

# Directed Cycles as Higher-Order Units of Information Processing in Complex Networks

Received: 12 August 2025

Accepted: 5 August 2026

Published online: 17 August 2026

**Cite this article as:** Rajpal, H., Expert, P., Vasiliauskaitė, V. Directed Cycles as Higher-Order Units of Information Processing in Complex Networks. *Commun Phys* (2026). <https://doi.org/10.1038/s42005-026-02820-3>

**Hardik Rajpal, Paul Expert & Vaiva Vasiliauskaitė**

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

# Directed Cycles as Higher-Order Units of Information Processing in Complex Networks

Hardik Rajpal<sup>1,2,3\*</sup>, Paul Expert<sup>4,+</sup>, and Vaiva Vasiliauskaitė<sup>5,6,+</sup>

<sup>1</sup>Centre for Complexity Science, Imperial College London, London, SW7 2AZ, United Kingdom

<sup>2</sup>Department of Mathematics, Imperial College London, London, SW7 2AZ, United Kingdom

<sup>3</sup>I-X Centre for AI in Science, Imperial College London, London, W12 0BZ, United Kingdom

<sup>4</sup>Global Business School for Health, Faculty of Population Health Sciences, University College London, WC1E 6BT, UK

<sup>5</sup>Laboratory of Biosensors and Bioengineering, Institute for Biomedical Engineering, Gloriastrasse 37/39, ETH Zürich, Zürich, 8092, Switzerland

<sup>6</sup>Computational Social Science, Stampfenbachstrasse 48, ETH Zürich, Zürich, 8092, Switzerland

\*h.rajpal15@imperial.ac.uk

+these authors contributed equally to this work

## ABSTRACT

Directed cycles are fundamental motifs in natural, social and artificial networks, yet their distinct roles in processing information remain poorly understood, particularly as higher-order structures. Furthermore, it is unclear how the function of a cycle depends on the wider network in which it is embedded. Here we show, using information-theoretic measures, that network size, sparsity and directionality critically shape how directed cycles process information. In networks with no preferred direction, feedforward cycles enable greater information flow, while feedback cycles support stronger information integration. The relative orientation of a feedforward cycle, and the structural incoherence it induces, tunes its capacity to generate higher-order behaviour. Introducing feedback into otherwise feedforward architectures increases the diversity of network activity while modifying the computational behaviour. In a supervised learning task, feedforward networks train faster and tolerate random node ablations, whereas feedback makes them robust to input noise. These findings reveal directed cycles as tunable computational units whose behaviour is set by their structural context.

## Introduction

Network science is a tool of choice to disentangle the relationship between the structure and the function of a complex system, across domains as diverse as social systems, technological infrastructures, and the human brain<sup>1</sup>. Graphs are mathematical objects representing pairwise relationships between the constituting elements of a system. Used as substrate for dynamics, the characterisation of their structural properties have helped to provide deep insight into processes such as contagion<sup>2</sup>, synchronization<sup>3,4</sup>, and resilience<sup>5</sup>. In the past ten years or so, the use of generalisations of graphs to represent more complicated interactions has increased dramatically, and are now collectively called as a group *higher-order network science*. The fundamental idea is to introduce a new object that explicitly encodes interactions among more than two elements, for example, simplicial complexes<sup>6</sup>, hypergraphs<sup>7</sup>, and higher-order Markov models<sup>8–11</sup>. These structures are then used to define new dynamics explicitly relying on these higher-order structures, with the aim of generating emergent behaviours that cannot be fully explained by dyadic interactions alone, higher-order behaviour, such as group synchronization<sup>12–15</sup>, collective decision-making<sup>16</sup>, and higher-order contagion<sup>17,18</sup>. However, the equivalence between higher-order structure and behaviour is not as clear as one might think<sup>19</sup>. Rosas *et al.*<sup>19</sup> show that higher-order behaviour can also emerge from specific arrangements of pairwise interactions, for example, signed pairwise networks can generate group-level frustration in systems of Ising spins. In particular, Rosas *et al.* argue for the use of information theoretical tools to detect the emergence of higher-order behaviour, in particular integrated information that measures the excess information a structure has about its future beyond what is available from its parts considered individually, and is thus a measure of emergent or higher-order behaviour.

Here, we aim to contribute to this discussion by showing that the ability of a network motif to generate higher-order behaviour and process information is not only intrinsic, but depends on how it is embedded in the wider network.

We consider directed networks as directness is a particularly important property in systems that process information, and the relative organisation of directed edges defines the function of a system<sup>20</sup>. Directed networks have been crucial in guiding our understanding of the informational architecture of the brain<sup>21</sup>, the propagation of social influence<sup>22</sup>, genetic regulatory pathways in living systems<sup>23</sup>, physiology<sup>24</sup> and the design of engineered systems such as feedforward and recurrent neural

networks<sup>25</sup>.

The role and importance of motifs, and in particular cycles, in networks cannot be understated, as they can play a crucial role in modulating the dynamics. Unlike undirected cycles, different classes of directed cycles can impact the flow of information on the network in fundamentally different ways, acting as minimal structures that generate higher-order *functional* behaviour - introducing memory, path dependencies, and non-Markovian dynamics, thereby shaping both local information processing and global system behaviour<sup>26</sup>. Directed cycles, in particular feedforward and feedback cycles, have been explored across biological, neural, social, and engineered systems. Much previous research on the functional roles of these cyclic motifs comes from analysis of the structure of gene regulatory networks, which capture interactions between DNA, RNA, and protein molecules that collectively control gene expression levels<sup>27</sup>. Studies in this domain have shown that acyclic motifs contribute to dynamic stability in these networks<sup>28,29</sup>, with coherent feedforward cycles in particular enhancing robustness against update-rule perturbations in Boolean models<sup>30</sup>. Feedforward cycles have also been shown to act as sign-sensitive filters, with coherent variants functioning as persistence detectors that ignore short input fluctuations, and incoherent variants generating accelerated or pulse-like responses<sup>23</sup>. Organisation also matters with coherent coupling between feedforward motifs enlarging the basin of attraction of stable states<sup>31</sup>. In contrast, networks enriched in feedback loops support a larger number of attractors, increased state-space complexity, and greater entropy<sup>29</sup>. Positive feedback cycles in particular can stabilise equilibria against perturbations to initial conditions, but this robustness comes at the cost of fragility: unexpected mutations at nodes embedded in many feedback cycles are more likely to destabilise the system<sup>32</sup>.

Cyclic motifs have also been extensively studied in neuroscience. Primary visual regions of the brain are organised in a feedforward manner, transmitting information from lower to higher areas of the visual cortex<sup>33</sup>. These feedforward networks allow creation of multiple copies of the same signal, ensuring robust information transfer through redundancy<sup>34</sup>. In contrast, higher-order visual areas exhibit stronger recurrent feedback cycles that support complex processes such as attention<sup>33</sup> and enable efficient visual inference in terms of space, time, energy, and data<sup>35</sup>.

Beyond biology, cyclic structures are also central to social and computational systems. For instance, extensions of social balance theory to directed networks provide a more nuanced view of balance in social groups<sup>36</sup>, while in scale-free networks, feedback cycles have been shown to facilitate echo chambers and polarization<sup>37</sup>. Simulation studies have shown that the presence of feedforward and feedback cycles impacts the synchronizability of a network<sup>38</sup>. In directed ring networks, an abundance of feedback cycles relative to feedforward cycles can facilitate synchronization at lower coupling strengths<sup>39</sup>. Some studies have also mapped the computational capabilities of directed cycles: the in-degrees of source and target nodes in a feedforward cycle can modulate the amount of information transfer<sup>40</sup>, and structurally balanced directed cycles can improve the performance of sparse recurrent neural networks<sup>41</sup>. Together, these findings suggest that cycles are not random by-products of network wiring but functional units whose type and embedding critically shape information processing and resilience of the graph. Yet so far, functional properties of cycles have been evaluated mainly through system-specific measures such as stability in gene regulatory models or synchronizability in coupled oscillator systems. This leaves open the question of how to characterise their computational role in a task- and model-independent way<sup>42</sup>.

In this paper, our contribution sits at the boundary between higher-order structure, higher-order behaviour and cycle information processing capacity in directed networks. We consider the four-node feedback and feedforward cycles that we consider as basic computational units<sup>43</sup> in processing information and giving rise to higher-order behaviour, making it natural to use an information-theoretical characterisation of their roles<sup>19,44,45</sup>, as it enables the quantification of their ability to store information about their past, transfer information, and integrate information from different sources. We find that feedback and feedforward cycles have intrinsically different capabilities for generating higher-order behaviour, but that the magnitude of these capabilities depends strongly on the properties of the environment in which they are embedded. In particular, by embedding the feedforward cycles in different relative directions and quantifying the disruption of directionality using a measure of local trophic coherence, we can modulate the information processing at these cycles. Our results hold for linear and non-linear dynamics, hinting that the information processing properties of the cycles are general rather than dynamics-dependent. In particular, we present a simple mechanism that allows us to tune the amount of higher-order behaviour a cycle can generate.

## Results

Our investigations proceed in two steps: we first characterise the intrinsic properties of feedback and feedforward cycles in terms of information-theoretic measures that quantify the capacity of transferring, storing and integrating information using transfer entropy (TE), active information storage (AIS) and integrated information ( $\Phi$ ). We then quantify how these characteristics are modulated by the network they are embedded in. The analytical component of our results enables the strict characterisation of the structural and expected capacity of each type of cycle, while the computational component enables the systematic exploration of how key network parameters shape information processing in feedforward and feedback cycles. We vary network size, density, and a global directionality parameter to assess their effects on both cycle types. By embedding cycles at different positions within the network, we examine how structural context modulates  $\Phi$ , AIS and TE. In particular, the relative difference

between the local directionality of a cycle and the global directionality of the host network can be used to tune the amount of higher-order behaviour produced. The relative difference in directionality leads to a natural refinement in the cycle typology, with 'aligned' and 'reversed' feedforward cycles, whose behaviours are strikingly different. Conversely, we also show that local variations in the cyclic structure can propagate back onto the global dynamics of the system by investigating the computational advantages that the cycles confer by comparing the task performance of different neural network architectures.

Together, these experiments allow us to directly compare the computational roles of feedforward and feedback cycles under controlled conditions. Our results hold for two different classes of dynamical systems: a linear vector autoregressive model and a non-linear leaky integrate-and-fire model, strongly suggesting the results are not due to an interaction between model and structure, but a genuine general mechanism.

## Experimental setup

We quantified spatiotemporal patterns of computation and information processing in networks using information-theoretic measures applied to simulated dynamics. Each directed network consisted of nodes  $i$ , each associated with a time-varying variable  $X^i(t)$  and connected to a set  $\mathcal{N}_i$  of neighbours via directed edges. We consider two types of dynamics: a linear Vector Autoregressive (VAR) process and a non-linear Leaky Integrate-and-Fire (LIF) process. The VAR process models a diffusion-like dynamics on the graph where each node updates its state as a function of a linear combination of inputs from its neighbours with some independent Gaussian noise. This process is computationally efficient to simulate using vectorised operations and also analytically tractable for steady-state calculations. The LIF process, on the other hand, models a non-linear biological process of spike propagation and integration in spiking neural networks. Here, the nodes (neurons) integrate the spikes from their neighbours over time and produce a spike once the integrated input exceeds a threshold. After spiking, the neurons go through a refractory period, introducing delays in the propagation of spikes. This model enables us to probe the role of cycles to process information in a more complex, non-linear setting where a closed-form solution is in general, not available.

To map the information-theoretic properties of cycles in a network, we consider active information storage (AIS) (Fig. 1Ai), which quantifies the predictive influence of a node's own past, and transfer entropy (TE) (Fig. 1Aii), which captures directed influences from neighbouring nodes. To capture higher-order collective behaviour within these cycles, we compute integrated information  $\Phi$  (Fig. 1Aiii), which quantifies the excess information that the cycle has about its joint future beyond what is available from its parts—edges—considered individually. This measure captures the extent to which a cycle in the network generates dynamics with higher-order behaviour, producing information that cannot be reduced to the sum of its pairwise parts.

We applied these measures to cycles embedded in a network, treating them as computational units. We focus on two directed cycle types—feedforward and feedback—which represent fundamentally different modes of activity propagation in directed networks<sup>26</sup>. Based on their topology, we expect feedforward cycles to promote robust information transfer through redundancy and parallel paths, whereas feedback cycles should support temporal integration by combining inputs from different network regions while retaining memory of past activity.

## Feedforward and feedback cycles exhibit contrasting information-processing capacities

We first compare the information-processing capabilities of feedforward and feedback cycles embedded in sparse Erdős–Rényi directed networks with no preferred directionality. Each network contains one embedded cycle with either feedforward or feedback directionality. For both VAR and LIF models, we generate an ensemble of 10,000 simulated trajectories and extract the activity of the nodes forming a cycle. This ensemble enables time-resolved estimation of the information-theoretic measures used in our analysis<sup>46</sup> (see Supplementary Note 1 and Supplementary Fig. S1). To isolate the specific contribution of the cycle, we compare these values to those obtained from a null model in which the cycle is replaced by a disconnected pair of edges (see Network models section in the Methods for description of the null model).

In the steady-state regime of the VAR model, AIS is significantly higher for the feedback cycle than for the feedforward cycle (Figure 1Ci). Within the integrated information decomposition framework<sup>47</sup>, AIS can be expressed in terms of partially overlapping modes of information processing, including TE, which quantifies directed information transfer, and  $\Phi$ , which quantifies information integration. Because the VAR model implements a diffusion-like process, both TE and  $\Phi$  peak during the transient regime, marking the most dynamic phase of information processing (Figure 1Cii–Ciii). Beyond this peak, the two cycle motifs diverge: in the feedforward cycle, TE and  $\Phi$  decay toward low steady-state values, whereas in the feedback cycle they stabilise at substantially higher levels. We therefore characterise the information-processing capacity of each cycle by the peak value of each measure during the transient. Note that AIS additionally decomposes into other modes of information processing, such as unique storage and erasure; consequently, TE and  $\Phi$  do not sum to AIS but rather correspond to the partially overlapping modes on which we focus in this work (see Supplementary Note 2 for details).

When comparing these temporal maxima in the VAR model, feedforward cycles exhibit higher TE (Figure 1Cii), whereas feedback cycles show greater  $\Phi$  (Figure 1Ciii). Across 1,000 independent network realisations, these differences are statistically significant for both measures (Figure 1Di,Dii), confirming that the two cycle types occupy distinct roles in network information processing.

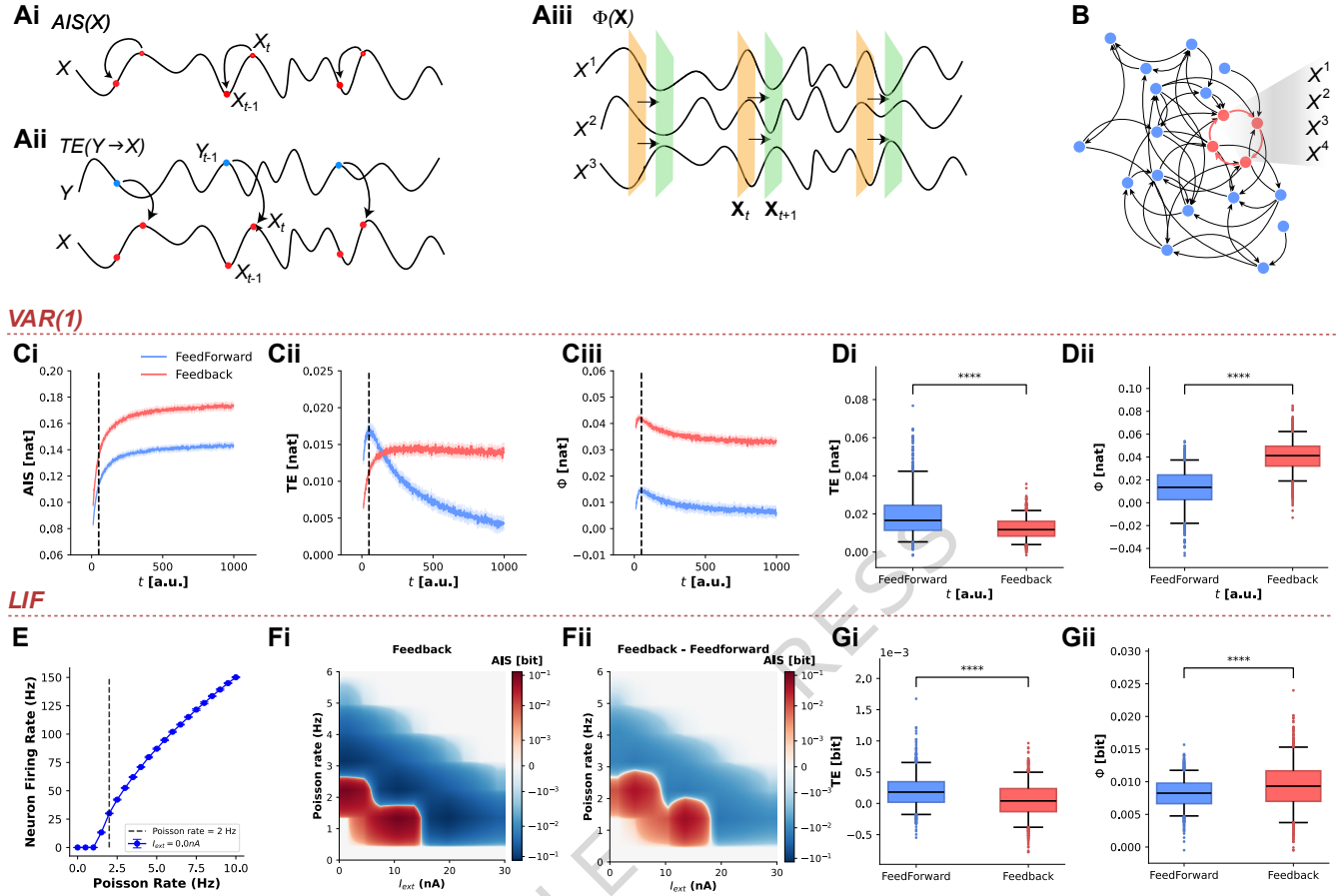


**Figure 1. Information-theoretic measures and their expression in cycles embedded into directed Erdős-Rényi graphs. A** Schematic representation of the information-theoretic measures used in this study. **Ai** *Active Information Storage* (AIS): predictive information stored in the past of a variable ( $X_{t-1}$ ) about its current state ( $X_t$ ). **Aii** *Transfer Entropy* (TE): information provided by the past of a source ( $Y_{t-1}$ ) about the present of a target ( $X_t$ ), beyond that stored in the target's own past. **Aiii** *Integrated Information* ( $\Phi$ ): excess information stored in the joint state of variables  $\mathbf{X}_t = \{X_t^1, X_t^2, X_t^3\}$  about their joint future  $\mathbf{X}_{t+1}$ , beyond the information contained in the individual parts ( $X_t^1, X_t^2, X_t^3$ ). **B** These measures are applied to time series extracted from a four-node embedded cycle in a network. **C** Estimates of each measure at every simulation time step, obtained from ensembles of directed Erdős-Rényi networks augmented with either a feedback or a feedforward cycle. AIS grows to a steady state (**Ci**), TE in the feedforward cycle peaks during the transient phase (**Cii**), and  $\Phi$  peaks during the transient in both cycle types (**Ciii**). **D** Information-processing characteristics at the transient peak: TE is significantly larger in feedforward cycles, whereas  $\Phi$  is larger in feedback cycles. **E** The Leaky Integrate-and-Fire model exhibits a phase transition as the external drive or the Poisson rate of incoming spikes increases. The activity of the network transitions from a quiescent state to a highly active state. We find that the information processing capabilities of the cycle are maximised in the vicinity of the phase transition. For our analysis, we stick to the spontaneous phase transition at  $I_{\text{ext}} = 0$  and Poisson rate  $\lambda = 2\text{Hz}$ . **Fi** In the LIF model, the difference in AIS between feedback cycle and the null model is maximised in the vicinity of the phase transition. The two dimensions of the phase transition are the external drive  $I_{\text{ext}}$  and the Poisson rate  $\lambda$  of incoming spikes. **Fii** Feedback cycles show higher AIS than feedforward cycles, this difference is also maximised in the vicinity of the phase transition. **G** In the LIF model, we observe similar trends to the VAR model, with feedforward cycles exhibiting higher TE and feedback cycles showing greater  $\Phi$  near the phase transition. All statistical comparisons were performed using a non-parametric permutation test with 10,000 permutations, and the significance level was set at  $p < 0.05$ , with  $n = 1,000$  independent network realisations for each cycle.

The LIF model, by contrast, generates stationary steady-state dynamics. The nature of these dynamics depends on model parameters, in particular the external drive  $I_{\text{ext}}$  and the Poisson rate  $\lambda$  of spontaneously generated spikes. A parameter sweep across  $(I_{\text{ext}}, \lambda)$  reveals that the difference in AIS between feedback and feedforward cycles is maximised in the vicinity of the phase transition between the quiescent and highly active states, with the sign of the difference flipping across the transition

(Figure 1Fi). We focus our analysis on the spontaneous-transition regime at  $I_{\text{ext}} = 0$  and  $\lambda = 2$  Hz, where both cycles show significantly higher AIS than the null. In this regime, we recover trends analogous to the VAR results, with feedback cycles showing higher AIS (Figure 1Fii) and  $\Phi$  (Figure 1Gii) than feedforward cycles. While feedforward cycles show higher transfer entropy than feedback cycles (Figure 1Gi).

These results indicate that feedforward and feedback cycles have distinct computational roles, with feedforward cycles primarily facilitating information transfer and feedback cycles supporting information integration. The fact that these differences are observed in both linear and non-linear models suggests that they may be fundamental properties of these cycle types, rather than artefacts of specific dynamics. However, the magnitude of these effects depends on the specific parameters of the network, such as size, density, and directionality, which we explore in the following sections.

### **Analytical results for the VAR(1) process on ER graph**

For the VAR(1) process, we can derive an analytical solution for the steady state-covariance matrix  $\Sigma$  of the dynamics of the cycle embedded within a sparse directed Erdős–Rényi graph in the mean field approximation (see the Network dynamics section in Methods for details). This allows us to compute the steady-state values of AIS, TE and  $\Phi$  for both feedforward and feedback cycles in the weak coupling regime ( $\alpha < \alpha_c \approx 0.22$ ). We find that the analytical results are in good agreement with the numerical simulations. Since the AIS can be written down as a function of the eigenvalues of the covariance matrix, we can identify the specific eigenmodes that enhance the AIS in feedback cycles.

$$AIS = \frac{1}{2} \sum_{i=1}^4 \ln \frac{\lambda_i(\Sigma)}{\sigma^2} \quad (1)$$

where  $\lambda_i(\Sigma)$  are the eigenvalues of the covariance matrix  $\Sigma$  and  $\sigma^2$  is the variance of the noise in the VAR(1) process. We find that AIS is enhanced in feedback cycles as the second largest eigenvalue of the covariance matrix becomes degenerate, while in feedforward cycles the eigenvalues are more spread out (Figure 2C). This essentially happens because the VAR process on a feedback cycle exhibits oscillatory dynamics and yields two complex conjugate eigenvalues with the same magnitude, which enhances the AIS. Furthermore, when comparing the node participation ratios in the eigenmodes of the covariance matrix, we find that the feedback cycle has all nodes participating in all the eigenmodes (NPR = 1), whereas the feedforward cycle has a more localised participation (NPR < 1) with the target node participating the most and the source node participating the least (Figure 2D). This spectral decomposition explains the increased AIS and  $\Phi$  in feedback cycles.

### **Strength of information processing depends on the cycle's positioning and the network's global structure**

The information-processing capacity of a cycle is determined not only by its own topology but also by the structural properties of the network in which it is embedded. To quantify these effects, we examine how active information storage (AIS) varies with network size ( $N$ ), density ( $p$ ), and relative global directionality ( $q$ ) in both VAR and LIF models. Across network sizes and densities, AIS decreases as  $N$  increases (Figure 3Dii, Fii), reflecting the fact that a cycle captures a larger fraction of the dynamics in smaller networks. AIS also decreases with increasing  $p$  (Figure 3Di, Fi), consistent with the cycle exerting a stronger influence on information flow in sparser networks.

Finally, to understand how the global directionality of the network affects the information processing capabilities of the cycles, we systematically vary the flow of edges on a random directed graph. The direction of each edge is assigned with probability  $q$  from the lower- to the higher-indexed node, such that  $q = 1$  yields a fully acyclic network and  $q = 0.5$  corresponds to no preferred direction. A feedforward cycle itself has a preferred flow direction, which implies that when embedded in a network with  $q \neq 1$ , it can be aligned with the network's preferred direction or reversed. The feedback cycle is not affected by this phenomenon due to its symmetry.

As  $q$  increases, AIS rises for both the reversed feedforward and feedback cycles, but remains comparatively unchanged for the aligned feedforward cycle (Figure 3Ei) in the VAR model. In the LIF model, however, the aligned feedforward cycle shows a significant decrease in AIS with increasing  $q$  (Figure 3Gi). Decomposing AIS shows that  $\Phi$  and TE increase with  $q$  for the reversed feedforward and feedback cycles (Figure 3E, G), whereas the aligned feedforward cycle exhibits negligible values of  $\Phi$  and a decline in TE in the VAR model (Figure 3Eii, Eiii). The decomposition for the LIF model shows significant deviations from the VAR model, especially in the aligned feedforward cycle. Here we observe a significant decline in  $\Phi$  with increasing  $q$  (Figure 3Gii), and an increase in TE with increasing  $q$  (Figure 3Giii). Furthermore, the TE of the aligned feedforward cycle is much higher than the feedback cycle, which is not observed in the VAR model. This deviation from the VAR model can be explained by the fact that the LIF model is a non-linear model with a discrete spiking dynamics, which can amplify the effects of changes in network structure on information transfer. These results indicate that a cycle's capacity to process information and act as a higher-order structure depends critically on its ability to disrupt and redirect the prevailing flow of information in the network.

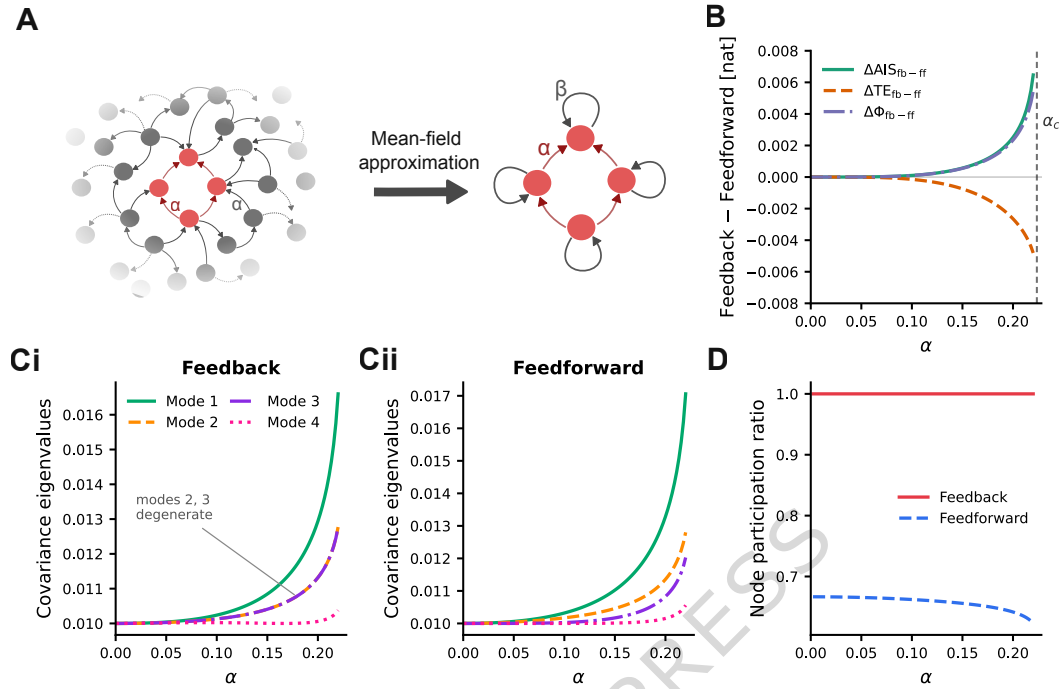


**Figure 2. Analytical results for the VAR(1) process on ER graph.** **A** Illustration of the mean field approximation, for a full derivation, see Supplementary Note 3. **B** Difference of information theoretical measures between feedback and feedforward cycles. Both AIS and  $\Phi$  are higher for feedback cycles, and the difference increases with  $\alpha$ . TE is higher for feedforward cycles, and the difference increases with  $\alpha$ . **C** Eigenvalues of the covariance matrix  $\Sigma$  as a function of  $\alpha$  for feedback (i) and feedforward (ii) cycles. The feedback cycle exhibits a pair of degenerate eigenvalues, which enhances AIS. **D** Node participation ratios (NPR) in the eigenmodes of the covariance matrix  $\Sigma$  for feedforward and feedback cycles. The feedback cycle has all nodes participating in all eigenmodes (NPR = 1), whereas the feedforward cycle has a more localised participation (NPR  $\approx 0.67$ ) which decreases with  $\alpha$ .

### Cycle-induced perturbations to trophic coherence shape information processing

To assess how embedded directed cycles disrupt the directional organisation of a network, we measure trophic coherence<sup>48</sup>. This measure is based on assigning each node a trophic level and counting the fraction of edges that violate the trophic ordering, defined as the global incoherence  $F_0$  (0 = fully coherent, 1 = maximally incoherent, see Methods).

For small networks, adding a single directed cycle markedly increases  $F_0$  (Figure 3A), but the effect diminishes as system size grows. To isolate cycle-specific effects independent of network size, we define a local incoherence  $F_C$  by counting only edges within the cycle that violate the trophic ordering. Here,  $F_C = 0$  means that the cycle is perfectly aligned with the network's trophic hierarchy, whereas  $F_C \approx 1$  means that, on average, each edge in the cycle violates the ordering once. Values  $F_C > 1$  indicate that multiple violations occur per edge when the cycle is considered within the network's trophic framework, corresponding to a strong local disruption of directional flow.

Across large networks,  $F_C$  is highest for the reversed feedforward cycle, followed by the feedback cycle, while the aligned feedforward cycle shows no deviation from the baseline (Figure 4Bi). Varying the global directionality  $q$  reveals that  $F_C$  increases linearly for the reversed feedforward cycle as it does for the feedback cycle, also with slope close to 0, and decreases linearly for the aligned feedforward cycle (Figure 4Bii).

Finally, relating  $F_C$  to the information-theoretic measures shows that AIS and  $\Phi$  both increase with  $F_C$  and reach their maximum at  $F_C \approx 1.0$ , whereas TE peaks at a slightly lower  $F_C$  (Figure 4Ci–Ciii) for the VAR model. Similar behaviour is observed for AIS and  $\Phi$  in the LIF model (Figure 4Di–Dii). However, TE in the LIF model peaks slightly below  $F_C \approx 1.0$  and then declines more sharply than in the VAR model (Figure 4Diii). This essentially reflects the reduced TE observed in feedback cycles in the LIF model compared to the VAR model, consistent with the results in Figure 3Giii.

This indicates that a cycle's capacity for information storage and integration is greatest when its edges collectively introduce, on average, one trophic violation—striking a balance between preserving and disrupting the global flow. In contrast, information transfer is optimal under a slightly more ordered configuration, short of being fully acyclic or tree-like. The non-monotonic trends further suggest that excessive disruption to the directional flow, approaching random connectivity, can be detrimental to

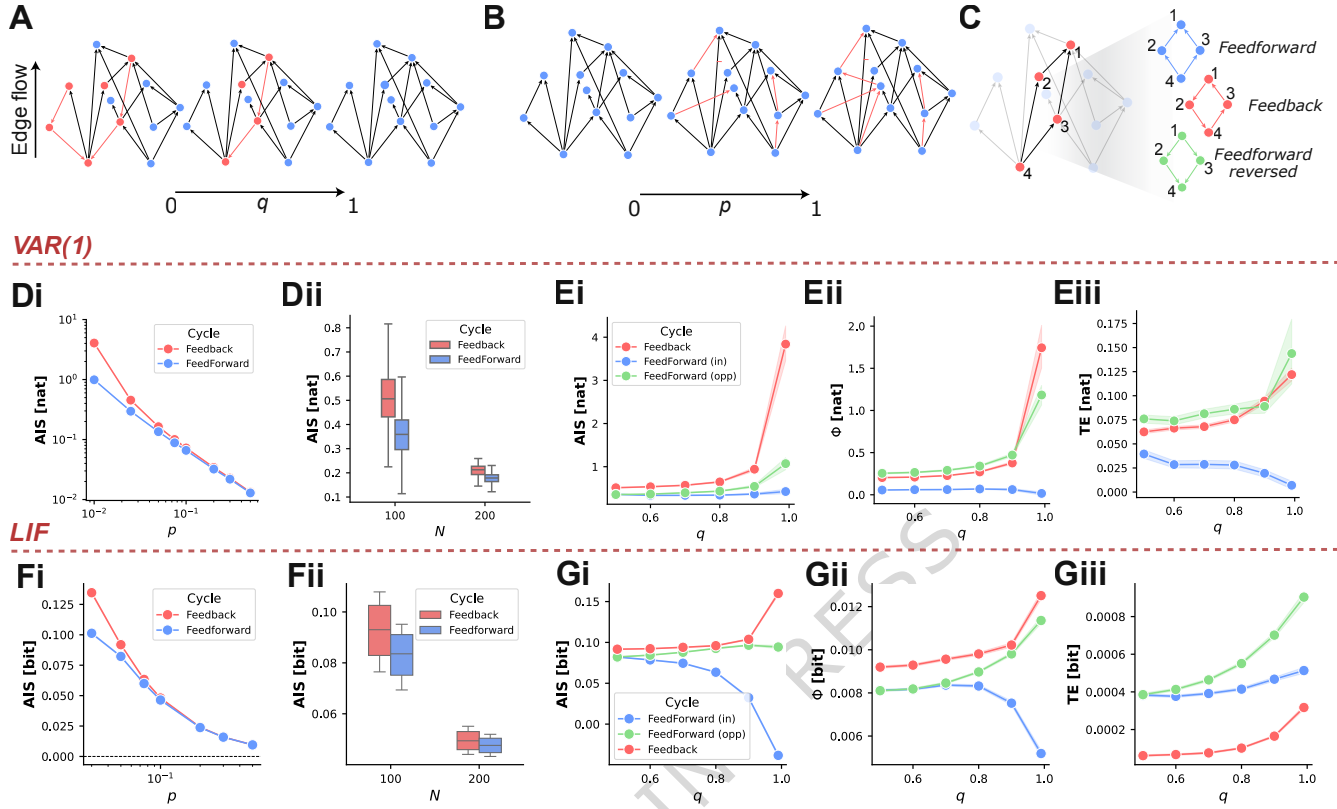


**Figure 3. Effect of network directionality on the information processing of embedded cycles.** We vary directed Erdős–Rényi graphs with a preferred directionality by systematically changing three parameters: network size ( $N$ ), global directionality ( $q$ ), and density ( $p$ ). **A** Probability  $q$  controls the alignment of edges with the global flow:  $q = 1$  corresponds to complete alignment, while  $q = 0.5$  indicates no preferred direction. **B** Increasing  $p$  adds more edges, producing denser directed networks. **C** The presence of a preferred direction in a network,  $q \neq 0.5$ , does not influence feedback cycles due to their symmetry, but feedback cycles can be embedded as aligned with the preferred direction as reversed with respect to the preferred direction. **D** AIS decreases with increasing  $p$  (**Di**, **Fi**) and  $N$  (**Dii**, **Fii**) in both VAR and LIF models. **E** AIS increases with  $q$  for feedback and reversed feedforward cycles, but remains unchanged for aligned feedforward cycles in the VAR model (**Ei**). Similar trends are observed in the LIF model, except the null corrected AIS decreases with  $q$  for the aligned feedforward cycle in the LIF model (**Gi**). Decomposing AIS into TE (**Eii**, **Gii**) and  $\Phi$  (**Eiii**, **Giii**) shows that both components increase with  $q$  for feedback and reversed feedforward cycles in both models. In contrast, the aligned feedforward cycle exhibits a slight decrease in  $\Phi$  and TE with increasing  $q$  in the VAR model, and a more pronounced decrease in  $\Phi$  in the LIF model. TE of aligned feedforward cycle increases with  $q$  and shows a much higher TE than the feedback cycle, which is not observed in the VAR model. Errorbands represent the 95% confidence interval of the mean across 1,000 independent network realisations for each cycle.

all forms of information processing.

### Local cycle structure modulates the diversity of global network dynamics and computation

Beyond their role in local information processing, directed cycles influence the repertoire of activity patterns a network can sustain. We quantify this global impact using *network harmonic decomposition*<sup>49</sup>, a Graph Signal Processing (GSP) approach that expresses the network's activity as a sum of spatial modes (Figure 5A,Bi). The spatial modes are ordered with increasing eigenvalues of the graph undirected Laplacian: lower-index modes capture broad, system-wide fluctuations, whereas higher-index modes reflect localised patterns restricted to smaller subsets of nodes<sup>50</sup>. The distribution of power across these modes provides a measure of the diversity of global activity patterns—higher diversity indicates that activity is distributed across many spatial scales rather than dominated by a few. Such diversity has been linked to functional richness in biological networks, including correlations with levels of consciousness<sup>51</sup> and the organisation of resting-state brain networks<sup>52</sup>, as well as to the characterisation of the phases of the XY model<sup>53</sup>. Further details of the network harmonic decomposition can be found

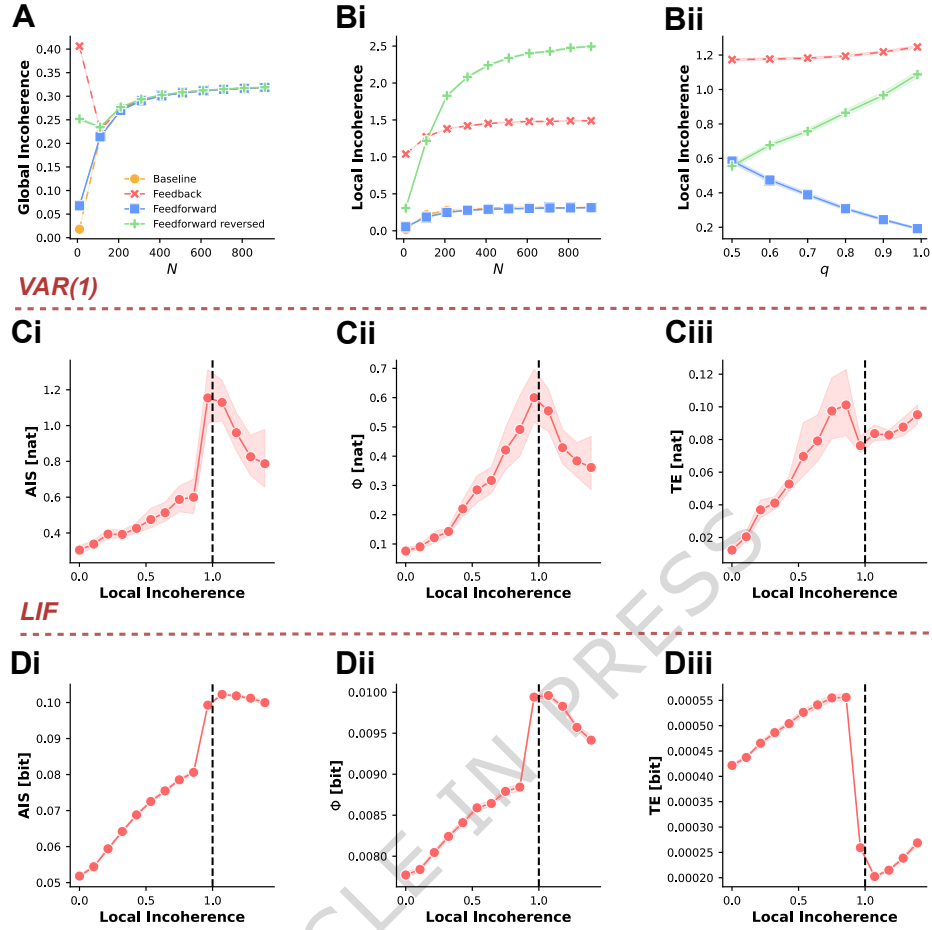


**Figure 4. Cycle-induced perturbations to trophic coherence and their impact on information processing.** **A** Global trophic incoherence ( $F_0$ ) for networks of varying size  $N$  with embedded feedback, feedforward, or reversed feedforward cycles, compared to baseline networks. Small networks show a marked increase in  $F_0$  when a cycle is added, but this effect diminishes with size. **Bi** Local incoherence ( $F_C$ ) for the same networks, computed using only the edges within the cycle. For large networks,  $F_C$  is highest for reversed feedforward cycles, followed by feedback cycles, with aligned feedforward cycles showing baseline-level  $F_C$ . **Bii** Dependence of  $F_C$  on global directionality  $q$ : reversed feedforward cycles increase linearly with  $q$ , feedback cycles increase sublinearly, and aligned feedforward cycles decrease linearly. **C** Relationship between  $F_C$  and **(Ci)** active information storage (AIS), **(Cii)** transfer entropy (TE), and **(Ciii)** integrated information ( $\Phi$ ) in a VAR(1) model. AIS and  $\Phi$  peak near  $F_C \approx 1.0$ , while TE peaks at slightly lower  $F_C$ . **D** The same behaviour is observed for AIS and  $\Phi$  in a LIF model (**Di-Dii**). However, similarly to the VAR model, TE (**Diii**) peaks slightly below  $F_C \approx 1.0$  but declines more sharply. Errorbands represent the 95% confidence interval of the mean across all independent network realisations across both cycles, binned by  $F_C$ .

in the Network Harmonic Decomposition section in Methods.

We assess how directed cycles modulate this diversity by embedding them into networks with systematically varied global directionality  $q$ —from fully aligned flows ( $q = 1$ ) to no preferred direction ( $q = 0.5$ )—and comparing two topologies: Erdős–Rényi graphs and multi-layer feedforward networks. In both cases, diversity increases as  $q$  decreases, reflecting the greater diversity of the signal with a larger prevalence of cycles opposing the dominant flow (Figures 5Ci,Ciii). We further confirm this trend by plotting diversity against the global trophic incoherence  $F_0$  of the network, which is modulated by the presence of directed cycles (Figure 5Cii). The effect is amplified in larger networks and particularly pronounced in multi-layer feedforward architectures, where reductions in directionality yield disproportionately large gains in diversity (Figure 5Cii).

Together, these results reveal that local disruptions to directional flow—such as feedback cycles—can substantially enrich the spectrum of spatial activity patterns available to a network, providing a mechanism to enhance its global dynamical repertoire.

Finally, to test if this increase in the diversity of global network dynamics conferred by the presence of directed cycles translates into computational advantages, we compare the performance of different neural network architectures on the MNIST

digit classification task. We initialised two sparse multi-layer feedforward networks with three layers of 100 nodes and low density of edges ( $p = 0.05$ ). One of the networks is fully feedforward ( $q = 1$ ), while the other has no preferred direction ( $q = 0.5$ ). We trained both networks using backpropagation for 30 epochs, while enforcing the same level of sparsity in the edges. We find that while both networks are able to learn the task (Fig. 5Ei), the feedforward network converges faster and has slightly higher accuracy (96% vs. 95% for feedback) on the test set (Figure 5Ei). However, to fully understand the distinct computational processes learned by the two trained networks, we deploy two robustness checks: (i) robustness to random bit flips in the input, and (ii) robustness to random node ablations. We find that the recurrent network is more robust to input noise, while the feedforward network is more robust to node ablations (Figure 5Eii). This suggests that the recurrent network is able to suppress the effect of noise by the enhanced integration of information across multiple time steps. On the other hand, the feedforward network is more robust to node ablations because of the enhanced transfer of information via multiple redundant feedforward paths. These results suggest that different directed cycles can confer distinct computational advantages to a network.

## Discussion

In this work, we investigate the role of directed cycles in local information processing, examine how the network structure influences their dynamical properties, and assess their impact on the global activity patterns of the network. We propose local trophic coherence first as a measure that explains the relationship between the relative directionality of a cycle and the network, and second as a parameter that determines the level of higher-order behaviour generated by feedforward cycles. These findings highlight the role of feedforward and feedback cycles as structures which impact both local and global network dynamics.

Although in this work, we focus on two types of directed cycles of length 4, other arrangements of directed edges can generate other cyclic motifs of any length<sup>26</sup> with distinct computational properties. The measures and tools presented in this paper can be extended to characterize the role of other higher-order structures in networks, as well as to infer local network structures from observed activities on nodes in real-world systems where the network structure is unobserved. Alternatively, in systems where the network structure is available, but the activity cannot be trivially measured, the measure of local trophic incoherence introduced in this paper can serve as a proxy for assessing the information processing capacity of cyclic motifs. Since this measure is computationally efficient to estimate, it provides a practical exploratory tool for detecting functionally relevant cycles in established networks, particularly in cases where calculating dynamical information measures is hindered by data sparsity or limited availability.

In order to assess the global impact of directed cycles, we adapted the framework of network harmonic decomposition, originally proposed for undirected networks. The framework, when applied to time series activity generated from directed graphs, is able to explain the spatial decomposition of the global activity; however, at the cost of directionality. While other frameworks have been developed<sup>54</sup>, general interpretability and applicability are still lacking. This allowed us to confirm that the differences in the information processing capacities of the two types of cycles is not just dynamical but also functional in the sense that they confer distinct computational advantages to a network. When trained as artificial neural networks on a digit classification task, the feedforward network converges faster and is more robust to node ablations, while the recurrent network is more robust to input noise. This suggests that the enhanced information transfer in feedforward cycles may facilitate faster learning and robustness to structural perturbations, while the enhanced information integration in feedback cycles may support robustness to input variability.

Our results hold for two very different categories of dynamics, pointing at a generalisability, at least qualitatively, of our findings. Thus, the distinct information-processing capacities of feedforward and feedback cycles, and their dependence on network directionality, provide a framework for reinterpreting the role of these motifs in biological systems. For example, in transcriptional networks, where directional biases often arise from environmental gradients or hierarchical regulatory cascades, the orientation and embedding of cycles may critically shape how information is transmitted or integrated across the system.

Our findings indicating the differences in behaviour of trained feedforward and recurrent networks are in line with the broader literature on the diverse set of partial recurrent neural networks<sup>55</sup>. The effect of the presence of cyclic motifs could, and our ability to predict it, allows theory crafting for the substrate of different types of systems. This could inform the design of a neuron or organoid-based ‘brain on a chip’ or reservoir computing architecture<sup>56</sup>. One could even imagine designing time-dependent networks where higher-order behaviour can be enhanced or reduced depending on the needs of the system, for example, in a synchronisation context or learning tasks.

In summary, the findings of the paper reveal that directed cycle motifs can be embedded in a network in a controlled way as computational primitives to generate higher-order behaviours in a network, which can significantly enhance the information processing capabilities of the network. This paper provides a foundation for future investigations into the role of directed cycles in complex networks and how they can be used to design more efficient and robust information processing systems.

Finally, our results confirm the clear distinction that needs to be made between higher-order structure and function<sup>19</sup>, i.e. that one doesn’t necessarily generate the other, but also highlights that simple alteration to the structure, in our case the relative

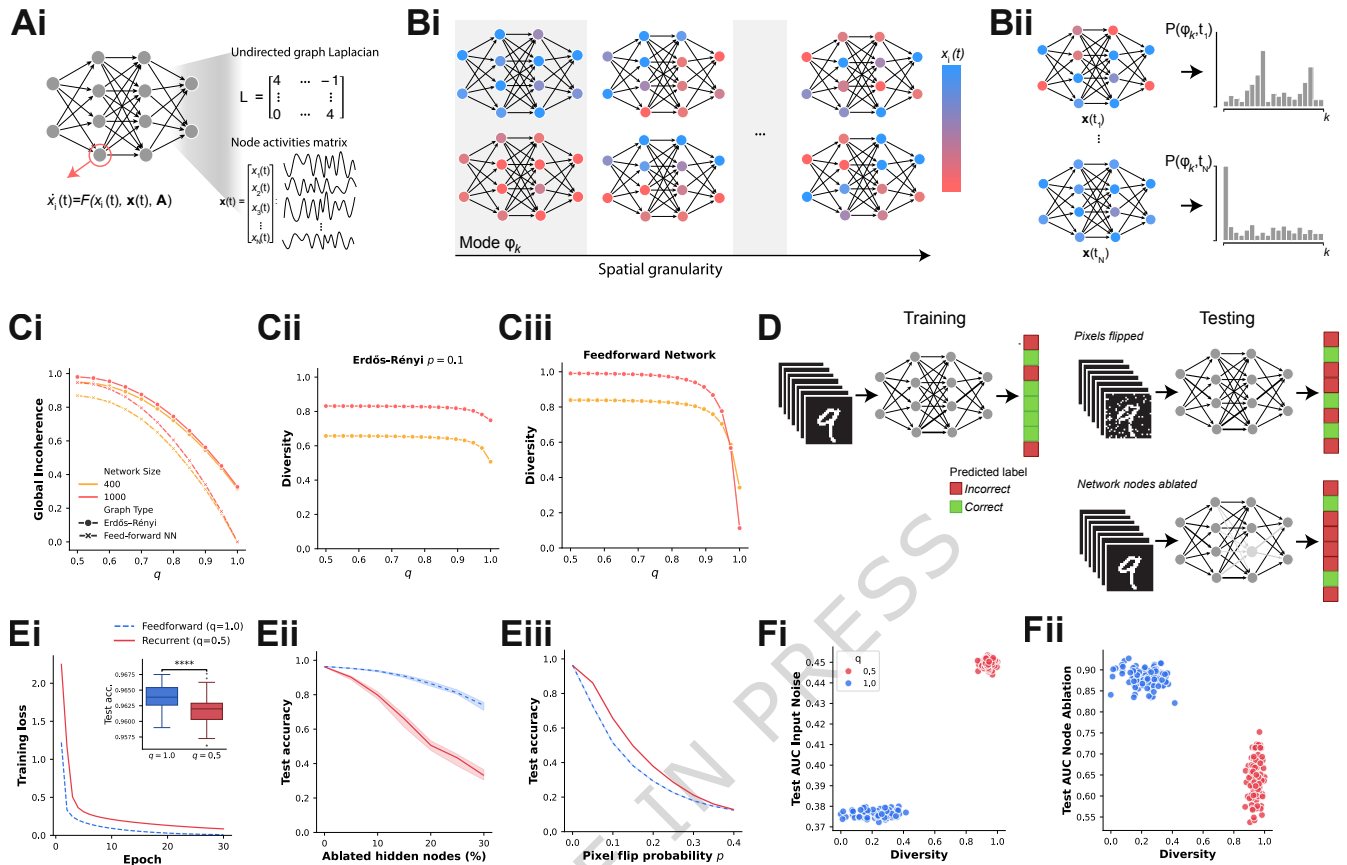


**Figure 5. Directed cycles enhance the diversity of global network dynamics.** (A) Network harmonic decomposition combines the eigenvectors of the symmetrised graph Laplacian with node activity time series from the VAR dynamics on a directed graph, yielding spatial modes of network activity. (Bi) Modes are ordered by eigenvalue  $k$ : low- $k$  modes correspond to global activity, while high- $k$  modes capture localized fluctuations. (Bii) The power spectrum across modes reveals how strongly each mode contributes to the network's dynamics at any time; the Shannon entropy of this distribution quantifies diversity. (Ci) In directed Erdős-Rényi graphs, diversity increases as the network becomes less directional ( $q \rightarrow 0.5$ ) and more feedback cycles appear; the effect is stronger in larger networks. (Cii) The diversity of global network dynamics increases with the global trophic incoherence  $F_0$  of the network, which is modulated by the presence of directed cycles. (Ciii) Multi-layer feedforward networks show a similar trend, with even greater gains in diversity as directionality decreases. (D) The setup for the MNIST digit classification task. We compare a feedforward network ( $q = 1$ ) and a recurrent network ( $q = 0.5$ ) with the same number of layers, nodes, and edges. Prediction accuracy is compared in two robustness tests: random pixel flip and node ablation. (E) Both networks achieve high accuracy ((Ei)), but the feedforward network converges faster and is more robust to node ablations (Eii), while the recurrent network is more robust to input noise (Eiii). (Fi) Test AUC under input noise scales with diversity — mixed networks ( $q = 0.5$ , red) cluster at high diversity and high noise robustness. (Fii) Test AUC under node ablation shows the inverse relationship — feedforward networks ( $q = 1$ , blue) cluster at low diversity but high ablation robustness. Errorbands represent the 95% confidence interval of the mean across 100 different trained networks for each architecture. Statistical comparisons were performed using a non-parametric permutation test with 10,000 permutations, and the significance level was set at  $p < 0.05$ .

directionality of a cycle, can provide explicit mechanisms to tune higher-order behaviour. While a better understanding of the interdependence between structure and function is necessary to further our understanding of higher-order behaviours in networked systems, we believe this work opens new avenues of research in that direction.

## Methods

### Simulation approach

Our analysis proceeds in four main stages: (i) generating directed network topologies with specified structural properties; (ii) embedding a tagged set of nodes into a directed cycle (feedforward, reversed feedforward, or feedback); (iii) simulating network activity using a first-order vector autoregressive (VAR(1)) process and the Leaky Integrate and Fire (LIF) model; and (iv) quantifying both local and global information-processing properties using information-theoretic and graph-spectral measures.

For each simulation, we generated a network with a prescribed size, density, and global directionality. A selected cycle was embedded by tagging four nodes and connecting them in a specific directed configuration. Network dynamics were simulated under the two processes from random initial conditions, producing ensembles of 10,000 trajectories per network. From these, we extracted the activity of the cycle nodes, along with the full network time series when required. The cycle trajectories were used to compute time-resolved measures of active information storage (AIS), transfer entropy (TE), and integrated information ( $\Phi$ ), and compared to a null model in which the cycle was replaced by a disconnected pair of edges. In parallel, we computed network-level metrics, including global and local trophic coherence, performed network harmonic decomposition and assessed their performance as Artificial Neural Networks (ANNs). Each of these steps is described in detail below, in the order required to reproduce the analysis.

### Network models

We considered two directed random network models, each parameterized by size ( $N$ ), density ( $p$ ), and global directionality ( $q$ ), along with a null model.

**Directed Erdős–Rényi model.** Starting from the classic Erdős–Rényi random graph  $G(N, p)$ , we assigned a unique integer index to each node and connected each unordered pair with probability  $p$ . For each existing edge  $(u, v)$ , direction was assigned from the lower to the higher index with probability  $q$ , and reversed otherwise. When  $q = 1$ , the graph is a directed acyclic graph with a perfectly ordered flow; when  $q = 0.5$ , edge directions are random with no preferred global orientation.

**Directed multi-layer feedforward model.** Nodes were arranged in  $L$  layers, with edges between successive layers assigned independently with probability  $p$ . Directionality was imposed by orienting each feedforward edge from the earlier to the later layer with probability  $q$ , and flipping its direction otherwise. Thus,  $q = 1$  yields a fully feedforward architecture, whereas  $q = 0.5$  produces a recurrent network with no preferred flow between layers.

**Null model.** For information-theoretic analyses, we constructed a null model in which the embedded cycle was replaced by a pair of disconnected edges with the same source and target. This preserves local degree and global topology, but removes the higher-order cyclic structure.

### Directionality of a graph

In a directed network, global directionality reflects the extent to which edges align with a consistent flow without forming feedback loops. A fully ordered directed acyclic graph (DAG) is maximally coherent, while a randomly oriented network is less coherent.

We quantify this property using the trophic coherence framework<sup>48</sup>. First, each node  $n$  is assigned a trophic level  $h_n$  by solving

$$\Lambda h = v, \quad (2)$$

where  $v_n = \sum_m A_{mn} - \sum_m A_{nm}$  and  $\Lambda = \text{diag}(u) - A - A^\top$  with  $u_n = \sum_m (A_{mm} + A_{nm})$ . The global trophic incoherence is

$$F_0 = \frac{\sum_{m,n} A_{mn} (h_n - h_m - 1)^2}{\sum_{m,n} A_{mn}}, \quad (3)$$

ranging from  $F_0 = 0$  (maximally coherent) to  $F_0 = 1$  (maximally incoherent).

**Local trophic coherence of a cycle** To quantify the extent to which an individual cycle disrupts global directionality, we define the local trophic incoherence

$$F_C = \frac{\sum_{m,n \in C} A_{mn} (h_n - h_m - 1)^2}{\sum_{m,n \in C} A_{mn}}, \quad (4)$$

where  $C$  is the set of edges in the cycle.  $F_C = 0$  indicates perfect alignment with the global trophic ordering; higher values correspond to greater local disruption. Unlike  $F_0$ ,  $F_C$  can exceed 1 because it is computed over a restricted subset of edges, allowing for stronger violations within the cycle than in the network as a whole.

## Network dynamics

We use two complementary models to simulate the dynamics of the network: a linear model based on a first-order vector autoregressive process (VAR(1)) and a non-linear model based on the Leaky Integrate and Fire (LIF) model. The VAR(1) model allows us to derive analytical results for the information-theoretic measures, while the LIF model provides a more biologically plausible setting to validate our findings.

**Vector autoregressive models (VAR)** have been extensively used in physical sciences and econometrics to quantify the interdependencies in a time-evolving multivariate system. Analytical traceability and efficient computational implementation make the VAR model a powerful tool for exploring multivariate linear dependencies in a complex system.

For our analysis, we consider a stationary first-order vector autoregressive (VAR (1)) process defined as

$$\mathbf{x}_{t+1} = \alpha \mathbf{x}_t \mathbf{A} + \tau \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_{t,i} \sim \mathcal{N}(0, 1) \quad (5)$$

with the stability condition,

$$\alpha = \frac{1}{\lambda_{\max} \mathbf{A}} \quad (6)$$

Here,  $\mathbf{A}$  represents the directed adjacency matrix of the network on which the dynamics are being simulated. For each simulation, we extract the trajectories of the four tagged nodes that form the feedback/feedforward cycle for the information-theoretic analysis.

In the mean-field limit, the activity on a four-node cycle embedded in a sparse directed Erdős–Rényi graph can be approximated by a VAR(1) process on the cycle with an effective self-coupling term ( $\beta$ ) that captures the influence of the rest of the network on the cycle. The full derivation of the mean-field approximation and the analytical solution for the covariance matrix of the cycle dynamics is provided in Supplementary Note 3.

We can then write the VAR(1) process on the cycle as

$$\mathbf{X}_{t+1} = P \mathbf{X}_t + \tau \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_{t,i} \sim \mathcal{N}(0, 1) \quad (7)$$

where  $P$  is the effective coupling matrix of the cycle, which depends on the original cycle adjacency  $W$  and the self-coupling term  $\beta$  as follows

$$P = \alpha W + \beta I_4 \quad (8)$$

Here,  $I_4$  is the identity matrix of size 4. Now, for this stable VAR(1) process, the covariance matrix  $\Sigma$  of the cycle dynamics is the positive-definite solution of the discrete Lyapunov equation,

$$\Sigma = P \Sigma P^\top + \tau^2 I_4 \quad (9)$$

This equation iteratively converges to a matrix power series solution, which can be expressed in closed form as

$$\Sigma = \tau^2 \sum_{k=0}^{\infty} (P^k)(P^k)^\top \quad (10)$$

This analytical solution allows us to derive closed-form expressions for the information-theoretic measures of interest. In particular, for a multivariate Gaussian system, the mutual information between two sets of variables can be computed directly from their covariance matrices.

**Leaky Integrate and Fire (LIF) model** is a widely used spiking neuron model that captures the essential features of neuronal dynamics, including the integration of inputs, generation of discrete spikes and the refractory period. The LIF model is defined by the following equations:

$$\tau_m \frac{dV_i}{dt} = -(V_i - V_{rest}) + R(I_{syn}(t) + I_i(t)), \quad (11)$$

where  $V_i$  is the membrane potential of neuron  $i$ ,  $\tau_m$  (set to 20 ms) is the membrane time constant,  $V_{rest}$  (set to -70 mV) is the resting potential,  $I_{syn}(t)$  is the synaptic input from other neurons, and  $I_i(t)$  is an external DC input.  $R$  is the membrane resistance, which we set to  $1M\Omega$ . When the membrane potential reaches a threshold  $V_{th}$  (set to -54 mV), the neuron emits a spike and the membrane potential is reset to a lower value  $V_{reset}$  (set to -80 mV). At every time step, the neuron can also fire spontaneously with a Poisson rate which is tuned close to the firing rate transition of the network. The synaptic input  $I_{syn}(t)$  is computed as a

weighted sum of the spikes from other neurons, where the weights are determined by the adjacency matrix of the network. The adjacency matrix is scaled by a factor of  $\alpha = 1/\lambda_{\max}$  to ensure that the network operates in a stable regime. In this study, we fix the time delay for all synapses to  $\tau_d = 1$  ms. The LIF model allows us to capture the non-linear dynamics of spiking neurons and validate our findings from the linear VAR model in a more biologically realistic setting.

For conducting the information-theoretic analysis, we extract the spike trains of the four tagged nodes that form the feedback/feedforward cycle. We then binarise these spike trains into discrete time bins to create binary time series, which are used to compute the information-theoretic measures using a discrete estimator of mutual information. The bin size is chosen by sweeping across a range of values and selecting the one that maximises the AIS of the cycle dynamics.

### Information-theoretic measures

Once we have the trajectories of tagged nodes from the cycles, we can analyze these to quantify various temporal information processing measures. For our analysis, we focus on three complementary metrics: *Information Storage*, *Integrated Information* and *Transfer Entropy*. These metrics are all based on calculating mutual information in multivariate systems. We briefly define these metrics below for a system  $X_t$ , where  $I(X;Y)$  refers to the mutual information between  $X$  and  $Y$ .

**Active Information Storage (AIS)**, quantifies the total amount of mutual information between the past and the future of the system. This information can further be decomposed into information integration, transfer and other information-specific metrics using the integrated information decomposition<sup>47</sup>. AIS can be simply written for a multivariate system  $X$  evolving in time  $t$  as,

$$AIS(X_t) = I(X_t; X_{t+1}) \quad (12)$$

For a multivariate Gaussian system with covariance matrix  $\Sigma$ , the AIS can be calculated in terms of the eigenvalues of the covariance matrix as,

$$\begin{aligned} I(X_t; X_{t+1}) &= \frac{1}{2} \ln \frac{|\Sigma(X_t)|^2}{|\Sigma(X_t, X_{t+1})|} \\ &= \frac{1}{2} \sum_{i=1}^4 \ln \frac{\lambda_i(\Sigma)}{\tau^2} \end{aligned} \quad (13)$$

where  $\lambda_i(\Sigma)$  are the eigenvalues of the covariance matrix  $\Sigma$  of the system  $X_t$ .

**Integrated information ( $\Phi$ )**, over the years, various measures of integrated information have been proposed to quantify the amount of excess information that a system possesses about its joint future activity as compared to its individual parts<sup>43,57</sup>. More recently, practical measures of integrated information have been proposed that can be computed in a computationally efficient manner on time-series data<sup>58</sup>. We use the integrated information measure proposed by Barrett and Seth<sup>58</sup>  $\Phi_E$  which measures the extra information that a system  $X$  possesses about its joint future activity, compared to the information that can be explained by the minimum information bipartition  $\mathcal{B} = M^1, M^2$  of the system. This measure can be written as,

$$\Phi_E(X_t, \mathcal{B}) = I(X_t; X_{t+1}) - \sum_{i=1}^2 I(M_t^i; M_{t+1}^i) \quad (14)$$

where  $M^i$  are the minimum information bipartitions of the system  $X$  at time  $t$ . The minimum information bipartition is selected by choosing the partition that minimises the normalised integrated information such that,

$$\mathcal{B} = \arg \min_{\mathcal{B}} \frac{\Phi_E(X_t, \mathcal{B})}{\min_i H(M_t^i)} \quad (15)$$

However, in systems with a lot of redundancy, there is a lot of shared information among the partitions. Thus, the  $\sum_i I(M_t^i; M_{t+1}^i)$  overcounts this redundant information. This overcounting often leads to a negative value of  $\Phi_E$  for systems with a lot of redundancy. Using integrated information decomposition<sup>47</sup>, a refined measure of integrated information is proposed where this redundancy is added back<sup>59</sup>, providing a better scaled measure of integrated information  $\Phi_R$ . Here,  $\Phi_R$  is defined as with respect to the redundancy function  $Red(X_t; X_{t+1})$ ,

$$\Phi_R = \Phi_E + (N - 1) * Red(X_t; X_{t+1}) \quad (16)$$

We use the minimum mutual information (MMI) redundancy function to estimate the redundant information among the parts. The MMI redundancy function has been shown to efficiently capture the redundancy in multivariate Gaussian systems<sup>60</sup>. Here, the MMI redundancy function is defined as,

$$Red(X_t; X_{t+1}) = \min_i I(M_t^i; X_{t+1}) \quad (17)$$

**Transfer Entropy (TE).** A conditional mutual information measure quantifying directional influence from a source to a target<sup>61,62</sup>. For a VAR(1) process, the past is the current timestep, and the delay  $\delta$  is chosen to maximise  $TE$ :

$$TE(X^i \rightarrow X^j) = I(X_t^i; X_{t+\delta}^j | X_t^j). \quad (18)$$

In our four-node cycles, source–target pairs are two edges apart, so information propagates in two timesteps; accordingly, we set  $\delta = 2$ , which matches the known peak of  $TE$  at the true coupling delay<sup>63</sup>.

### Network Harmonic Decomposition

Fourier transform has enabled the decomposition of a time series in various frequency modes to characterise the activity at different timescales. Similarly, Graph Signal Processing (GSP) has enabled Network Harmonic decomposition, which enables the characterisation of dynamics at different scales of the network, from local (node level) to global (whole network) and all levels of organisation in between. For full reviews and introductions to the field, see<sup>50,64</sup>. This Network Harmonic Decomposition has enabled the understanding of structure-function correspondence in network dynamics, especially in neuroscience, where it has uncovered scale-specific effects in relation to different levels of consciousness<sup>51,65</sup>.

Here, we use the network harmonic decomposition to decompose the power of the signal into different harmonic modes on the network. Then we use Shannon entropy to calculate the diversity of different harmonic modes active on the network. This diversity metric provides a measure of how many different scales of network organisation are implicated in the overall dynamics. The diversity is high when the power is distributed across a lot of different scales of organisation and low when only a few modes are active. The power of each network harmonic mode is calculated as described in the previous works<sup>51,65</sup>, using the eigendecomposition.

First, the normalised Laplacian matrix ( $\tilde{L}$ ) and its eigenvectors are calculated for the network with adjacency matrix  $\mathbf{A}$ . The out-degree diagonal matrix of  $\mathbf{A}$  is defined as  $\mathbf{D} = \text{Diag}(\sum_j A_{i,j})$ , where  $\mathbf{D}$  is a matrix ( $N \times N$ ), with diagonal values as out-degrees of the nodes and the off-diagonal elements are zero.

$$\begin{aligned} \tilde{L} &= D^{-1/2} L D^{1/2} \text{ where } L = D - A \\ \tilde{L} \varphi_k &= \lambda_k \varphi_k \end{aligned} \quad (19)$$

After solving the eigenvalue problem to extract the eigenvector matrix  $\varphi_k$ , it is convolved with the whole network activity  $Y_t$  to estimate the contributions of each independent eigenmode ( $\omega_k$ ) using a dot product.

$$\begin{aligned} Y_t &= \omega_1(t) \varphi_1 + \omega_2(t) \varphi_2 + \dots + \omega_N(t) \varphi_N = \sum_{k=1}^N \omega_k(t) \varphi_k \\ \omega_k(t) &= \langle Y_t, \varphi_k \rangle \end{aligned} \quad (20)$$

The magnitude or power of the contribution of each harmonic mode  $k$  can be calculated from  $\omega_k(t)$  at each timestep as,

$$P_k(t) = |\omega_k(t)| \quad (21)$$

At each timestep, the distribution of power over the different harmonic modes is normalised,  $\tilde{P}_k(t) = \frac{P_k(t)}{\sum_{k=1}^N P_k(t)}$ , and the Shannon entropy of the normalised distribution is calculated as a measure of the harmonic diversity of the dynamics. The measure of diversity  $H$  used here is written as

$$H(t) = - \sum_{k=1}^N \tilde{P}_k(t) \log \tilde{P}_k(t) \quad (22)$$

The measure of diversity is calculated at each timestep, and the results present an average value during the steady state. Since the harmonic decomposition requires the estimation of real eigenvalues, we use the normalised Laplacian matrix  $\tilde{L}$  of the symmetrised adjacency matrix of the directed network. However, the activity of the nodes is generated using the directed adjacency matrix  $\mathbf{A}$ , which is used to simulate the VAR dynamics.

### Task performance of feedforward and recurrent networks

To compare the functional advantages the two different kinds of cycles can confer to a network, we test how a completely feedforward multi-layer network ( $q = 1$ ) and a completely recurrent multi-layer network ( $q = 0.5$ ) perform on the MNIST digit classification task. Both networks have three layers of 100 nodes and a low density of edges ( $p = 0.05$ ). We train the weights of both networks using backpropagation for 30 epochs, while using a sparsity mask to enforce the same level of density of edges in the networks.

The MNIST dataset consists of 70,000 images of handwritten digits (0-9), with 60,000 used for training and 10,000 for testing. Each image is a 28x28 pixel grayscale image, which we flatten into a 784-dimensional input vector. The output layer consists of 10 nodes, each representing one of the digit classes. We use a softmax activation function in the output layer to produce class probabilities and optimise the cross-entropy loss function during training. After training, we evaluate the accuracy of both networks on the test set and perform robustness checks by introducing random bit flips in the input and random node ablations to assess their resilience to perturbations.

### References

1. Vasilaiuskaite, V. & Rosas, F. E. Understanding complexity via network theory: a gentle introduction. *arXiv preprint arXiv:2004.14845* (2020).
2. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *science* **342**, 1337–1342 (2013).
3. Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y. & Zhou, C. Synchronization in complex networks. *Phys. reports* **469**, 93–153 (2008).
4. Gómez-Gardenes, J., Moreno, Y. & Arenas, A. Paths to synchronization on complex networks. *Phys. review letters* **98**, 034101 (2007).
5. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *nature* **406**, 378–382 (2000).
6. Petri, G. *et al.* Homological scaffolds of brain functional networks. *J. The Royal Soc. Interface* **11**, 20140873 – 20140873, DOI: [10.1098/rsif.2014.0873](https://doi.org/10.1098/rsif.2014.0873) (2014).
7. Bianconi, G. & Dorogovtsev, S. N. Theory of percolation on hypergraphs. *Phys. Rev. E* **109**, 014306 (2024).
8. Lambiotte, R., Rosvall, M. & Scholtes, I. From networks to optimal higher-order models of complex systems. *Nat. physics* **15**, 313–320 (2019).
9. Battiston, F. *et al.* Networks beyond pairwise interactions: Structure and dynamics. *Phys. Reports* **874**, 1–92, DOI: <https://doi.org/10.1016/j.physrep.2020.05.004> (2020). Networks beyond pairwise interactions: Structure and dynamics.
10. Bick, C., Gross, E., Harrington, H. A. & Schaub, M. T. What are higher-order networks? *SIAM Rev.* **65**, 686–731, DOI: [10.1137/21M1414024](https://doi.org/10.1137/21M1414024) (2023). <https://doi.org/10.1137/21M1414024>.
11. Majhi, S., Perc, M. & Ghosh, D. Dynamics on higher-order networks: A review. *J. Royal Soc. Interface* **19**, 20220043 (2022).
12. Millán, A. P., Torres, J. J. & Bianconi, G. Explosive higher-order kuramoto dynamics on simplicial complexes. *Phys. Rev. Lett.* **124**, 218301 (2020).
13. Arnaudon, A., Peach, R. L., Petri, G. & Expert, P. Connecting Hodge and Sakaguchi-Kuramoto through a mathematical framework for coupled oscillators on simplicial complexes. *Commun. Phys.* **5**, 211, DOI: [10.1038/s42005-022-00963-7](https://doi.org/10.1038/s42005-022-00963-7) (2022).
14. Nurisso, M. *et al.* A unified framework for simplicial Kuramoto models. *Chaos: An Interdiscip. J. Nonlinear Sci.* **34**, 053118, DOI: [10.1063/5.0169388](https://doi.org/10.1063/5.0169388) (2024).
15. Carletti, T., Giambagli, L. & Bianconi, G. Global Topological Synchronization on Simplicial and Cell Complexes. *Phys. Rev. Lett.* **130**, 187401, DOI: [10.1103/physrevlett.130.187401](https://doi.org/10.1103/physrevlett.130.187401) (2023). [2208.14783](https://doi.org/10.1103/physrevlett.130.187401).
16. Burgio, G., Matamalas, J. T., Gómez, S. & Arenas, A. Evolution of cooperation in the presence of higher-order interactions: From networks to hypergraphs. *Entropy* **22**, 744 (2020).
17. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 2485, DOI: [10.1038/s41467-019-10431-6](https://doi.org/10.1038/s41467-019-10431-6) (2019). [1810.07031](https://doi.org/10.1038/s41467-019-10431-6).
18. Ferraz de Arruda, G., Aleta, A. & Moreno, Y. Contagion dynamics on higher-order networks. *Nat. Rev. Phys.* **6**, 468–482 (2024).

19. Rosas, F. E. *et al.* Disentangling high-order mechanisms and high-order behaviours in complex systems. *Nat. Phys.* **18**, 476–477 (2022).
20. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
21. Bassett, D. S. & Sporns, O. Network neuroscience. *Nat. neuroscience* **20**, 353–364 (2017).
22. Borgatti, S. P., Mehra, A., Brass, D. J. & Labianca, G. Network analysis in the social sciences. *science* **323**, 892–895 (2009).
23. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci.* **100**, 11980–11985 (2003).
24. Alon, U. *Systems Medicine: Physiological Circuits and the Dynamics of Disease* (Chapman and Hall/CRC, 2023).
25. Mocanu, D. C. *et al.* Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. *Nat. communications* **9**, 2383 (2018).
26. Vasiliauskaite, V., Evans, T. S. & Expert, P. Cycle analysis of directed acyclic graphs. *Phys. A: Stat. Mech. its Appl.* **596**, 127097 (2022).
27. Badia-i Mompel, P. *et al.* Gene regulatory network inference in the era of single-cell multi-omics. *Nat. Rev. Genet.* **24**, 739–754 (2023).
28. Prill, R. J., Iglesias, P. A. & Levchenko, A. Dynamic properties of network motifs contribute to biological network organization. *PLoS biology* **3**, e343 (2005).
29. Oosawa, C., Takemoto, K. & Savageau, M. A. Effects of feedback and feedforward loops on dynamics of transcriptional regulatory model networks. *arXiv preprint arXiv:0711.2730* (2007).
30. Le, D.-H. & Kwon, Y.-K. A coherent feedforward loop design principle to sustain robustness of biological networks. *Bioinformatics* **29**, 630–637 (2013).
31. Kwon, Y.-K. & Cho, K.-H. Coherent coupling of feedback loops: a design principle of cell signaling networks. *Bioinformatics* **24**, 1926–1932 (2008).
32. Kwon, Y.-K. & Cho, K.-H. Quantitative analysis of robustness and fragility in biological networks based on feedback dynamics. *Bioinformatics* **24**, 987–994 (2008).
33. Markov, N. T. *et al.* Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J. comparative neurology* **522**, 225–259 (2014).
34. Luppi, A. I. *et al.* A synergistic core for human brain evolution and cognition. *Nat. neuroscience* **25**, 771–782 (2022).
35. van Bergen, R. S. & Kriegeskorte, N. Going in circles is the way forward: the role of recurrence in visual inference. *Curr. Opin. Neurobiol.* **65**, 176–193 (2020).
36. Bingjie Hao, I. A. K., Elijah Platnick. Social balance in directed networks. *Commun. Phys.* **8** (2025).
37. Rajpal, H., Rosas, F. E. & Jensen, H. J. Tangled worldview model of opinion dynamics. *Front. Phys.* **7**, 163 (2019).
38. Lizier, J. T., Bauer, F., Atay, F. M. & Jost, J. Analytic relationship of relative synchronizability to network structure and motifs. *Proc. Natl. Acad. Sci.* **120**, e2303332120 (2023).
39. Brede, M. Synchronization on directed small worlds: Feed forward loops and cycles. *Europhys. Lett.* **84**, 40004 (2008).
40. Novelli, L., Atay, F. M., Jost, J. & Lizier, J. T. Deriving pairwise transfer entropy from network structure and motifs. *Proc. Royal Soc. A* **476**, 20190779 (2020).
41. Zhang, X.-J., Moore, J. M., Yan, G. & Li, X. Universal structural patterns in sparse recurrent neural networks. *Commun. Phys.* **6**, 243 (2023).
42. Cramer, B. *et al.* Control of criticality and computation in spiking neuromorphic networks with plasticity. *Nat. communications* **11**, 2853 (2020).
43. Mediano, P. A., Seth, A. K. & Barrett, A. B. Measuring integrated information: Comparison of candidate measures in theory and simulation. *Entropy* **21**, 17 (2018).
44. Lizier, J. T., Atay, F. M. & Jost, J. Information storage, loop motifs, and clustered structure in complex networks. *Phys. Rev. E—Statistical, Nonlinear, Soft Matter Phys.* **86**, 026110 (2012).
45. Lizier, J. T. *The local information dynamics of distributed computation in complex systems* (Springer Science & Business Media, 2012).

46. Gómez-Herrero, G. *et al.* Assessing coupling dynamics from an ensemble of time series. *Entropy* **17**, 1958–1970 (2015).
47. Mediano, P. A. *et al.* Toward a unified taxonomy of information dynamics via integrated information decomposition. *Proc. Natl. Acad. Sci.* **122**, e2423297122 (2025).
48. MacKay, R. S., Johnson, S. & Sansom, B. How directed is a directed network? *Royal Soc. open science* **7**, 201138 (2020).
49. Atasoy, S., Deco, G., Kringelbach, M. L. & Pearson, J. Harmonic brain modes: a unifying framework for linking space and time in brain dynamics. *The Neurosci.* **24**, 277–293 (2018).
50. Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A. & Vandergheynst, P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **30**, 83 – 98, DOI: [10.1109/msp.2012.2235192](https://doi.org/10.1109/msp.2012.2235192) (2013).
51. Luppi, A. I. *et al.* Distributed harmonic patterns of structure-function dependence orchestrate human consciousness. *Commun. biology* **6**, 117 (2023).
52. Atasoy, S., Donnelly, I. & Pearson, J. Human brain networks function in connectome-specific harmonic waves. *Nat. communications* **7**, 10340 (2016).
53. Expert, P., Nigris, S. D., Takaguchi, T. & Lambiotte, R. Graph spectral characterization of the XY model on complex networks. *Phys. Rev. E* **96**, 012312 – 8, DOI: [10.1103/physreve.96.012312](https://doi.org/10.1103/physreve.96.012312) (2017).
54. Marques, A. G., Segarra, S. & Mateos, G. Signal processing on directed graphs: The role of edge directionality when processing and learning from network data. *IEEE Signal Process. Mag.* **37**, 99–116, DOI: [10.1109/MSP.2020.3014597](https://doi.org/10.1109/MSP.2020.3014597) (2020).
55. Ghosh, M. & Goodman, D. F. Partial recurrence enables robust and efficient computation. *bioRxiv* 2025–07 (2025).
56. Cai, H. *et al.* Brain organoid reservoir computing for artificial intelligence. *Nat. Electron.* **6**, 1032–1039, DOI: [10.1038/s41928-023-01069-w](https://doi.org/10.1038/s41928-023-01069-w) (2023).
57. Balduzzi, D. & Tononi, G. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS computational biology* **4**, e1000091 (2008).
58. Barrett, A. B. & Seth, A. K. Practical measures of integrated information for time-series data. *PLoS computational biology* **7**, e1001052 (2011).
59. Luppi, A. I. *et al.* What it is like to be a bit: an integrated information decomposition account of emergent mental phenomena. *Neurosci. consciousness* **2021**, niab027 (2021).
60. Barrett, A. B. Exploration of synergistic and redundant information sharing in static and dynamical gaussian systems. *Phys. Rev. E* **91**, 052802 (2015).
61. Schreiber, T. Measuring information transfer. *Phys. review letters* **85**, 461 (2000).
62. Bossomaier, T. *et al.* *Transfer entropy* (Springer, 2016).
63. Wibral, M. *et al.* Measuring information-transfer delays. *PloS one* **8**, e55809 (2013).
64. Stanković, L. *et al.* Vertex-frequency graph signal processing: A comprehensive review. *Digit. Signal Process.* **107**, 102802, DOI: [10.1016/j.dsp.2020.102802](https://doi.org/10.1016/j.dsp.2020.102802) (2020).
65. Atasoy, S. *et al.* Connectome-harmonic decomposition of human brain activity reveals dynamical repertoire re-organization under lsd. *Sci. reports* **7**, 17661 (2017).

## Acknowledgements

We acknowledge computational resources and support provided by the Imperial College Research Computing Service (<http://doi.org/10.14469/hpc/2232>).

## Funding statement

HR is supported by the Statistical Physics of Cognition project funded by the EPSRC (Grant No. EP/W024020/1) and the Schmidt Sciences LLC., through the AI in Science Fellowship. VV acknowledges financial support from the Swiss National Science Foundation and from the European Union Horizon 2020 program (“INFRAIA-01-2018-2019 – Integrating Activities for Advanced Communities,” Grant Agreement No. 871042, SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics, <http://www.sobigdata.eu>).

### Author contributions statement

All authors conceived and designed the study. HR and VV developed and implemented the network models. HR performed simulations and analysed data. PE provided code for graph signal processing. HR and VV generated figures. All authors analysed and interpreted the results, wrote and reviewed the manuscript.

### Code Availability Statement

The python scripts for simulations, information-theoretic calculators and graph measures are available on [GitHub](#) and archived on Zenodo. Rajpal, H., Expert, P. & Vasiliauskaitė, V. Directed Cycles as Higher-Order Units of Information Processing in Complex Networks. Zenodo <https://doi.org/10.5281/zenodo.21468752> (2026).

### Data Availability Statement

No empirical datasets were generated or analyzed during the current study. All simulated data supporting the findings of this manuscript were generated using the simulation scripts, which are available in the GitHub repository linked above.

### Competing interests

Authors declare no competing interests.

### Generative AI statement

Generative AI tools were used solely for language editing and improving clarity of expression. All scientific content, data analysis, figures, and interpretations were conceived, developed, and validated entirely by the authors.