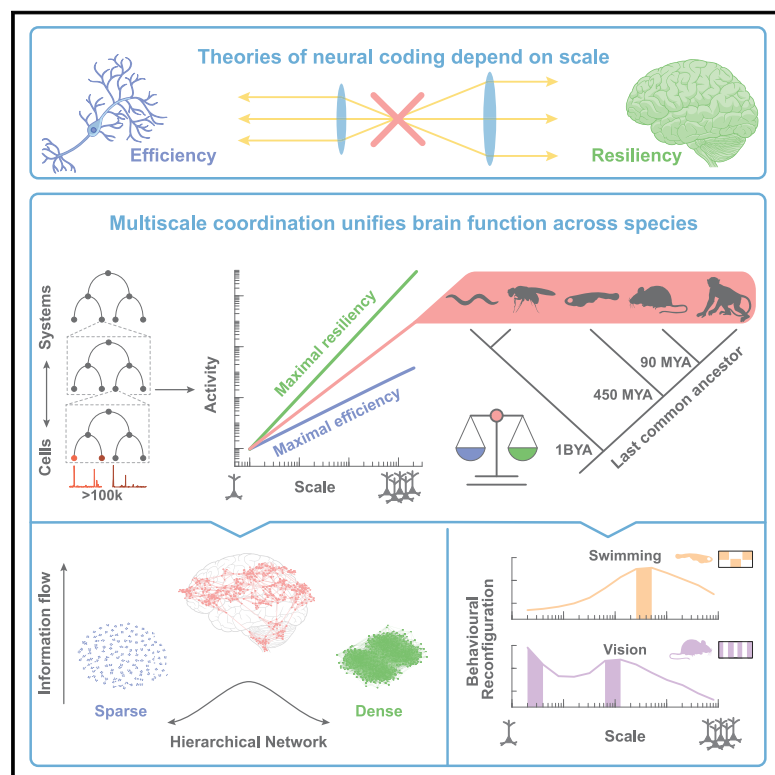


Multiscale organization of neuronal activity unifies scale-dependent theories of brain function

Graphical abstract



Authors

Brandon R. Munn, Eli J. Müller, Itia Favre-Bulle, Ethan Scott, Joseph T. Lizier, Michael Breakspear, James M. Shine

Correspondence

brandon.munn@sydney.edu.au

In brief

Analysis of cellular recordings from five phylogenetically diverse species uncovers a conserved, multiscale organization of neuronal activity that resolves disparate theories of brain function. This hierarchical structure allows the brain to operate across multiple timescales, enhances information flow, flexibly reconfigures activity during behavior, and balances functional resiliency with efficiency across scales.

Highlights

- Conserved multiscale brain organization across five phylogenetically diverse species
- Unifies scale-dependent theories of brain function, balancing efficiency and resiliency
- Hierarchical structure permits multiple timescales and enhances information flow
- Enables flexible cross-scale reconfigurations of neural activity during behavior

Article

Multiscale organization of neuronal activity unifies scale-dependent theories of brain function

Brandon R. Munn,^{1,2,8,*} Eli J. Müller,^{1,2} Itia Favre-Bulle,^{3,4} Ethan Scott,⁵ Joseph T. Lizier,^{2,6} Michael Breakspear,⁷ and James M. Shine¹

¹Brain and Mind Centre, School of Medical Sciences, The University of Sydney, Sydney, NSW, Australia

²Centre for Complex Systems, The University of Sydney, Sydney, NSW, Australia

³Queensland Brain Institute, The University of Queensland, St Lucia, QLD, Australia

⁴School of Mathematics and Physics, The University of Queensland, St Lucia, QLD, Australia

⁵Department of Anatomy and Physiology, The University of Melbourne, Parkville, VIC, Australia

⁶School of Computer Science, The University of Sydney, Sydney, NSW, Australia

⁷School of Psychology, College of Engineering, Science and the Environment, School of Medicine and Public Health, College of Health and Medicine, University of Newcastle, Callaghan, NSW, Australia

⁸Lead contact

*Correspondence: brandon.munn@sydney.edu.au

<https://doi.org/10.1016/j.cell.2024.10.004>

SUMMARY

Brain recordings collected at different resolutions support distinct signatures of neural coding, leading to scale-dependent theories of brain function. Here, we show that these disparate signatures emerge from a heavy-tailed, multiscale functional organization of neuronal activity observed across calcium-imaging recordings collected from the whole brains of zebrafish and *C. elegans* as well as from sensory regions in *Drosophila*, mice, and macaques. Network simulations demonstrate that this conserved hierarchical structure enhances information processing. Finally, we find that this organization is maintained despite significant cross-scale reconfiguration of cellular coordination during behavior. Our findings suggest that this nonlinear organization of neuronal activity is a universal principle conserved for its ability to adaptively link behavior to neural dynamics across multiple spatiotemporal scales while balancing functional resiliency and information processing efficiency.

INTRODUCTION

Brain function emerges from the functional activity of neurons, which is coordinated across multiple spatial and temporal scales (e.g., 1/f spectra).¹ However, the brain is typically studied through localized apertures—e.g., cellular scale in electrophysiology vs. systems scale in neuroimaging—owing to a lack of large-scale neuronal recordings² and multiscale analytical techniques capable of characterizing how neuronal coordination changes across spatiotemporal resolutions.^{3,4} Understanding the cross-scale principles of neuronal activity can resolve the inconsistencies inherent within theories of brain function derived from scale-dependent analyses.

One such tension involves opposing theories of neural coding that argue for either efficiency or resiliency through inversely modifying redundant (shared) information between cells.⁵ A sparse and efficient (minimally redundant) “neuronal” code can represent many distinct signals,^{6,7} whereas a resilient (largely redundant) “population” code can readily generalize and robustly maintain function despite significant neuronal loss.^{8–10} While the brain must balance these trade-offs, the extent to which functional neuronal activity favors either regime remains unclear.

Curiously, these opposing theories of brain function both have substantial empirical support, albeit from recordings of the brain

oriented at contrasting spatiotemporal scales.^{11,12} At the micro-scale, neuronal activity is uncorrelated, suggesting coding principles of efficiency.¹³ Conversely, in macroscopic recordings, coarse-grained population activity—pooled across thousands to billions of neurons^{14,15}—appears strongly spatiotemporally correlated,^{14,15} which advocates for resiliency. How can the brain be organized to reconcile these perspectives?

Recent advances in calcium imaging have made it possible to simultaneously interrogate the brain at the microscale (neurons) and macroscale (whole-brain systems). Using iterative coarse-graining (ICG)¹⁶—a computational method inspired by statistical physics¹⁷ for analyzing activity across increasing spatiotemporal scales—we resolved the theoretical schism between efficient and resilient dynamics. We applied ICG to whole-brain recordings in zebrafish (*Danio rerio*)¹⁸ and nematode (*Caenorhabditis elegans*)¹⁹, as well as to the sensory cortex in fly (*Drosophila melanogaster*),²⁰ mouse (*Mus musculus*),²¹ and monkey (*Macaca mulatta*),^{22,23} across multiple spatial and temporal scales. We first demonstrate that efficient and resilient activity emerges from a phylogenetically consistent, multiscale organization principle. Next, we simulate neuronal dynamics atop a hierarchical-modular structural network that recapitulates the observed self-similar scaling. Finally, we show that the multiscale coordination is conserved across behaviors—despite significant

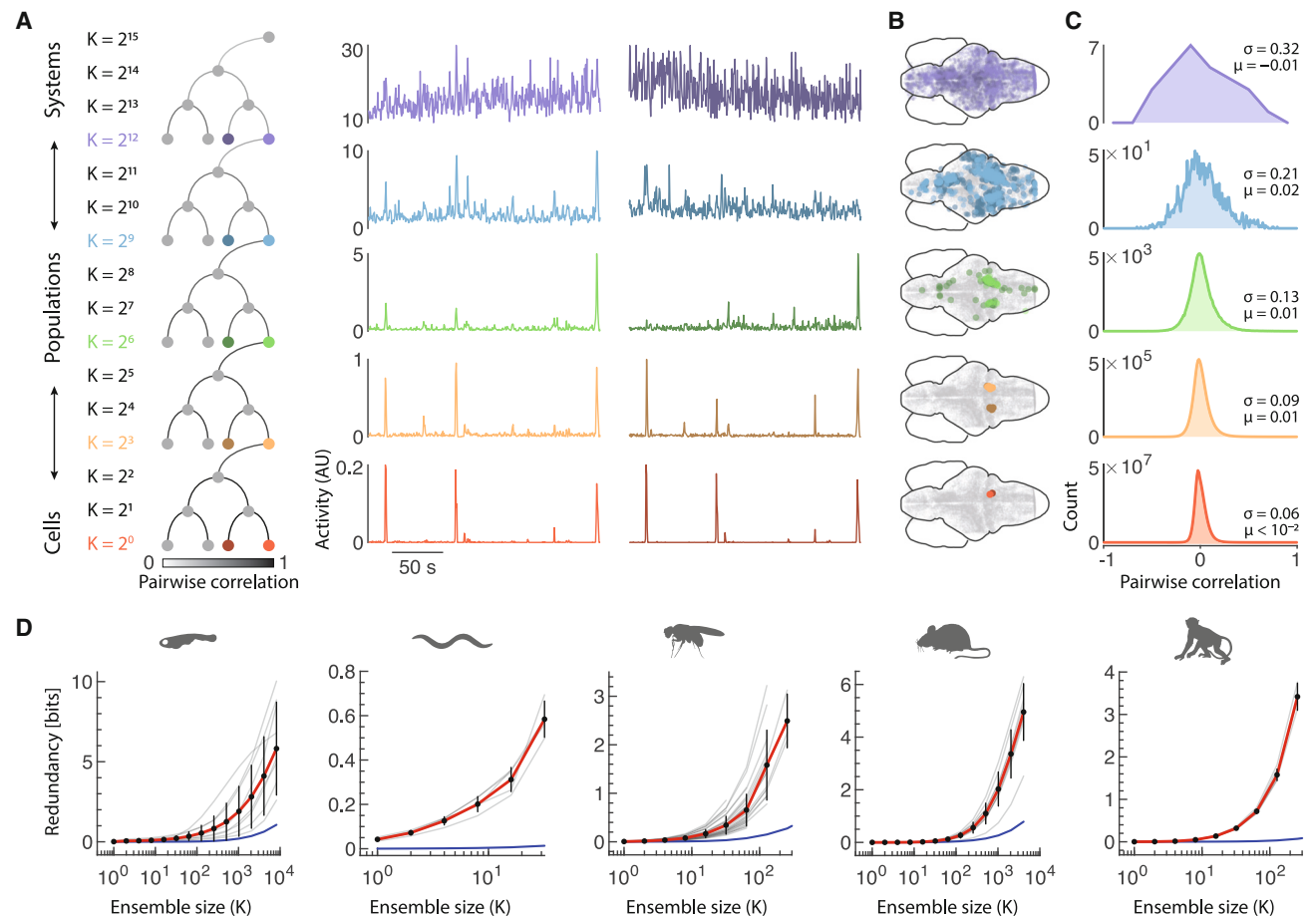


Figure 1. Signatures of neural coding change across scales

(A) Iterative coarse-graining (ICG) of neuronal activity is achieved by greedily pairing neurons by their pairwise correlation and summing their activity. This process is repeated on the coarse-grained variables, iterating across larger ensembles of coordinated neurons (K ensemble size), and increasing the spatial scale (l). This approach links individual neurons to systems-level whole-brain neuronal populations across dyadically nested scales. The coarse-grained neuronal activity changes continuously from small and fast fluctuations at the neuronal scale (red) to large and slow fluctuations (purple). Pairs of combined neuronal ensemble activity at levels $l = 1, 4, 7, 10$, and 13 across an example branch of a zebrafish ICG hierarchy.

(B) Examples of spatially distributed, dyadically increasing neuronal ensembles following ICG in the same zebrafish.

(C) At the neuronal scale (bottom red, $l = 1$), the pairwise correlation distribution is tightly distributed around zero with few strongly correlated cells ($\mu < 10^{-2}$; $\sigma = 0.06$), whereas at the macroscale (top purple, $l = 13$), the coarse-grained neuronal ensemble correlations are broadly distributed with positive and negative correlations ($\mu = -0.01$; $\sigma = 0.32$).

(D) Net redundancy in shared information increases with neuronal ensemble size in extensive recordings from zebrafish, nematode, fruit fly, mouse, and macaque. Faint gray lines are individual recordings, red lines are species averages, and blue lines are data-matched independent null models. Error bars are $\mu \pm \sigma$. Mean (μ) and standard deviation (σ).

See also Figure S1.

reconfiguration at microscale and mesoscale—suggesting multiple, scale-dependent neural codes. Our results provide a common framework for simplifying the analysis of large-scale neuronal recordings and unifying findings across scales, phylogenetically diverse species, and behaviors.

RESULTS

Cross-scale transition from efficient to resilient neuronal activity

ICG begins by calculating the pairwise correlation between all neurons, N , then greedily pairing neurons by the temporal simi-

larity (Pearson correlation) of their calcium fluorescence and coarse-graining by summing activity pairwise before iterating at increasingly coarser scales, l . ICG thus finds dyadically increasing, coarse-grained, functionally coordinated “ensembles” of neurons of size $K = 2^{l-1}$ (neuronal cluster size; Figure 1A). The dyadic approach provides a consistent ensemble size for comparison between recordings, while the method reveals spatially diverse ensembles (Figure 1B).

To demonstrate this approach, we analyze neurons from the whole brain of larval zebrafish during spontaneous (quiescent) conditions. Analyzing the ICG of neuronal activity, we can systematically track neural dynamics as they transition from

microscale fast “spikes” (Figure 1A, red) to slower macroscale fluctuations (Figure 1A, purple).¹⁸ Zebrafish neuronal activity shows substantial spatial coordination, with ICG revealing functional neuronal ensembles of expanding size that respect known anatomical boundaries at microscales (Figure 1B, orange) yet integrate across regions at coarser scales (Figure 1B, blue). The mean inter-cluster distance is more significant than expected if clustering by spatial proximity ($p < 10^{-4}$), yet less than that expected of randomly paired clusters ($p < 10^{-4}$; Figure S1), supporting our choice to pair by temporal similarity.

ICG recapitulates known microscale-macroscale differences in pairwise correlations.^{13,14} At the neuronal resolution, the correlation distribution is tightly distributed around zero (i.e., uncorrelated; example from a single zebrafish $l = 1$; mean (μ) = 10^{-2} ; standard deviation (σ) = 0.06; Figure 1C, red). However, with coarsening, this distribution incrementally broadens such that at the systems level, ensembles of spatially distributed neurons display broad correlations ($l = 13$; $\mu = -0.01$; $\sigma = 0.32$; Figure 1C, purple). While the ICG algorithm greedily selects positively correlated ensembles, it is notable that the method still identifies anticorrelated populations at macroscales.

What does this cross-scale coordination imply for neural coding? To explore this, we calculated the net redundant shared information in pairwise neural activity estimated as the overlapping local entropy exceeding joint synergistic information (see STAR Methods).²⁴ Across all recordings and species, we observed a strong emphasis of low redundancy at the neuronal scale ($l = 1, K = 1$ neurons cross-species mean ~ 0 bits; Figure 1D). In contrast, we found a shift to increased net redundancy at coarser scales (e.g., $l = 13, K = 4,096$ neurons cross-species mean ~ 4 bits), significantly exceeding that expected from the coarse-graining procedure and finite sampling (Figure 1D, blue; demonstrated with data-matched independent null models; $p < 0.05$ 95% confidence interval [CI]). In this way, our information-theoretic decomposition supports both efficient (neuronal) and resilient (population) neural codes, albeit at opposing sizes of neuronal ensembles. How can the same neuronal activity support these opposing functional principles?

To explain this increase in redundancy, we compare changes in the extreme values (“tailedness”) of the pairwise correlation distribution. We calculated the kurtosis, $(\mu_4/\sigma^4)_l$ (where μ_4 is the fourth central moment; Figure 2A), across scales and found a transition from heavy-tailed correlations at the microscale to Gaussian-like distributed (Figure 2A, green line; $\mu_4/\sigma^4 = 3$) correlations at the macroscale. Interestingly, the scale that marked the transition to Gaussian-like kurtosis coincides with the shift from efficient to resilient coding regimes (Figures 1D and 2B–2F). The decay in kurtosis across scales is slower in fish (Figure 2B), nematode (Figure 2C), and fruit fly (Figure 2D), while it is faster in mammals, i.e., mouse (Figure 2E) and primate (Figure 2F), owing to their larger cellular heavy-tailedness ($\mu_4/\sigma^4 > 20$). We found that the few initially strongly correlated neuronal pairs drove this population-level transition—as their interdependence persisted across scales, becoming emphasized at the coarsest resolutions (Figure S2).^{4,25} Moreover, the smooth decay is not a general feature of coarse-graining (Figure S3), and leptokurtosis is not observed in trivial coarse-graining of independent or perfectly correlated signals (Figure S4).

This relation supports the use of a multiscale analysis as these rare pairwise correlations are invisible through the lens of the typically reported first two moments. These findings suggest a multiscale coordination of neuronal activity.

Functional neuronal activity displays self-similar scaling across species

Next, we use scaling analysis to quantify the rate of activity changes across different spatiotemporal resolutions. We compared changes in functional signatures by calculating the mean-variance (σ_l^2 ; Figure 2G) and the mean-timescale (τ_l) estimated as the autocorrelation integral over 3 s; Figure 2H) of the coarse-grained activity at each scale (see STAR Methods). These measures quantify static (σ_l^2) and dynamic (τ_l) properties that can characterize the transition from sparse and fast activity at the neuronal scale to slow and variable activity at the macroscale (Figure 1A). The rate at which σ_l^2 and τ_l change across scales reflects the internal organization of a multiscale system. We can quantify this relationship through the linear slope of the *logarithmically* plotted signatures—i.e., for variance scaling as $\sigma_l^2 = \sigma_0^2 K^\alpha$, α can be obtained from the logarithmic $\log \sigma_l^2 = \alpha \log K + \log \sigma_0^2$.

Mean-variance and mean-timescale were chosen because their scaling is theoretically constrained, such that they can discriminate trivial coarse-graining. For instance, if ICG is applied to an independent time series, then the mean-variance scales linearly, $\sigma_l^2 \propto K^1$, proportional to the number of independent neurons (Figure 2G, blue), and $\tau_l \propto K^0$ remains constant across coarser levels (Figure 2H, blue; see STAR Methods). Another trivial case arises if the neuronal activity is perfectly correlated with equivalent variability: in this case, $\sigma_l^2 \propto K^2$ (Figure 2G, green)—i.e., the mean-variance scales quadratically—and again $\tau_l \propto K^0$ as the timescale remain constant (Figure 2H, green; see STAR Methods). Alternatively, if there is a privileged scale, K_p (Figures 2G and 2H, orange star), at which neuronal activity transitions between regimes, the scaling will display a kink. In contrast, power-law scaling of $\sigma_l^2 \propto K^\alpha$ with an intermediate slope $1 < \alpha < 2$ (Figure 2G, red) and $\tau_l \propto K^\beta$ where $\beta > 0$ (Figure 2H, red) indicates a nontrivial, nonlinear increase in variance by α and timescale by β with increasing neuronal ensemble size.

We find strong evidence supporting nontrivial, self-similar static and dynamic coordination of spontaneous neuronal activity across all species (demonstrated by a correlation coefficient of logarithmic ensemble size and variance/timescale of $r > 0.9$ across all recordings, where $r = 1$ indicates a power law). In zebrafish, we observed power-law scaling up to five orders of magnitude ($\alpha = 1.50 \pm 0.02$, $\mu \pm 95\%$ CI; Figures 2I, $\beta = 0.22 \pm 0.04$, and 2J). Similarly, we also found near whole-brain scaling in nematode ($\alpha = 1.52 \pm 0.02$; Figures 2K, $\beta = 0.21 \pm 0.02$, and 2L). Furthermore, we observed power-law scaling in sensory regions of fruit fly mushroom body ($\alpha = 1.51 \pm 0.01$; Figures 2M, $\beta = 0.20 \pm 0.06$, and 2N) and within the primary visual system of mice ($\alpha = 1.51 \pm 0.02$; Figures 2O, $\beta = 0.19 \pm 0.04$, and 2P) and visual and motor regions of primate ($\alpha = 1.50 \pm 0.01$; Figures 2Q, $\beta = 0.20 \pm 0.04$, and 2R). The absence of a privileged scale and self-similar coherence suggests that this transition in function emerges from a systematic coordination of neuronal interdependence across coarser resolutions.

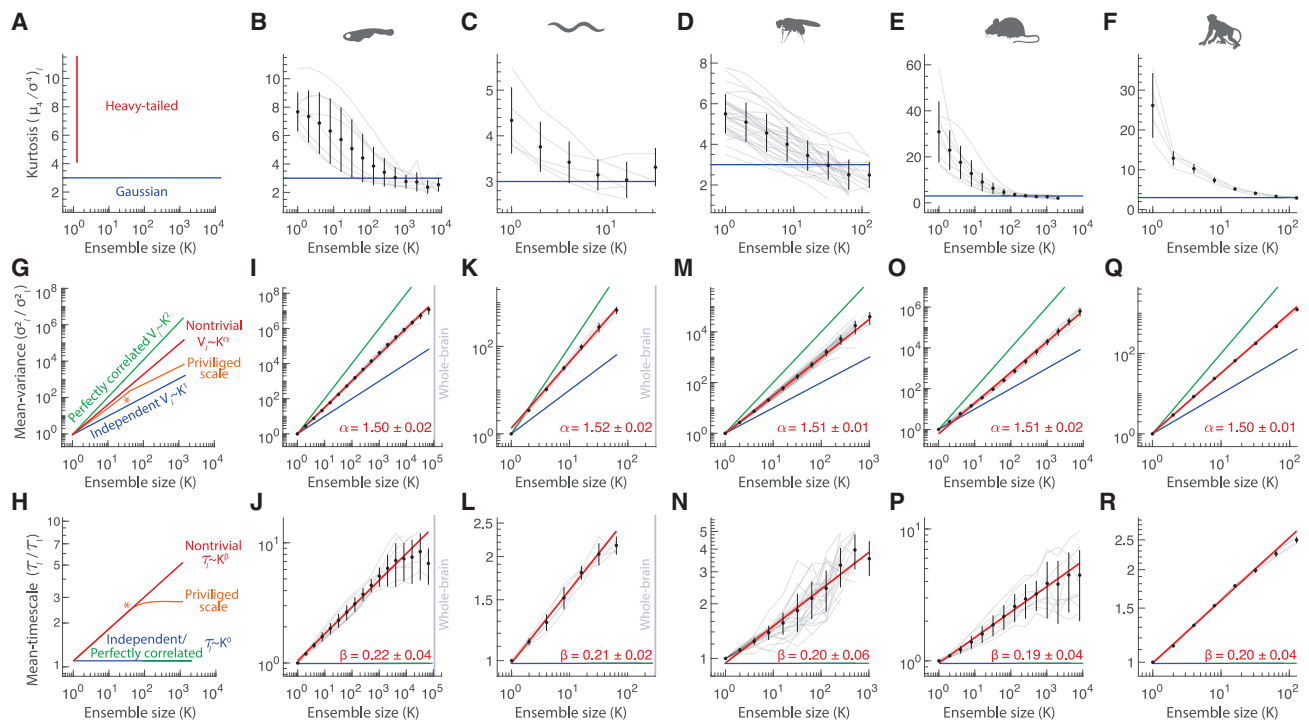


Figure 2. Phylogenetically preserved self-similar scaling of neuronal activity

(A) The kurtosis (or tailedness) of the pairwise correlation distribution shift across the iterative multiscale procedure from leptokurtotic “heavy-tailed” (red line; $\mu_4/\sigma^4 > 3$) to matching that of the Gaussian distribution (blue line).

(B and C) We calculated these measures across (B) zebrafish and (C) nematode *C. elegans* recordings with near whole-brain coverage.

(D–F) We also applied ICG to regional recordings from (D) fruit fly *Drosophila* central brain, (E) mouse primary visual cortex, and (F) macaque primary visual and motor cortices.

(G) The mean-variance, σ_t^2 , is a static measure that captures the average activity variability at each stage of the iterative multiscale procedure. The approach can differentiate between the presence of a scale-dependent transition in the neuronal organization (i.e., “privileged scales,” orange) and trivial scaling of either perfectly correlated (quadratic scaling, green) or independent (i.e., uncorrelated; linear scaling, blue) neural dynamics. Nontrivial self-similar scaling ($1 < \alpha < 2$) is denoted by a power law $\sigma_t^2 \propto K^\alpha$ (following a straight line on logarithmic axes).

(H) The mean-timescale, τ_t , is a dynamic measure calculated as the average autocorrelation integral (between a log of 0 and 3 s) across scales. Trivial coarse-graining of perfectly correlated (green) or independent (blue) signals does not alter the timescale. Nontrivial self-similar scaling ($\beta > 0$) is denoted by a power law $\tau_t \propto K^\beta$ (red).

(I–R) σ_t^2 and τ_t were calculated across: (I and J) zebrafish; (K and L) nematode; (M and N) fruit fly; (O and P) mouse; and (Q and R) macaque. Error bars are $\mu \pm 1.96\sigma$.

See also [Figures S2, S3, S4, and S5](#).

To ensure the scaling findings are not biased by the ICG algorithm and that the technique can correctly identify the trivial conditions, we repeated the analysis using generative and data-matched null models. First, we asked whether a data-matched covariance structure can explain the scaling. While this null preserved the static scaling, dynamic scaling was destroyed ([Figure S3A](#)). Next, we tested the sensitivity of the algorithm to scale dependencies by repeating the analysis after removing higher-order correlations outside of artificially generated scale dependencies (e.g., $K_p = 4, 32, 128$), leading to scale-dependent kinks in mean-variance ([Figure S3A](#)). We then asked whether the time-scale analysis is sensitive to scaling changes by circularly time shifting neuronal activity independently between neurons and among ICG-identified neuronal ensembles ($K_p = 2, 8, 32$) before reapplying ICG, leading to trivial scaling between the matched scale-dependent departures ([Figure S3B](#)).

We also asked if the intermediate static scaling emerges from a sparse sampling of a strong global coordination²⁶ that can result in

weak correlations at the neuronal scale. We created this generative null model by simulating neuronal activity using an inhomogeneous Poisson process with a nonstationary rate matched to an empirical coarse-grained global activity. Due to the sparse Poisson sampling, the neuronal activity appeared weakly correlated ($\mu \sim 0.05$; [Figure S4B](#)). Nevertheless, the ICG algorithm analysis quickly revealed the global coordination ($\alpha = 2$ and $\beta = 0$) within four iterations of the algorithm ([Figure S4B](#)). Thus, neuronal activity contains a self-similarity that cannot be explained by perfectly correlated ($\alpha = 2$; $p < 1 \times 10^{-4}$, paired t test) or independent ($\alpha = 1$; $p < 1 \times 10^{-4}$, paired t test) activity, and ICG scaling analysis can discriminate spurious noise correlations, independent samplings of correlated global fluctuations, and the presence of a privileged scale dependence.

The scaling was consistent across all animals ($p > 0.05$, one-sided ANOVA with Bonferroni correction for multiple comparisons testing the null hypothesis that the scaling exponent is equal; [Table S1](#)). All scaling exponents were consistent across

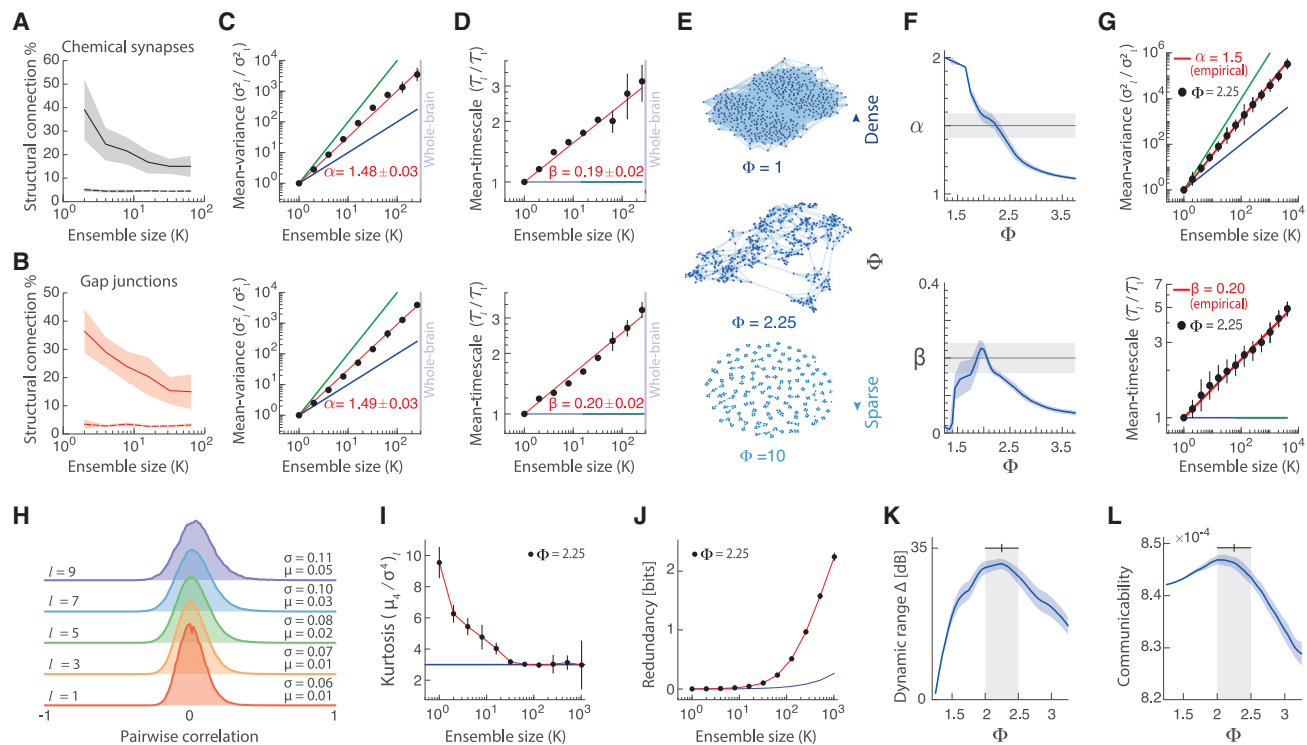


Figure 3. Multiscale structure and function and its benefits

(A–D) In *C. elegans*, neurons that are connected—either via chemical synapses (A, black) or gap junctions (B, red)—are likely to be functionally correlated, particularly at the microscale. Solid line: empirical recordings $\mu \pm 95\%$ CI; dashed line: 100 random pairings. Simulations atop these physical connectivities recapitulate the empirical nematode self-similar static (C) and dynamic (D) scaling exponents (see Figures 2K and 2L).

(E) Schematic, a hierarchical-modular network can be organized into topologies ranging from dense ($\Phi = 1$) to sparse ($\Phi = 10$) fractal gradients.

(F) Simulations on the hierarchical topologies display varying self-similar static scaling and dynamic scaling, which overlap with the empirical exponents (gray box). Error bars are $\mu \pm \sigma$.

(G) Example scaling for the fractal topology ($\Phi = 2.25$ black dots; $\mu \pm 1.96\sigma$ over 100 simulations) with empirically observed scaling overlaid (red).

(H) The empirically matched fractal network recapitulates the transition in pairwise correlation distribution from being tightly distributed around zero to broadening at coarser scales (see Figure 1C).

(I) The model also recapitulates the pairwise correlation distribution transition from heavy-tailed (high kurtosis) to normally distributed (kurtosis ~ 3). Error bars are $\mu \pm 1.96\sigma$.

(J) Similarly, this simulation displays increasing net redundancy with neuronal population ensemble size (see Figure 1D). Error bars are $\mu \pm \sigma$.

(K) The network structures that match the empirical multiscale dynamics possess a maximal dynamic range, Δ , the range (in dB) where population activity can discernibly encode variations in the stimulus intensity. Error bars are $\mu \pm \sigma$.

(L) The network structures that match the empirical multiscale dynamics also possess peaked communicability, a measure of the capacity to propagate information across the system. Error bars are $\mu \pm \sigma$.

See also Figure S7.

the first and second halves of their respective recordings (Figure S5) and robust across alternative approaches to measure similarity, such as the Spearman coefficient and Jaccard index (Figure S6). Therefore, these results demonstrate that a precise self-similar scaling of neuronal activity is preserved across diverse species, facilitating a breadth of functionally beneficial timescales that the system can operate across, and may reflect a whole-brain organizing principle (stretching across five orders of magnitude in the zebrafish). Nevertheless, it remains to be seen where this multiscale organization emerges from and whether this permits any information processing benefits that may explain its preservation across broadly evolutionarily separated species.

Multiscale structure and function and their efficiencies

How does functional self-similar scaling relate to the underlying connections between neurons? To address this, we analyzed the ICG neuronal pairings of the *C. elegans* recordings, as both the recorded cells have been uniquely identified, and the nematode's structural connectivity (chemical synapse and gap junction) is fully characterized.²⁷ Across all scales, the mean likelihood of a functional ensemble grouping of neurons possessing a synaptic (Figure 3A) or gap junction (Figure 3B) connection is more significant than chance (Figures 3A and 3B, dashed lines). In particular, the grouping of ICG at finer scales correlates most strongly with structural connectivity, with $40\% \pm 13\%$ of initial neuronal pairings possessing a synaptic connection. At coarser scales, the

likelihood of all neurons possessing a structural connection is halved ($\sim 17\% \pm 7\%$) as functional couplings become driven by higher-order (indirect) effects. These results suggest that the structural connectivity between neurons significantly informs the emergent, multiscale functional coordinated activity.

To investigate the structural origin of multiscale neuronal dynamics, we next simulated cellular activity using a stochastic three-state cellular automaton neuronal model.^{28,29} This branching process is a canonical model of neuronal activity^{30,31} that reduces biological complexity, which improves computational efficiency (facilitating simulations across multiple topological families), and it is routinely utilized to explore universal dynamics (following that universality is independent of system details).^{16,17,29,32} Briefly, cells are simulated in discrete time steps and can exist in either of three states: quiescent-active-refractory (QAR). Cells enter the active (“spiking”) state from the quiescent state either stochastically via an active presynaptic neuron or through an external input (Poisson drive), after which the active-state cells enter the refractory state before ultimately returning to quiescence (see STAR Methods). The cellular model was able to generate rich (synchronous and asynchronous) temporal patterns of coordinated neural activity and the correlation between structure-function that decays with scale (significantly exceeding chance levels, $p < 0.05$, 95% CI; Figure S7). Furthermore, the model recapitulated the empirical static (Figure 3C) and dynamic (Figure 3D) scaling when simulated atop both the nematode’s chemical synapse and gap junction structural connectivities. That is, physical connections play an integral role in generating multiscale coordination. These results confirm a multiscale structure-function relationship within the nematode and support the utility of the QAR model as a generative process to investigate the structural control of functional activity scaling relationships.

Following the relationship between structure and function within the brain, we theorized that a hierarchical-modular network could reproduce the empirically observed functional activity scaling.^{33,34} Hierarchical-modular networks possess multiple scales that are strongly connected at the base level and that decrease exponentially with $P_l = 1/\Phi^l$, where P_l is the probability of a connection between ensemble pairs at scale l , and modules are iteratively connected (see STAR Methods). This fractal gradient can be transitioned between dense ($\Phi = 0$) and sparse ($\Phi = 10$) hierarchical networks (Figure 3E). We found that dynamics simulated on this network yield self-similar static and dynamic scaling that smoothly changes from strongly correlated ($\Phi = 0 \rightarrow \alpha = 2$) to independent ($\Phi = 4 \rightarrow \alpha = 1$) activity (Figure 3F). Within a restricted range of hierarchical-modular networks (between $2 < \Phi < 2.5$), the simulated dynamics recapitulate both the empirically observed static and dynamic scaling exponents (Figure 3G). Consistent with empirical findings (Figure 1D), the correlation distribution is tightly distributed around zero at the finest scale ($l = 1$) and broadens at coarser scales ($l = 9$; Figure 3H). Further, this model recapitulates the empirical transition in kurtosis from heavy-tailed to normally distributed pairwise correlations (Figure 3I), which is aligned with the transition to increasing redundancy (Figure 3J).

Next, we simulated activity across alternative topological families to identify the relationship between structure-function and multiscale dynamics. To contrast simulations across topologies,

we rescaled network connectivity strength or external input to produce a biophysically consistent maintained spike rate of ~ 7 Hz (see STAR Methods). Nevertheless, we found that other topologies typically used to simulate brain network structure—such as modular, scale-free degree, geometric spatially coupled, and small-world (Figure S8)—could not simultaneously reproduce both the static and dynamic scaling (e.g., scale-free degree and small-world), or they displayed privileged scales (e.g., modular and geometric spatially coupled). These generative models emphasize that the observed scaling is nontrivial and suggest that a limited range of hierarchical-modular topologies may explain neural multiscale dynamics.

What functional benefits for information processing do this precise scaling and transition between efficiency and resiliency facilitate? This transition may reflect the capacity of the network to trade off distinct responsivity with persistent activity.³⁵ To explore this possibility, we drove the hierarchical-modular networks with a broad range of transient inputs of increasing frequency. We measured the resulting population spiking output over 1s, and from this, we calculated the dynamic range, Δ (dB), defined as the 10%/90% discernible stimulus range of the input-output curve.²⁹ The networks reproducing empirical functional scaling coincided with a peaked dynamic range (Figure 3K). To further verify that the network regime that matched empirical scaling facilitated increased information processing signatures, we calculated the network communicability—an estimation of diffuse information flow³⁵—across the hierarchical-modular topologies. We found that the networks that match empirical scaling also possess peaked communicability (Figure 3L).^{36,37} These results suggest that empirically observed functional multiscale activity can facilitate increased information flow and a susceptibility to a wide range of stimuli.³⁸

Flexible multiscale behavioral reconfigurations

The analysis thus far has focused on empirical spontaneous conditions to understand the ongoing coding regimes that can constrain adaptive behavior. To examine behavior directly, we next investigated the functional impact of a multiscale organization during evoked behaviors in zebrafish and mice. First, we compared ICG activity between spontaneous (Figures 4A, 4B, 4E, and 4F, top) and evoked behaviors (visually evoked fictive swimming while head-fixed fish were presented with left vs. right phototactic stimuli and mice viewed moving visual gratings; Figures 4A, 4B, 4E, and 4F, bottom).^{18,39} Our analysis found that static variance scaling is consistently self-similar across spontaneous and behavioral states ($\alpha \sim 1.5$; Figures 4A and 4E). In contrast, we found that the zebrafish dynamic scaling (Figure 4B) displays a characteristic kink (Figure 4B, orange star) consistent with the induced timescale of the phototactic stimulation. These findings suggest that a benefit of multiscale neuronal organization is the capacity for neural ensembles to adapt their timescales to behavioral demands.

We hypothesized that behavioral demands require substantial neuronal ensemble reconfigurations. To test this hypothesis, we quantified cross-scale reconfigurations by first calculating the adjusted mutual information (AMI) between neuronal ensembles (at each scale) in spontaneous and behavioral sessions (see STAR Methods). An AMI of 1 indicates identical ensemble

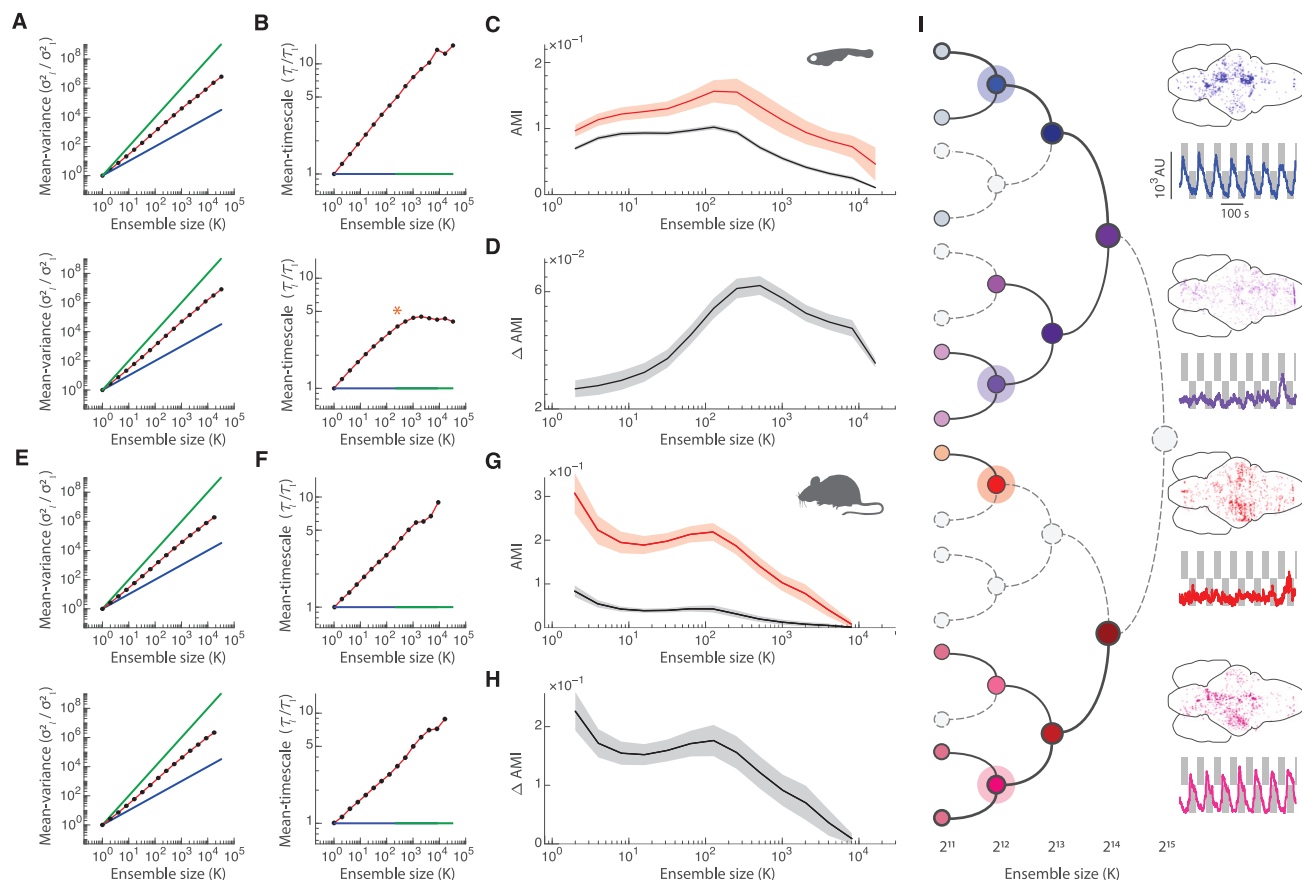


Figure 4. Preserved multiscale organization during behavior facilitates flexible scale-dependent functional reconfigurations

(A and B) Comparing zebrafish multiscale neuronal activity static (A) and dynamic (B) scaling in spontaneous (top) and visually evoked phototactic “swimming” (bottom). Despite conserved variance scaling, the time-locked stimuli entrain a temporal dependence (orange star).

(C) Ensembles are consistent within a behavioral state demonstrated by the adjusted mutual information (AMI) that is higher for within (red) vs. between (black) behavior comparisons for $l = 3-14$ ($p < 0.001$; calculated from an independent sample t test).

(D) The scale-dependent difference in AMI (Δ AMI; within minus between behaviors) shows that the strongest behavioral reconfiguration is at the mesoscale ($K = 512$).

(E–H) Static (E) and dynamic (F) scaling is preserved in mice visual cortexes when viewing a blank screen (top) and moving sinusoidal stimuli (bottom). Mice display a peak in ensemble stability (G) and behavioral reconfiguration (H) at the mesoscale; however, there is also significant reconfiguration at the microscale. Error bars are $\mu \pm 95\%$ CI.

(I) Left: An ICG hierarchical tree denoting whether the zebrafish ensemble activity significantly differentiated left vs. right phototactic stimuli: filled, colored circles with thick lines denote significant differences between left vs. right at $p < 0.001$ (5,000 permuted null distributions), those with thin lines denote $0.001 < p < 0.05$, and gray circles with dashed lines denote a lack of significant behavioral coordination ($p > 0.05$). Right: Whole-brain maps and activity detailing the spatial coordinates of the cells that formed significant functional clusters at level $l = 11/K = 2^{12}$ (outlined circles)—note the unique and spatially distributed mesoscale ensembles. Each ensemble’s activity (colored line) demonstrates distinct temporal relationships with the phototactic stimulus timing (shown in gray).

pairings, whereas an AMI of 0 indicates the mutual information as expected from chance pairings.⁴⁰ We found pairings are more stable within (Figures 4C and 4G, red) rather than between (Figures 4C and 4G, black) behaviors across split-half recordings ($p < 0.05$ $1.96\sigma / \sqrt{N}$).

We next calculated the reconfiguration between behaviors (Δ AMI; see STAR Methods), which provides an estimate of behavior discriminability (i.e., ensemble information content). Within the zebrafish, we found the largest Δ AMI at the mesoscale ($K = 512$; Figure 4D). In contrast, within mice we found both significant mesoscale reconfigurations and microscale reconfigurations during visual stimuli (Figure 4H). Interestingly,

this cross-scale reconfiguration was bimodal, with distinct peaks at the microscale and mesoscale. These findings support evidence that the multiscale organization is conserved across behaviors and species, and it may facilitate simultaneous distinct neural coding principles (e.g., at the microscale and mesoscale in mice).

To investigate the coordinated mesoscale ensembles during the zebrafish fictive-swimming task, we conducted a cross-scale analysis in which activity at different scales was used to predict the timing of the left vs. right visual stimuli.¹⁸ We observed significant differentiation of left vs. right swimming behavior across scales, with a rising proportion of significant

ensembles with coarser spatial scale (increasing from 18% at $K = 1$ to 37% explained variance at $K = 4,096$) that then diminished at levels coarser than $l = 13$ (i.e., peak explanatory power in ensembles of thousands of cells).

Tracking the branches of this hierarchy demonstrates how neurons reconfigure their coordination as a function of phototaxis swimming. At $l = 11$, 4 (out of 16) ensembles significantly differentiated left vs. right visual input, yet each involved unique anatomical regions and characteristic dynamics associated in distinct ways with the task. One ensemble (Figure 4I, blue) was dominated by hypothalamic, thalamic, and brainstem cells, displaying distinct left vs. right differentiation throughout the task. Another ensemble involving the telencephalon and thalamus (Figure 4I, purple) exhibited left vs. right differentiation later in the task. A third ensemble (Figure 4I, pink) was aligned with switches between task blocks, and it involved the telencephalon. The final behaviorally significant ensemble (Figure 4I, red) showed increasing responses to the task and predominantly contained brainstem neurons. Taken together, these patterns reveal the intricate and widespread neuroanatomical reconfigurations that occur in response to adaptive behavior, which fundamentally differ across scales, suggesting potential targets for future experiments.

DISCUSSION

We applied ICG to phylogenetically diverse and extensive calcium-imaging recordings to explore neuronal activity across multiple spatiotemporal scales. Notably, this approach revealed a self-similar organizing principle that supports scale-dependent theories of brain function: at the neuronal scale, activity correlations were tightly distributed around zero with minimal redundancy, enhancing information efficiency. In contrast, when coarse-grained, the same neuronal activity exhibited broad correlations with increased redundancy, supporting a resilient population code. The transition between these two information processing regimes is marked by a shift from heavy-tailed to Gaussian-like distributed correlations. This change could be explained by a phylogenetically conserved and nontrivial self-similar coordination of functional neuronal activity and intrinsic time-scales. This functional activity scaling is closely linked to the underlying structural connectivity (i.e., multiscale structure and function). Neuronal simulations on a subset of hierarchical-modular networks recapitulated the empirical scaling, which also increased the network's information communicability and susceptibility to a broad range of input. Exploration across diverse behaviors suggests that multiscale neuronal coordination may facilitate simultaneous neural coding schemas at distinct scales.

What are the benefits of a brain operating across multiple neural scales, each with differing statistical properties? One advantage of a redundant population code lies in its ability to encode function through ensemble activity, such as a broad dynamic range capable of discriminating between slight changes in input within the network. The observation that neural coordination during fictive swimming peaks at the mesoscale supports theories that movement is redundantly encoded across large populations of neurons, providing a necessary resilience to cellular variability for adaptive behavior.⁴¹ Moreover, in mouse V1, we observed evidence for distinct scale-dependent reconfigurations occur-

ring simultaneously at both the microscale and mesoscale. The minimally redundant (i.e., maximally efficient) peak at the microscale supports theories of cellular efficiency derived from well-known sparse cellular responses to grating visual stimuli.⁷ However, the secondary mesoscale peak in functional affiliation reconfigurations suggests an extension to this theory, highlighting a cross-scale interplay between efficiency and redundancy in visual coding. In this way, the multiscale coordination bridges distinct theories of neural coding.

These analyses suggest that multiscale neuronal activity is a preserved principle of neural organization consistent within vertebrates (last common ancestor ~450 million years ago) and invertebrates (last common ancestor ~1 billion years ago).⁴² We hypothesize that the simple, functional scaling represents an evolutionary pressure of beneficial information processing—e.g., balancing neuronal efficiency and population resiliency, the susceptibility to a broad range of internal and external time-scales, and the capacity to substantially reconfigure as a function of shifting behavioral demands. Future experiments should be designed to explore regional divergences in scaling exponents for insight into coding specialization (e.g., differential scaling in the hippocampus vs. sensory regions). Further, there is evidence of conserved scaling in electrophysiological recordings (across hundreds of cells), albeit with slightly shallower scaling,⁴³ suggesting the need for more direct comparisons across recording modalities. Finally, the structure-function relationship we observed may explain the persistence in allometric scaling in brain-body size across species.⁴⁴

Understanding how these cross-scale principles bridge cellular-level details and whole-brain dynamics is an important open question for neuroscience. The consistency of an invariant multiscale structure across species and simulations (even in our simplified model) provides new footholds for further theoretical modeling. For example, quantifying how the moments of neuronal correlations and activity (i.e., their mean, standard deviation, and kurtosis) change across scales provides summary measures of system behavior to triangulate and invert multiscale neural models. The relaxation of neural statistics from sparse and heavy-tailed correlations (leptokurtotic) at fine scales that accumulate to redundant and variable at coarse scales is a central feature of neuronal activity in all species we analyzed. This shift in activity regimes motivates the need for new population models that violate the “diffusion approximation” upon which classic mean-field theoretical models depend.⁴ Notably, identifying fixed points under ICG can elucidate critical mean-field phenomena.¹⁶ However, the observed macroscale activity patterns are inconsistent with classic avalanches—or crackling-noise—where dynamics remain bursty and sparse and display avalanche-like cascades across all scales, unlike the observed coarse whole-brain population activity.⁴⁵ This supports findings showing that avalanche criticality is sufficient⁴⁶ but not necessary for the observed multiscale dynamic.⁴⁷ Likewise, our generative cellular model rests upon a close coupling of functional activity and structure (a hierarchical network) and is not the type of microscopic branching process that yields classical avalanches. These considerations align with the framework that adaptive neural processes are a slightly subcritical process^{48–50} and that avalanche dynamics are more frequently associated with pathological states

such as burst suppression.⁵¹ Therefore, we hope that more complex biophysical modeling approaches will extend our findings to understand healthy³ and pathological⁵² multiscale brain states. For example, pharmacological manipulation can test the role of the neuromodulatory ascending arousal system control of activity scaling and its disengagement from the underlying physical structural connectivity.^{3,53}

Our results highlight the benefit of utilizing a cross-scale approach to analyze the rapidly increasing scope of neuronal recordings.^{1,54–56} Through this approach, we provide evidence for universal self-similar scaling of neuronal activity across species and identify a functional benefit for this organization—offsetting neuronal-coding efficiency and population-coding resiliency⁵—thus reconciling scale-dependent theories of brain function.^{6,8} We believe that future studies incorporating this methodology—especially in rich behavioral experiments with sophisticated causal manipulations—will hasten the identification of the cross-scale neural code.

Limitations of the study

The mammalian recordings analyzed in this study are of substantial size for contrasting dynamics across multiple orders of cell populations, yet (unlike zebrafish and *C. elegans* recordings) it remains to be demonstrated whether this also applies across the mammalian whole brain. Similarly, temporal scaling analysis is limited by the recording duration and Ca²⁺ sensor resolution demonstrated by the within-species and cross-species variability. As Ca²⁺ sensor and imaging technology advances, so will our capacity to test these hypotheses in whole-brain sampling^{2,57} and their consistency within recordings made with alternative modalities.⁵⁸ Likewise, while the behaviors analyzed in this study are limited to visually evoked head-fixed conditions, we expect that freely moving recordings will enrich our understanding of multiscale neural reconfigurations as a function of nuanced behaviors and behavioral information processing. Further, to isolate the role of topology constraining neural dynamics in our study, we did not utilize a biophysically complex model, which significantly reduced the number of parameters in the model and allowed testing for detail-invariant universal scaling. We anticipate that more biophysically detailed models capturing neuronal nonlinearities,⁵⁹ the balance of excitation and inhibition,¹³ or the role of subcortical projections⁶⁰ will refine our understanding of additional factors that contribute to instantiating multiscale neuronal activity and further functional benefits.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Brandon Munn (brandon.munn@sydney.edu.au).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Zebrafish calcium imaging has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).

- All original code has been deposited at Zenodo and is publicly available as of the date of publication; code can also be obtained at <https://www.github.com/Bmunn/ICG>. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

This work is dedicated to Richard Munn. We thank the scientists and laboratories committed to open science, particularly those used in this study. I.F.-B. acknowledges funding from the NHMRC (APP2012140) and ARC (DP220103812 and DE230100972). M.B. acknowledges funding from the NHMRC (APP2008612). J.M.S. acknowledges funding from the NHMRC (APP2033162 and APP1193857) and ARC (DP240101295).

AUTHOR CONTRIBUTIONS

Conceptualization, B.R.M.; methodology, B.R.M., E.J.M., and J.L.; writing, B.R.M. and J.M.S.; editing, B.R.M., E.J.M., I.F.-B., E.S., J.L., M.B., and J.M.S.; investigation, B.R.M.; analysis, B.R.M. and J.M.S.; resources, E.S.; supervision, M.B. and J.M.S.; funding, J.M.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
 - Animals
- [METHOD DETAILS](#)
 - Iterative Coarse Graining
 - Neuronal network simulations
 - Topological families
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
 - Kurtotic
 - Static
 - Dynamic
 - Estimation of scaling exponent
 - Redundancy estimation
 - Limitations of the ICG algorithm
 - Statistical null models

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2024.10.004>.

Received: August 5, 2023

Revised: August 9, 2024

Accepted: October 3, 2024

Published: October 30, 2024

REFERENCES

1. He, B.J. (2014). Scale-free brain activity: past, present, and future. *Trends Cogn. Sci.* 18, 480–487. <https://doi.org/10.1016/j.tics.2014.04.003>.
2. Machado, T.A., Kauvar, I.V., and Deisseroth, K. (2022). Multiregion neuronal activity: the forest and the trees. *Nat. Rev. Neurosci.* 23, 683–704. <https://doi.org/10.1038/s41583-022-00634-0>.
3. Shine, J.M., Müller, E.J., Munn, B., Cabral, J., Moran, R.J., and Breakspear, M. (2021). Computational models link cellular mechanisms of

- neuromodulation to large-scale neural dynamics. *Nat. Neurosci.* 24, 765–776. <https://doi.org/10.1038/s41593-021-00824-6>.
4. Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nat. Neurosci.* 20, 340–352. <https://doi.org/10.1038/nn.4497>.
5. Panzeri, S., Moroni, M., Saffari, H., and Harvey, C.D. (2022). The structures and functions of correlations in neural population codes. *Nat. Rev. Neurosci.* 23, 551–567. <https://doi.org/10.1038/s41583-022-00606-4>.
6. Barlow, H.B. (1961). Possible principles underlying the transformations of sensory messages. In *Sensory Communication*, W.A. Rosenblith, ed. (The MIT Press), pp. 216–234. <https://doi.org/10.7551/mitpress/9780262518420.003.0013>.
7. Olshausen, B.A., and Field, D.J. (2004). Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* 14, 481–487. <https://doi.org/10.1016/j.conb.2004.07.007>.
8. Tononi, G., Sporns, O., and Edelman, G.M. (1999). Measures of degeneracy and redundancy in biological networks. *Proc. Natl. Acad. Sci. USA* 96, 3257–3262. <https://doi.org/10.1073/pnas.96.6.3257>.
9. Hebb, D.O. (1949). *The Organization of Behavior: A Neuropsychological Theory* (Wiley).
10. Harris, K.D. (2005). Neural signatures of cell assembly organization. *Nat. Rev. Neurosci.* 6, 399–407. <https://doi.org/10.1038/nrn1669>.
11. Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.* 14, 811–819. <https://doi.org/10.1038/nn.2842>.
12. Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* 7, 358–366. <https://doi.org/10.1038/nrn1888>.
13. Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587. <https://doi.org/10.1126/science.1179867>.
14. Shine, J.M., Breakspear, M., Bell, P.T., Ehgoetz Martens, K.A., Shine, R., Koyejo, O., Sporns, O., and Poldrack, R.A. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat. Neurosci.* 22, 289–296. <https://doi.org/10.1038/s41593-018-0312-0>.
15. Cunningham, J.P., and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* 17, 1500–1509. <https://doi.org/10.1038/nn.3776>.
16. Meshulam, L., Gauthier, J.L., Brody, C.D., Tank, D.W., and Bialek, W. (2019). Coarse graining, fixed points, and scaling in a large population of neurons. *Phys. Rev. Lett.* 123, 178103. <https://doi.org/10.1103/PhysRevLett.123.178103>.
17. Kadanoff, L.P., Götzke, W., Hamblen, D., Hecht, R., Lewis, E.A.S., Palciauskas, V.V., Rayl, M., Swift, J., Aspnes, D., and Kane, J. (1967). Static phenomena near critical points: theory and experiment. *Rev. Mod. Phys.* 39, 395–431. <https://doi.org/10.1103/RevModPhys.39.395>.
18. Chen, X., Mu, Y., Hu, Y., Kuan, A.T., Nikitchenko, M., Randlett, O., Chen, A.B., Gavornik, J.P., Sompolsky, H., Engert, F., et al. (2018). Brain-wide organization of neuronal activity and convergent sensorimotor transformations in larval zebrafish. *Neuron* 100, 876–890.e5. <https://doi.org/10.1016/j.neuron.2018.09.042>.
19. Kato, S., Kaplan, H.S., Schröder, T., Skora, S., Lindsay, T.H., Yemini, E., Lockery, S., and Zimmer, M. (2015). Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. *Cell* 163, 656–669. <https://doi.org/10.1016/j.cell.2015.09.034>.
20. Pacheco, D.A., Thiberge, S.Y., Pnevmatikakis, E., and Murthy, M. (2021). Auditory activity is diverse and widespread throughout the central brain of *Drosophila*. *Nat. Neurosci.* 24, 93–104. <https://doi.org/10.1038/s41593-020-00743-y>.
21. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brain-wide activity. *Science* 364, 255. <https://doi.org/10.1126/science.aav7893>.
22. Tang, S., Zhang, Y., Li, Z., Li, M., Liu, F., Jiang, H., and Lee, T.S. (2018). Large-scale two-photon imaging revealed super-sparse population codes in the V1 superficial layer of awake monkeys. *eLife* 7, e33370. <https://doi.org/10.7554/eLife.33370>.
23. Trautmann, E.M., O’Shea, D.J., Sun, X., Marshel, J.H., Crow, A., Hsueh, B., Vesuna, S., Cofer, L., Bohner, G., Allen, W., et al. (2021). Dendritic calcium signals in rhesus macaque motor cortex drive an optical brain-computer interface. *Nat. Commun.* 12, 3689. <https://doi.org/10.1038/s41467-021-23884-5>.
24. Finn, C., and Lizier, J. (2018). Pointwise partial information decomposition using the specificity and ambiguity lattices. *Entropy* 20, 297. <https://doi.org/10.3390/e20040297>.
25. Buzsáki, G., and Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* 15, 264–278. <https://doi.org/10.1038/nrn3687>.
26. Schneidman, E., Berry, M.J., Segev, R., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 1007–1012. <https://doi.org/10.1038/nature04701>.
27. Cook, S.J., Jarrell, T.A., Brittin, C.A., Wang, Y., Bloniarz, A.E., Yakovlev, M.A., Nguyen, K.C.Q., Tang, L.T.-H., Bayer, E.A., Duerr, J.S., et al. (2019). Whole-animal connectomes of both *Caenorhabditis elegans* sexes. *Nature* 571, 63–71. <https://doi.org/10.1038/s41586-019-1352-7>.
28. Haber, A., and Schneidman, E. (2022). Learning the architectural features that predict functional similarity of neural networks. *Phys. Rev. X* 12, 021051. <https://doi.org/10.1103/PhysRevX.12.021051>.
29. Kinouchi, O., and Copelli, M. (2006). Optimal dynamical range of excitable networks at criticality. *Nat. Phys.* 2, 348–351. <https://doi.org/10.1038/nphys289>.
30. Larkum, M.E., and Nevian, T. (2008). Synaptic clustering by dendritic signalling mechanisms. *Curr. Opin. Neurobiol.* 18, 321–331. <https://doi.org/10.1016/j.conb.2008.08.013>.
31. Schiller, J., Schiller, Y., and Clapham, D.E. (1998). NMDA receptors amplify calcium influx into dendritic spines during associative pre- and postsynaptic activation. *Nat. Neurosci.* 1, 114–118. <https://doi.org/10.1038/363>.
32. Stauffer, D., and Aharony, A. (2018). *Introduction to Percolation Theory* (Taylor & Francis) <https://doi.org/10.1201/9781315274386>.
33. Moretti, P., and Muñoz, M.A. (2013). Griffiths phases and the stretching of criticality in brain networks. *Nat. Commun.* 4, 2521. <https://doi.org/10.1038/ncomms3521>.
34. Sporns, O. (2006). Small-world connectivity, motif composition, and complexity of fractal neuronal connections. *Biosystems* 85, 55–64. <https://doi.org/10.1016/j.biosystems.2006.02.008>.
35. Estrada, E., and Hatano, N. (2008). Communicability in complex networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 77, 036111. <https://doi.org/10.1103/PhysRevE.77.036111>.
36. Tononi, G., Sporns, O., and Edelman, G.M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA* 91, 5033–5037. <https://doi.org/10.1073/pnas.91.11.5033>.
37. Shine, J.M., Bissett, P.G., Bell, P.T., Koyejo, O., Balsters, J.H., Gorgolewski, K.J., Moodie, C.A., and Poldrack, R.A. (2016). The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* 92, 544–554. <https://doi.org/10.1016/j.neuron.2016.09.018>.
38. Munn, B., and Gong, P. (2018). Critical dynamics of natural time-varying images. *Phys. Rev. Lett.* 121, 058101. <https://doi.org/10.1103/PhysRevLett.121.058101>.
39. Stringer, C., Michaelos, M., Tsybolski, D., Lindo, S.E., and Pachitariu, M. (2021). High-precision coding in visual cortex. *Cell* 184, 2767–2778.e15. <https://doi.org/10.1016/j.cell.2021.03.042>.
40. Vinh, N.X., Epps, J., and Bailey, J. (2010). Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. *J. Mach. Learn. Res.* 11, 2837–2854.

41. Georgopoulos, A.P., Schwartz, A.B., and Kettner, R.E. (1986). Neuronal population coding of movement direction. *Science* 233, 1416–1419. <https://doi.org/10.1126/science.3749885>.
42. Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819. <https://doi.org/10.1093/molbev/msx116>.
43. Morales, G.B., di Santo, S., and Muñoz, M.A. (2023). Quasiuniversal scaling in mouse-brain neuronal activity stems from edge-of-instability critical dynamics. *Proc. Natl. Acad. Sci. USA* 120, e2208998120. <https://doi.org/10.1073/pnas.2208998120>.
44. Holloway, R.L. (1996). *Evolution of the Human Brain. Handbook of Human Symbolic Evolution* (Wiley), pp. 74–116.
45. Sethna, J.P., Dahmen, K.A., and Myers, C.R. (2001). Crackling noise. *Nature* 410, 242–250. <https://doi.org/10.1038/35065675>.
46. Nicoletti, G., Suweis, S., and Maritan, A. (2020). Scaling and criticality in a phenomenological renormalization group. *Phys. Rev. Research* 2, 023144. <https://doi.org/10.1103/PhysRevResearch.2.023144>.
47. Morrell, M.C., Sederberg, A.J., and Nemenman, I. (2021). Latent dynamical variables produce signatures of spatiotemporal criticality in large biological systems. *Phys. Rev. Lett.* 126, 118302. <https://doi.org/10.1103/PhysRevLett.126.118302>.
48. Cocchi, L., Gollo, L.L., Zalesky, A., and Breakspear, M. (2017). Criticality in the brain: A synthesis of neurobiology, models and cognition. *Prog. Neurobiol.* 158, 132–152. <https://doi.org/10.1016/j.pneurobio.2017.07.002>.
49. Müller, E.J., Munn, B.R., and Shine, J.M. (2020). Diffuse neural coupling mediates complex network dynamics through the formation of quasi-critical brain states. *Nat. Commun.* 11, 6337. <https://doi.org/10.1038/s41467-020-19716-7>.
50. Munn, B.R., Müller, E.J., Aru, J., Whyte, C.J., Gidon, A., Larkum, M.E., and Shine, J.M. (2023). A thalamocortical substrate for integrated information via critical synchronous bursting. *Proc. Natl. Acad. Sci. USA* 120, e2308670120. <https://doi.org/10.1073/pnas.2308670120>.
51. Roberts, J.A., Iyer, K.K., Finnigan, S., Vanhatalo, S., and Breakspear, M. (2014). Scale-Free Bursting in Human Cortex following Hypoxia at Birth. *J. Neurosci.* 34, 6557–6572. <https://doi.org/10.1523/JNEUROSCI.4701-13.2014>.
52. Jirsa, V.K., Sporns, O., Breakspear, M., Deco, G., and McIntosh, A.R. (2010). Towards the virtual brain: network modeling of the intact and the damaged brain. *Arch. Ital. Biol.* 148, 189–205.
53. Munn, B.R., Müller, E.J., Wainstein, G., and Shine, J.M. (2021). The ascending arousal system shapes neural dynamics to mediate awareness of cognitive states. *Nat. Commun.* 12, 6016. <https://doi.org/10.1038/s41467-021-26268-x>.
54. Müller, E.J., Munn, B., Hearne, L.J., Smith, J.B., Fulcher, B., Amatičević, A., Lurie, D.J., Cocchi, L., and Shine, J.M. (2020). Core and matrix thalamic sub-populations relate to spatio-temporal cortical connectivity gradients. *NeuroImage* 222, 117224. <https://doi.org/10.1016/j.neuroimage.2020.117224>.
55. Munn, B., Zeater, N., Pietersen, A.N., Solomon, S.G., Cheong, S.K., Martin, P.R., and Gong, P. (2020). Fractal spike dynamics and neuronal coupling in the primate visual system. *J. Physiol.* 598, 1551–1571. <https://doi.org/10.1113/JP278935>.
56. Buzsáki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron* 80, 751–764. <https://doi.org/10.1016/j.neuron.2013.10.002>.
57. Manley, J., Lu, S., Barber, K., Demas, J., Kim, H., Meyer, D., Traub, F.M., and Vaziri, A. (2024). Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number. *Neuron* 112, 1694–1709.e5. <https://doi.org/10.1016/j.neuron.2024.02.011>.
58. Lake, E.M.R., Ge, X., Shen, X., Herman, P., Hyder, F., Cardin, J.A., Higley, M.J., Scheinost, D., Papademetris, X., Crair, M.C., et al. (2020). Simultaneous cortex-wide fluorescence Ca²⁺ imaging and whole-brain fMRI. *Nat. Methods* 17, 1262–1271. <https://doi.org/10.1038/s41592-020-00984-6>.
59. Munn, B.R., Müller, E.J., Medel, V., Naismith, S.L., Lizier, J.T., Sanders, R.D., and Shine, J.M. (2023). Neuronal connected burst cascades bridge macroscale adaptive signatures across arousal states. *Nat. Commun.* 14, 6846. <https://doi.org/10.1038/s41467-023-42465-2>.
60. Shine, J.M., Lewis, L.D., Garrett, D.D., and Hwang, K. (2023). The impact of the human thalamus on brain-wide information processing. *Nat. Rev. Neurosci.* 24, 416–430. <https://doi.org/10.1038/s41583-023-00701-0>.
61. Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B.L., Koay, S.A., Taxis, J., Najafi, F., Gauthier, J.L., Zhou, P., et al. (2019). CalmAn an open source tool for scalable calcium imaging data analysis. *eLife* 8, e38173. <https://doi.org/10.7554/eLife.38173>.
62. Ercsey-Ravasz, M., Markov, N.T., Lamy, C., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., and Kennedy, H. (2013). A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron* 80, 184–197. <https://doi.org/10.1016/j.neuron.2013.07.036>.
63. Vértés, P.E., Alexander-Bloch, A.F., Gogtay, N., Giedd, J.N., Rapoport, J.L., and Bullmore, E.T. (2012). Simple models of human brain functional networks. *Proc. Natl. Acad. Sci. USA* 109, 5868–5873. <https://doi.org/10.1073/pnas.1111738109>.
64. Henderson, J.A., and Robinson, P.A. (2013). Using geometry to uncover relationships between isotropy, homogeneity, and modularity in cortical connectivity. *Brain Connect.* 3, 423–437. <https://doi.org/10.1089/brain.2013.0151>.
65. Chung, F., Lu, L., and Vu, V. (2003). Spectra of random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA* 100, 6313–6318. <https://doi.org/10.1073/pnas.0937490100>.
66. Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., et al. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* 17, 1661–1663. <https://doi.org/10.1038/nn.3862>.
67. Cover, T.M., and Thomas, J.A. (1991). Entropy, relative entropy and mutual information. *Elem. Inf. Theor.* 2, 12–13.
68. Ince, R.A.A. (2017). The Partial Entropy Decomposition: Decomposing Multivariate Entropy and Mutual Information via Pointwise Common Surprisal. <https://doi.org/10.48550/ARXIV.1702.01591>.
69. Finn, C., and Lizier, J.T. (2020). Generalised measures of multivariate information content. *Entropy (Basel)* 22, 216. <https://doi.org/10.3390/e22020216>.
70. Crofts, J.J., and Higham, D.J. (2009). A weighted communicability measure applied to complex brain networks. *J. R. Soc. Interface* 6, 411–414. <https://doi.org/10.1098/rsif.2008.0484>.
71. Gautam, S.H., Hoang, T.T., McClanahan, K., Grady, S.K., and Shew, W.L. (2015). Maximizing sensory dynamic range by tuning the cortical state to criticality. *PLoS Comput. Biol.* 11, e1004576. <https://doi.org/10.1371/journal.pcbi.1004576>.
72. Shew, W.L., Yang, H., Petermann, T., Roy, R., and Plenz, D. (2009). Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. *J. Neurosci.* 29, 15595–15600. <https://doi.org/10.1523/JNEUROSCI.3864-09.2009>.
73. Gollo, L.L. (2017). Coexistence of critical sensitivity and subcritical specificity can yield optimal population coding. *J. R. Soc. Interface* 14, 20170207. <https://doi.org/10.1098/rsif.2017.0207>.
74. Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25. <https://doi.org/10.1002/hbm.1058>.
75. Légaré, A., Lemieux, M., Desrosiers, P., and De Koninck, P. (2023). Zebrafish brain atlases: a collective effort for a tiny vertebrate brain. *Neurophotonics* 10, 044409. <https://doi.org/10.1117/1.NPh.10.4.044409>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
HuC:H2B-GCaMP6f	Zfin	ZDB-ALT-150916-4
Deposited data		
Zebrafish	Current study	https://doi.org/10.5281/zenodo.13283803
Zebrafish (phototactic swimming)	Chen et al. ¹⁸	https://doi.org/10.25378/janelia.7272617
Nematode	Kato et al. ¹⁹	https://doi.org/10.1016/j.cell.2015.09.034
Fruit fly	Pacheco et al. ²⁰	https://doi.org/10.34770/gv6w-5351
Mouse	Stringer et al. ²¹	https://doi.org/10.1126/science.aav7893
Mouse (visual stimuli)	Stringer et al. ³⁹	https://janelia.figshare.com/articles/dataset/Recordings_of_20_000_neurons_from_V1_in_response_to_oriented_stimuli/8279387
Macaque (visual cortex)	Tang et al. ²²	https://doi.org/10.7554/eLife.33370
Macaque (motor cortex)	Trautmann et al. ²³	https://doi.org/10.1038/s41467-021-23884-5
Software and algorithms		
ICG algorithm and modeling	Current study	https://doi.org/10.5281/zenodo.13283803
Analysis and code	Current study	https://github.com/Bmunn/ICG
MATLAB r2023B	Mathworks	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

Zebrafish

The HuC:H2B-Gcamp6f zebrafish (n = 2) transgenic line was used for these experiments, targeting the calcium indicator GCaMP6f to the nuclei of all neurons. 6-day post-fertilisation larval zebrafish (*Danio rerio*) were set in a 2% low melting point agarose (Progen Biosciences), dorsal side up, and placed into the custom-made chamber. The 3D-printed 24mm² chamber had a square plastic base and a 0.2mm² post in each corner. 20mm² glass coverslips (ProSciTech) were fixed on each side using waterproof glue (Liquid Fusion Clear Urethane Adhesive). The glass coverslips enabled light-sheet illumination from the front and one side of the animal with minimal light distortions. Agarose-set fish were mounted onto the chamber platform with additional agarose to prevent the fish from moving throughout the experiment. Once the agarose had been set, the chamber was filled with E3 medium. Whole brain calcium fluorescence imaging was performed using a custom-built selective plane illumination microscope (SPIM) to determine the neural responses. The fish were simultaneously illuminated with two planes from the front and one side and imaged at 10μm increments in the dorsoventral axis with an exposure time of 10ms. This imaging produced a 25-slice volumetric representation of the entire brain, with a 4 Hz volumetric imaging rate. Images were captured, videos were cropped, and the transverse slices were segmented to identify ROIs corresponding to individual neurons. Each of the 25 planes was motion corrected using the NoRMCorre algorithm, and fluorescent traces generated by calcium transients in each ROI were extracted, demixed, and denoised using the *Calman* package.⁶¹ All Zebrafish housing, breeding, larval maintenance, and experiments were performed with approval from the University of Queensland Animal Ethics Committee. The remaining zebrafish utilised in this work were obtained from the open dataset.¹⁸ The zebrafish GCaMP6f spontaneous (n = 6) and a phototactic swimming behavioural recording are described in <https://doi.org/10.25378/janelia.7272617>. The zebrafish were imaged with a mean imaging rate of ~2.11 Hz.

C. elegans

The *C. elegans* GCaMP imaging data consists of previously recorded spontaneous activity (n = 5; full experimental detail¹⁹). Neurons detected in each worm span the head ganglia, motor/sensory neurons and interneurons with a mean imaging rate of ~2.9 Hz.¹⁹ The *C. elegans* multiscale connectivity analysis utilised the Cook et al.,²⁷ hermaphrodite gap-junction and chemical synapse connectivity. Briefly, this connectivity structure was generated using electron microscope reconstruction, which led to weighted connectivity based on the sizes of the synapses.

Drosophila

Two-photon imaging from fruit flies (*Drosophila melanogaster*) using GCaMP6s and spontaneous activity was analysed ($n = 33$); recordings are described in detail in Pacheco et al.²⁰ and were obtained from (<https://doi.org/10.34770/gv6w-5351>).

Mice

The spontaneous two-photon calcium recordings in mouse visual cortex using GCaMP6s in layers 2/3 excitatory neurons spontaneous ($n = 9$) recordings are described in detail in Stringer et al.²¹ and were obtained from (https://figshare.com/articles/dataset/Recordings_of_ten_thousand_neurons_in_visual_cortex_during_spontaneous_behaviors/6163622). Briefly, the recordings were performed using multiplane acquisition controlled by a resonance scanner at either 2.5 or 3 Hz and the mice were free to run on an air-floating ball. Spontaneous activity was recorded in darkness or with a grey background on surrounding monitors. Mouse visual stimuli (static square grating; $n = 3$) recordings are described in detail in Stringer et al.³⁹ and obtained from (https://figshare.com/articles/dataset/Recordings_of_20_000_neurons_from_V1_in_response_to_oriented_stimuli/8279387).

Macaque

Awake macaque ($n = 2$) two-photon calcium recordings from neurons expressing GCaMP5 in the visual cortex are described in Tang et al.²² available from (<https://github.com/leelabcnbc/sparse-coding-elife2018>). Monkeys performed a visual fixation evoked (natural images) task and spontaneous (blank screen) recordings. The recordings were made from a metal head post implanted over the dorsal portion of V1, a region representing the lower contralateral visual field at eccentricities of 2–5 deg. As these recordings did not possess continuous temporal activity (i.e., not permitting timescale analysis), we also analysed GCaMP6f recordings from the motor cortex of awake macaques ($n = 2$) that were obtained from (<https://datadryad.org/stash/dataset/doi:10.5061/dryad.cnp5hqc4k>).²³

METHOD DETAILS

Iterative Coarse Graining

We utilised an Iterative Coarse-Graining (ICG) approach to systematically cluster pairs of neurons to obtain reduced descriptions of neural activity that respected the discrete nature of individual neurons. Specifically, we clustered neurons using the similarity of their temporal activity $\{\rho_i\}$, calculated as the Pearson's product-moment correlation between their activity over time (temporal resolutions vary with recording),

$$C_{ij} = \frac{\langle \rho_i \rho_j \rangle - \langle \rho_i \rangle \langle \rho_j \rangle}{\sqrt{\langle \rho_i \rho_i \rangle - \langle \rho_i \rangle^2} \sqrt{\langle \rho_j \rho_j \rangle - \langle \rho_j \rangle^2}}.$$

Following previous approaches,¹⁶ we greedily paired neurons by the magnitude of their correlations with removal, i.e., pairing the most positively correlated neurons followed by subsequent pairings in descending order of their correlations, such that a given neuron i possesses a single pairing j . We then coarse-grained the data by summing activity to create a new variable,

$$\rho_i^2 = \rho_i^1 + \rho_j^1,$$

where ρ_i^1, ρ_j^1 represents the neuronal activity at level 1, the new variable ρ_i^2 is at a coarser resolution (level 2), and $i = 1, 2, \dots, \lfloor N/2 \rfloor$, where N is the total number of neurons to be iteratively coarse-grained. This process is iterated by calculating the correlations, C_{ij} , of the new variables, greedily clustering pairs, and repeating the coarse-graining,

$$\rho_i^l = \rho_i^{l-1} + \rho_j^{l-1},$$

where the variable ρ_i^l is the summed activity of $K = 2^{l-1}$ neurons at ICG level l . This approach pairs neurons and thus leads to increasingly larger clusters of size $K = 1, 2, 4, \dots, 2^{l-1}$. We confirmed that alternative approaches to measure similarity, such as the Spearman coefficient and the Jaccard index, lead to consistent pairings (Figure S6).

Neuronal network simulations

All networks were simulated using a stochastic spiking (cellular automata) neuronal model (Figure S7A) where a neuron, i , is in either one of three-states, s : quiescent (Q; $s_i = 0$), spiking or active (A; $s_i = 1$), or refractory (R; $s_i = -1$). The QAR stochastic model represents the basic neural model of spiking dynamics reducing biophysical complexity into three states.^{29,33} This stochastic cellular model has a rich history in physics where it is used to study universal critical dynamics (as universal dynamics are independent of model details).^{16,17,32} It is computationally efficient allowing the exploration of multiple topological families (connectivity matrix W_{ij}) within large-scale networks ($N = 4096$) to test the hypothesis that neuronal connectivity (structure) constrains multiscale activity (function).

In the model, a quiescent neuron, (Q, $s_i = 0$), will stochastically spike (A; $s_i = 0 \rightarrow 1$) due to either: a stochastic external stimulus modelled as a Poisson process with transition probability $P_{\text{spont}}(s_i = 0 \rightarrow 1)$ or evoked with probability $P_{\text{evoked}}(s_i = 0 \rightarrow 1) = W_{ij}$ following all presynaptic cells, j , that spiked in the previous time step. Following the Active state, a cell deterministically enters the Refractory period ($s_i = 1 \rightarrow -1$) where it remains for two Δt (so as to prevent self-reinforcing loops between mutually connected neurons) before returning to the Quiescent state ($s_i = -1 \rightarrow 0$). All neurons are updated synchronously in discrete

timesteps ($\Delta t = 33$ ms or a sampling frequency of $f_s = 30$ Hz) and simulated for 11 seconds, with the final 10 s of spiking activity analysed with ICG (discarding initial 1 s of transience). This model simulates N-methyl-D-aspartate activity, which are slow currents persisting for tens of milliseconds that can nonlinearly trigger cellular spiking (acting as a superposition of many integrate and fire cells)³⁰ and significantly contribute to calcium dynamics.³¹ The activity was simulated on both weighted topologies and discrete binary topologies. For all simulations, we constrained the network model to have a mean firing rate of $f_r \sim 7$ Hz by rescaling the probabilistic connectivity W_{ij} (modifying P_{evoked}) to preserve the topological structure, however, in simulations with sparse connectivities (e.g., modular, hubless degree distributed, sparse fractal), we increased stochastic baseline drive, P_{spont} , which is a rescaling of the external Poisson drive, f_{spont} , as $P_{\text{spont}} = 1 - \exp(-f_{\text{spont}}\Delta t)$. It should be noted that other normalisation techniques may be adopted (such as preserving degree or random pruning of connections); however, we opted to preserve mean spiking dynamics as this approach minimally altered the features of the topologies that we were specifically interested in exploring. We verified that pruning and randomly adding fixed strength connections with a fixed drive does not alter the qualitative findings. Nevertheless, a key limitation of this simplified model is that potentially integral biophysical constraints are not explored (e.g., the balance of excitation and inhibition as all cells are excitatory or spike-adaptation as cells possess a fixed refractory period). However, our findings demonstrate that these are not essential for reproducing the universal scaling properties observed in the simplified model and empirical data.

Topological families

We studied five generative network topological families, each with a unique control parameter, allowing a broad analysis of topology. We considered modular networks, random rewiring networks, geometric spatially connected networks, scale-free degree distributed networks, and hierarchical-modular networks. Each distinct model parameter was simulated with 100 repetitions to determine if there was an overlap with the experimentally observed scaling. The topologies were simulated on networks of 4096 nodes, balancing sufficient iterations of ICG (over 12 levels) with computational limitations.

Modular networks

These networks were created by defining internally connected modules with a characteristic size m , and we present results for $m = 2, 4, \dots, 512$. $m = 2$ implies that only pairs of neurons are connected, and the network contains 2048 modules, whereas $m = 512$ implies that the network consists of four modules. Connectivity was generated stochastically, with 95% connectivity within modules and 5% outside. Activity simulated on this network topology possessed a kink in the scaling measures at the module size (i.e., the privileged scale K_p); we fitted the scaling exponents to scales below the module size (Figure S8A).

Random rewiring networks

These networks were created by randomly rewiring local connections within a 4096-node ring topology with mean degree = 4 and rewiring probability λ , which we varied in steps of 0.01 between 0 to 1. This shifts the network from an invariant ring topology ($\lambda = 0$) to a random topology ($\lambda = 1$; i.e., an Erdos-Renyi network)³⁴ a rewiring probability of $\lambda \sim 0.2$ corresponds to a small-world network (Figure S8B).

Geometric stochastic distance dependent networks

These networks were created by stochastically connecting nodes on a grid (distance of 2 between adjacent nodes and $2\sqrt{2}$ between diagonally adjacent) with a distance-dependent exponential probability, $P(x) \propto e^{-\alpha x}$. We presented results for $d = 0.05$ to 1.6 in intervals of 0.05. This shifted the network from local connections ($d = 1.6$; i.e., a fast drop-off in connectivity) to long-range connections ($d = 0.05$; i.e., a slow drop-off in connectivity). Interestingly, we observed a smooth transition in the initial scaling from $\alpha = 2$ for $d = 0.05$ to $\alpha = 2$ for $d = 1.6$, before evidence of a characteristic scale (kink in the variance) for larger clustering (Figure S8C; due to the characteristic scale imposed by the range of the exponential decay). Nevertheless, this initial scaling, hinted that a distant dependent exponent may be able to map between dynamical and structural scaling. Furthermore, the early match between this model and neuronal activity is promising, given the model's important role of spatially embedded networks to generate multiscale dynamics while minimising wiring costs.^{62–64}

Scale-free degree-distributed networks

These networks were created by generating a scale-free degree, d_g , distribution with scaling exponent γ , $P(d_g) \propto d_g^{-\gamma}$, sampling 1,024 positive integers ($0 < d_g \leq 1,024$), then generating the network using the Chung-Lu model.⁶⁵ We presented results for $\gamma = 1.5$ to 4.5 in steps of 0.1. This shifts the degree distribution from a heavy-tailed ($\gamma = 1.5$) to a rapidly decaying ($\gamma = 4.5$) short-tailed distribution (Figure S8D). Of note, the Barabasi-Albert model of preferential attachment⁷ possesses a scaling exponent of $\gamma = 3$.

Hierarchical-modular networks

These networks were created by defining a modular pair of connected nodes connected to other pairs with a hierarchical connection probability following $P_l = 1/\Phi^l$, where P_l is the probability of a connection between different clusters at level l and Φ determines the scaling.^{33,34} At the beginning ($l = 0$), distinct pairs of nodes are connected, and then two groups are interconnected with a connection density $P_l = 1/\Phi^l$. The connection density of these nodes forms the connectivity matrix, W_{ij} , referring to the probabilistic likelihood of a connection between these node groups. This process is repeated for increasing l and dyadically increasing group size. The value of Φ controls the hierarchical gradient. A value of $\Phi \sim 1$ yields dense connectivity (i.e., all node-pairs are strongly connected across all levels), whereas $\Phi > 5$ yields sparse connectivity as the probability of connecting node-pairs decreases exponentially across levels. We presented results for $\Phi = 1.25$ to 3.75 in steps of 0.05. Simulations were run by converting the weighted probabilities into discrete and directed connections (randomised across iterations); nevertheless, we confirmed that the simulated results are robust when simulated using weighted connections.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical properties of the reduced variables at each resolution, l , of the system and their flow across scales are informative of the underlying mechanisms. Statistical tests, error bar types and sample sizes are stated in the results and figure/table legends.

Kurtotic

We calculated the kurtosis of the pairwise correlation distribution, C_{ij} , at each scale, l , as:

$$\left(\frac{\mu_4}{\sigma^4}\right)_l = \frac{\langle (C_{ij} - \langle C_{ij} \rangle)^4 \rangle}{\left(\langle (C_{ij})^2 \rangle - \langle C_{ij} \rangle^2\right)^2},$$

where μ_4 is the fourth central moment, the kurtosis of a normal distribution is 3, and a heavy-tailed distribution has a kurtosis exceeding 3 (i.e., leptokurtic).

Static

We calculated the mean-variance of the coarse-grained variables at each scale l by,

$$\sigma_l^2 = \frac{1}{N_l} \sum_{i=1}^{N_l} \left(\langle (\rho_i^l)^2 \rangle - \langle \rho_i^l \rangle^2 \right),$$

where we calculate the mean-variance across N_l clusters where $N_l = \lfloor \frac{N}{2^{l-1}} \rfloor$. The way the variance scales across resolutions is informative of the underlying organisation of the system. If the coarse-graining procedure summed independent variables then $\sigma_l^2 \propto K^1$, while if all variables were perfectly correlated then $\sigma_l^2 \propto K^2$. Thus, scaling following $\sigma_l^2 \propto K^\alpha$ with $1 < \alpha < 2$ is nontrivial and suggests the correlations possess an underlying self-similar coordination.

Dynamic

We calculated the mean-timescale of the multiscale neural activities, quantifying the temporal similarity of a signal via the mean-auto-correlation curve.⁶⁶ We first calculated the mean-normalised autocorrelation, $A_l(t')$ as

$$A_l(t') = \frac{1}{N_l} \sum_{i=1}^{N_l} \left(\frac{\langle \rho_i^l(t) \rho_i^l(t+t') \rangle_{t'} - \langle \rho_i^l \rangle^2}{\langle (\rho_i^l)^2 \rangle - \langle \rho_i^l \rangle^2} \right)$$

for zero and positive lags, t' and at each scale l . We then calculated the timescale as the area under the curve (integral) estimated using the trapezoidal method $\tau_l = \int_0^{t_{\max}} A_l(t) dt$ (making τ_l in units of time [s]), and we present results for a maximum lag of 3 seconds.

We found that the results are robust for a maximum lag between 0.5 s to 30 s. We then tested for scaling, $\tau_l \propto K^\beta$. Coarse-graining two perfectly correlated signals with the same timescale or two uncorrelated signals will lead to scaling with $\beta = 0$. Hence, an exponent of $\beta > 0$ suggests nontrivial temporal scaling.

Estimation of scaling exponent

To assess for self-similarity in the statistical properties, we fit a power-law to the static and dynamic measures across levels of l . Specifically, the power-law was fit from the second to the penultimate level (i.e., across $l = 2$ to $l_{\max} - 1$). The slope fit and uncertainty were computed by linear regression in log-log space from the logarithm of the statistical measures and the logarithm of K across the points $K = 2, 4, \dots, 2^{l_{\max}-1}$ which are evenly spaced in log-log coordinates. The quality of fit to a power-law was assessed by the correlation coefficient (r) between the logarithm of the statistical measures and the logarithm of K , where $r = 1$ indicates a power-law and $r < 0.8$ indicates a departure from scale-free scaling such as a privileged scale if accompanied with a characteristic kink. Note that for the most extensive dataset used in this work, the power-law fits span ~ 5 decades, from single neuron to whole-brain zebrafish ($10^5 = 100,000$ neurons over $k = 16$ dyadic scales).

Redundancy estimation

We leveraged insights from information theory to calculate the redundant information between neuronal activity. In particular, a signal's entropy (such as neuronal or ensemble activity) can be derived from first principles as a pointwise 'local' quantity, the Shannon information content of individual neuronal (or ensemble) instantaneous activations rather than entire sequences of activity.⁶⁷ The Shannon information content (or pointwise local entropy) of a signal X with distinct states $x \in X$ for state x is $h_x = -\log P(x)$, which quantifies the information content of a single event x , given the states probability $P(x)$. Taking the expectation of this measure over all events recovers the global entropy, $H_X = \langle h_x \rangle = \sum_{x \in X} P(x) h_x$. Similarly, this can be extended to two signals, X & Y , with states x & y ,

as $h_{x,y} = -\log P(x,y)$. We can then apply the formalism of a joint entropy decomposition into components that are redundant, R , unique, U , and synergistic, S ,⁶⁸ defined as

$$H_{X,Y} = U_X + U_Y + R + S,$$

$$H_X = U_X + R, \&$$

$$H_Y = U_Y + R.$$

From the local measure of entropy, we can calculate an estimate of the local redundant information, $R = \langle \min\{h_x(t), h_y(t)\} \rangle_t$, which is the minimum information or specificity that any two neurons (or ensemble of neurons) can jointly provide, averaged across the recording.^{24,69} This estimate captures the information-theoretic basis that redundant information is that which can be extracted at minimum from either of two neural signals and it is nonnegative. In the manuscript, we present the net redundancy, $\hat{R}$ calculated as the redundant information exceeding the synergistic information, $\hat{R} = R - S$, that is, the extent to which joint activity is coordinated and does not provide extra complementary information. Moreover, to improve probability estimation, we analyse the ICG activity, generated from initially binarised cellular activity (ρ_i^1). We obtain binarised activity from the non-zero/baseline timesteps in recordings where a baseline is present (e.g., zebrafish and mice) or via thresholding timesteps that exceed 1.96σ (i.e., most extreme 2.5% of states) in continuous activity. The scale-dependent average net redundancy, $\hat{R}_l$, is either calculated over all pairwise variables, ρ_i^l , or estimated from 10^5 samplings in large recordings (when the number of variables > 512). We present up to the antepenultimate level (i.e., third from last). As the information-theoretic measure is biased by the finite properties of the recording (i.e., recording duration and resolution) and of the variables ρ_i^l (specifically their range and distribution of values) we contrasted the empirical net redundancy against data-matched null models. The data-matched nulls were generated by repeating the analysis on white-noise thresholded activity (i.e., independent signals) with the same duration (100 repeated trials across species). In this way, we can establish significant increases in information theoretic redundancy beyond that expected from the iterative coarse-graining procedure or empirical limitations.

Limitations of the ICG algorithm

Pearson correlation

Our algorithm utilised a greedy approach based on dynamical similarity measured using the Pearson correlation of neuronal activity. As such, we are limited in the ability of the Pearson correlation to detect similarity, constrained to whether there is a linear association and lacking a causal (or directional) relationship. Furthermore, the greedy approach meant we were limited to pairwise clustering at each resolution. While both of these limitations were considered justified given that the computational efficiencies benefited from the fast calculations of pairwise correlations and greedy sorting (allowing the analysis of up to 10^5 neurons). We also explored alternative similarity measures using the Jaccard index or Spearman correlation and found consistent scaling of our dynamics (Figure S6).

Pairwise clustering: Dyads to triads

Another primary reason for pairwise clustering, instead of triadic clustering, is the nontriviality in defining a triple correlation. Nevertheless, we explored a modified triadic ICG algorithm where we sorted for the triplets of neurons with the triple correlation defined as:

$$\text{corr}(A, B, C) = \text{corr}(A, B)\text{corr}(A, C)\text{corr}(B, C),$$

This definition is imperfect, as a single correlation near zero would significantly diminish the overall correlations, or two negative correlations would lead to an overall positive correlation. Furthermore, even this simple extension rapidly blows up the computation, taking the length of the list to be sorted from $NP = (N^2 - N)/2$ where N is the number of neurons and NP is the number of unique neuronal pairings to NP choose three or $\frac{NP!}{(NP-3)!3!}$. Nevertheless, utilising this algorithm on the largest sample size computationally feasible, we observe consistent static scaling between dyadic (black dots; Figure S6) and triadic (red dots; Figure S6) ICG after 2 iterations of triadic coarse-graining. We expect this relationship to extend to larger clustering sizes.

Statistical null models

Preserved scaling and artificial privileged scale data-driven null models

To ascertain whether the behaviour of these statistical properties was nontrivial and if the ICG algorithm was sensitive to signatures of a privileged scale, we generated null models from the data itself, preserving the covariance or temporal dynamics of all neurons. In one null model, we simulated white-noise activity with a matched covariance structure and shuffled external correlations outside paired ensembles. This null recapitulated the static scaling up to an artificial privileged scale (Figure S3A). In a further null model, we randomised the pairwise correlations by performing random circular shifts (temporal rotations) of individual neuronal activity within ensembles of neurons. These shifts rendered each cell (or ensembles of cells) independent while preserving temporal autocorrelation. This null model destroyed scaling beyond a privileged scale (Figure S3B).

Independent vs perfectly correlated null models

To demonstrate the theoretical scaling of independent and perfectly correlated neurons, we simulated stochastic Poisson spikes with independent rates (Figure S4A) and spikes from an inhomogeneous Poisson process sampling a global signal between neurons (i.e., perfectly correlated rate but stochastic sampling) at low and intermediate spike rates (Figure S4B).

Communicability

Communicability, G , is a network measure that captures how efficiently information may pass between all network nodes by pathways of all lengths (including shortest-path and higher-order connections that may be indirect between nodes).⁷⁰ The communicability, G_{ij} , between nodes i and j in a weighted network is given by:

$$G_{ij} = \left(e^{D^{-\frac{1}{2}} W_{ij} D^{-\frac{1}{2}}} \right)_{ij},$$

where the $m \times m$ connectivity matrix, W_{ij} , is normalised by $D = \text{diag}(d_i)$ via the product $\sqrt{d_i d_j}$ where $d_i = \sum_{k=1}^m W_{ik}$ is the generalized weighted degree of node i . We present the summary statistic of the mean-communicability of the network, $G = \langle G_{ij} \rangle$.

Dynamic range

To probe the information processing properties of the model given stimuli, we calculated the dynamic range, Δ_s , from the range of discernible responses to the range of stimulatory input. We calculated the response function, F , as $F = \sum_{t=1}^T \rho(t)$, that is the spiking activity generated over T , where $T = 1$ s (results are robust for varying $T = 100$ ms to $T = 2$ s), in response to a transient stimulus ($T/2$ s) of strength S , ranging from $S = 10^{-5}$ to $S = 10^{0.5}$ m s⁻¹, where the stimulus is modelled as afferent spikes generated as a Poisson process to each neuron at a stimulus rate S . Finally, F was averaged across 100 trials for each stimulus intensity. After calculating the average elicited response, F , the dynamic range, $\Delta = 10 \log_{10} \frac{S_{0.9}}{S_{0.1}}$, was calculated as the stimulus range (in dB) where variations in S can be robustly encoded by variations in F , after discarding elicited responses that are too small to distinguish from baseline, F_0 , or network saturation, F_{max} .²⁹ The stimulus range $[S_{0.1}, S_{0.9}]$ is calculated from the elicited response range $[F_{0.1}, F_{0.9}]$, where $F_x = F_0 + x(F_{max} - F_0)$ which is the standard range reported.^{29,71–73}

Adjusted Mutual Information

The adjusted mutual information quantifies the similarity between multiple partitioning of a set while correcting for the expected increase in mutual information for larger clusterings due to chance.⁴⁰ Given a set S of N data points, partitioned into two distinct clusterings, $U = \{U_1, U_2, \dots, U_R\}$ with R clusters, or $V = \{V_1, V_2, \dots, V_C\}$ with C clusters, where $\bigcup_{i=1}^R U_i = \bigcup_{j=1}^C V_j = S$ and $U_i \cap U_j = V_i \cap V_j = \emptyset$ for $i \neq j$. The information about their overlap can be summarised from the $R \times C$ contingency table of elements $[n_{ij}]_{i=1, \dots, R}^{j=1, \dots, C} = |U_i \cap V_j|$, where n_{ij} denotes the number of objects that are common to clusters U_i and V_j , and the partial sums of the table are $a_i = \sum_{j=1}^C n_{ij}$ and $b_j = \sum_{i=1}^R n_{ij}$. The entropy, H , associated with the partitioning U is $H(U) = - \sum_{i=1}^R \frac{a_i}{N} \log \frac{a_i}{N}$ where $\frac{a_i}{N}$ denotes the probability that a data point of S belongs to cluster U_i (and v.v. for $H(V)$) and joint entropy $H(U, V) = - \sum_{i=1}^R \sum_{j=1}^C \frac{n_{ij}}{N} \log \frac{n_{ij}}{N}$. The mutual information, MI , between the partitions is:

$$MI(U, V) = H(U) + H(V) - H(U, V),$$

where $MI(U, V)$ measures the reduction in uncertainty knowing a clustering tells us about the other, note that $0 < MI(U, V) \leq \max\{H(U), H(V)\}$. The adjusted mutual information corrects the MI for the similarity between random clustering by removing the baseline expected mutual information, $\langle MI(U, V) \rangle$, under a generalised hypergeometric distribution model of randomness, which holds that the expected value of $MI(U, V)$ due to finite size effects and chance non-uniform joint distributions is,

$$\langle MI(U, V) \rangle = \sum_{i=1}^R \sum_{j=1}^C \sum_{n_{ij}=\max(0, a_i+b_j-N)}^{\min(a_i, b_j)} \frac{n_{ij}}{N} \log \left(\frac{N n_{ij}}{a_i b_j} \right) \times \frac{a_i! b_j! (N - a_i)! (N - b_j)!}{N! n_{ij}! (a_i - n_{ij})! (b_j - n_{ij})! (N - a_i - b_j + n_{ij})!}.$$

This relationship leads to the adjusted mutual information, $AMI(U, V)$, which is calculated as,

$$AMI(U, V) = \frac{MI(U, V) - \langle MI(U, V) \rangle}{\max(H(U), H(V)) - \langle MI(U, V) \rangle},$$

which is upper-bounded by 1, for identical partitions and equals 0, when the unadjusted MI equals its expected value (i.e., it is normalised in a stochastic sense).

We also calculated the normalised behavioral reconfigurations across states, ΔAMI , as,

$$\Delta AMI = (1 - \langle AMI_b \rangle) - (1 - \langle AMI_w \rangle),$$

$$\Delta AMI = \langle AMI_w \rangle - \langle AMI_b \rangle,$$

which is the difference of reconfigurations between behavioural states, $1 - \langle AMI_b \rangle$, and within behavioural states, $1 - \langle AMI_w \rangle$. We demonstrate results where, AMI_w and AMI_b are calculated using ICG clustering's calculated on the first and second halves of recordings, across different behavioural states.

General Linear Model

A mass-univariate general linear model was constructed to predict ensemble activity (at each level, l) during a zebrafish phototactic fictive swimming behaviour as a function of left vs. right visual signals convolved with a delayed exponential kernel (half-time = 0.4 s; peak delay 0.08 s) to mimic the temporal signature of the GCaMP6f indicator (as in¹⁸). We generated a null distribution from the regression coefficients of 5,000 iterations of a circularly shifted dataset to estimate significance non-parametrically.⁷⁴ Significant results ($p < 0.002$) were plotted onto a tree denoting each level and ensemble defined using ICG on the task time series and anatomical templates warped to the Z-brain atlas.⁷⁵

Supplemental figures

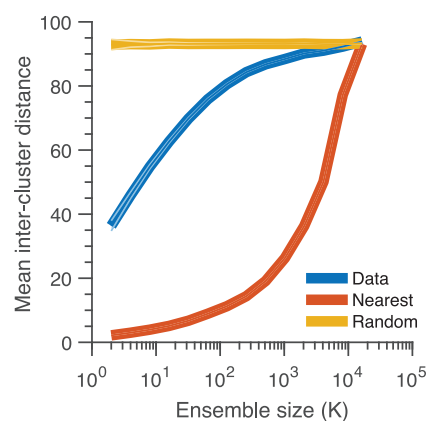


Figure S1. Contrasting functional, random, and geometric iterative coarse-graining, related to Figure 1

Mean inter-cluster distance for ICG ensemble size K (blue) and when clustering by spatial proximity (orange) or randomly (yellow). Error bars represent 95% CI. This result indicates that pairings display far larger spatial separations than expected for a geometric pairing.

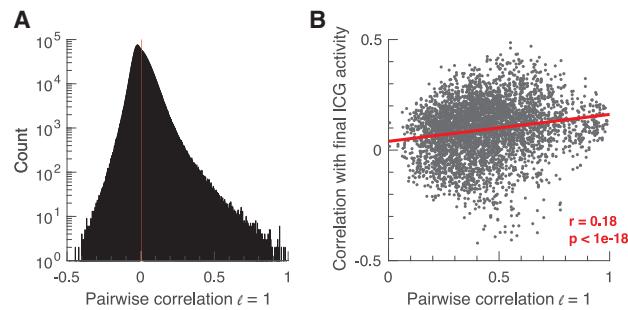


Figure S2. Heavy-tailed correlations are preserved across coarse-graining, related to Figure 2

Left: Pairwise correlation distribution of neurons from a zebrafish recording display a tight mean around zero with a heavy tail (note logarithmic y-axis). The red line indicates the smallest correlation between paired neurons following the ICG greedy algorithm. Right: Individual neuronal correlations for initial ICG pairings (at $l = 1$) vs. the correlation between the cell and the global activity (i.e., the final ICG activity $l = 15$) suggest that the extreme cellular activity is preserved up to the whole-brain scale.

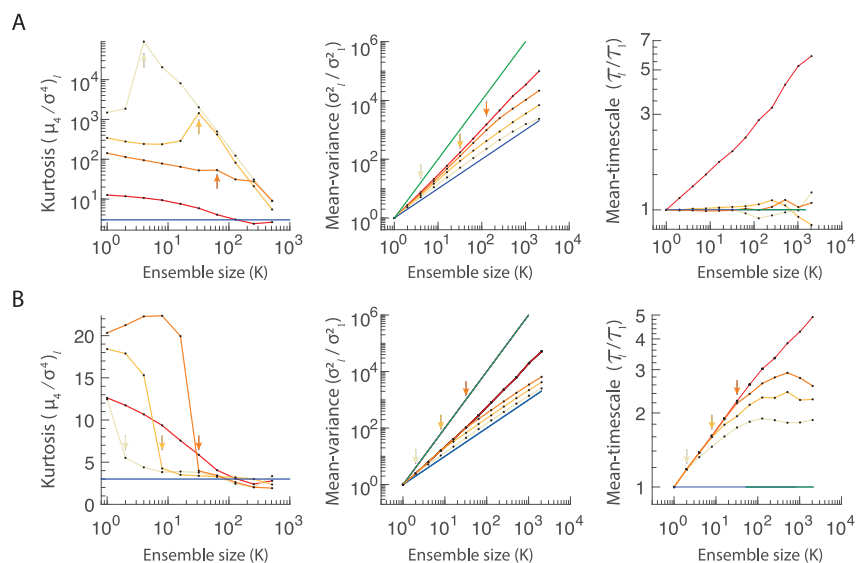


Figure S3. Comparison of ICG with scale-dependent data-matched null models, related to Figure 2

(A) Empirical activity (red) and artificial scale-dependent activity; curved lines examples for ICG ensemble sizes of $K = 4$ (cream), 32 (yellow), 128 (orange). Activity is simulated atop empirically matched covariance structure lacking correlations outside of artificial scale-dependent ensemble size. Left: Kurtosis displays kinks at artificial scale and is significantly increased with increasing removal of higher-order interactions. Middle: Variance scales up to artificial scale dependence before scaling with a slope of unity corresponding to independent signals. Right: The artificially generated signals do not display dynamic scaling.

(B) We simulated null models by randomly temporally permuting (circularly shifting) empirical data within their ICG ensembles of the artificial scale dependence — i.e., the neurons are temporally coordinated within ensembles and match empirical data, yet independent outside the ensembles; examples shown for circular shifting within ensembles of size $K = 2$ (cream), 8 (yellow), 32 (orange). This data-matched null model displays a large deviation in kurtosis beyond the artificial scale (left) and recapitulates the static (middle) and dynamic (right) empirical scaling up to the artificially imposed scale dependency.

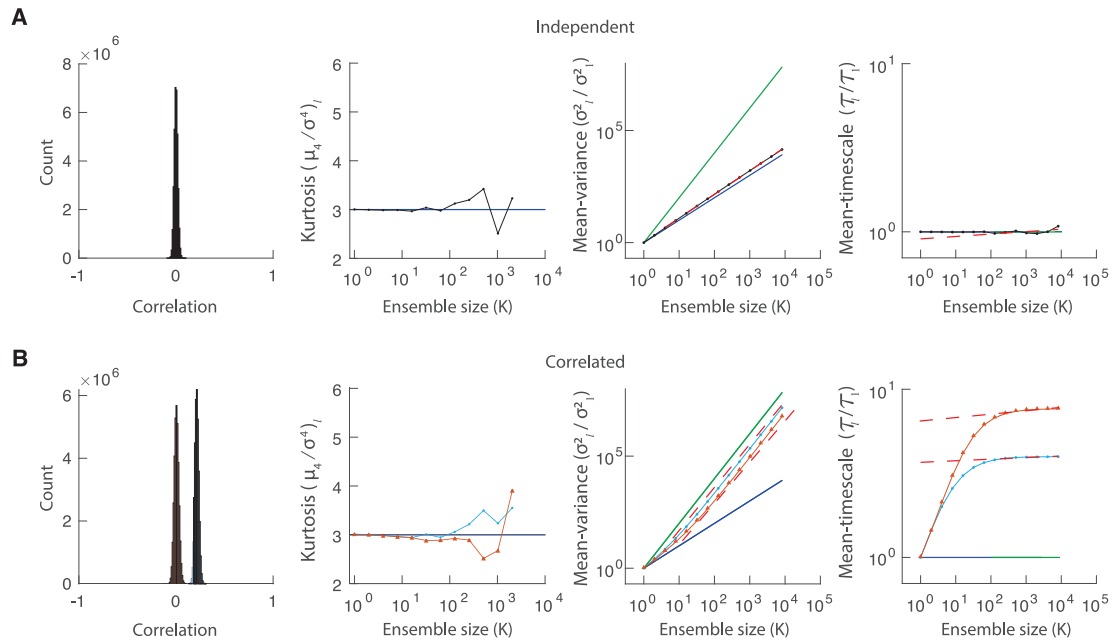


Figure S4. Comparison of ICG with Poisson null models, related to Figure 2

(A) For independent Poisson-generated spikes, the pairwise correlation distribution is tightly distributed around zero (first column) with Gaussian kurtosis (second column), linear static scaling exponent $\alpha = 1$ (third column), and constant dynamic exponent $\beta = 0$ (fourth column), as theoretically predicted.

(B) Poisson spikes generated with perfectly correlated nonstationary fluctuations (final coarse-grained activity of an example zebrafish) display a pairwise correlation distribution approaching zero, with increasing sparse sampling of global signal (first column) increasing mean pairwise correlation with mean Poisson firing rate 3.5 (orange) vs. 7 Hz (blue). Small noise fluctuations lead to Gaussian kurtosis (second column), while mean-variance $\alpha = 2$ (third column) and mean-timescale $\beta = 0$ (fourth column) scale as theoretically expected (3.5 Hz triangles, orange; 7 Hz circles, blue). In the dynamic scaling, an initial transient is broader in sparse samplings.

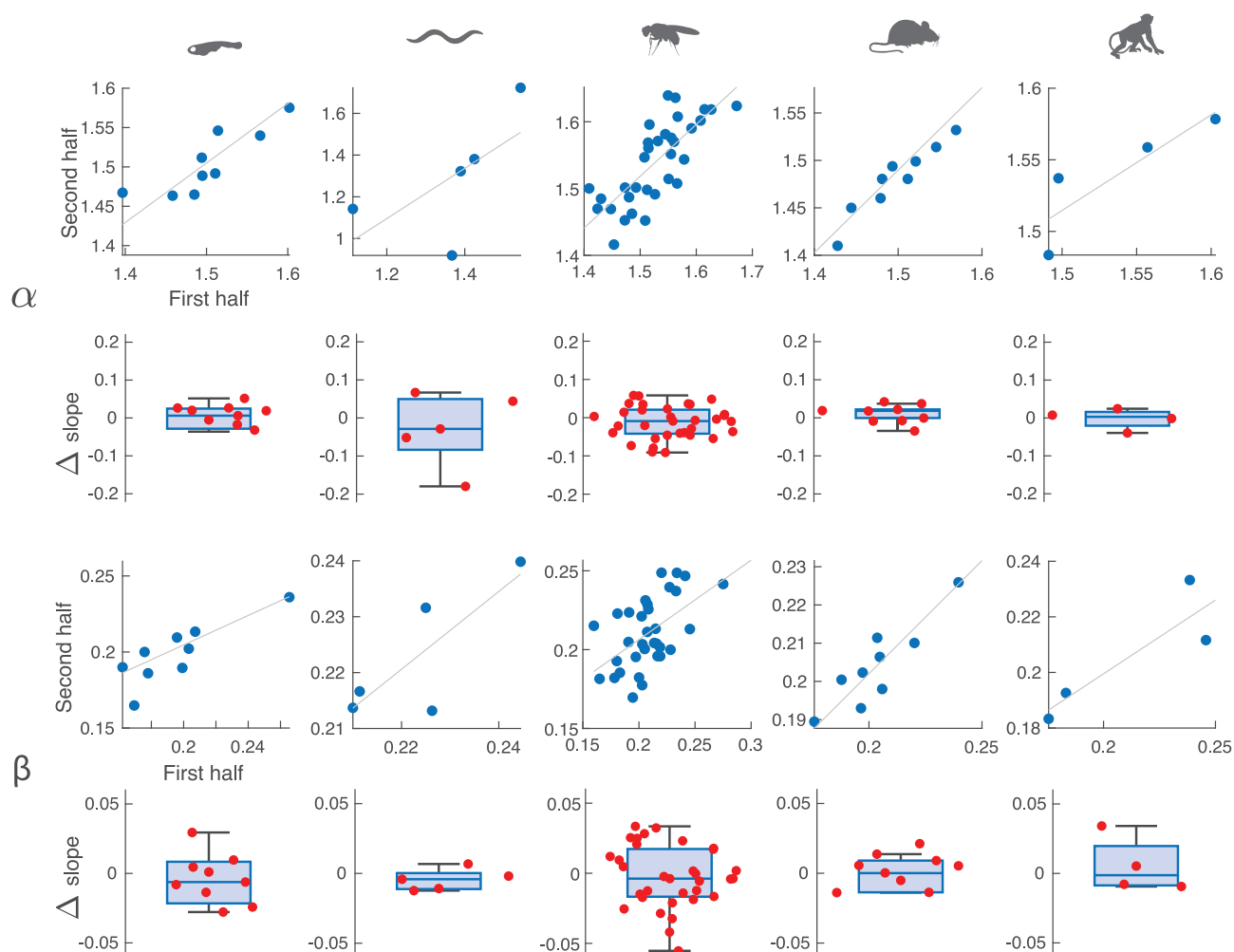


Figure S5. Cross-validation of scaling relationships within a recording, related to Figure 2

Mean-variance (top) and mean-timescale (bottom) scaling exponents calculated in the first half and second half of each recording. The static and dynamic scaling differences (second-first) are distributed around zero, indicating an intra-difference of $\alpha \pm 0.15$ and $\beta \pm 0.05$. Thus, the scaling is consistent within recordings of the same animal, between animals of the same species, and between species.

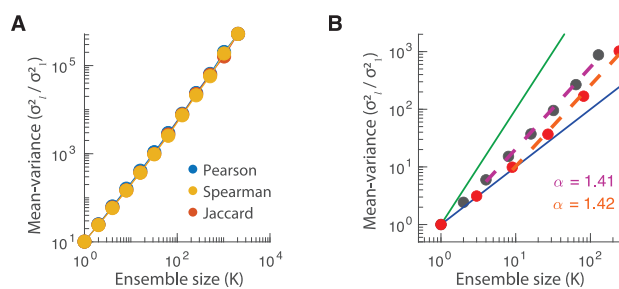


Figure S6. Comparison of ICG scaling using alternative similarity measures and triadic clustering, related to STAR Methods

(A) The self-similar static scaling is consistent across all three similarity measures (blue, Pearson correlation; orange, Jaccard index and Pearson correlation; and yellow, Spearman).

(B) The dyadic (black) and triadic (red) aggregating procedures produce similar scaling (dashed lines represent scale-free slope estimate purple-dyadic and orange-triadic).

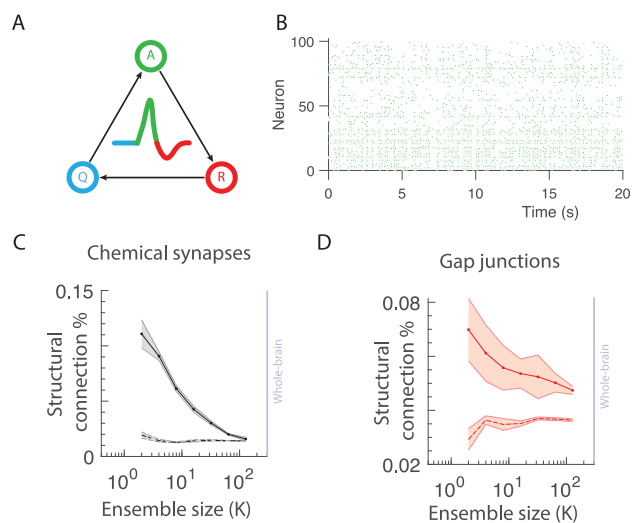


Figure S7. Three-state cellular model atop *C. elegans* structural connectome, related to Figure 3

(A) QAR model recapitulates the three stages of an action potential (spike): quiescence (able to spike), active (spiking), and refractory period (unable to spike). (B–D) Example spiking activity (B) from the model simulated atop the *C. elegans* chemical synapse (examples of simultaneous synchronous and asynchronous activity). Likelihood of simulated activity functional connections obtained using ICG possessing a physical (C) chemical synaptic or (D) gap junction connection. Error bars are $\mu \pm 95\%$ CI.

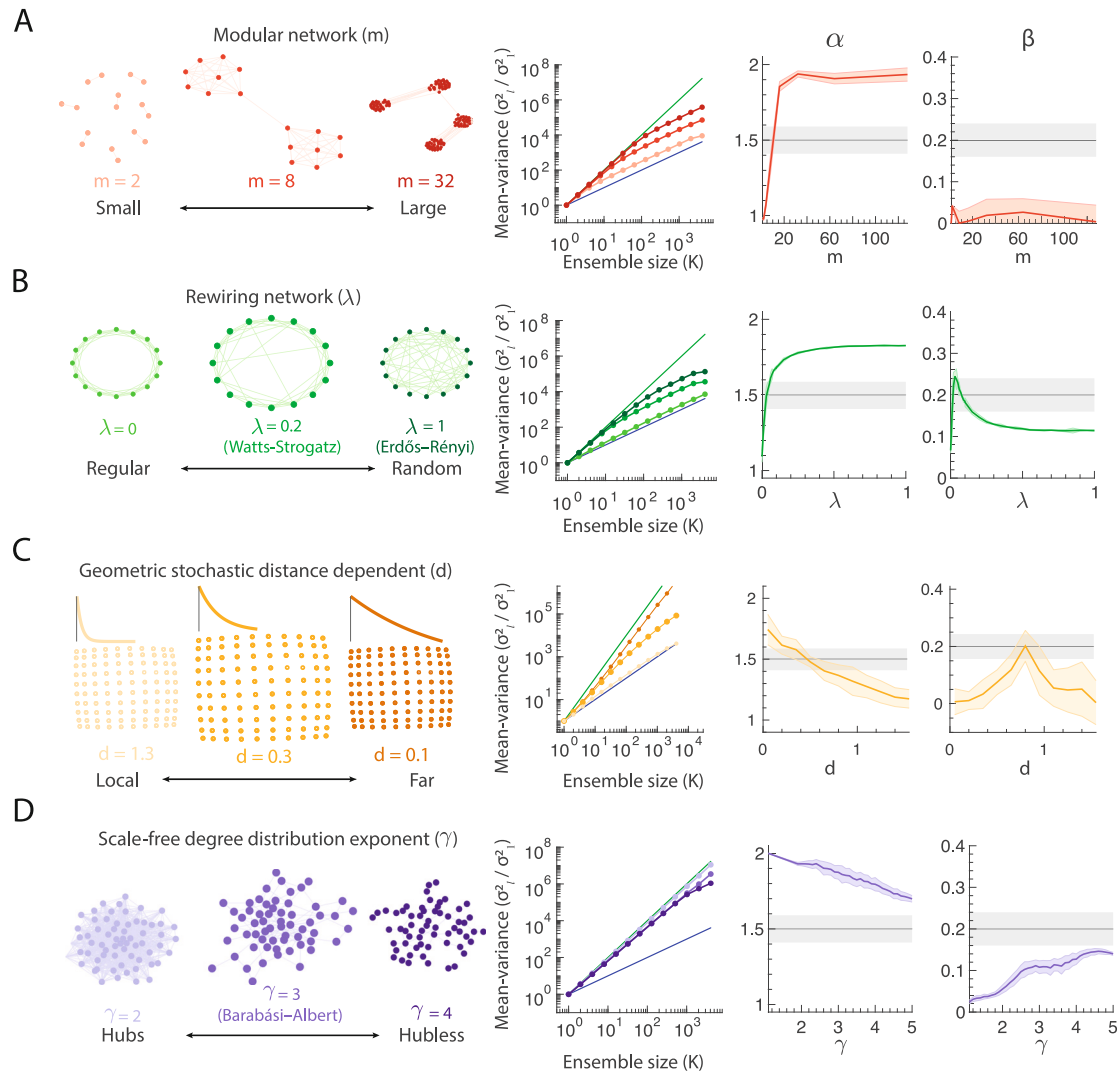


Figure S8. Simulated neuronal dynamics atop different topological families, related to STAR Methods

(A) Simulations on a modular network topology do not display power-law scaling, possessing a kink above the module size. The dominant scaling exponent for module size (m ; right $\mu \pm \sigma$ across 100 simulations, red-shaded bar) does not coincide with the empirically observed variance exponent (black line $\pm$ gray error bars).

(B) Rewiring topologies extending from regular (rewiring probability $\lambda = 0$) to random ($\lambda = 1$; Erdős-Rényi) via a Watts-Strogatz small-world connectivity ($\lambda = 0.2$) support self-similar scaling for minor rewirings that do not coincide with the empirically observed exponents.

(C) Geometric stochastic distance-dependent networks where spatially embedded neurons are stochastically connected proportional to a distance-dependent exponential distribution; the slope of this decay extends the networks from locally connected ($d = 1.3$) to broadly connected to spatially far neurons ($d = 0.1$). This network recapitulates the empirical scaling from efficient to robust, albeit for a transient regime, before scaling independently beyond a finite ensemble size proportional to the range of the geometric coupling.

(D) Scale-free degree distributed networks with degree exponent ($-\gamma$), including the Barabási-Albert network ($\gamma = 3$), support self-similar scaling but with strong correlations even following connectivity reweighting, which do not coincide with the empirically observed exponents.