


OPEN

Predicting collapse of adaptive networked systems without knowing the network

Leonhard Horstmeyer^{1,2,5}, Tuan Minh Pham^{1,2,5}, Jan Korbel^{1,2,5}  & Stefan Thurner^{1,2,3,4,5*}

The collapse of ecosystems, the extinction of species, and the breakdown of economic and financial networks usually hinges on topological properties of the underlying networks, such as the existence of self-sustaining (or autocatalytic) feedback cycles. Such collapses can be understood as a massive change of network topology, usually accompanied by the extinction of a macroscopic fraction of nodes and links. It is often related to the breakdown of the last relevant directed catalytic cycle within a dynamical system. Without detailed structural information it seems impossible to state, whether a network is robust or if it is likely to collapse in the near future. Here we show that it is nevertheless possible to predict collapse for a large class of systems that are governed by a linear (or linearized) dynamics. To compute the corresponding early warning signal, we require only non-structural information about the nodes' states such as species abundances in ecosystems, or company revenues in economic networks. It is shown that the existence of a single directed cycle in the network can be detected by a "quantization effect" of node states, that exists as a direct consequence of a corollary of the Perron–Frobenius theorem. The proposed early warning signal for the collapse of networked systems captures their structural instability without relying on structural information. We illustrate the validity of the approach in a transparent model of co-evolutionary ecosystems and show this quantization in systems of species evolution, epidemiology, and population dynamics.

Complex networked systems are prone to collapse. This is the case for species evolution^{1–3} where collapse may lead to mass extinctions in ecosystems^{4,5}, but it equally holds for economic systems^{6,7}, the financial sector⁸, social networks⁹, or societies as a whole¹⁰. The prediction of collapse in large networked systems is extremely challenging, since stability and resilience may depend on specific details of the system that are hard to monitor. Collapse might arise as a consequence of the extinction of a single species in an ecosystem^{11–14}, or the default of a single financial player^{15–18}. Often systems are composed of many thousands of constituents and even more interactions. Whether the elimination of any node or interaction in a network leads to collapse depends to a large extent on the structural details of the interaction network.

Let us consider a minimal version of a networked dynamical system composed of N components, given by

$$\frac{d}{dt}X_i = \sum_{j=1}^N M_{ij}X_j - \Phi X_i, \quad (1)$$

where X_i is the state of the i -th component. Depending on the context, X_i could be populations of species, revenues of companies, abundances of molecules, or the degree of contagiousness of individuals. The growth of each X_i depends on the other states via the interaction matrix, M , whose entries M_{ij} determine to what extent j influences i . Self-interactions are excluded, $M_{ii} = 0$; ΦX_i is a depletion or decay term. We consider the interactions described by a Metzler matrix, i.e., $M_{ij} \geq 0$. When representing M as an interaction network, a directed link is drawn from j to i if and only if $M_{ij} \neq 0$. The influence of individual nodes can spread through the entire network via (1), which results in interdependencies that depend on the global network structure. More complex behaviour arises when the network M is co-evolving or when constraints for the dynamics are in place. Many systems are

¹Section for the Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, A-1090, Vienna, Austria. ²Complexity Science Hub Vienna, Josefstädterstrasse 39, A-1080, Vienna, Austria. ³Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM, 87501, USA. ⁴IIASA, Schlossplatz 1, 2361, Laxenburg, Austria. ⁵These authors contributed equally: Leonhard Horstmeyer, Tuan Minh Pham, Jan Korbel and Stefan Thurner. *email: stefan.thurner@meduniwien.ac.at

well described by adaptive linear models as in (1), however, the matrix M also evolves. Examples include catalytic networks, chemical reaction networks, aging processes, or economical models^{19–21}. Many systems often behave effectively linear, especially when they evolve around an equilibrium or a quasi-stationary state. In ecological modelling, one often considers the slowly evolving degrees of freedom as effectively fixed and derives linear models for the fast variables²². Also linear epidemiological models are used when modeling situations close to an endemic state or just before the outbreak²³.

It is a well known fact that the existence of cycles in the interaction network, i.e. directed closed paths whose only repeated vertices are the first and the last, is a crucial stability factor for systems of type (1) since they constitute self-reinforcing and self-sustaining structures. Conversely, the absence of cycles indicates a lack of self-sustaining mechanisms. For the minimal linear model (1) it is known that all states decay to zero for $\Phi > 0$, when cycles are absent. A foundational underpinning of this argument was put forward by Manfred Eigen, who proposed autocatalytic cycles as a guiding principle for self-replication in nature, and a possible scenario for the origin of life²⁴. These ideas were further developed in²⁵ and later investigated in^{26–28}, and were tested experimentally²⁹. Other famous examples of self-sustaining cycles include the autocatalytic Bethe-Weizsäcker cycle of hydrogen burning in the sun³⁰, or the Calvin cycle³¹ of photosynthesis in plants. The interaction structure of complex systems often evolves over time (co-evolution), and cycles may get created and destroyed in this process. This structural evolution is generally the result of an adaptation process³², and happens at a much slower rate than the dynamics of the nodes. For example, evolutionary processes are much slower than the population dynamics of the species, and new ties between companies are formed at a slower rate than the revenue generation from these ties. The dynamics in (1) constitutes a fast process on a static or quasi-static interaction network, given by the interaction matrix M . In this view, collapse is tightly related to the rupture of all cycles. A system with a single cycle is the most unstable because the deletion of any cycle-node or link breaks the sustaining feedback mechanism. The systemic relevance of a node can therefore be seen as the number of cycles that pass through it. Generally, the more cycles there are in a networked system, the more stable it becomes. If we understand the collapse of a network as a massive change of topology, accompanied by the extinction of a macroscopic fraction of nodes, we can define a *collapse* of a networked dynamical system as the transition from a system with cycles to one without cycles. This definition follows the definition of collapse in the Jain-Krishna model³³, which is a prominent example of an adaptive linear networked model.

The central idea of this paper is to develop an early warning signal that detects the last surviving cycle in a networked dynamical system. We will show that—surprisingly—this is possible from the observation of the node states for systems of type (1) with binary (0, 1) entries in M . This result can be phrased as a mathematical theorem that extends a special case of the Perron–Frobenius theorem³⁴. Its significance in this context has hitherto not been recognized. Intuitively, the detection of cycles requires the full knowledge of the interaction network, since a cycle could pass through only two nodes or through the entire network. The fact that the proposed warning signal only requires information about the node states has two advantages. First, for a networked system with N nodes, the information about the entire interaction network could require up to $\mathcal{O}(N^2)$ observations, which usually becomes unfeasible for large systems. Second, detailed network information is often not available or accessible. For the case it is available, there are many algorithms^{35–38} to find the cycles.

This type of structural precursor signal, that only relies on non-structural observations, is different from previously used precursors. Precursor signal theory has mostly focused on changes in time series close to a critical transition or a collapse point. In particular, classical precursor signals exploit the increasing auto-correlations³⁹, increasing volatility, increasing relaxation times from perturbations⁴⁰, and their scaling behaviour close to the transition⁴¹. These methods typically probe the stability of the fast changing processes and can not capture the presence of cycles. However, for systems with complex network structure the stability of the network itself is crucial, i.e. the stability with respect to a structural perturbation such as the deletion of nodes or links. Precursor signals that take into account the network structure directly are typically based on motif analysis⁴². Network motifs are local structures that do not capture the influence of the entire network. The risk of collapse can be severely misjudged from this partial knowledge, as has been recently demonstrated in the context of adaptive epidemiology⁴³.

Results

Consider a system that evolves as (1). We state our main result in the simplest formulation of a directed unweighted network M , which means that $M_{ij} = 1$, if the change of i depends on the change of j . In such systems, the normalized vector $x_i = X_i / \sum_j X_j$ “quantizes” if there is one single cycle remaining in M . This means that all entries in x_i are multiples of some minimum value x_{\min} . This is captured in the following theorem:

Theorem 1 (Eigenvector Quantization). *Let M be a binary matrix with entries $M_{ij} \in \{0, 1\}$ and diagonal entries $M_{ii} = 0$ for all $i \in \{1, \dots, N\}$. Let G be the directed network with directed adjacency matrix M . Let $X(t) = (X_1(t), \dots, X_N(t))$ be an N -dimensional state vector, whose components $X_i(t)$ evolve according to (1). For all initial conditions $X(0)$, except for a set of Lebesgue-measure zero, the normalised vector $x(t)$, defined component-wise as $x_i(t) = X_i(t) / \sum_j X_j(t)$, converges to a stable fixed point $x = \lim_{t \rightarrow \infty} x(t)$, for which the following holds:*

Eigenvector Quantization: *Suppose G contains only one single cycle. Then any component x_i can be expressed as*

$$x_i = n_i x_{\min}, \quad (2)$$

where x_{\min} is the minimal non-zero component and n_i is a natural number. The value of x_{\min} is taken by the cycle-nodes, and the integer $n_i \geq 0$ is the number of directed paths that lead from cycle-nodes to node i . If there are no paths from cycle nodes, then $x_i = 0$.

For the proof of the theorem, see (SI). Note that the theorem gives a sufficient condition for the eigenvector quantization of the node states. A graphical illustration of the theorem is shown in Fig. 1. We compare two networks that are identical except for one link. The network in Fig. 1 on the left-hand side contains two cycles,

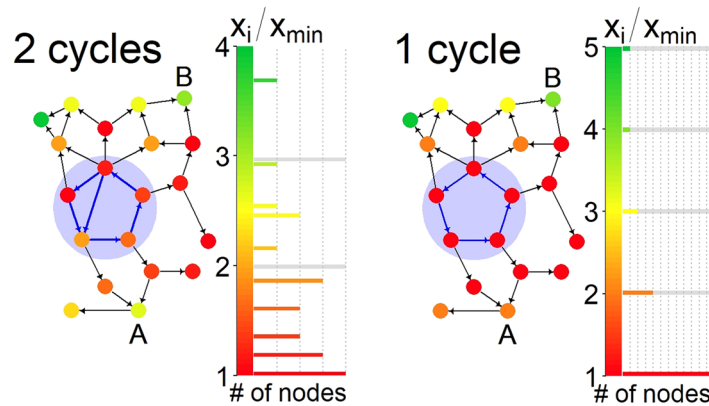


Figure 1. Graphical demonstration of the quantization theorem. Directed networks M containing two cycles (left) and one cycle (right). Cycles are in the shaded area. The node color indicates the state, x_i , (value of the component of the state vector, x) in units of the minimal value x_{\min} . The histograms show the number of nodes in a given state. In the right-hand figure we see the quantization of states due to the presence of a single (!) remaining cycle. The occurrence of quantization at the last remaining cycle can be used as a precursor signal. The state x_i/x_{\min} in the single-cycle network (right) coincides with the number of directed paths from the cycle to node i . Node A can be reached through two paths from the cycle, while node B can be reached by four. For the multi-cycle case (left), the number of paths no-longer coincides with the states.

indicated by blue arrows, the network in Fig. 1 right-hand side has only one last remaining cycle, again indicated by blue arrows. The node color corresponds to the respective components of the normalized state vector, x , rescaled by the minimum value, x_{\min} . The histogram to the right of each network figure indicates the number of nodes, whose components share a given value of x/x_{\min} . It can be seen from the histogram in Fig. 1 right that the state vector takes only discrete, equidistant values when the network has one remaining cycle. Once the quantization occurs, one can immediately determine the number of paths by which a node can be reached from the cycle, x_i/x_{\min} . Node A can be reached via two paths from the cycle, while node B can be reached via four. For the multi-cycle case (a), the number of paths no-longer coincides with the states.

The key mechanism behind quantization can be explained as follows. The state of a given node does not only depend on the nodes which feed directly into it, but through them, indirectly also depends on a large number of other nodes, and possibly the entire network. This global interdependence is best understood through the equation that is satisfied by the stable fixed point of the normalized state-vector $x = \lim_{t \rightarrow \infty} x(t)$, namely the eigenvalue equation (Lemma 1 in SI):

$$\lambda_1 x_i = \sum_j M_{ij} x_j \quad (3)$$

where λ_1 is the maximal eigenvalue of M . This equation couples all node states. It is a well-known fact that $\lambda_1 = 1$ if the network contains only one cycle³⁴. In this case $x_i = \sum_j M_{ij} x_j$. The state of any node without incoming links must therefore be zero, and any node-state that receives only in-links from such zero-nodes also vanishes. Consider, on the other hand a single cycle in isolation. Each of its nodes inherits the states of their respective parent nodes, and therefore all cycle-nodes must have the same non-zero state. Now, if the cycle is embedded in a network, it has outgoing and incoming links. The incoming links cannot originate from the cycle, or else there would be nodes that are part of more than one cycle. Hence, they must come from vanishing nodes and cannot contribute to the cycle states. Therefore, also the nodes in the non-isolated cycle still have the same states. Finally, consider a node that has multiple in-links. Its state is simply the sum of the node-states pointing to it. Consequently, if a node has two, three, or four paths leading to it from the cycle (see node A or B in Fig. 1 right), its state is two times, three times, or four times the state of a cycle-node.

The magnitude of state x_i also captures the eigenvector centrality of the corresponding node i in the network. The eigenvector centrality is either defined iteratively or directly via Eq. (3), with λ_1 equal to the largest eigenvalue⁴⁴. The centrality x_i does not just measure the number of neighbors that i receives links from, but also accounts for their importance. So, a node with a high eigenvector centrality is one that receives links from very central nodes. The definition depends on the convention of the arrows for a given adjacency matrix. Equation (1) therefore gives rise to the eigenvector centrality in a dynamical way. Likewise, one may ask to which extent the quantization of the state vector changes for other centrality measures.

In practice, it is more useful to consider a reverse direction of the theorem, i.e., to predict one cycle from the eigenvector quantization. Indeed, the reverse direction is not true in general. When $\lambda_1 = 2, 3, \dots$, it may happen that we observe the quantization too. However, for sparse networks, where λ_1 is typically close to one, the reverse is true. In particular, this is the case for the Jain–Krishna model, which is presented in the following. In this case the eigenvector quantization can be used to determine the structural stability of and can provide a warning signal for impending collapse. Further, one may determine the path-distance of a given node to the last remaining cycle. In the following, we discuss several examples in which the quantization can be detected. In the first example, we focus on the aforementioned Jain–Krishna model where we demonstrate explicitly how well it works to actually predict a collapse.

Co-evolutionary complex systems — Jain–Krishna model. We consider a simple model for the evolutionary dynamics of a toy ecosystem³³. It was used to explain the rapid changes (punctuated equilibria) in species diversity and the network mechanisms behind it. In particular, the model shows large-scale collapses in diversity. We briefly describe the model, explain how our main result applies, and show how predictions about the average time to collapse can be obtained.

The Jain–Krishna model is a co-evolving network model that distinguishes between a fast time scale, at which populations evolve on a fixed interaction network M , and a slow time scale, at which the network evolves through selection and mutation of its species. The population X_i of species i is a proxy for its abundance. It evolves according to the linear catalytic dynamics given by (1). The relative abundance $x_i(t)$ converges to a fixed point x_i . Thus, x_i is endogenously defined by the interaction network M and can be used as a measure for the fitness of species. The network is updated on the slow timescale as follows: the species with the smallest value of x_i (along with all its links) is eliminated. A new species is introduced, again labelled by i , which forms relations (links) with, on average, m randomly selected already existing species j . This is realised by independently assigning in-links to ($M_{ji} = 1$), and out-links from ($M_{ij} = 1$) the new species, i , both with the same probability $m/(N - 1)$. In³³, a range of m from 0.05 to 2 is considered. For $m < 1$, the model shows collapses. The Jain–Krishna model depends on two characteristic parameters: the average connectivity m , and the maximal diversity N . The observables in the model are the abundances x_i and the diversity given by the number of species that have a positive abundance,

$$S = \sum_{i=1}^N \Theta(x_i),$$

where Θ is the Heaviside step function that equals 1 whenever its argument is greater than 0.

We briefly describe the mechanisms that lead to species diversification and to collapse. If there are no cycles in the network, most species must die out. We show this formally in the SI (case $\lambda = 0$ in Lemma 1), but here we provide an illustrative example. Consider only two species connected by a directed link. The receiving species grows proportional to the giving species. Asymptotically the former species gains a positive relative abundance, whereas the latter vanishes in relative terms. This argument can be extended to all cycle-free graphs, for which it can be shown that the only the species at the end of chains attain a positive relative abundance and the rest have asymptotically vanishing relative abundance (Lemma 1 of the SI). If there is at least one cycle, many species have positive abundance, they exist. Existing species are either part of a cycle, or have incoming paths from a cycle. In the network updating process the system may change, if a hitherto unfit species dies out and upon its re-introduction attaches itself to a fit (very abundant) species, thus becoming fit itself. The network updating process can also lead to the creation of new cycles, which can lead to a strengthening of the existing system³³. Eventually, cycles can also break, whenever cycle-nodes become the least abundant species. This inevitably happens, since those species which are not part of a cycle either are already well adapted by influx from highly abundant species, or keep being replaced until they become well adapted. When all cycles break, all self-sustaining structures disappear, and the ecosystem collapses.

Every collapse is always preceded by a phase where only one remaining cycle exists in the network, see (SI). This is demonstrated in Fig. 2 left, where a typical example of the Jain–Krishna collapse mechanism is shown. The detection of this last cycle thus becomes a warning signal for the impending collapse. Since the population dynamics is governed by (1), the quantization theorem applies for the Jain–Krishna model. Whenever the species vector x quantizes a warning signal is produced; the network now has only one cycle left and is potentially very close to collapse.

We can estimate the expected time-to-collapse once quantization is observed. This calculation is shown in detail in the (SI). Here we sketch the main idea: First, the probability of removing a cycle-node in the critical regime is $p = p_w p_c$, where p_w is the probability that the cycle-node belongs to the set of the weakest nodes, and p_c is the probability of choosing it from this set. The former can be approximated for sufficiently large sparse graphs ($m < 1$ and $N \gg 1$) by the average network connectivity m , $p_w \sim m$. This is true because a new species only survives in the ecosystem if it receives at least one in-link, which happens with probability $1 - (1 - m/(N - 1))^{N-1} \sim m$. Otherwise, it remains unpopulated and becomes the weakest species with vanishing abundance and will be replaced again. The latter, p_c , can be expressed as the probability that a cycle of length L_c is formed among the L weakest nodes. This probability assumes its maximum and hence most likely value $p_c \sim 1/e$ at $L_c \sim L/e$, as shown in the (SI) (e is Euler's constant).

In summary, we obtain that the probability of removing a cycle-node is $p = m/e$. It is now possible to estimate the average life-time of the ecosystem once it enters the critical phase. The probability that the system collapses within the next T steps after the appearance of the quantization-based warning signal is given by $P(T) = (1 - p)^{T-1} p$. The expected time-to-collapse in the critical phase can then be expressed as

$$\langle T \rangle = \sum_{T=1}^{\infty} T P(T) = \frac{1}{p} \equiv \frac{e}{m}. \quad (4)$$

We explain the intuition behind this result. Since m is the rate at which a populated species is chosen for extinction, the factor $1/m$ gives the average time that it takes to remove a populated node. Further, since there is a chance $1/e$ that this node is a cycle-node, one needs to repeat the network update on average e times to pick out one cycle-node. Figure 2 right shows the comparison of the average times-to-collapse $\langle T \rangle$ for numerical simulations of the model and the theoretical predictions from (4). We run the model on R for a range of connectivities m and system sizes N , always for 10^7 time steps. At each collapse, we record the time between the first observation of the quantization and the collapse. We average these times for all collapses in the run. The simulation results follow the theoretical predictions very well, irrespective of system size and m .

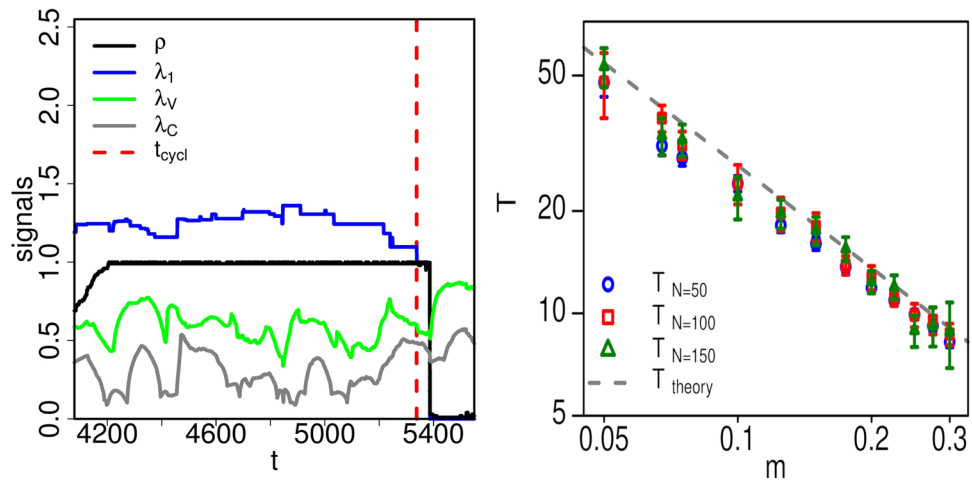


Figure 2. Left: sample run of the Jain-Krishna model. ρ is the fraction of populated nodes, λ_1 is the largest eigenvalue. The time point t_{cycl} denotes the moment of entering the one-cycle phase (where the eigenvector quantization appears). We see that the collapse is preceded by the one-cycle phase. For comparison, we add two other precursor measures — volatility spectral radius λ_V and cross-correlation spectral radius λ_C (their definitions are in SI). We observe no signal *before* the collapse, the signal appears *after* the collapse. Right: comparison between numerical simulations (symbols) and the theoretical prediction of (4) (dashed line) of the expected time-to-collapse, $\langle T \rangle$, for various system sizes $N \in \{25, 100, 200\}$, and connectivities, m between 0.05 and 0.375. Simulation results follow the theoretical prediction closely, independent of system size and m .

Finally, we compare the eigenvector quantization with other types of precursor signals. Well known precursor signals based on critical slowing down near the collapse point are given by the spectral radius of the lagged cross-correlation matrix λ_c and the spectral radius of volatility λ_V ⁴⁰. For their definitions, see SI. We show the various precursor signals in Fig. 2 left, the cross-correlation spectral radius (green), and volatility spectral radius (grey). We observe no characteristic signals for the spectral radius before the collapse. Signals appear only after the collapse. We have run several independent simulation runs and never observed any signal for any of these two measures.

Epidemic spreading in the SIS model. Let us now discuss a model where traces of the eigenvector quantization can be observed. We consider the classical SIS model on a sparse, directed (static) network M . It describes the spreading of a disease without immunity. Nodes can be either infected (I) or susceptible (S). An infected node can recover at a rate, r , or transmit the infection along a directed link to a susceptible neighbour at a rate, β . According to the individual-based mean-field approximation, the time evolution of the probability, p_i , that a node i is infected is given by²³

$$\frac{d}{dt}p_i = \beta \sum_j M_{ij} p_j (1 - p_i) - r p_i. \tag{5}$$

When the average disease prevalence, p_i , is small, i.e., $1 - p_i \approx 1$, the dynamics of (5) becomes effectively linear and can be cast into the form of (1) by rescaling time $t' = \beta t$, and setting $\Phi = r/\beta$. Therefore, in this regime the theorem applies. However, because of these two approximations, the linearity condition is not exactly met. We expect a deviation of p from the exact eigenvector quantization, which appears in the limiting vector x of the linear dynamics (1), whenever the network contains just one single cycle.

We run numerical simulations of the SIS model to compare the situation of single-cycle with the multi-cycle networks. The probability that a given node is infected, p_i , is approximated by the average time it spends in the infected state, as a fraction of the total simulation time. We draw 10^3 single- and multi-cycle graphs from the ensemble of Erdős–Rényi graphs of size $N = 200$ at the critical connectivity parameter, where the giant component would appear almost surely in the $N \rightarrow \infty$ limit. We set $\beta = 1$, which corresponds to fixing a time-scale. Since for any finite value of r the infection dies out at some point (In the language of probability theory, this event happens almost surely, i.e., with probability one), we have re-initialized the model randomly, whenever this happened. For each network the duration of the simulation, including re-initializations, was $10^4 \times (r + \beta)$, which ensures the same average number of events per simulation for a fixed network size. The average infection time, p_i , is then recorded for every node and rescaled by a factor p_{min} , that shifts the location of the first maximum to 1.

In Fig. 3 left we show the comparison between the rescaled values p/p_{min} , obtained from numerical simulations, and the rescaled eigenvector x/x_{min} for one randomly sampled Erdős–Rényi graph with a single cycle. The level structure of the infection duration is clearly visible and strongly correlated with the exact level structure of the eigenvector. In Fig. 3 right we show the histogram of p/p_{min} generated by sampling 10^3 networks from the single-cycle ensemble and we compare it to the histogram of the multi-cycle ensemble in the inset. We choose a recovery rate $r = 0.5$. For r larger than $\beta = 1$, the infection has no chance to propagate through the cycle and the effect becomes less pronounced when compared to the $r < \beta$ case, yet it is still visible. We color-coded the

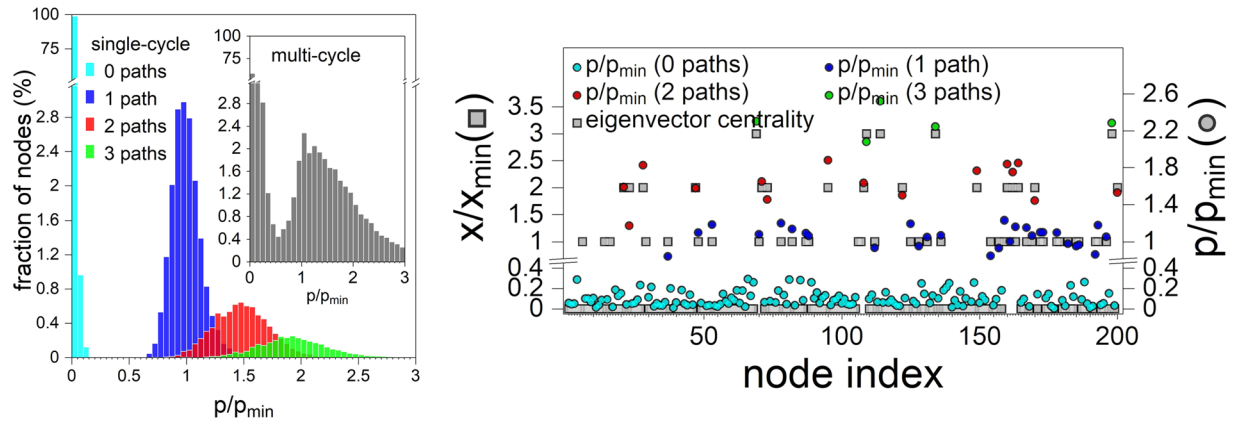


Figure 3. Comparison between the rescaled state vector, x/x_{\min} , and the rescaled average time of node infection, p/p_{\min} , which is taken as a proxy for the infection probability of a given node. Left: situation for a single-cycle network with $N=200$ species that was randomly sampled from the Erdős–Rényi ensemble. Circles represent the average time of infection of the respective nodes, squares are the components of the rescaled eigenvector. The strong correlation between the two is apparent. Right: histogram of the average infection times for 10^3 different networks of size $N=200$ with single (main) and multiple cycles (inset). Note the presence of four peaks that correspond to nodes with zero, one, two, and three paths from the cycle. They are colored with cyan, blue, red, and green, respectively. No clustering around integer values is present for multi-cycle networks (inset). All networks are drawn from the Erdős–Rényi ensemble with at least one node being linked to the cycle via three paths.

contributions according to the number of paths that lead to the contributing node from the cycle. One can clearly see the signature of the eigenvector quantization with four peaks in the single-cycle histogram, which indicates the average time of infection for nodes with zero, one, two, or three paths from the cycle. Quantization is not exact, as expected. Higher levels are not at integer values, but are shifted toward lower values. The reason for this is that the linearity assumption (5) breaks down. A node with two or three incoming paths from the cycle has a higher average duration of infection, because it can receive the infection in multiple ways. It cannot, however, receive the infection when it is already infected, which is reflected in the non-linear term $p_j(1-p_i)$ in (5) that forces the value of the infection probability p_i to be less than the multiplicity of paths.

Generalizations of eigenvector quantization. Here we demonstrate that the eigenvector quantization is still approximately observable for mild generalizations of eigenvector centrality of binary networks. We focus on three examples — first is the Katz centrality, second the eigenvector centrality of (moderately) weighted networks and third a logarithmic variant of the eigenvector quantization in weighted networks.

Katz centrality. Apart from the eigenvector centrality (3), one often considers the Katz centrality which is defined as

$$x_i^{(K)} = \alpha \sum_j M_{ij} x_j^{(K)} + \beta, \quad (6)$$

where $0 < \alpha < (\lambda_1)^{-1}$ and $\beta > 0$. The main difference between the Katz and the eigenvector centralities is the second constant term which gives any vertex (even those without any in-link) a score β in addition to the contribution from the centrality of its neighbours. We would like to show the possibility to observe an approximate eigenvector-quantization-like picture for the Katz centrality. In fact, the histograms of single-cycle networks Fig. 4(b) still show the level structure, but as long as α decreases, these levels are shifted from integer values and split up into sublevels.

Slightly weighted networks. We consider a moderately weighted network, by drawing the weights of the links from the interval $[1-\varepsilon, 1]$ uniformly at random, where ε is a noise parameter. When $\varepsilon=0$ we recover the unweighted case. We investigate to which extent it is possible to distinguish the single-cycle networks from the multicycle networks. The noise parameter ε is varied from zero, the baseline model, to one, the fully random model.

The histograms of single-cycle networks still show the quantization peaks as one departs from the baseline scenario, but as ε increases, these peaks broaden and gradually lose their equidistance. For the single-cycle network in Fig. 4 left, we have looked at the histograms of x of weighted networks for various values of ε . The examples are seen in Fig. 4 right.

Example of the weighted network — Leslie model of population aging. The Leslie model is used in population ecology to understand age distributions by means of a linear reproduction dynamics. Even though it is mathematically trivial, it is relevant ecologically⁴⁵. For reproductions in ecosystems, it is important that cycles are present. The life-cycles of animals that undergo metamorphosis and significant transformations are examples of these

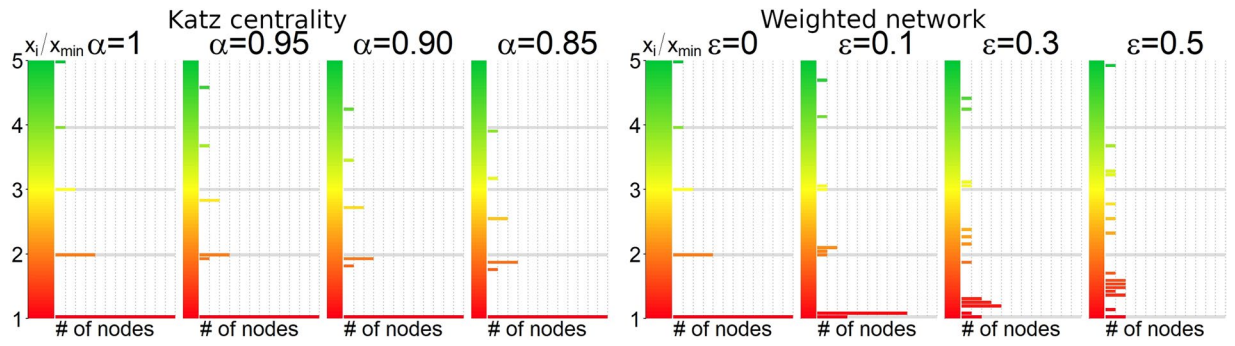


Figure 4. Left: histograms of the Katz centrality for single-cycle networks. The histograms show the number of nodes with a given value of $x^{(K)}/x_{\min}^{(K)}$ for different values of $\alpha \in \{1, 0.95, 0.9, 0.85\}$ and $\beta = 1$. For $\alpha \rightarrow 1$, the same quantization can be seen. For $\alpha \neq 1$, while the level structure is still observed, these levels are no longer at integer values and some of them can be split up into sub-levels. Right: histograms of weighted networks. The single-cycle network from Fig. 1 right has been subjected to a weighting of its links. The weights are drawn from the interval $[1 - \varepsilon, 1]$ uniformly at random. The histograms show the number of nodes with a given value of x/x_{\min} for various values of $\varepsilon \in \{0, 0.1, 0.3, 0.5\}$. For $\varepsilon = 0$ the pure quantization can be seen. For $\varepsilon > 0$ the signature deteriorates.

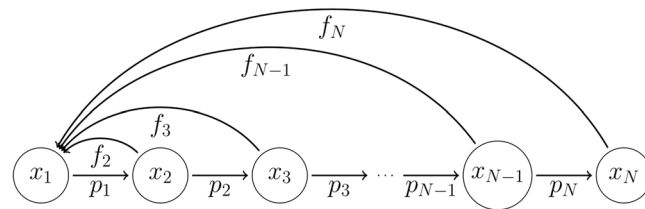


Figure 5. Illustration of the transitions in the Leslie model. There are N stages in the population (circles), each characterized by a population size x_i . A percentage $p_i \leq 1$ of the population in the i^{th} stage survives to the next stage p_{i+1} , while the number of new individuals born into stage 1 for a given individual at stage i is f_i .

reproductive cycles. Salmon, for instance, pass through various stages that involve the adaptation from sweet water to saltwater environments, and large migrations⁴⁶. If at any of these stages there is a bottleneck, e.g. closed migration routes, extensive fishery, or a major ecological catastrophe, then the reproduction cycle is broken and the entire population is at risk of extinction.

We now describe how the theorem extends to the case of weighted links by means of the Leslie model. The population is divided into N stages, with population sizes $x_1(t), \dots, x_N(t)$ at time t . The population at time $t + 1$ is obtained from the population at t by (see also Fig. 5)

$$\begin{aligned} x_1(t + 1) &= \sum_{j>1} f_j x_j(t) \\ x_j(t + 1) &= p_{j-1} x_{j-1}(t) \quad \text{for } j > 1. \end{aligned} \tag{7}$$

The first equation represents reproduction, where f_j is the number of offspring per individual at stage j . The second set of equations determines the survival (or aging) process, where p_j is the fraction of individuals that survive to the next stage. In matrix form, $x(t + 1) = Lx(t)$, the components are given by $L_{ij} = f_j \delta_{i,1} + p_{j-1} \delta_{i-1,j}$, where δ_{ij} is the Kronecker delta. Here L is called the *Leslie matrix* and (7) is a discrete version of (1).

Since the Leslie matrix L is weighted, one cannot use the quantization criterion directly. However, often one is interested to know whether there are one or more cycles: certain species reproduce only in a certain stage of their life. This means that there is only one non-zero f_i , and $f_j = 0$, for $j \neq i$. Let us also consider the case where all survival percentages are similar, such that $p_i \sim p, \forall i$.

Without loss of generality, we choose the reproductive stage to be the last one, $i = N$. If this is not the case, we focus only on the part of the system, which creates the reproduction cycle; the remaining stages do not feed back into the cycle and thus don't contribute to the reproduction. For long times, the population is determined by the eigenvector equation, $Lx = \lambda_1 x$. In this case, it is possible to calculate λ_1 and the corresponding eigenvector, $\lambda_1 = p(f/p)^{\frac{1}{N}}$, and

$$x = \left(\left(\frac{f}{p} \right)^{\frac{N-1}{N}}, \left(\frac{f}{p} \right)^{\frac{N-2}{N}}, \dots, \left(\frac{f}{p} \right)^{\frac{1}{N}}, 1 \right).$$

The logarithm of the population vector then reads

$$\log x = \frac{1}{N} \log \frac{f}{p} \cdot (N - 1, N - 2, \dots, 1, 0), \quad (8)$$

where we see a *logarithmic quantization*. Log-quantization also holds for more general networks, where we have all links with weight p , and a single reproduction link with weight, f . Moreover, each node should have only one in-link. For the unweighted network this corresponds to a situation where all links have the same population, $x_i = x_c$. This scenario disappears for two or more reproduction cycles.

Discussion

In many networked dynamical systems, directed cycles are crucial for self-sustaining feedback, autarchy, autocatalysis²⁴, and consequently, are responsible for the structural stability of these systems. The breakdown of the last cycle in such systems is often associated with a drastic change of dynamical behavior—often collapse. We found that for linear networked dynamical systems the presence of a single cycle in an evolving network can be detected by a quantization phenomenon in the dynamical variables of the node states; the detailed information about the entire network is not necessary. Our method can be used to probe whether the system is in a highly vulnerable and unstable state, prone to collapse in the near future. Since structural features often evolve much slower than node features, this method can be used as an early warning signal for the collapse of networked dynamical systems. The advantage of this method is that it probes relevant structural features (last cycle) without relying on any information about the interaction network. We demonstrated the predictive quality of the method in a transparent model of co-evolutionary dynamics.

We expect to see traces of these cycles in data from a diverse range of applications, ranging from autocatalytic sets of RNA molecules²⁹, gene regulatory networks⁴⁷, small-sized ecosystems, to life-cycles of migrating animals, and those undergoing metamorphosis. Applications might also extend to socio-economic situations. Recently, it has been argued that autocatalytic cycles appear in technology⁴⁸, and production networks⁴⁹. A detailed data analysis of these systems might further confirm the relevance of the presented results. Note that it might not be the vector of populations, abundances, or revenues that exhibits the quantization in the single-cycle phase; it could show in the vector of deviations from its mean since the dynamics of many systems is linear around a given equilibrium.

The presented approach has limitations. We discuss the three most significant ones. For weighted networks the exact result does no longer hold. However, extensions and generalizations can be developed. The first extension to networks with link weights has been discussed in the last example of the previous section, which treats the population-age dynamics in the Leslie model with uniform survival and reproduction fractions. There we showed that a variant of the eigenvector quantization holds, where the logarithm of the node states is quantized in the presence of a single-cycle network. The second extension to moderately weighted networks, with weights taken at random from a small interval around one, has been discussed in the second example of the previous section. The general case of arbitrarily weighted networks requires a more thorough treatment in further research. The other crucial assumption behind our approach is the linearity of the dynamics (1). We have shown in the epidemic spreading example that systems still retain signatures of quantization, even though the dynamics is not strictly linear. In this example, nonlinearity broadens the quantization peaks and for all positive levels ($n \geq 1$) it reduces the inter-peak distance while retaining equidistance. However, even in linear systems, eigenvector quantization may cause false alarms, since in exceptional cases there may be an apparent quantization effect for very specific network-topologies with multiple cycles.

The linearity assumption implies that the eigenvector centrality is a reasonable observable of the system. In linear systems, the asymptotic relative state vector coincides with the eigenvector centrality. It is worth investigating to what extent other centrality measures, that arise from different (non-linear) dynamics, such as in the epidemic example, show an equivalent behaviour to the eigenvector quantization. We took first steps in this direction by investigating the Katz centrality for parameters that are close to the limiting case of the eigenvector centrality. A more comprehensive analysis is necessary in future work.

Finally, note that no assumptions are necessary about the dynamics of the network evolution, its rewiring, the deletion or addition of links or nodes. This fact is due to an immediate corollary of the Perron-Frobenius theorem — it follows that for linear dynamics, as in (1) the absence of cycles corresponds to a collapsed state because the vast majority of nodes cannot sustain themselves; their respective states vanish. Of course, one can conceive scenarios in which the no-cycle regime does not correspond to collapse states, or in which also other configurations are associated with collapse. It can also be that in such cases the collapsed regime can be reached directly without passing through the single-cycle phase.

Received: 4 September 2019; Accepted: 27 December 2019;
Published online: 27 January 2020

References

1. Jackson, J. B. *et al.* Historical overfishing and the recent collapse of coastal ecosystems. *Science* **293**, 629–637 (2001).
2. Bellwood, D. R., Hughes, T. P., Folke, C. & Nyström, M. Confronting the coral reef crisis. *Nature* **429**, 827 (2004).
3. Hanel, R. & Thurner, S. Solution of the unanimity rule on exponential, uniform and scalefree networks: A simple model for biodiversity collapse in foodwebs. *European Physical Journal B*. **62**, 327–330 (2008).
4. Wake, D. B. & Vredenburg, V. T. Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences* (2008).
5. Barnosky, A. D. *et al.* Has the earth's sixth mass extinction already arrived? *Nature* **471**, 51 (2011).
6. Schumpeter, J. A. *Business cycles*, vol. 1 (McGraw-Hill New York, 1939).
7. Klimek, P., Hausmann, R. & Thurner, S. Empirical confirmation of creative destruction from world trade data. *PloS one* **7**, e38924 (2012).

8. Haldane, A. G. & May, R. M. Systemic risk in banking ecosystems. *Nature* **469**, 351 (2011).
9. Murase, Y., Jo, H.-H., Török, J., Kertész, J. & Kaski, K. Modeling the role of relationship fading and breakup in social network formation. *PLoS one* **10**, e0133005 (2015).
10. Tainter, J. *The collapse of complex societies*. (Cambridge University Press, 1990).
11. Naem, S. *et al.* Declining biodiversity can alter the performance of ecosystems. *Nature* **368**, 734–737 (1994).
12. Rizzo, D. M. & Garbelotto, M. Sudden oak death: endangering california and oregon forest ecosystems. *Frontiers in Ecology and the Environment* **1**, 197–204 (2003).
13. Sekerciöglu, Ç. H., Daily, G. C. & Ehrlich, P. R. Ecosystem consequences of bird declines. *Proceedings of the National Academy of Sciences* **101**, 18042–18047 (2004).
14. Hallmann, C. A. *et al.* More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS one* **12**, e0185809 (2017).
15. Elliott, M., Golub, B. & Jackson, M. O. Financial networks and contagion. *American Economic Review* **104**, 3115–53 (2014).
16. Eisenberg, L. & Noe, T. H. Systemic risk in financial systems. *Management Science* **47**, 236–249, <https://doi.org/10.1287/mnsc.47.2.236.9835> (2001).
17. Poledna, S. & Thurner, S. Elimination of systemic risk in financial networks by means of a systemic risk transaction tax. *Quantitative Finance* **16**, 1599–1613 (2016).
18. Klimek, P., Poledna, S., Farmer, J. D. & Thurner, S. To bail-out or to bail-in? Answers from an agent-based model. *Journal of Economic Dynamics and Control* **50**, 144–154 (2015).
19. Leontief, W. *Input-output Economics*. (Oxford University Press, 1986).
20. Leontief, W. Quantitative input and output relations in the economic systems of the united states. *The Review of Economics and Statistics* **18**, 105–125 (1936).
21. Ghosh, A. Input-output approach in an allocation system. *Economica* **25**, 58–64 (1958).
22. Rinaldi, S. & Scheffer, M. Geometric analysis of ecological models with slow and fast processes. *Ecosystems* **3**, 507–521 (2000).
23. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Review of Modern Physics* **87**, 925–979, <https://doi.org/10.1103/RevModPhys.87.925> (2015).
24. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523, <https://doi.org/10.1007/BF00623322> (1971).
25. Eigen, M. & Schuster, P. The hypercycle. *Naturwissenschaften* **65**, 7–41 (1978).
26. Kauffman, S. A. Autocatalytic sets of proteins. *Journal of Theoretical Biology* **119**, 1–24 (1986).
27. Kauffman, S. A. *The Origins of Order Self-Organization and Selection in Evolution*. (Oxford University Press, 1993).
28. Farmer, J. D., Kauffman, S. A. & Packard, N. H. Autocatalytic replication of polymers. *Physica D* **2**, 50–67 (1986).
29. Arsène, S., Ameta, S., Lehman, N., Griffiths, A. D. & Nghe, P. Coupled catabolism and anabolism in autocatalytic rna sets. *Nucleic Acids Research* **46**, 9660–9666, <https://doi.org/10.1093/nar/gky598> (2018).
30. Bethe, H. A. Energy production in stars. *Physical Review* **55**(103–103), 434–456, <https://doi.org/10.1103/PhysRev.55.103> (1939).
31. Bassham, J. A. & Calvin, M. The path of carbon in photosynthesis. In *Die CO₂-Assimilation/The Assimilation of Carbon Dioxide*, 884–922 (Springer, 1960).
32. Gross, T. & Sayama, H. *Adaptive Networks: Theory, Models and Applications*. (Springer Publishing Company, Incorporated, 2009).
33. Jain, S. & Krishna, S. Autocatalytic sets and the growth of complexity in an evolutionary model. *Physical Review Letters* **81**, 5684–5687, <https://doi.org/10.1103/PhysRevLett.81.5684> (1998).
34. Meyer, C. D. *Matrix Analysis and Applied Linear Algebra*. (Society for Industrial and Applied Mathematics, 2000).
35. Tarjan, R. Enumeration of the elementary circuits of a directed graph. *SIAM Journal on Computing* **2**, 211–216 (1973).
36. Mateti, P. & Deo, N. On algorithms for enumerating all circuits of a graph. *SIAM Journal on Computing* **5**, 90–99 (1976).
37. Frieze, A. M. An algorithm for finding hamilton cycles in random directed graphs. *Journal of Algorithms* **9**, 181–204 (1988).
38. Evans, J. *Optimization algorithms for networks and graphs* (Routledge, 2017).
39. Ramos, O. Criticality in earthquakes. good or bad for prediction? *Tectonophysics* **485**, 321–326, <https://doi.org/10.1016/j.tecto.2009.11.007> (2010).
40. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**(7260), 53–9, <https://doi.org/10.1038/nature08227> (2009).
41. Scheffer, M. *et al.* Anticipating Critical Transitions. *Science* **338**, 344, <https://doi.org/10.1126/science.1225244> (2012).
42. Squartini, T., Van Lelyveld, I. & Garlaschelli, D. Early-warning signals of topological collapse in interbank networks. *Scientific Reports* **3**, 3357 (2013).
43. Horstmeyer, L., Kuehn, C. & Thurner, S. Network topology near criticality in adaptive epidemics. *Physical Review E*. **98**, 042313, <https://doi.org/10.1103/PhysRevE.98.042313> (2018).
44. Newman, M. *Networks: An Introduction*. (Oxford University Press, 2010).
45. Leslie, P. H. On the use of matrices in certain population mathematics. *Biometrika* **33**, 183–212, <https://doi.org/10.1093/biomet/33.3.183> (1945).
46. Groot, C. & Margolis, L. *Pacific salmon life histories*. (UBC press, 1991).
47. Stokić, D., Hanel, R. & Thurner, S. Inflation of the edge of chaos in a simple model of gene interaction networks. *Physical Review E* **77**, 061917, <https://doi.org/10.1103/PhysRevE.77.061917> (2008).
48. Napolitano, L., Evangelou, E., Pugliese, E., Zeppini, P. & Room, G. Technology networks: the autocatalytic origins of innovation. *Royal Society open science* **5**, 172445 (2018).
49. Thurner, S., Klimek, P. & Hanel, R. Schumpeterian economic dynamics as a quantifiable model of evolution. *New Journal of Physics* **12**, 075029 (2010).

Acknowledgements

We acknowledge support from the Austrian science foundation FWF projects under P 29252 and I 3073, and FFG project 857136.

Author contributions

L.H. and S.T. conceptualized the work. L.H., T.M.P., and J.K. did analytic and computational work. All authors analyzed, interpreted results and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-57751-y>.

Correspondence and requests for materials should be addressed to S.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020