta Math. Acad. Sci. H.* **28**, 397–400 (1976).
59. C.-G. Esseen, On the Liapunoff limit of error in the theory of probability. *Ark. Mat. Astron. Fys.* **28A**, 1–19 (1942).
60. I. G. Shevtsova, An improvement of convergence rate estimates in the Lyapunov theorem. *Dokl. Math.* **82**, 862–864 (2010).
61. E. Zakon, *Mathematical Analysis I* (The Saylor Foundation, 2011).
62. F. Bullo, *Lectures on Network Systems* (CreateSpace Independent Publishing Platform, 2018).
63. J. Cao, M. Olvera-Cravioto, Connectivity of a general class of inhomogeneous random digraphs. *Random Struct. Algorithms* **56**, 722–774 (2020).
64. Y. Wang, F. Gao, Deviation inequalities for an estimator of the conditional value-at-risk. *Oper. Res. Lett.* **38**, 236–239 (2010).
65. E. Çınlar, *Probability and Stochastics* (Springer Science and Business Media, 2011).
66. M. E. J. Newman, S. H. Strogatz, D. J. Watts, Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001).

Acknowledgments

Funding: This research was supported by ARO grant no. W911NF-19-1-0383. D.E. also acknowledges support from TÜBİTAK grant no. 119F125. **Author contributions:** All authors contributed to the design of the research. C.D., T.N., and D.E. processed the network data and performed the simulations. C.D., T.N., and A.E.M. led the modeling, analyzed the results, and wrote the paper. All authors approved the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper, in the Supplementary Materials, and/or at https://doi.org/10.5281/zenodo.5964372.

Submitted 14 October 2021
Accepted 2 June 2022
Published 15 July 2022
10.1126/sciadv.abm8310