bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Topological insights into the neural basis of flexible behavior

Tevin C. Rouse, Amy M. Ni, Chengcheng Huang, Marlene R. Cohen

September 24, 2021

1 Abstract

It is widely accepted that there is an inextricable link between neural computations, biological mechanisms, and behavior, but there exists no framework that can simultaneously explain all three. Here, we show that topological data analysis (TDA) provides that necessary bridge. We demonstrate that cognitive processes change the topological description of the shared activity of populations of visual neurons. These topological changes provide uniquely strong constraints on a mechanistic model, explain behavior, and, via a link with network control theory, reveal a tradeoff between improving sensitivity to subtle visual stimulus changes and increasing the chance that the subject will stray off task. These discoveries provide a blueprint for using TDA to uncover the biological and computational mechanisms by which cognition affects behavior in health and disease.

2 Main Text

Perhaps the most remarkable hallmark of the nervous system is its flexibility. Cognitive processes including visual attention have long been known to affect both behavior (e.g. performance on visual tasks) and virtually every measure of neural activity in visual cortex and beyond ((*1*), (*2*)). The diversity of changes associated with cognitive processes like attention makes it unsurprising that very simple, common measures of neural population activity provide limited accounts of how those neural changes affect behavior.

Arguably, the most promising simple link between sensory neurons and behavior is correlated variability (often quantified as noise or spike count correlations, or r_{SC}, which measure correlations between trial-to-trial fluctuations in the responses of a pair of neurons to repeated presentations of the same stimulus; (*3*)). Correlated variability in visual cortex is related to the anatomical and functional relationships between neurons ((*4*); (*3*)). We demonstrated previously that the magnitude of correlated variability predicts performance (see Fig. 1D) across experimental sessions and the animal's choices on individual trials ((*5*)). This early success explaining simple behaviors means that correlated variability is

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a foundation on which to build efforts to explain more complex aspects of flexible behavior and the concomitant neural computations.

However, our efforts to use correlated variability to explain a wider variety of sensory and cognitive phenomena and to constrain mechanistic models reveal a need for more sophisticated ways to characterize neuronal population activity. For example, although low correlations are associated with better performance in the case of attention and learning ((*5*), (*6*)), they are associated with worse performance when modulated by adaptation or contrast ((*7*)). Even in the case of cognitive processes like attention or task switching, good performance is associated with increases in correlation among particular subgroups of neurons ((*8*); (*7*)). And although mean correlated variability places much stronger constraints on cortical circuit models of cognition than measures of single neuron responses ((*9*), (*10*)), these models remain under-constrained.

These results highlight the need to use holistic methods to investigate the relationship between noise correlations and behavior. We focused on topological data analysis (TDA; (*11*); (*12*)), which is an emerging area in mathematics and data science that leverages groundbreaking advances in computational topology to summarize, visualize and discriminate complex data based on topological data summaries. These approaches, which have mostly been used in fields like astrophysics or large scale neural measurements (see, e.g., (*13*); (*14*); (*15*)), are able to identify features in the data that are qualitatively distinct from those highlighted using traditional analytic methods.

2.1 Topological signatures of correlated variability

We used TDA (specifically persistent homology, (*16*)) to quantify the structure of higher-order interactions among simultaneously recorded neurons from area V4 of rhesus monkeys performing a difficult visual detection task with an attention cue (Figure 1A; different aspects of these data have been presented previously; (*5*)). We analyzed the structure of noise correlations in a population of neurons in the visual cortex by constructing a space in which the distance measure between a pair of neurons is 1 – their noise correlation (Fig. 1B). In this space, highly correlated neurons are near each other, and anticorrelated neurons are far apart.

As is typical of the persistence homology approach, we iterate through a distance threshold (left to right in Fig. 1B) to understand the topological features of the correlation structure. For each threshold, we consider a pair of neurons to be functionally connected if their distance is less than the threshold. As the threshold increases, we thus include functional connections between pairs that are less correlated.

For each distance threshold, we use established TDA methods to identify "holes" in the correlation structure, which correspond to the lack of connections between a subset of neurons and have implications for the organization and function of the network ((*17*); (*18*)).

We use TDA in a slightly different way than past work, which focuses on persistent features (holes that persist through a large range of distance thresholds; (*19*), (*20*), (*21*), (*22*)). In our data sets, we simply did not observe persistent features (Figure 1C). Instead,

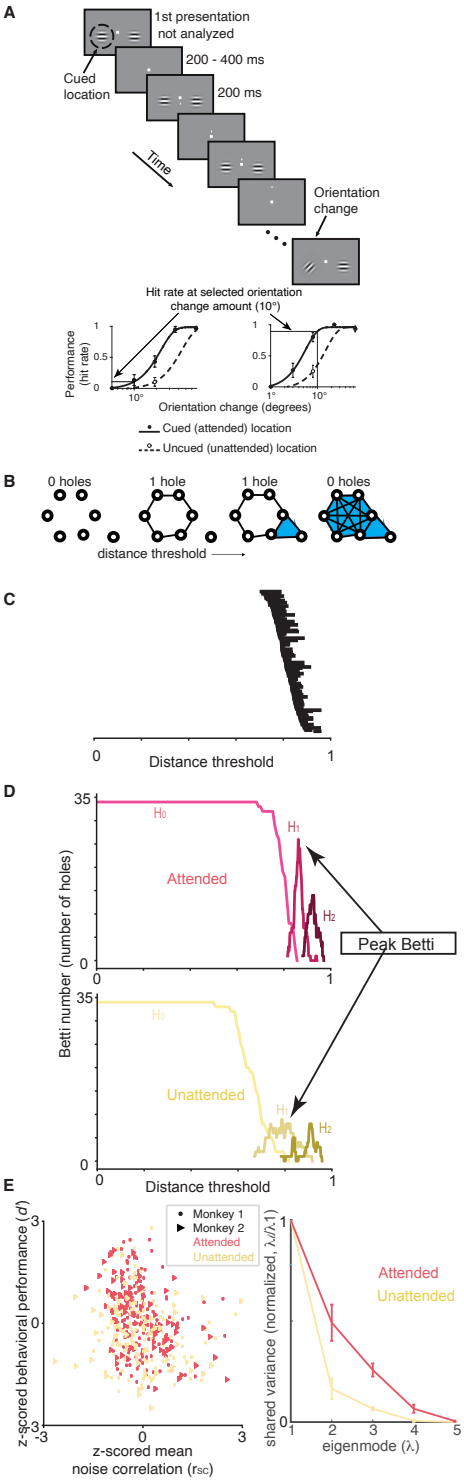
bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

we observed large numbers of holes that did not persist, and the number and distance threshold of those holes flexibly depended on attention and other cognitive processes.

We therefore summarize the topology of the correlation matrix as the peak Betti number, which is the maximal number of holes that appeared at any threshold for a given dimensionality (Figure 1D, (*23*)). We focus here on holes detected by the first homology group (equivalent to circles) and second homology group (equivalent to spheres), because these can be estimated using data sets of experimentally tractable size. In our data and models, focusing on the peak Betti curve led to qualitatively similar conclusions as other common topological summaries ((*24*); (*25*); Supplementary Figures 1 and 2).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 1



bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1: Experimental and topological methods. (A). Orientation change detection task with cued attention ((*5*)). The lower panels are psychometric curves (hit rate as a function of orientation change amount) for two example recording sessions to illustrate how we calculated performance at one selected orientation change amount on every recording session. (B). Illustration of topological data analysis methods. Each circle represents a neuron, and the distance between each pair is 1-their pairwise noise correlation (note that in real networks, more than two dimensions are typically required to represent all of the pairwise interactions). This analysis method iterates through distance thresholds (going from small to large from left to right). When the distance between two points is less than the threshold, they are considered connected. The shaded regions indicate groups of points that are fully interconnected, which indicates a higher order interaction between that subgroup of neurons. We summarize this structure by counting the number of holes at each threshold (which constructs the Betti curve in (D). (C). Topological description (persistence barcode) of an example recording session showing the distance thresholds (x-axis) at which each hole exists (holes are ordered by the threshold at which they appear). Many data sets are characterized by a small number of persistent topological features, which would show up as long horizontal lines in this plot. Instead, our neural data are characterized by a large number of holes that persist only for a small range of distance thresholds (many short horizontal lines in this plot). (D) Example Betti curves (plots of the number of holes as a function of distance threshold, which corresponds to the number of lines present at each threshold in (C)) for the $0th$ (lightest color; 'holes' correspond to fully shaded regions), $1st$ (middle color; holes are equivalent to circles), and $2nd$ homology groups (darkest color; holes are equivalent to spheres) for the two attention conditions (top/reds and bottom/yellows) for an example recording session. Our analyses focus on the peak of the Betti curve for the 1st and 2nd dimension (see supplement for other topological descriptions). (E) We focus our topological analyses of neural populations on correlated variability because there is a strong relationship between r_{SC} and performance on our psychophysical task (quantified as sensitivity or $d^{'}$). The plot shows $d^{'}$ as a function of r_{SC} (both values are z-scored for each animal and computed from responses to the stimulus before the orientation change). The colors represent the trials when attention was directed inside ('attended', red) or outside the receptive fields of the recorded V4 neurons ('unattended', yellow). The correlation between $d^{'}$ and r_{SC} was significant for each attention condition (attended: $r = -0.11$, $p = 0.11$; unattended $r = -0.25$, $p = 4.58e - 5$). (E) Factor analysis is a common linear method to assess the dimensionality of the correlated variability. The plot shows the shared variance (first five eigenvalues of the shared covariance matrix with private variance removed using Factor analysis) normalized by the shared variance in the first (dominant mode).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Here, we demonstrate that topological descriptions of correlated variability are an effective bridge between behavioral, physiological and theoretical approaches to studying neuronal populations. The peak Betti number is flexibly modulated by cognition, explains performance on a visually guided task, and gives novel insights into neuronal mechanisms and the function of real and artificial neural networks in different cognitive conditions.

2.2 Topology as a bridge to behavior

The primary reason for focusing on noise correlations is that the magnitude of noise correlations in visual cortex has been strongly linked to performance on visually guided tasks ((*5*), (*26*)). To justify our use of persistent homology to study neuronal networks, we tested the hypothesis that topological signatures of network activity capture key properties of the relationship between correlated variability and behavior.

Four observations suggest that the peak Betti number captures the aspects of noise correlations that explain performance. First, across recording sessions, there was a strong negative relationship between peak Betti number and the average noise correlation (Figure 2A, B)), meaning that sessions in which the average noise correlation was low tended to have a higher peak Betti number. Second, consistent with the observation that attention reduces noise correlations ((*5*); (*27*); (*28*); (*2*)), attention changes the peak Betti number (Figure 2). Third, the peak Betti number was higher on trials in which the animal correctly detected a change in a visual stimulus compared to trial in which the animal missed the stimulus change (Attended condition average peak Betti NumberH_1 correct:14.47,incorrect:13.46;H_2 correct:8.38,incorrect:7.41; Paired T test(peak Betti Number H_1) $p = 0.014$, (peak Betti Number H_2) $p = 0.015$). Finally, and there was a positive correlation between peak Betti number and behavioral performance (Figures 2 C, D). Together,these results show that peak Betti number is a good description of the aspects of correlated variability that correspond to changes in behavior.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 2

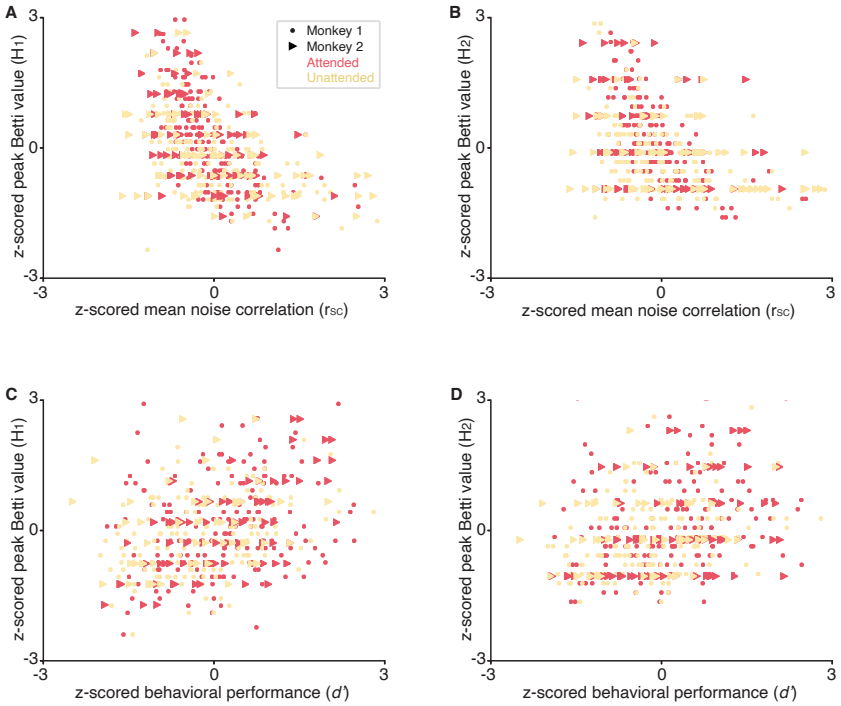


Figure 2: TDA reveals the relationship between neurons and behavior. (A) There is a strong relationship between the peak of the Betti curve from the $1st$ homology group and mean r_{sc} in both attention conditions (both are z-scored for each animal; (Attended: $r = -0.52$, $p = 1.3e-18$; Unattended: $r = -0.42$, $p = 7.9e-12$; Paired T-test (Attended and Unattended,peak of the Betti curve):$p = 8.41e-5$) (B) Same, for the $2nd$ homology group (Attended: $r = -0.44$, $p = 4.9e-13$; Unattended: $r = -0.36$, $p = 5.3e-9$; Paired T-test (Attended and Unattended,peak of the Betti curve):$p = 3.09e-6$) (C) There is a strong relationship between the peak of the Betti curve from the $1st$ homology group and and behavioral performance (d' or sensitivity calculated for a single orientation change for each session; both measures are z-scored; Attended: $r = 0.32$, $p = 2.34e-7$; Unattended: $r = 0.3$, $p = 1.41e-5$) (D) Same, for the second homology group (Attended: $r = 0.27$, $p = 2.06e-5$; Unattended: $r = -0.28$, $p = 4.67e-5$).

2.3 Topology as a bridge to mechanism

The magnitude of correlated variability places strong constraints on circuit models of the neuronal mechanisms underlying attention (Figure 3A; (*9*); (*10*)). In particular, network models are constrained by the observation that attention changes correlated variability in

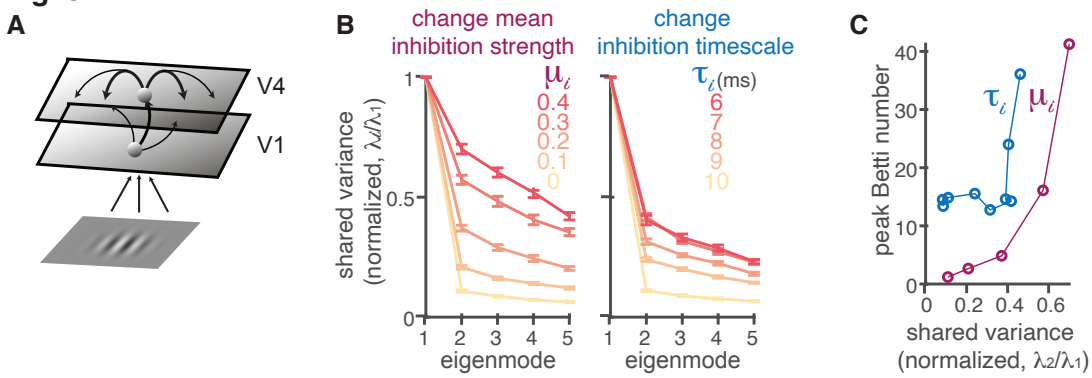
bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

essentially a single dimension of neuronal population space in area V4 (Figure 2, (*29*); (*30*); (*10*), (*7*)).

Topological descriptions of simulated networks can distinguish between competing models in situations when the magnitude of shared variability, even in the most relevant dimension, fails to do so. We analyzed the outputs of our spatially extended network of spiking neuron models, which internally generate correlated variability through spatiotemporal dynamics ((*10*)). In the model, the magnitude of correlated variability can be changed by modulating inhibition in two distinct ways: either increasing the input drive to the inhibitory neurons (μ_i in Figure 3B) or decreasing the timescale of inhibition (τ_i in Figure 3B) changes correlated variability in a low rank way.

These two mechanisms have very different effects on the topology of the correlated variability, even when the mean variability is equivalent. For most parameter values, changing the input drive to the inhibitory neurons has a much greater effect on the peak Betti number than changing the timescale (Figure 3C). While changing the timescale of inhibition is extremely common in circuit models (for review, see (*10*)), in real neural networks, the timescale of inhibition is longer than excitation and is inflexible ((*31*), (*32*); (*33*)). Both the biology and the topology are consistent with the idea that attention instead acts by increasing the input drive to the inhibitory neurons ((*9*); (*10*)).

Fig. 3



bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3: TDA can distinguish between different mechanistic models of attention. (A). Model schematic of a two-layer network of spatially ordered spiking neurons modeling primary visual cortex (V1) and area V4, respectively. The visual inputs to the model are the same Gabor stimuli used in our experiments. (B). Two distinct attention mechanisms can decrease correlated variability in a low rank way that is similar to linear descriptions of our data. We can reduce correlations either by increasing the currents to all inhibitory neurons (μ_i) or decreasing the decay timescale of inhibitory currents (τ_i^d). The plots depict the shared variance in each mode (the top five eigenvalues from the shared covariance, with private variance removed using Factor analysis), normalized by the shared variance in the first mode, for different values of μ_i (left) or τ_i^d (right; error bars represent standard error of the mean; SEM). The two mechanisms appear indistinguishable using linear methods. (C). The two mechanisms cause different changes in the topological descriptions of the modeled V4 populations. As μ_i increases, so does the shared variance present outside the first mode (x-axis) as well as the peak Betti number (shown for the first homology dimension in the y-axis). Changes in τ_i^d result in a different relationship between the peak Betti number and the shared variance in higher dimensions, affecting the peak Betti number only at very short (unrealistic) timescales. The peak Betti number is computed from the same simulated responses as in (B).

These results demonstrate that topological signatures of correlated variability provide unique constraints on mechanistic models that are unavailable using linear measures of neural activity. Changes in the mean or dimensionality of correlated variability are not necessarily coupled with changes in the topological signatures of the network. These results highlight the value of using circuit models as a platform on which to test and generate hypothesized mechanisms underlying perception and cognition.

2.4 Topology as a bridge to network function

The past decade has seen an explosion in the number of studies demonstrating that correlated variability depends on a wide range of sensory, cognitive, and motor conditions that change behavior (for review, see (*2*)). Despite much effort from the experimental and theoretical neuroscience communities ((*34*), (*35*), (*10*), (*9*), (*29*), (*36*), (*37*)), there is no satisfying explanation of exactly how changes in correlated variability might affect behavior. Here, we demonstrate that, through connections with network control theory ((*38*), (*39*)), TDA can provide unique insight into the relationship between correlated variability, the function of a network, and behavior.

We reasoned that we could gain insight into the function of our neuronal networks using measures from the field of network control theory, which seeks to quantify the ability of

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interventions (in our case, visual stimuli, cognitive processes, or random fluctuations) to alter the state of a network ((*40*)). While network control theory is primarily used in engineering, recent work has used controllability to quantify the flexibility of large neural systems, constrained by fMRI data ((*41*)).

These methods focus on quantifying the energy required to move between states of the neural population. We define a state as the vector of neural population activity on a given trial, and moving between states is constrained by the noise correlation matrix (e.g. in Figure 4A). For example, if the responses of all of the neurons are highly positively correlated, then reaching a state in which the response of some are high while the others are low is unlikely and therefore requires significant energy.

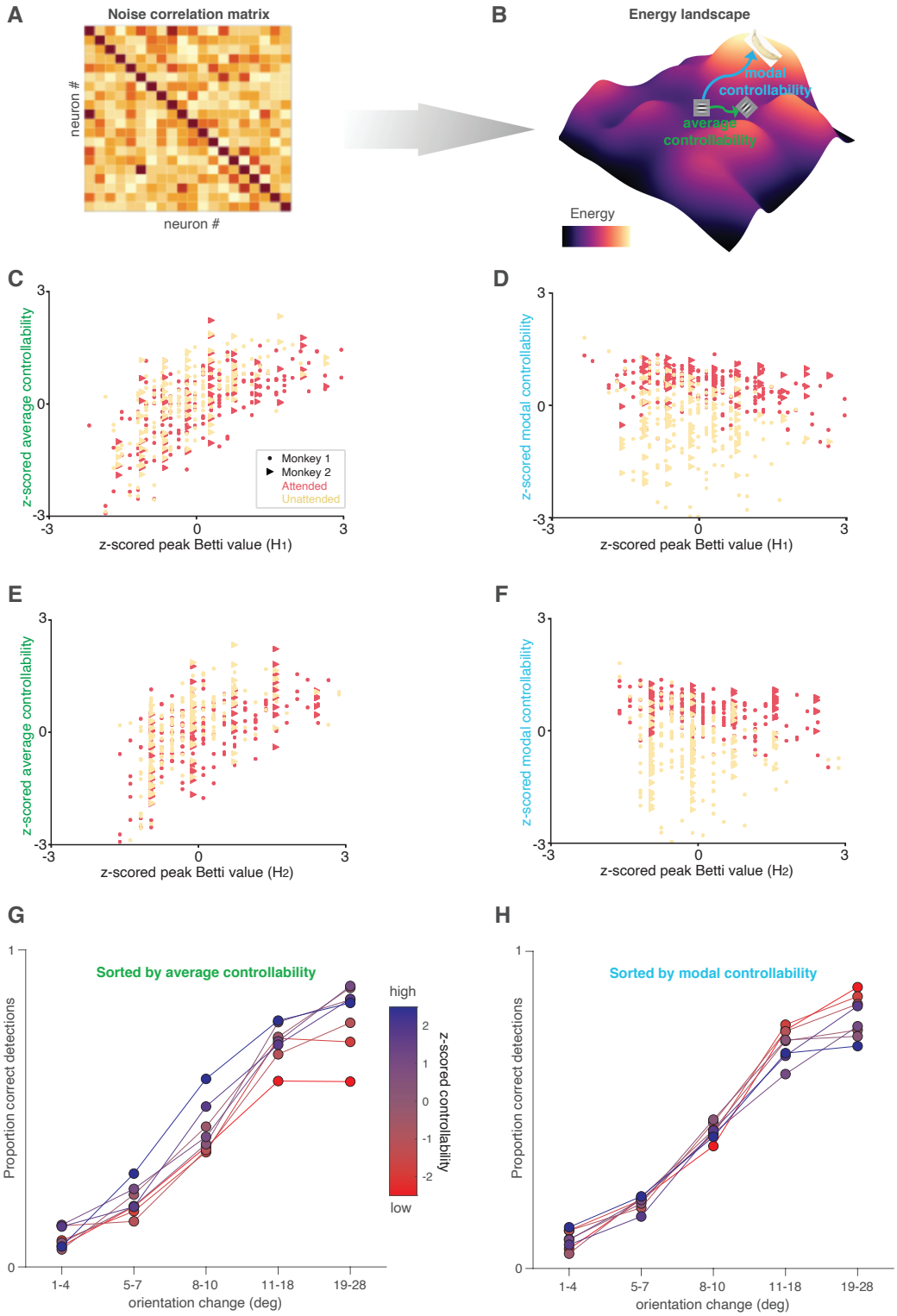
If our starting point (e.g. the horizontal Gabor in Fig. 4B) is the population response to a horizontal Gabor stimulus presented before the orientation change in our task (see Fig. 1A), a nearby state might be the population response to the changed stimulus (e.g. the oblique Gabor in Fig. 4B). A distant state might be a population response when the monkey is concentrating on something very different and task-irrelevant (e.g. thinking about the banana in Figure 4B). Average controllability quantifies how readily the population moves from the starting point to nearby states while modal controllability quantifies how readily the population moves to distant states.

There is no mathematical relationship between average and modal controllability. Indeed, average and modal controllability were uncorrelated across sessions in our data (R=0.008; p=0.9).

However, we discovered that the topological descriptions of neuronal population are strongly related to both average and modal controllability, and both are related to attention. High peak Betti value (which occurs more readily in the attended state) is associated with decreases in the energy required to drive the system to nearby states (high average controllability; Fig. 4C, E). In contrast, there is a negative relationship between peak Betti value and the energy required to drive the system to distant states (modal controllability; Fig. 4D, F) but changing attention conditions increases modal controllability (compare the red and yellow points in Fig. 4D, F).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. 4



bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4: TDA and controllability provide new insight into network function. (A,B). Illustration of our controllability calculation. We consider the noise correlation matrix (A) as a functional connectivity matrix, and use this to calculate an energy landscape (illustrated for a hypothetical situation in B; colors indicate energy). Average controllability is defined as the energy required to move from a starting point (e.g. a response to a horizontal Gabor stimulus) to nearby states (e.g. a response to an oblique Gabor), and modal controllability is defined as the energy required to move to distant states (e.g. thinking about a banana). (C). High average controllability is associated with high peak Betti number in the first homology group (both measures were z-scored for each animal, and the lines were fit for each attention condition; attended: $r = 0.65$, $p = 1.03e - 31$; unattended: $r = 0.68$, $p = 1.28e - 33$; Paired T-test (Attended and Unattended,average controllability):$p = 7.5e - 61$). (D) High modal controllability is associated with low peak Betti number in the first homology group (conventions as in A; (attended: $r = -0.37$, $p = 7.7e{-}10$; unattended: $r = -0.23$, $p = 3.2e - 4$; Paired T-test (Attended and Unattended,modal controllability):$p = 1.5e - 38$) E) Relationship between average controllability and peak Betti number in the $2nd$ homology group (attended: $r = 0.59$, $p = 3.5e - 25$; unattended: $r = 0.59$, $p = 9.5e - 24$). Conventions as in (C). (F). Relationship between modal controllability and peak Betti number in the $2nd$ homology group (attended: $r = -0.38$, $p = 3.12e - 10$; unattended: $r = -0.18$, $p = 4.7e{-}3$). Conventions as in (D). (G). High average controllability is associated with better performance at all orientation change amounts. Colors represent z-scored average controllability (the experimental sessions were split into six equally sized bins by average controllability) and the plot shows proportion correct detections (hit rate) as a function of orientation change amount. (H). High modal controllability (bluer colors) is associated with a worse lapse rate (worse performance on easier trials). Conventions as in (H).

The different relationships between topology and average and modal controllability observations give new insight into the tradeoffs associated with attention. Attending to a stimulus improves the network's ability to respond to subtle interventions, which explains attention-related improvements in the animal's ability to detect a subtle change in the visual stimulus (Fig. 1A), but it has complex effects on the ability of the network to change states dramatically, which may mean that attention reduces cognitive flexibility. In future work, it would be interesting to study whether changes in controllability can account for change blindness and other behavioral demonstrations that attention reduces the ability of observers to notice very unexpected stimuli, such as the classic example of failing to notice a gorilla walking through a basketball game ((*42*)).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Indeed, average and modal controllability have distinct relationships with the monkeys' performance in our task. We sorted the experimental sessions by average controllability (colors in Fig. 4G) or modal controllability (colors in Fig. 4H). Increased average controllability was associated with improvements in the monkeys' ability to detect all orientation change amounts (except the smallest changes in which a floor effect meant that they were rarely detected). In contrast, modal controllability was unrelated to the monkeys' ability to detect subtle orientation changes and was anticorrelated with the ability to detect large, easy orientation changes. One interpretation is that when modal controllability is high, the monkeys' minds wander more easily to distant, potentially task-irrelevant states, increasing the lapse rate on easy trials. Together, these results demonstrate that in addition to linking to behavior and mechanism, topological signatures of the structure of noise correlations can provide insight into the function of the network and the behavioral trade-offs associated with changes in correlated variability.

2.5 Implications for topology

Although TDA has been used for many scientific applications ((*17*)), our use of TDA differs substantially from previous work. The prevailing paradigm used in virtually all TDA applications, including many in neuroscience ((*18*); (*13*)), focuses on identifying persistent topological features, such as holes that persist across many thresholds (Figure 1B). These persistent features are appealing because they can reveal the structure of a simple network. However, applying these methods to analyze neural circuits may not lead to any scientific discoveries, since persistent features are not typically present in the complex networks for which linear methods fail.

However, we demonstrated that TDA is very useful, even in the absence of persistent features. The link that we demonstrated between the topology of noise correlations, which have been shown to reflect both cognition and the anatomy of the system ((*3*)), and the controllability of the network on individual trials (which are what matter for guiding behaviors) therefore has implications far beyond neuroscience. Throughout the natural and physical sciences, natural systems are complex and call for sophisticated data analytics pipelines. In astronomy, for example, TDA has been used to understand the relationship between planets, stars, and galaxies on a huge range of spatial scales. Our novel TDA approach based on non-persistent topological features will be a bridge between neuroscience and other fields. These tools for analyzing and interpreting complex networks can be deployed in many other scientific domains.

2.6 Implications for neuroscience

We demonstrated here that TDA is a unique analytical bridge between behavior, neurons, computations, and mechanisms. This sort of bridge between different levels of investigation has the potential to be broadly transformative. In an age of massive improvements in experimental technologies and tools, perhaps the greatest barrier to success understanding the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

neural basis of behavior is that it is different to compare and integrate results from experiments using different methods in different model systems. TDA can reveal relationships between neural networks, computations, and behaviors that are robust to the differences in neuronal responses that occur between every different experimental system((*43*)). These tools will bridge levels of investigation, making it possible to leverage the unique and complementary strengths of each approach.

A holistic view of neuronal populations is necessary for understanding any neural computation. Essentially every normal behavioral process or disorder of the nervous system is thought to involve the coordinated activity of large groups of neurons spanning many brain areas. Tools for understanding and interpreting large populations have lagged far behind tools for measuring their activity. Standard linear methods have provided a limited view, and the field is in dire need of a new, holistic window into population activity. Our results demonstrate a hopeful future for using the topology of neural networks to fulfill that need.

3 Methods

3.1 Experimental methods

Different analyses of these data have been presented previously ((*5*)). Briefly, two adult rhesus monkeys (Macaca mulatta) performed an orientation change detection task with a spatial attention component ((*27*)). The monkeys fixated a central spot while two peripheral Gabor stimuli flashed on (for 200 ms) and off (for a randomized period between 200 and 400 ms). At a random and unsignaled time, the orientation of one stimulus changed, and the monkey received a liquid record for making an eye movement to the changed stimulus within 500 ms. We cued attention in blocks of 125 trials, and the orientation change occurred at the cued location on 80% of trials. Our analyses are based on responses to the stimulus presentation before the change, which was the same on every trial within a recording session. The location, contrast, and spatial frequency of the Gabor stimuli were the same during every recording session, but the orientation differed across sessions. The location of one stimulus was within the receptive fields of the recorded neurons and the other stimulus was in the opposite hemifield.

While the monkey performed the task, we recorded simultaneously from a chronically implanted 96-channel microelectrode array (Blackrock Microsystems) in the left hemisphere of visual area V4. We include both single units and sorted multiunit clusters (mean 34 and 15 units for Monkeys 1 and 2, respectively). The average number of simultaneously recorded pairs of units (for computing noise correlations) was 561 for Monkey 1 and 105 for Monkey 2. The data presented are from 42 recording sessions from Monkey 1 and 28 recording sessions from Monkey 2.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.2 Data preparation

To examine how the topology of networks of neurons in visual cortex or outputs of spiking models depend on attention, we constructed distance matrices from noise correlation matrices ((*3*)) constructed from binned spike train data. We based our analyses on spike count responses between 60-260 ms after the onset of the visual stimulus to allow for the latency of visual responses in area V4. We used responses to the stimulus before the orientation change because those are the same on every trial. We focused on trials when the monkey correctly identified the changed stimulus and compared responses in the two attention conditions.

Many measures of neuronal activity depend on experimental details like the number of recorded neurons or their mean firing rates, which were different for the two monkeys (see (*5*) for details). To allow us to combine across animals, we z-scored the results for each animal (across both attention conditions) and plot those normalized measures in the figures.

3.3 Behavioral measures

To analyze the relationship between neuronal responses and behavior, we adopted a signal detection framework ((*44*); (*45*)) to assess how behavior depends on neurons and attention ((*46*); (*47*); (*48*); (*49*); (*50*); (*51*), (*52*)). Criterion is defined as Sensitivity is defined as $d^{'}$, Different orientation change amounts were used in different recording sessions. To compare across sessions, we fit the psychometric curve using a Weibull function and computed performance at a single, fixed orientation change for each session (see Figure 1A and (*5*)).

3.4 Topological measures

We examine the topology of the noise correlations in each attention condition using a Vietoris Rips construction. This consists of defining a distance matrix (better understood as a weighted adjacency matrix), which we constructed from the noise correlation matrix r_{SC}, to define pairwise (and higher order) connections between the vertices (representing neurons) in the simplicial complex. The distance matrix was chosen to be $1 - r_{SC}$, so that higher weighted interactions (i.e. those neurons which are strongly correlated) are defined as shorter distance and therefore entered the simplicial complex first. For brevity, we refer to this matrix as the distance matrix of the space while acknowledging that our measure does not satisfy the axioms of a distance function.

We consider a distance threshold which defines those pairwise interactions that are permitted to be considered in the simplicial complex. Such a process allows us to examine the evolution of the simplicial complex across different distance thresholds. We assess several properties of the simplicial complex at each threshold value, including the existence of holes (or higher dimensional voids) of a given dimensionality (termed homology dimension). A hole signifies a lack of connections (i.e. differences in the degree of correlation) between a subset of neurons at the current distance threshold. We focus our analysis on the first and second homology dimensions (which correspond to holes that are topologically equivalent to

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

circles and spheres, respectively) because they can be estimated reliably given the size of our data sets.

We then examined properties of the generated Betti curves and observed how they relate to common measures of attention like average noise correlations and behavioral performance. Because noise correlations are varied and require substantial data to estimate ((*3*)), we focus on the distribution of topological features rather than on looking for long persistent cycles (see Figure 1C for an example of observed features).

3.5 Spatial model construction

Our spatial model is a variation of the two layer network of neurons discussed in ((*10*)). Neurons in this network are arranged uniformly on a $[0,1] \times [0,1]$ grid. The first layer (i.e. the feed-forward layer) consist of $N_x = 2500$ excitatory neurons that behave as independent Poisson processes. The second layer consist of $40,000$ excitatory and $10,000$ inhibitory neurons that are recurrently coupled. The second layer receives input from the first layer. The network's connectivity is probabilistic but dependent on a Gaussian of width σ_*. Thus neurons that are further away from each other on the grid are less likely to connect. The parameters are the same as the two-layer network in the (*10*), except the following. The the feedforward connection strength from layer 1 to layer 2 is $J_{ex} = 140$ and $J_{ix} = 0$ for excitatory and inhibitory neurons, respectively. Fig 3B left: $\mu_i = 0, 0.1, 0.2, 0.3, 0.4$ and τ_i is 10. Fig 3B right: $\mu_i = 0$, and $\tau_i = 6, 7, 8, 9, 10$ ms. There were a total of 15 simulations of 20 seconds each for each parameter conditions. The first 1 second in each simulation was removed. The spike counts were computed using 140 ms time window. We implemented this model using EIF neurons. The voltage dynamics of these neurons are governed by the following equation((*10*)):

$$C_m \frac{dV_j^\alpha}{dt} = -g_L(V_j^\alpha - E_L) + g_L \Delta_T e^{\frac{(V_j^\alpha - V_T)}{\Delta_T}} + I_j^\alpha(t) \tag{1}$$

where $\tau_m = \frac{C_m}{g_L} = 15ms$, $E_L = -60mV$,$V_T = -50mV$, $V_{th} = -10mV$,$\Delta_T = 2mV$,$V_{re} = -65mV$,$\tau_{ref} = 1.5ms$ and the total current $I_j^\alpha(t)$ obeys the following equation:

$$\frac{I_j^\alpha(t)}{C_m} = \sum_{k=1}^{N_F} \frac{J^{\alpha F}}{\sqrt{N}} \sum_n \eta_F(t - t_n^{F,k}) + \sum_{\beta=e,i} \sum_{k=1}^{N_\beta} \frac{J_{jk}^{\alpha\beta}}{\sqrt{N}} \sum_n \eta_\beta(t - t_n^{\beta,k}) + \mu_\alpha \tag{2}$$

where N is the total number of neurons in the second layer and μ_α is the static current to the $\alpha(\in \{E, I\})$ population. η_β is the postsynaptic current given by the following equation

$$\eta_\beta(t) = \frac{1}{\tau_{\beta d} - \tau_{\beta r}} \begin{cases} e^{\frac{-t}{\tau_{\beta d}}} - e^{\frac{t}{\tau_{\beta r}}}, t \geq 0 \\ 0, t < 0 \end{cases} \tag{3}$$

where the rise time constant $\tau_{\beta r} = 5$. We consider multiple values of the decay time constant $\tau_{\beta d}$. For both dimensionality (see Factor Analysis section) and topological (see Topological

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Measures section) related comparisons, we also considered a range of values of the μ_I parameter which correspond to the overall depolarization of the inhibitory population and which has been shown to affect the dimensionality of the generated data. The data generated from

3.6 Factor analysis

To assess the low dimensional nature of the data generated using the spatial model, we used factor analysis ((*37*)). Using the spike count matrix generated from the excitatory population data, we compute the spike count covariance matrix . Factor analysis entails separating the spike count covariance matrix into a shared component that represents how neurons covary together and an independent component that captures neuron-specific variance. In the notation of ((*37*)), we refer to the shared component as LL^T and the independent component as Ψ. With this partition, the dimensionality of the network's activity is assessed by looking at the 5 largest eigenvalues generated by LL^T. To improve the clarity of the analysis that can arise from magnitude changes in the eigenvalues, we normalized the data by dividing each eigenvalue by the largest eigenvalue. We performed this analysis on spike count data randomly sampled from 500 neurons. All results shown were generated from analysis that was cross-validated. Errorbars shown in the above figures come about by applying this analysis over random instances of the network with fixed model parameters.

3.7 Controllability measures

The goal of our controllability measures is to understand how the noise correlation matrix in each attention state constrains estimates of the function of the network. We consider a hypothetical (possibly nonlinear) dynamical system whose dynamics can be linearized and whose effective connectivity is defined by the noise correlation matrices. We analyze the properties of the system to assess the amount of effort it takes to change the system's state using external input. We summarize these calculations using two standard measures of controllability ((*53*), (*40*)): average controllability, which relates to the ability to push the system into nearby states or states with little energy and modal controllability, which relates to the ability to push the system into distant states or states that require more energy.

The average controllability ϱ_c is defined as

$$\varrho_c = Trace(W_c) \tag{4}$$

where $W_c = \int_0^{t_f} e^{A_\tau} BB^T e^{A_\tau^T} d\tau$ is the controllability gramian.

The modal controllability ϕ_i is defined as

$$\phi_i = \sum_{j=1}^{N} = (1 - e^{\lambda_j(A)}) v_{ij}^2 \tag{5}$$

where λ_j is the jth eigenvalue of the effective connectivity matrix A.

To determine the relationship between controllability and behavior (Figures 4G,H), we computed average and modal controllability for each session, z-scored those measures for

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

each monkey, and divided the sessions into six equally sized bins for each controllability measure.

References

1. M. Carrasco, *Vision research* **51**, 1484 (2011).
2. J. H. Maunsell, *Annual Review of Vision Science* **1**, 373 (2015).
3. M. R. Cohen, A. Kohn, *Nature neuroscience* **14**, 811 (2011).
4. H. Ko, *et al.*, *Nature* **473**, 87 (2011).
5. A. M. Ni, D. A. Ruff, J. J. Alberts, J. Symmonds, M. R. Cohen, *Science* **359**, 463 (2018).
6. Y. Gu, *et al.*, *Neuron* **71**, 750 (2011).
7. D. A. Ruff, C. Xue, L. E. Kramer, F. Baqai, M. R. Cohen, *Proceedings of the National Academy of Sciences* **117**, 29321 (2020).
8. D. A. Ruff, M. R. Cohen, *Nature neuroscience* **17**, 1591 (2014).
9. T. Kanashiro, G. K. Ocker, M. R. Cohen, B. Doiron, *Elife* **6**, e23978 (2017).
10. C. Huang, *et al.*, *Neuron* **101**, 337 (2019).
11. G. Carlsson, *Acta Numerica* **23**, 289 (2014).
12. P. Y. Lum, *et al.*, *Scientific reports* **3**, 1 (2013).
13. A. E. Sizemore, *et al.*, *Journal of computational neuroscience* **44**, 115 (2018).
14. M. Saggar, *et al.*, *Nature communications* **9**, 1 (2018).
15. C. T. Ellis, M. Lesnick, G. Henselman-Petrusek, B. Keller, J. D. Cohen, *Network Neuroscience* **3**, 695 (2019).
16. H. Edelsbrunner, J. Harer, *Computational topology: an introduction* (American Mathematical Soc., 2010).
17. C. Giusti, R. Ghrist, D. S. Bassett, *Journal of computational neuroscience* **41**, 1 (2016).
18. C. Giusti, E. Pastalkova, C. Curto, V. Itskov, *Proceedings of the National Academy of Sciences* **112**, 13455 (2015).
19. G. Singh, *et al.*, *Journal of vision* **8**, 11 (2008).
20. R. Chaudhuri, B. Gerçek, B. Pandey, A. Peyrache, I. Fiete, *Nature neuroscience* **22**, 1512 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

21. E. Rybakken, N. Baas, B. Dunn, *Neural computation* **31**, 68 (2019).

22. R. J. Gardner, *et al.*, *bioRxiv* (2021).

23. J.-B. Bardin, G. Spreemann, K. Hess, *Network Neuroscience* **3**, 725 (2019).

24. A. Adcock, E. Carlsson, G. Carlsson, *arXiv preprint arXiv:1304.0530* (2013).

25. M. Rucco, F. Castiglione, E. Merelli, M. Pettini, *Proceedings of ECCS 2014* (Springer, 2016), pp. 117–128.

26. D. A. Ruff, A. M. Ni, M. R. Cohen, *Annual review of neuroscience* **41**, 77 (2018).

27. M. R. Cohen, J. H. Maunsell, *Nature neuroscience* **12**, 1594 (2009).

28. J. F. Mitchell, K. A. Sundberg, J. H. Reynolds, *Neuron* **63**, 879 (2009).

29. N. C. Rabinowitz, R. L. Goris, M. Cohen, E. P. Simoncelli, *Elife* **4**, e08998 (2015).

30. A. S. Ecker, G. H. Denfield, M. Bethge, A. S. Tolias, *Journal of Neuroscience* **36**, 1775 (2016).

31. J. R. Geiger, J. Lübke, A. Roth, M. Frotscher, P. Jonas, *Neuron* **18**, 1009 (1997).

32. Z. Xiang, J. R. Huguenard, D. A. Prince, *the Journal of Physiology* **506**, 715 (1998).

33. M. C. Angulo, J. Rossier, E. Audinat, *Journal of neurophysiology* **82**, 1295 (1999).

34. A. S. Ecker, A. S. Tolias, *Nature neuroscience* **17**, 750 (2014).

35. R. L. Goris, J. A. Movshon, E. P. Simoncelli, *Nature neuroscience* **17**, 858 (2014).

36. J. D. Semedo, A. Zandvakili, C. K. Machens, M. Y. Byron, A. Kohn, *Neuron* **102**, 249 (2019).

37. R. C. Williamson, *et al.*, *PLoS computational biology* **12**, e1005141 (2016).

38. S. F. Muldoon, *et al.*, *PLoS computational biology* **12**, e1005076 (2016).

39. E. Tang, D. S. Bassett, *Reviews of modern physics* **90**, 031003 (2018).

40. K. P. Szymula, F. Pasqualetti, A. M. Graybiel, T. M. Desrochers, D. S. Bassett, *arXiv preprint arXiv:2006.14565* (2020).

41. S. Gu, *et al.*, *Nature communications* **6**, 1 (2015).

42. B. T. Hutchinson, *Neuroscience & Biobehavioral Reviews* **104**, 87 (2019).

43. F. J. Luongo, *et al.*, *bioRxiv* (2021).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

44. D. M. Green, J. A. Swets, *et al.*, *Signal detection theory and psychophysics*, vol. 1 (Wiley New York, 1966).

45. N. A. Macmillan, C. D. Creelman, *Detection theory: A user's guide* (Psychology press, 2004).

46. H. S. Bashinski, V. R. Bacharach, *Perception & psychophysics* **28**, 241 (1980).

47. H. J. Müller, J. M. Findlay, *Perception & Psychophysics* **42**, 383 (1987).

48. C. J. Downing, *Journal of Experimental Psychology: Human perception and performance* **14**, 188 (1988).

49. H. L. Hawkins, *et al.*, *Journal of Experimental Psychology: Human Perception and Performance* **16**, 802 (1990).

50. V. Wyart, S. Dehaene, C. Tallon-Baudry, *Frontiers in Human Neuroscience* **6**, 16 (2012).

51. T. Z. Luo, J. H. Maunsell, *Neuron* **97**, 1382 (2018).

52. T. Z. Luo, J. H. Maunsell, *Proceedings of the National Academy of Sciences* **116**, 26187 (2019).

53. T. M. Karrer, *et al.*, *Journal of neural engineering* **17**, 026031 (2020).

4 Acknowledgments

We thank Douglas Ruff, Joshua Alberts, and Jen Symmonds for contributions to the previously published experiments and Alessandro Rinaldo and Ralph Cohen for helpful conversations. This work was supported by NIH grants 2R01EY022930 (MRC, TCR), R01NS121913 (MRC, CH), and K99NS118117 (AMN), and grant 542961SPI from the Simons Foundation (MRC).

5 Author contributions

TCR and MRC designed the project; TCR developed the analysis techniques and analyzed the data and simulations; AMN collected the electrophysiological data; CH performed the model simulations; MRC supervised the project; TCR, AMN, CH, and MRC contributed to writing the manuscript.

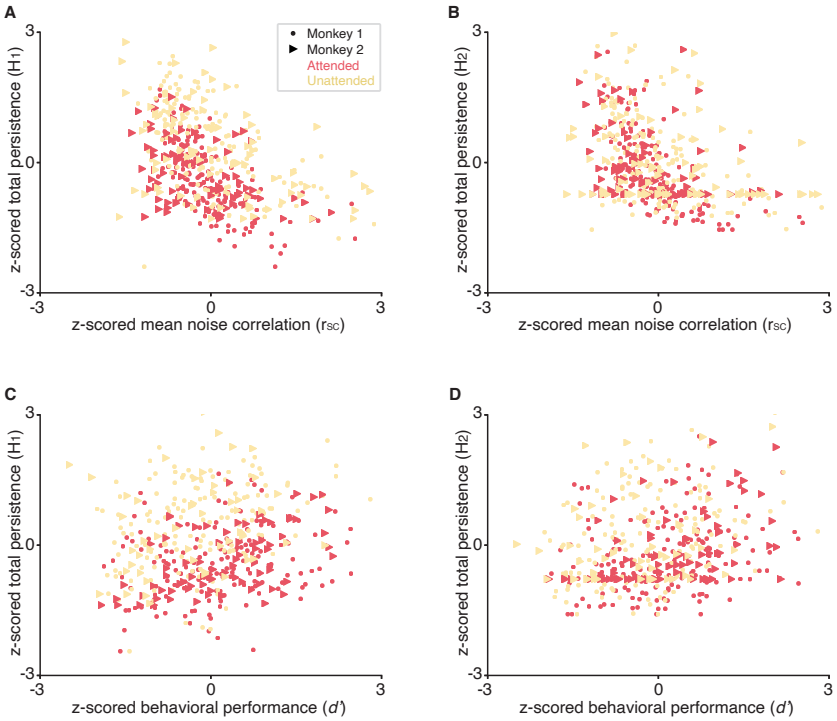
6 Competing interests

The authors declare no competing financial interests.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

7 Supplemental Figures

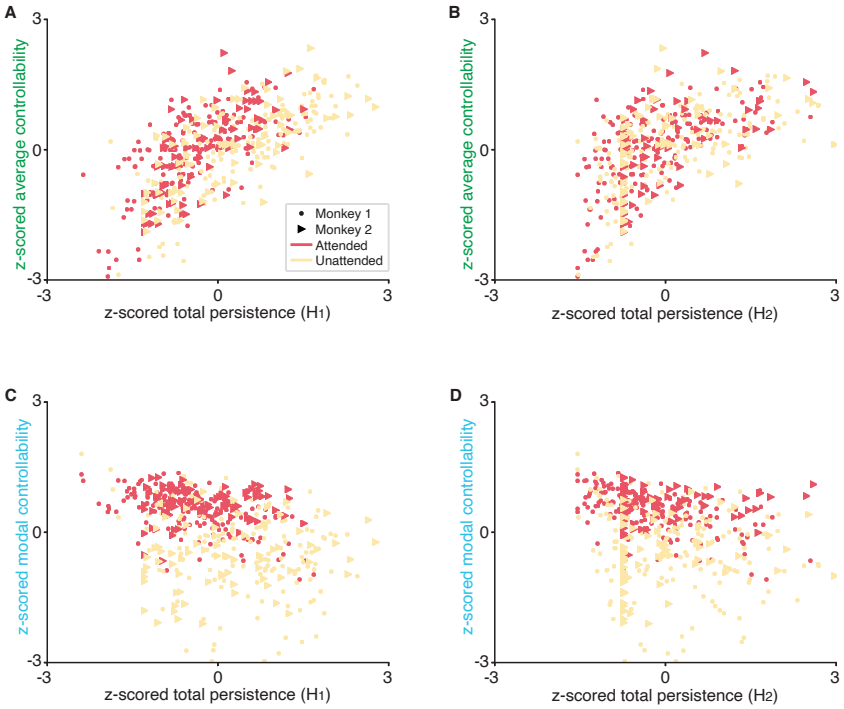
Fig. S1



Supplemental Figure 1: Other topological descriptions of correlated variability are also strongly related to mean noise correlations and behavior. (A). Total persistence for the $1st$ homology group as a function of mean noise correlation. Total persistence is defined as the sum divided by the difference between the distance thresholds at which each hole begins and ends (termed birth and death times of the persistence features in the language of algebraic topology). Both the noise correlation values and total persistence values were z-scored for each animal, and the lines were fit for each attention condition. Symbols were assigned based on monkey. (attended: $r = -0.54$, $p = 5.14e - 20$; unattended: $r = -0.5$, $p = 1.3e - 16$; Paired T-test (Attended and Unattended,total persistence):$p = 7.3e - 26$). (B). Same, for the $2nd$ homology group (attended: $r = -0.45$, $p = 2.8e - 14$; unattended: $r = -0.41$, $p = 2.7e{-}11$; Paired T-test (Attended and Unattended,total persistence):$p = 1.5e - 7$). (C) Relationship between total persistence and behavioral sensitivity $d^{'}$ for the $1st$ homology group. Conventions and z-scoring as in (A). (attended: $r = 0.32$, $p = 2.7e - 7$; unattended: $r = 0.27$, $p = 5.7e - 5$) (D) Same, for the $2nd$ homology group (attended: $r = 0.31$, $p = 1.5e - 6$; unattended: $r = 0.41$, $p = 6.8e - 6$).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.24.461717; this version posted September 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig. S2



Supplemental Figure 2: Total persistence has similar relationships to average and modal controllability as the peak Betti value. A). Total persistence for the $1st$ homology group as a function of average controllability. Z-scoring and other conventions the same as Supplementary Fig. 1. (attended: $r = 0.68$, $p = 5.19e - 35$; unattended: $r = 0.7$, $p = 6.86e - 37$). (B) Same, for the $2nd$ homology group (attended: $r = 0.59$; $p = 1.64e - 24$; unattended: $r = 0.57$, $p = 4.36e - 22$) (C). Total persistence for the $1st$ homology group as a function of modal controllability. Conventions as in (A). (attended: $r = -0.38$, $p = 3.91e - 10$; unattended: $r = -0.18$, $p = 5.2e - 3$). (D). Same, for the $2nd$ homology group. (attended: $r = -0.37$, $p = 2.07e - 9$; unattended: $r = -0.18$, $p = 4.7e - 3$)