Task-evoked activity quenches neural correlations and variability in large-scale brain systems

Authors: Takuya Ito[1,2,*], Scott L. Brincat[3], Markus Siegel[4,5,6], Ravi D. Mill[1], Biyu J. He[7], Earl K. Miller[3], Horacio G. Rotstein[1,8,9], Michael W. Cole[1]

[1]Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102
[2]Behavioral and Neural Sciences PhD Program, Rutgers University, Newark, NJ 07102
[3]The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139
[4]Centre for Integrative Neuroscience & MEG Center, University of Tübingen, 72076 Tübingen, Germany
[5]Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany
[6]MEG Center, University of Tübingen, 72076 Tübingen, Germany
[7]Neuroscience Institute, Langone Medical Center, New York University, New York, NY 10016
[8]Federated Department of Biological Sciences, Rutgers University, Newark, NJ 07102
[9]Institute for Brain and Neuroscience Research, New Jersey Institute of Technology, Newark, NJ 07102

Corresponding Author:
Takuya Ito
Center for Molecular and Behavioral Neuroscience, Rutgers University
197 University Avenue, Newark, NJ 07102
taku.ito1@gmail.com

Classification: Biological Sciences - Neuroscience

Conflict of Interest: None

Acknowledgements: We thank Drew B. Headley, Bart Krekelberg, Olaf Sporns, and Mauricio Delgado for helpful feedback on earlier versions of the manuscript. The authors acknowledge support by the US National Institutes of Health under awards R01 AG055556 and R01 MH109520, by the European Research Council (ERC StG335880), and the Centre for Integrative Neuroscience (DFG, EXC 307). Some of the data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Abstract

Many studies of large-scale neural systems have emphasized the importance of communication through increased inter-region correlations ("functional connectivity") during task states relative to resting state. In contrast, local circuit studies have demonstrated that task states reduce correlations among local neural populations, likely enhancing their information coding. Here we sought to adjudicate between these conflicting perspectives, assessing whether large-scale system correlations tend to increase or decrease during task states. To establish a mechanistic framework for interpreting changes in neural correlations, we conceptualized neural populations as having a sigmoidal neural transfer function. In a computational model we found that this straightforward assumption predicts reductions in neural populations' dynamic output range as task-evoked activity levels increase, reducing responsiveness to inputs from other regions (i.e., reduced correlations). We demonstrated this empirically in large-scale neural populations across two highly distinct data sets: human functional magnetic resonance imaging data and non-human primate spiking data. We found that task states increased global neural activity, while globally quenching neural variability and correlations. Further, this global reduction of neural correlations led to an overall increase in dimensionality (reflecting less information redundancy) during task states, providing an information-theoretic explanation for task-induced correlation reductions. Together, our results provide an integrative mechanistic account that encompasses measures of large-scale neural activity, variability, and correlations during resting and task states.

Keywords: functional connectivity, neural correlations, neural variability, dynamical systems

Significance Statement

Statistical estimates of neural correlations and variability are widely used to characterize neural systems during different states. We provide empirical evidence from two highly distinct data sets – human functional magnetic resonance imaging and non-human primate neural spiking data – to illustrate the generality with which task-evoked activity is associated with the quenching of neural variability and correlations. To explain empirical results, we use dynamical systems modeling to demonstrate the deterministic relationship between task-evoked activity, neural variability and correlations. Together, we provide an integrative account, showing that task-evoked activity quenches neural variability and correlations in large-scale neural systems.

Introduction

Measures of neural correlations and variability are widely used in neuroscience to characterize neural processes. Neural variability has been consistently shown to be reduced during task states across human functional magnetic resonance imaging (fMRI) (1, 2), local spiking populations (3, 4), and both local circuit (4–6) and large-scale models (7, 8). Despite this convergence in the neural variability literature, there are disparities in the use and interpretation of neural correlations. In the human fMRI literature, neural correlations are often estimated by measuring the correlation of blood oxygenated level-dependent (BOLD) signals and is commonly referred to as functional connectivity (FC) (9). In the animal spiking literature, neural correlations have been measured by computing the correlation between the spike rate of two or more neurons and is commonly referred to as the spike count correlation (or noise correlation) (10, 11). Yet despite the use of different terms, the target statistical inference behind these two techniques is consistent: to identify interacting neural units.

In the human fMRI literature, studies have identified large-scale functional brain networks through clustering sets of correlated brain regions using resting-state activity (12–14). During task states, the FC structure has been demonstrated to dynamically reconfigure (15–17). Previous studies have suggested that during tasks, the functional brain network organization transitions to a state of global integration (from segregation at rest) by increasing FC across networks (15, 18). Though it has been suggested that correlation increases and decreases respectively facilitate and inhibit inter-region communication (17, 19, 20), the underlying causes of these FC changes remain unclear.

Studies in the local circuit literature using electrophysiological recordings in animals have characterized the correlation structure between neuron spikes across a range of task demands. These studies have found that the spike count correlation between neuron spikes generally decrease during task states (21–23). Though the reduction in spike count correlations during task states has been linked to better information coding by the local neuronal population via reduced information redundancy (10, 11), the mechanism underlying these reduced correlations remains unclear.

These differing perspectives of the role of neural correlations in (large-scale) human imaging and (local circuit) animal neurophysiology studies appear to be incompatible. To further complicate this issue, we recently found that task activations can inflate task functional connectivity estimates in human fMRI data (24), suggesting some previous neuroimaging studies may have erroneously reported correlation increases. Here we sought to adjudicate the differing perspectives in the neural correlation literature while simultaneously confirming and extending previous findings on task state neural variability reduction. We use two data sets that span distinct data modalities and animal species (human fMRI and non-human primate (NHP) spiking data). In addition, to provide a single, mechanistic account capable of explaining changes in neural variability and correlations in task and rest states, we use simple neural mass models to describe the relationship between the underlying network properties and measures of neural variability and correlations using dynamical systems theory. Dynamical systems theory

allowed us to provide a detailed description of the deterministic neural processes underlying changes in emergent network statistics such as neural variability and correlations.

Neural processes are widely acknowledged to be emergent properties of a dynamical system (25, 26). Despite this recognition, it remains unclear how changes to the underlying neural attractor dynamics relate to changes in descriptive neural statistics. A recent study provided empirical and theoretical evidence that stimulus- or task-evoked activity alters the underlying dynamical processes of neural populations (4). Hennequin and colleagues suggested that following a stimulus onset, widespread neural variability is quickly reduced likely due to strong inhibitory feedback quenching variability around a steady state attractor. However, that study focused on the dynamical regime within local neural populations (in sensory cortex), rather than across large-scale neural systems. Building on those findings, we establish rigorous dynamical systems mechanisms that underlie neural variability and correlation suppression in large-scale neural populations (brain systems and functional areas).

We report multiple sources of empirical and theoretical evidence demonstrating that task-evoked activity quenches neural correlations and variability in large-scale neural systems. We first demonstrate how a sigmoid activation function, which is widely used to characterize neural population activity, can explain how evoked activity quenches output variability (Fig. 1). We then characterize task-evoked neural variability and correlations in empirical data using two highly distinct data sets: human fMRI and NHP electrophysiology. This allowed us to test whether there were consistent variability and correlation changes during task states independent of data modality or species in large-scale neural populations. Finally, we provide theoretical insight using dynamical systems theory, demonstrating that task-evoked activity strengthens the system's attractor dynamics around a stable fixed point in neural mass models, quenching neural correlations and variability. Together, our results provide a mechanistic framework for interpreting how task-evoked activity alters global neural variability and correlations.

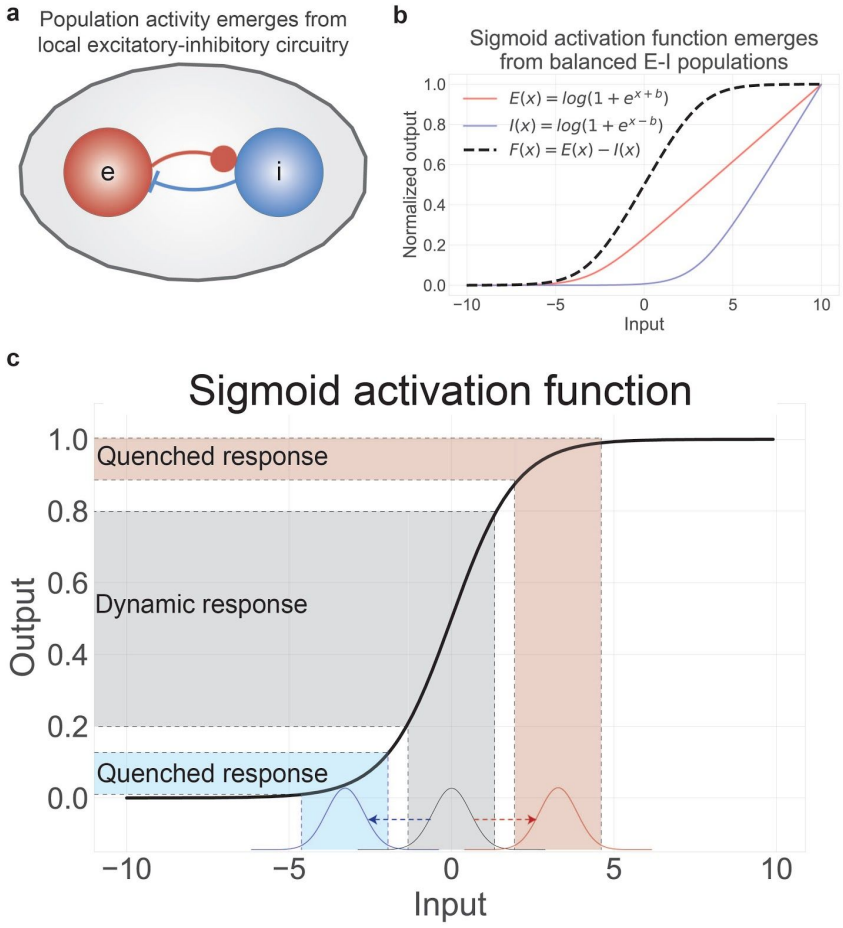

Figure 1. Activation dynamics of neural populations. a) We conceptualize the overall neural population activity as being derived from a loosely balanced excitatory-inhibitory circuit (27). **b)** The neural activation dynamics of an excitatory population subject to loose, inhibitory regulation approximates the activation dynamics of the well-known sigmoid activation function. Excitatory (red) and inhibitory (blue) activation functions are modeled with a softplus function, offset by some bias, *b*, to allow for a dynamic range. The composite function, F(x) = E(x) - I(x), approximates a sigmoid function (black). **c)** A sigmoid activation function produces an optimal dynamic input-output response for a narrow range of inputs (gray). The same input distribution mean shifted by some external stimulation produces a quenched response range, either due to subthreshold (blue) inputs or saturating (red) responses (caused by inhibitory feedback, for example).

Results

Neurophysiologically, a sigmoid activation function implicitly models the neural transfer function of a loosely-balanced E-I population (Fig. 1A,B) (4, 27). The saturation at high levels of input captures the dynamic regimes in which excitatory cells are sensitive to inputs yet runaway excitation is prevented via inhibitory feedback (27). Further, the choice of a sigmoid activation function is intuitive, given that insufficient input into a neural population will fail to excite many (or any) neurons, while extremely strong levels of input will either excite all neurons within the population or activate feedback inhibition (28) (Fig. 1C). This allowed us to focus our modeling

efforts on simplified excitatory networks across large neural populations, mimicking long-range excitatory projections across cortical areas (29).

Careful consideration of the implications of the sigmoid activation function suggests that high or low mean input activity (inputs red and blue, respectively, in Fig. 1C) would quench overall output variability. In turn, this is expected to reduce shared variability (e.g., covariance). Based on these considerations, we hypothesized that during periods in which global neural activity levels are elevated, such as task-evoked states, both neural variability and correlations would be globally quenched.

We first show empirically that task-evoked activity reduces neural correlations and variability in two highly distinct neural data sets: human fMRI and NHP spiking data. This confirms previous findings showing quenched neural variability during task states in both humans and NHPs (1–4), while going beyond those previous studies to report globally quenched inter-region task-state neural correlations. Previous studies have also studied spike count correlations within and between local circuits and neuron pairs (10, 21, 30, 31), but have not characterized these correlations across more than two cortical areas. Following our empirical results, we provide a mechanistic framework using computational simulations and detailed dynamical systems analyses to explain the quenching of neural variability and correlations during task-evoked states.

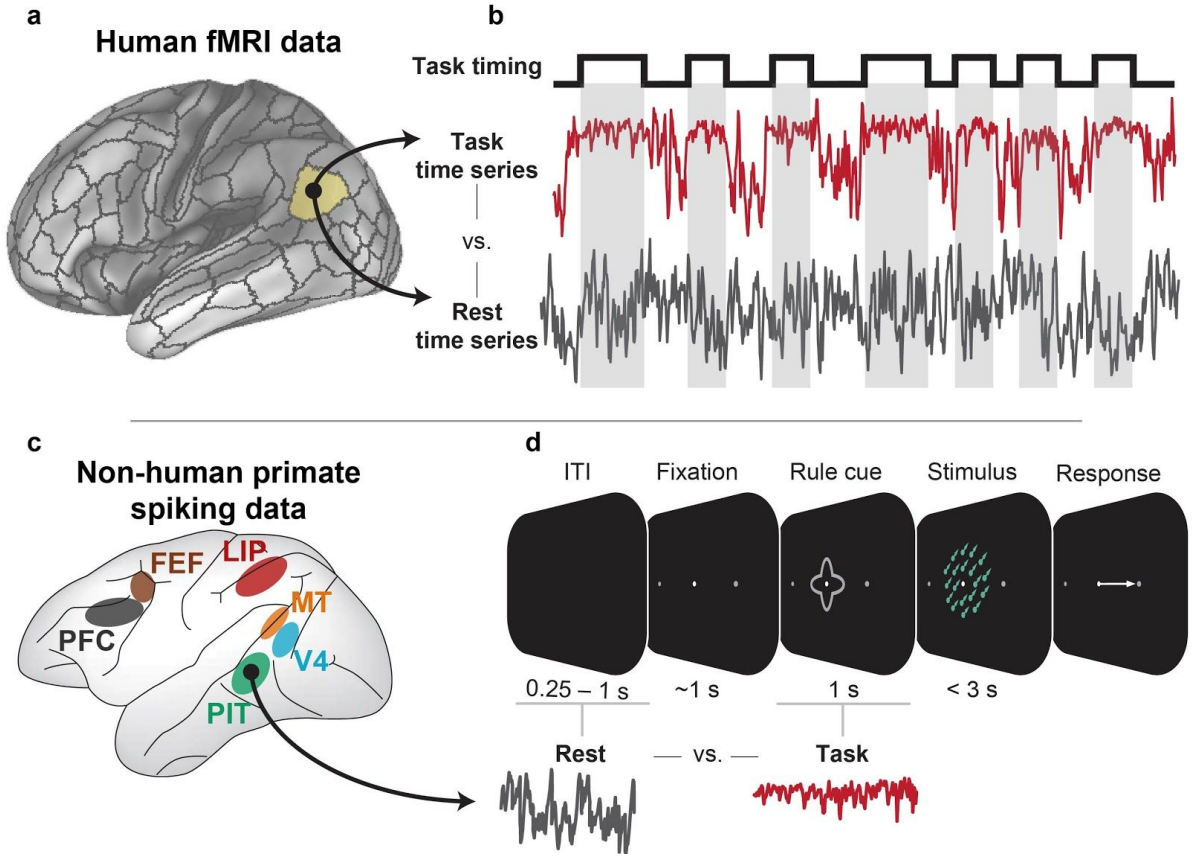

Figure 2. Empirical evidence of task-evoked neural variability and correlation quenching using human fMRI and monkey electrophysiology (spiking) data sets. We used two highly distinct data sets

to support the core hypothesis that task-evoked activity at the population-level quenches neural variability and correlations. **a,b)** Using human fMRI data obtained from the Human Connectome Project (32), we compare the neural variability and correlations (i.e., FC) of the BOLD signal during task block intervals to equivalent intervals during a separate resting-state scan. We compare the task-evoked BOLD variability and correlations to resting-state across seven highly distinct cognitive tasks. Note, time series and task timings illustrated here are used only to demonstrate the analysis approach, and do not reflect actual data (data is generated from simulations). **c,d)** Using multi-unit spiking data collected simultaneously from six different cortical areas (33, 34), we compared the cross-trial spiking variability and spike count correlations between task-state (i.e., following task cue onset) and rest-state spiking activity. We defined rest state as the inter-trial interval (ITI) directly preceding the trial. This was performed using the spike rate averaged within each cortical area, allowing us to target the activity of large neural populations.

Task-evoked activity quenches variability across a wide battery of cognitive tasks in human fMRI data

Consideration of the sigmoid activation function suggests that increased activity associated with task-evoked states would likely quench neural variability (Fig. 1C). Though evidence for these effects have been previously reported in humans and NHPs (1–3), we extend those findings to evaluate variability quenching across seven distinct cognitive tasks in humans using data from the Human Connectome Project (HCP) (32, 35). We leveraged the task block design to calculate the average variability (estimated using the SD across each region's time series) within all task blocks, averaged across tasks and across regions. Consistent with our hypothesis, we found that the variability during task blocks was significantly lower than the variability during equivalent periods of resting-state activity, averaged across the entire brain (exploratory cohort SD difference = -0.018, $t_{175} = -22.36$, $p < 0.0001$; replication cohort SD difference = -0.011, $t_{175} = -11.21$, $p < 0.0001$; Fig. 3A). Moreover, we found that this effect was consistent in 5/7 of the HCP tasks when analyzed independently (5/7 tasks FDR-corrected $p < 0.0001$; Fig. S9A). For the two tasks that did not show global reductions in variability during task states (the Emotion and Motor tasks), task-evoked neural activations was negatively correlated with task-evoked variability reductions across spatial regions (Emotion task, $rho = -0.18$, Motor task, $rho = -0.16$, both FDR-corrected $p < 0.01$; Fig. S9B). This suggests that regional task-evoked activity is associated with task-evoked variability reductions, consistent with our hypothesis and previous accounts (2, 7).

To better understand how global this phenomenon was, we plotted the change in variability from rest to task for each brain region separately. We found that almost all brain regions significantly reduced their variability from rest to task, suggesting that variability reduction is global (cortical maps are thresholded using an FDR-corrected threshold of $p < 0.05$; Fig. 3B,C). This finding extends the work of a previous study in human fMRI data during a finger tapping task (1, 7), suggesting that task-induced variability reduction is a general phenomenon consistent across most cortical regions, and across a wide variety of cognitive tasks.

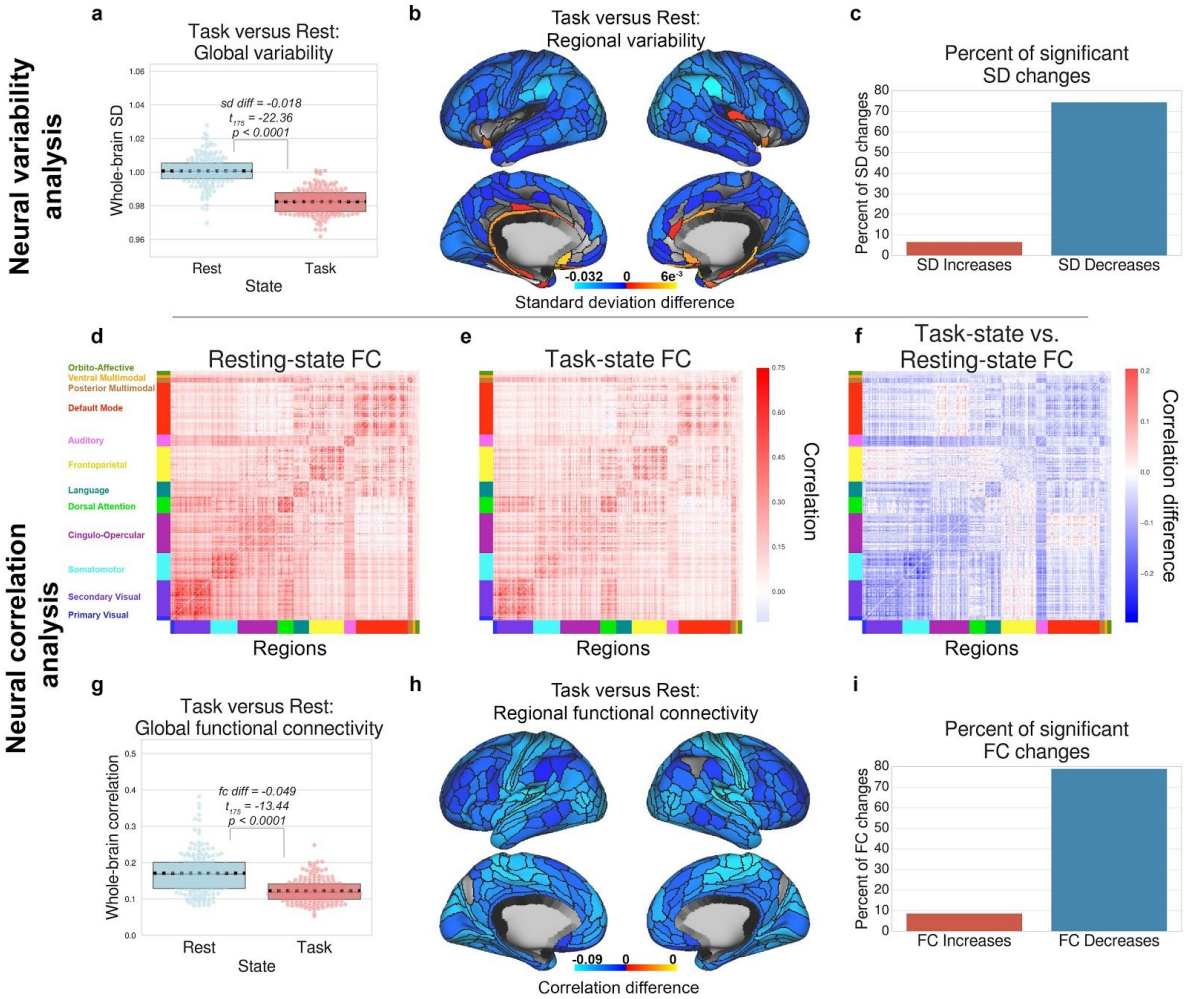

Figure 3. Neural variability and correlations decrease during task states in human fMRI data. Figures for the replication cohort can be found in Fig. S2. Figures for each task separately can be found in Fig. S9-S10. **a-c)** We compared the variability (estimated using the standard deviation of a time series) between task-evoked time windows with equivalent resting-state time windows. We first compare the global variability, which is averaged across all brain regions **(a)**, and then compute the task-state versus resting-state variability difference for each brain region **(b)**. **c)** The percentage of brain regions with significant increases/decreases of neural variability from rest to task states. **d-f)** We compared the neural correlation (i.e., FC) between the exact same task-state time windows used in the variability analyses with equivalent resting-state time windows between all pairs of brain regions. We visualize the correlation matrices for resting state **(d)**, task state **(e)**, and the differences between task state versus resting state **(f)**. **g)** We find that the average FC between all pairs of brain regions is significantly greater during resting state than during task state. **h)** We find that the average correlation for each brain region, or its weighted degree centrality (36), decreases for each brain region from rest to task states. **i)** We calculated the percentage of correlation increases/decreases for all pairs of brain regions. For panels **b, c, e, f, h,** and **i**, plots were thresholded and tested for multiple comparisons using an FDR-corrected $p<0.05$ threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the distribution is visualized using a swarm plot.

Task-evoked activity reduces global FC across a wide battery of tasks in human fMRI data

Despite the accumulation of evidence noting the existence of task-induced FC changes (15–17, 37), the mechanisms for how FC can change remain unclear. Findings in the electrophysiology and computational modeling literature consistently show that spike count correlations decrease during task-evoked states (3, 10, 11, 21, 23, 38). Compounding this contrast, we recently found that task-evoked activations can inflate task FC estimates due to the hemodynamic lag of the BOLD signal (24). Given that a neural population's output variability should be reduced during states with high and low mean activity (Fig. 1C), we hypothesized that when properly accounting for the mean-evoked response (as is commonly done in the spiking literature), shared variability, or FC, would also be globally reduced during task-evoked states.

We first calculated the average FC across all pairwise correlations across all cortical regions for both task and rest states (Fig. 3D,E,F). We found that during task states, the global FC was significantly reduced, relative to resting-state fMRI (exploratory cohort FC diff = -0.049, $t_{175} = -13.44$, $p < 0.0001$; replication cohort FC diff = -0.046, $t_{175} = -13.76$, $p < 0.0001$; Fig. 3G). Recent studies have suggested that the use of correlation provides an ambiguous description of how shared brain signals (relative to unshared variability) change between brain areas (39, 40). Thus, to generalize these results, we also calculated FC using covariance rather than correlation, finding that covariance also globally decreases (covariance diff = -131.94, $t_{351} = -19.23$, $p < 0.0001$ Fig. S6) (39). Task-evoked global FC reductions were also present in each of the 7 HCP tasks separately (all tasks FDR-corrected $p < 0.0001$; Fig. S10A). To identify exactly how global this phenomenon was, we plotted the average task versus rest FC change for each brain region (Fig. 3H,I). We found that nearly all cortical regions significantly reduced their correlation with the rest of cortex during task states.

We hypothesized that decreases in neural variability and correlations are related to task-evoked activity. To provide evidence for this hypothesis, we computed the global brain activity (averaging across all regions). We found that the global activity was significantly greater than baseline across different task states (exploratory cohort, $t_{175} = 6.46$, $p < 0.0001$; replication cohort, $t_{175} = 12.63$, $p < 0.0001$), demonstrating that decreases in neural variability and FC were accompanied by global increases in evoked activity.

Previous work has shown that regions that are more active tend to reduce their variability during task states in fMRI data (2). We sought to replicate this effect in the current data set, while extending those results to demonstrate that more active regions tend to reduce their global FC during task states. We first correlated regional task-evoked activity with task-evoked variability reduction (task SD minus rest SD) across regions at the group-level. We found that regions with more task-evoked activity (averaged across tasks) exhibited greater variability reductions during task states, confirming previous findings in a finger tapping task (exploratory cohort rho=-0.33, *p*<0.0001; replication cohort rho=-0.63, *p*<0.0001; Fig. S1A,C) (2). This negative relationship was also observed in 6/7 of the HCP tasks when analyzed separately (FDR-corrected *p*<0.01; Fig. S9B). To explicitly link regional task activations with FC decreases, we correlated regional task-evoked activity with the regional global FC change during task

states for each region. Consistent with our hypothesis, we found that regions with greater task-evoked activity (averaged across tasks) reduced their average FC more during task states (exploratory cohort rho=-0.21, *p*<0.0001; replication cohort rho=-0.19, *p*=0.0002; Fig. S1B,D). When tasks were analyzed separately, this negative correlation was observed in 4/7 of the HCP tasks (FDR-corrected *p*<0.05; Fig. S10B). Thus, brain areas with increased levels of task-evoked activity tend to reduce both their task-evoked variability and FC.

Previous studies have identified connections that tend to increase FC during task states (37, 41, 42). While we primarily observed global decreases in FC, a small portion of connections increased their FC during task states (exploratory cohort, 8.74% of all connections; replication cohort, 9.85% of all connections; FDR-corrected *p*<0.05) (Fig. 3I, Fig. S2I). However, FC increases were typically limited to cross-network connections with task-positive networks (e.g., the frontoparietal network), where baseline resting-state FC is already quite low (Fig. 3D,E,F). Notably, the sigmoid activation function predicts that neural populations resting with very low activity and/or low connectivity between them (resting in the lower end of the input range) would move up to the linear input range during task, increasing output variability and shared signal variability (e.g., correlations). This is consistent with the hypothesis that task-positive control networks typically have very low activity (and low between-network connectivity) during resting state, but activate during task states. The activation of task-positive networks during task states shift inputs into the dynamic/linear portion of the population's sigmoid activation function (Fig. 1C), increasing shared variability with other brain regions.

Task onset reduces neural variability and correlations across spiking populations in NHPs

Any neural population with dynamics that can be described by a sigmoid activation function, such as large brain regions in NHPs, would likely exhibit quenched neural variability during task-evoked states. Thus, we estimated the spiking variability and spike count correlations of cortical populations in NHPs following task cue onset (task periods) and during the inter-trial intervals (ITI) (rest periods).

We found that across trials, global spiking variability and spike count correlations (r_{sc}) decreased during task as compared to rest (exploratory subject, SD diff = -0.06, $t_{1542} = -49.77$, $p < 0.0001$; r_{sc} diff = -0.02, $t_{14030} = -7.97$, $p < 0.0001$ Fig. 4B,C; replication subject, SD diff = -0.06, $t_{555} = -33.02$, $p < 0.0001$; r_{sc} diff = -0.01; $t_{555} = -2.72$; $p = 0.008$ Fig. S3B,C). These global reductions were also observed using spike count covariance, rather than correlations (Fig. S7). Moreover, we were able to demonstrate variability and correlation decreases within trial (across time points within a trial), demonstrating that task state quenching also occurs on a moment-to-moment basis with spontaneous activity, rather than on just a slower trial-to-trial timescale (Fig. S8). We also measured the spiking variability for each cortical area separately, finding that each cortical area reduced its spiking variability during task states (all six areas for both exploratory and replication subjects, FDR-corrected $p < 0.0001$). Similarly, we found that during task states, the spike count correlation significantly decreased for a majority of cortical areas (FDR-corrected $p < 0.05$; Fig. 4D,E,F, Fig. S3D,E,F).

Though we observed some task-state correlation increases in the exploratory monkey (Fig. 4F), we were unable to replicate any spike count correlation increases in the replication monkey (Fig. S3F). Moreover, we could not demonstrate these increases using covariance (Fig. S7F). This suggests that these correlation increases are not due to increases in shared signal dynamics, but rather are due to decreases in both shared and unshared signal variances (covariance is correlation without variance normalization; Fig. S7) (39, 40, 43). These results are consistent with the results observed in the fMRI data, demonstrating that the variability and correlations of spiking populations largely decrease during task states.

To ensure that correlation and variability decreases were associated with increases in activity (rather than just the task period), we estimated the global spike rate during the task cue interval and the equivalent rest interval. Indeed, we found that the global spike rate during task states was significantly greater than the mean spike rate during rest (exploratory subject, $t_{1542} = 26.16,\ p < 0.0001$, replication subject, $t_{555} = 28.99$, $p < 0.0001$). These findings suggest that task states increase the global neural activity, thereby quenching the global spiking variability reducing neural correlations.

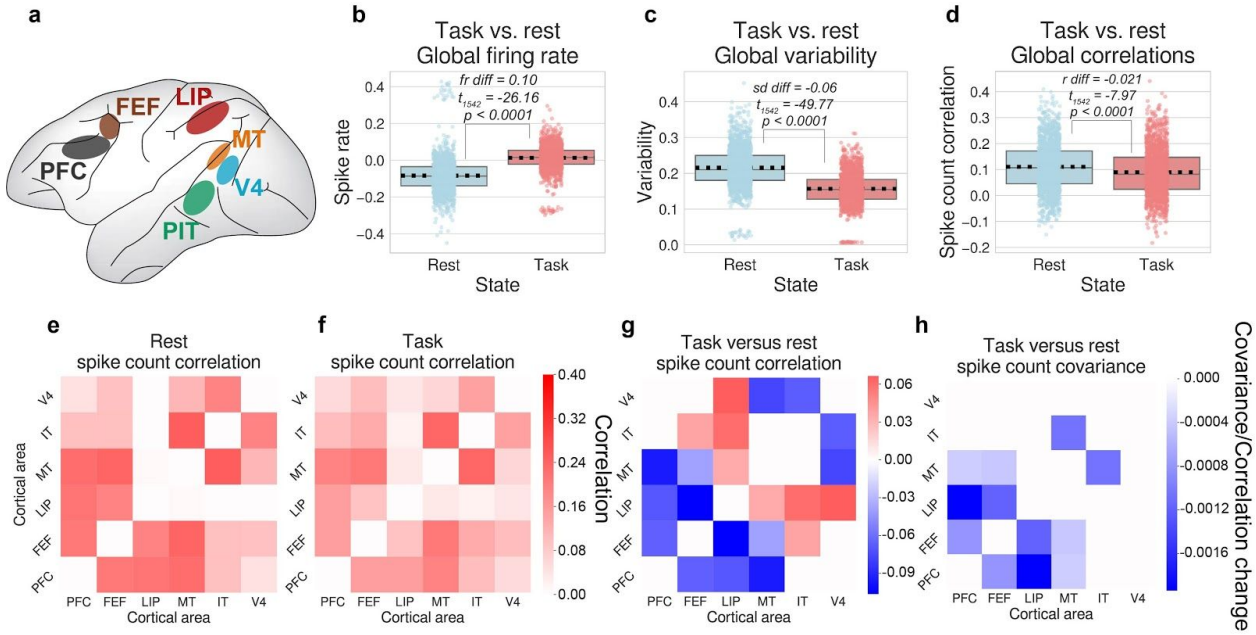

Figure 4. Large-scale neural variability and correlations mostly decrease during task states relative to rest in spiking data. Results for the replication subject are reported in Fig. S3. Full results using covariance (correlation without variance normalization) are reported in Fig. S7. **a)** Multi-unit spike recordings from six different cortical regions fed into our analyses. **b)** We calculated the average spike rate across all 6 recordings during the rest period (ITI) and task period (task cue), across trials. **c)** As in our empirical fMRI data set, we calculated the global variability during task and rest state averaged across regions and estimated using the cross-trial standard deviation. **c)** We calculated the global neural correlation (i.e., the spike count correlation across trials) for task and rest states between all pairs of recorded brain regions. (Spike rates were averaged within each cortical area.) **e-g)** For each pair of brain regions, we visualize the correlation matrices between each recording site for the averaged rest period, task period, and the differences between task versus rest state spike count correlations. Though we found

instances of task-state correlation increases **(h)**, these increases were not found when using covariance (see also Fig. S7). This suggests that these increases are due to reductions in unshared signal variance rather than increases in shared variability (39, 40, 43). For panels **e-h**, plots were thresholded and tested for multiple comparisons using an FDR-corrected p<0.05 threshold. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the full distribution is visualized using a strip plot.

The information-theoretic relevance of task state reduction of neural correlations

Results from our empirical data converge across imaging modalities and species to suggest that increased activity during task states reduces neural variability and correlations globally. However, the theoretical implication of a decreased correlated task state remains unclear. Here we sought to better characterize the information-theoretic implication of a global reduction in neural correlations. In particular, we hypothesized that reductions in neural correlations increase the collective information across units by reducing information redundancy, supporting robust information representations (11). This predicts that the dimensionality of neural activity at the global scale would increase during task states, increasing the collective information in large-scale systems.

We measured the dimensionality (44, 45) of the neural activity (for human fMRI and monkey spiking data) during rest and task states. Intuitively, this is similar to identifying the number of components needed to explain variance greater than some fixed threshold, with more needed components indicating greater dimensionality. Consistent with our hypothesis, we found that task states increased their overall dimensionality relative to rest states (fMRI task versus rest, exploratory cohort difference = 15.35, $t_{175} = 18.74$, $p < 0.0001$, replication cohort difference = 22.48, $t_{175} = 22.48$, $p < 0.0001$; NHP task versus rest, exploratory subject difference = 0.18, $t_{1542} = 13.94$, $p < 0.0001$, replication subject difference = 0.05, $t_{555} = 2.99$, $p = 0.003$) (Fig. 5A,B). This suggests that during task states, the patterns of neural activity reorganize to increase information capacity by reducing redundant information across neural populations. We also found that when analyzing each of the 7 HCP tasks separately, dimensionality increased in all 7 tasks relative to resting-state fMRI (all tasks FDR-corrected $p < 0.0001$; Fig. S11). The present results suggest that task states are associated with a decrease in neural variability and correlations, which increase the effective dimensionality of neural activity. This offers an information-theoretic explanation underlying the reduction in global neural correlations. It will be important for future studies to expand on this result to investigate the behavioral and cognitive relevance of increasing neural dimensionality during task states.

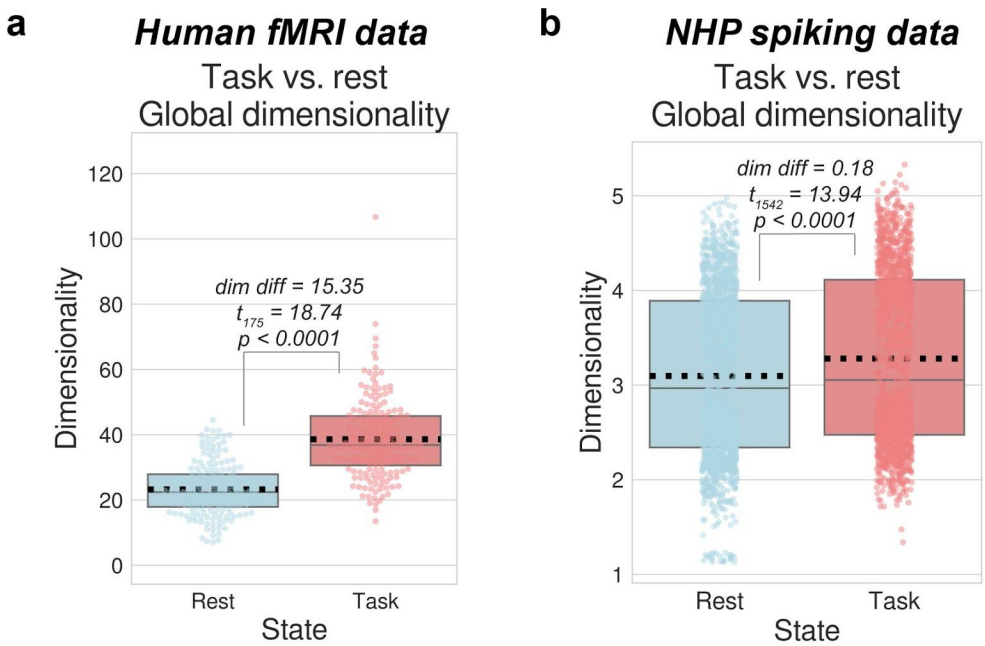

Figure 5. Dimensionality of neural activity increases during task periods relative to resting-state activity. a) For each subject, we calculated the dimensionality of fMRI activity during equivalent resting-state and task state periods, and found that during task states, dimensionality significantly increased. **b)** We calculated the dimensionality of spiking activity across trials and found that during task states, dimensionality significantly increased. These findings provide a potential information-theoretic interpretation of neural correlation reduction during task states. Boxplots indicate the interquartile range of the distribution, dotted black line indicates the mean, grey line indicates the median, and the full distribution is visualized using a swarm plot (in panel A) or strip plot (in panel b).

Neural variability is quenched during task-evoked states in a neural mass model

We proposed – and found empirical evidence supporting – the possibility that especially high/low mean input activity reduces overall output variability due to the nature of a sigmoid neural transfer function (Fig. 1C). Here we rigorously ground those intuitions using neural mass modeling and dynamical systems theory. A recent study by Hennequin and colleagues provided evidence that an evoked stimulus drives neural populations in sensory cortex around a stable fixed point attractor, identifying the dynamical regime in sensory cortex (4). Using a simplified neural mass model, we first extended these findings to identify generalizable dynamical systems mechanisms underlying changes to observed neural variability. This allowed for a comprehensive dynamical systems analysis that is mathematically difficult in higher dimensions. Additionally, this enabled a simpler theoretical approach to investigating changes in neural dynamics that are generalizable across large neural populations (i.e., populations with sigmoidal neural transfer functions, such as brain regions in fMRI and spiking data).

We first characterized the relationship between task-evoked and spontaneous activity in a large-scale neural population using a single neural mass unit. We simulated the neural population's dynamics across a range of fixed input strengths (Fig. 6A). We ran our simulation

across input amplitudes ranging from -5 to 5 in 0.01 increments, finding a nonlinear relationship between stimulus strength and the observed variability of the neural population (Fig. 6C). We found that variability was highest when there was no stimulation, while variability decreased for any type of evoked stimulation (e.g., negative or positive input amplitudes). Despite the model's simplicity, these findings are consistent with previous (and the current) empirical studies demonstrating that task states quench time series variability in both human and animal data (1, 3, 4). Further, this nonlinear relationship between the amplitude of the stimulation and the observed variability of the neural population is consistent with previous results in empirical data and network models (1, 4, 5, 7, 8). However, previous models investigating variability reduction focus on large-scale network mechanisms, which are difficult to analyze using dynamical systems tools. Therefore, we leveraged the simplicity of our model to determine the causes of neural variability reduction in terms of dynamical systems mechanisms.

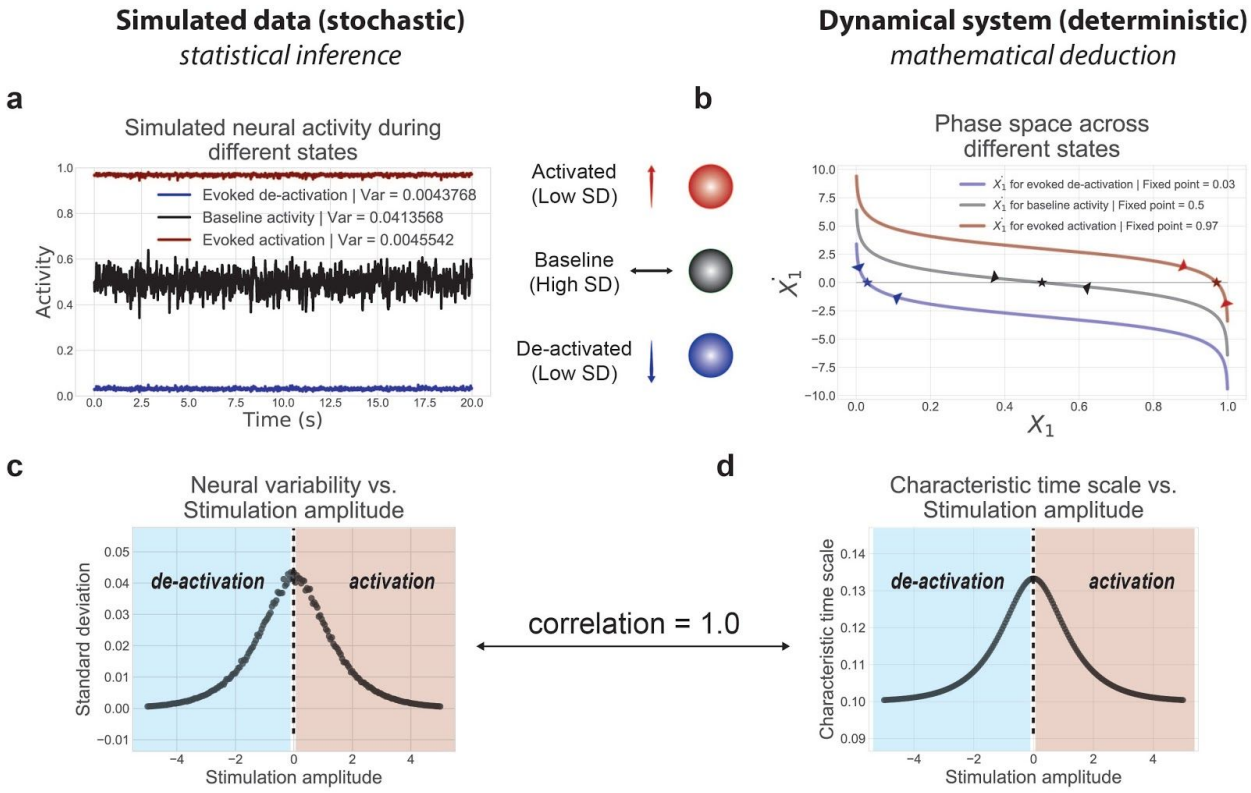

Figure 6. Evoked activity induces changes in neural variability and the underlying attractor dynamics. Our minimal modeling approach leverages the mathematical formalisms of firing rate networks to directly link descriptive neural statistics from the stochastic simulation (e.g., time series variability) with rigorous dynamical systems analysis (e.g., attractor dynamics). **a)** During different evoked states (i.e., fixed inputs), there is a reduction in the observed time series variability (measured by SD). This is directly related to how input-output responses change due to the changing slope in the nonlinear sigmoid transfer function (Fig. 1C). **b)** To more rigorously analyze this, we visualized the phase space for each of the neural populations according to state by plotting the derivative of X_1 denoted by $\dot{X}_1$. For each state, we calculated the fixed point attractor (plotted as a star), denoting the level of mean activity the system is

drawn to given some fixed input (or absence thereof). Arrows denote the direction/vector toward each fixed point, which specify the characteristic time scale (i.e., the speed) the system approaches the fixed point (46). **c)** We ran simulations across a range of stimulation amplitudes, ranging from -5 to 5 in .01 increments. For each stimulation amplitude, we calculated the SD across time to measure the neural variability. **d)** We characterized the shifting attractor dynamics for each stimulus by computing the characteristic time scale at the fixed point for each stimulation amplitude. The characteristic time scale across all fixed points is perfectly correlated with the neural variability of the simulated time series across all fixed inputs (rank correlation = 1.0).

We sought to leverage the model's simplicity to characterize dynamical systems properties governing the emergence of the observed neural variability. Identifying a dynamical systems mechanism underlying the observed change in descriptive statistics would provide rigorous evidence that shifting the underlying attractor dynamics alters the observed neural signals. We first performed a state space analysis in one dimension to identify the stable fixed point attractor (i.e., the equilibrium level of activity the system is drawn to during a particular state in the absence of spontaneous activity) for the intrinsic and evoked states (Fig. 6B). The state space view enabled visualization of the system's full dynamics across different evoked states (Fig. 6B). For example, dynamics around the fixed point attractor in the intrinsic baseline state appeared to approach equilibrium slowly. This can be identified by observing the angle where the curve intersects 0 on the y-axis (i.e., when $\dot{x} = 0$; Fig. 6B). The angle of this curve corresponds to the *characteristic time scale*, a dynamical property characterizing the speed with which the system approaches the attractor (a higher value reflects a slower approach; see Methods) (46).

To quantify this more rigorously, we performed a linear stability analysis around the fixed point attractor of the system across the same range of stimulation amplitudes, and analytically calculated the characteristic time scale at each fixed point. Again, we found a nonlinear relationship between the amplitude of the stimulus and the characteristic time scale of the neural population (Fig. 6D), and found that the characteristic time scale explained 100% of the variance of the simulated stimulus-evoked variability (Fig. 6C). This is consistent with previous empirical accounts of NHP spiking data (4), rodent calcium imaging data (47), and human fMRI data (1, 48), suggesting that intrinsic states are characterized by high-amplitude slow fluctuations, while evoked states are characterized by low-amplitude fast fluctuations. These results demonstrate that changes in observed neural variability can be directly attributed to changes in the underlying attractor dynamics. Previous work has suggested that such changes are biologically implemented by feedback inhibition in a loosely balanced E-I circuit (4, 6, 49).

To ensure that the model explanation would generalize to data obtained on a slower time scale (e.g., fMRI BOLD data), we transformed the simulated neural activity into fMRI BOLD activity using the Balloon-Windkessel model (50). The Balloon-Windkessel is a nonlinear transformation of neural activity to model the BOLD signal that takes into account the normalized blood volume, blood inflow, resting oxygen extraction fraction, and the normalized deoxyhemoglobin content. Consistent with previous accounts (2), we found that the characteristic time scale around the fixed point attractor was still strongly correlated with BOLD signal variability (rho=0.97; $p<0.0001$; Fig. S4).

The dynamical mechanisms identified in our model suggest that attractor dynamics are slow during rest states but fast during task states. In addition, the biological plausibility of the sigmoid activation function for large neural populations enables biological interpretation. Specifically, evoked activity brings the neural population closer to a state of quenched output variability due to saturation of the sigmoid function (see Fig. 1C). Similarly, evoked de-activation brings a neural population's activity closer to a subthreshold state, which also quenches the output variability due to low input activity preventing enough elements in the population to activate. The quenching of output variability can potentially emerge due to different biological mechanisms such as a tightening of E-I balance due to inhibition, neural adaptation, and/or irregular synaptic vesicle release (4, 6, 38, 51). The manifestation of these biological mechanisms can be summarized by the reduction of the response variability due to the decreased slope in the sigmoid transfer function during highly active or inactive states (Fig. 1C).

Neural correlations are quenched during task states in a network model

We generalize the dynamical systems analysis in one dimension to two dimensions, allowing us to focus on correlations between neural populations. We show illustrations of the state space for intrinsic and activated states (Fig. 7B,D), as well as the corresponding time series (Fig. 7A,C) for our model. Induced negative activity produces qualitatively similar results to the activated state (Fig. 7D) due to subthreshold levels of activity rather than saturating levels of activity.

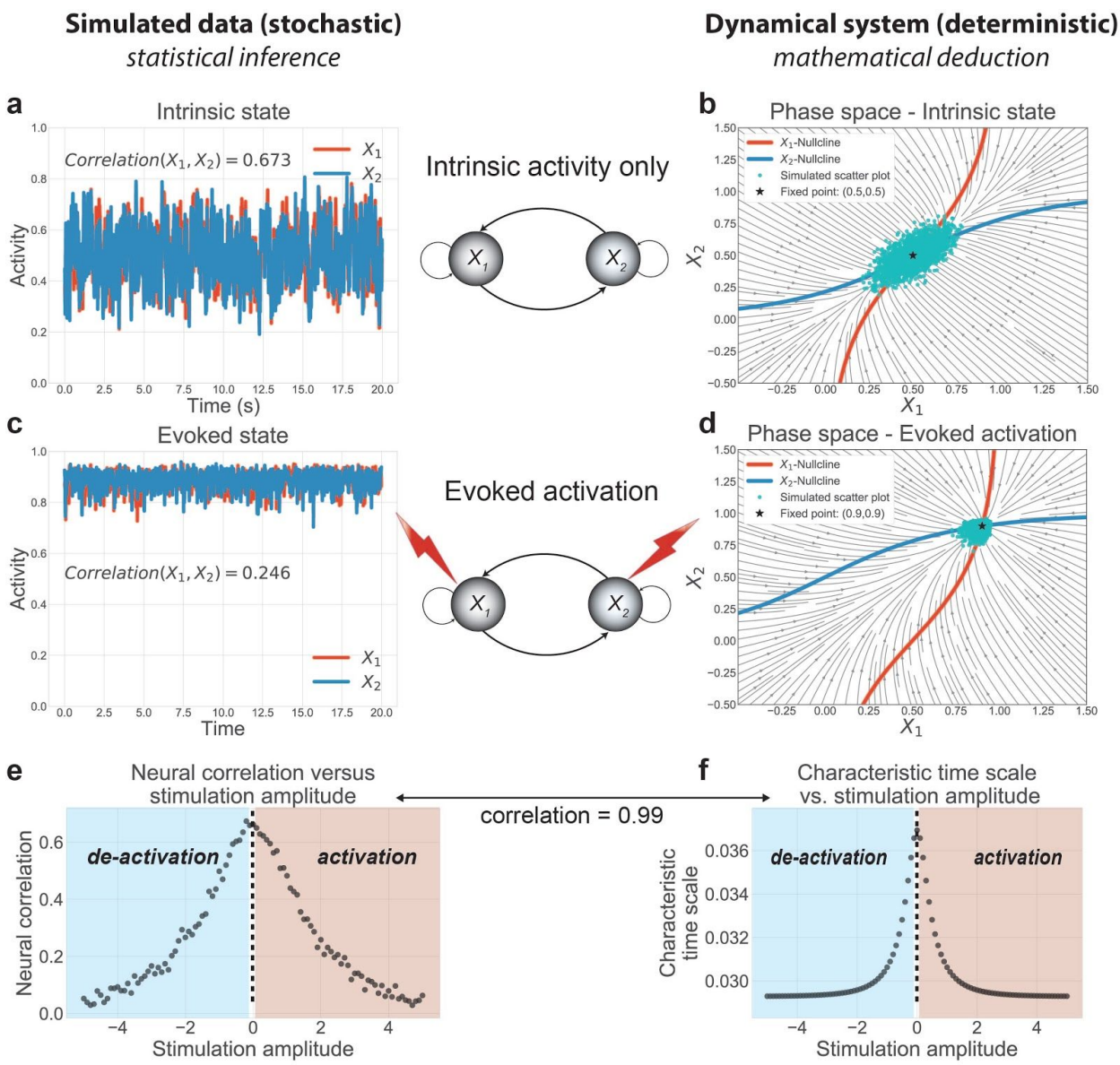

Figure 7. Evoked activity induces changes in neural correlations and the underlying attractor dynamics. We use a two unit network model, the minimal model necessary to study dynamic changes in neural correlations. **a)** At baseline, we observe slow, high amplitude fluctuations and high neural correlations, consistent with findings across data sets and animal models (1, 3, 5, 48). **b)** To characterize the underlying attractor dynamics, we visualized the two-dimensional state space, visualizing the flow field and the nullclines (0-growth isoclines, where the rate of change is 0) for each unit. The intersection of the two nullclines (blue and red curves) denote the fixed point attractor. We overlay the simulated scatter plot (cyan dots) to demonstrate the correspondence between the attractor dynamics and simulation. **c)** We inject a fixed input stimulation, shifting the network to an 'evoked' state, causing a decrease in neural variability and correlation. **d)** The external input transiently moves the fixed point, altering the underlying attractor dynamics and the corresponding scatter plot. **e)** As in Fig. 6, we systematically vary the input strength for a range of inputs, calculating the change in the neural correlations. Neural correlations are optimal with no external stimulation, and decrease with any external stimulation. **f)** For a given stimulation strength, we calculated the generalized characteristic time scale (see Methods) at the fixed point, thereby

characterizing the attractor dynamics in the absence of noise. We found that the system's characteristic time scale across stimulation strengths explained 98% of the neural correlation variance, providing a direct association between the network's attractor dynamics and observed neural correlations.

The state space analysis (Fig. 7B,D) allowed us to track the simultaneous evolution of the two neural masses, providing a geometric interpretation of the system (52, 53). We observed qualitatively that shifts in the underlying dynamics due to stimulation (i.e., changes to the flow field) were directly associated with changes to the observed scatter plot, resulting in correlation changes between the two neural masses. Specifically, we observed that intrinsic state dynamics supported slower, elongated trajectories along a diagonal axis, consistent with the emergence of correlated neural activity between the two populations (Fig. 7B). This is due to a slower characteristic timescale near the fixed point attractor, which corresponds to eigenvalues with smaller magnitudes. In contrast, during evoked states, the system approached the fixed point attractor at a faster speed (i.e., faster characteristic timescale), quenching trajectories in state space that support correlated variability between the two neural masses (Fig. 7D). Thus, the visualization of the state space flow field demonstrates that changes in neural correlations are associated with changes to the underlying flow field dynamics around the fixed point attractor.

To more carefully test the relationship between state-dependent neural correlations, we simulated our network model across a range of stimulation amplitudes. Despite no changes to the network's connectivity structure, we found that neural correlations systematically changed (primarily decreased) as a function of evoked stimulation (Fig. 7E). Further, using dynamical systems analysis, we found that a generalization of the characteristic time scale in higher dimensions could account for changes in neural correlations (rho=0.99; p<0.0001; Fig. 7F). In other words, we could analytically determine that evoked stimulation shifted the underlying attractor dynamics, giving rise to changes in neural correlations in a network model with fixed connections. We found consistent results after transforming the neural activity to fMRI BOLD activity using the Balloon-Windkessel model (50), finding that changes to the characteristic time scale could still account for changes in BOLD correlations (rho=0.97; *p*<0.0001; Fig. S5). These results were also reproduced using mutual information instead of correlation, a nonlinear measure of statistical dependence (54). Specifically, evoked states decreased the amount of mutual information between the two neural populations relative to baseline, and was associated with a faster characteristic time scale around the fixed point attractor (rho=0.94; p<0.0001). This suggests that the quenching of shared variance encompasses both linear and nonlinear measures of statistical dependencies.

The computational and analytical results presented here provide a rigorous theoretical basis that explain neural variability and correlation quenching associated with task-evoked activity observed in empirical data.

Discussion

The present results suggest that task-evoked neural activity globally quenches neural time series variability and correlations. Based on evidence that large neural populations activate

in a sigmoid-like fashion (28), we hypothesized that task-evoked activity would quench global neural variability and correlations. We showed in two distinct empirical data sets that task-evoked neural activity was associated with global neural variability and correlation reductions, illustrating the generality of the phenomena across human fMRI and NHP electrophysiology data. We subsequently provided a dynamical systems model to demonstrate that evoked activity strengthened the system's attractor dynamics by decreasing the characteristic time scale (a deterministic property capturing the speed at which the system approaches a fixed point attractor), which quenched the observed neural variability and correlations. This provided a mechanistic framework to interpret the empirical results. Importantly, the use of a sigmoid activation function revealed a simple interpretation underlying neural variability and correlation suppression widely applicable to many types of neural data. During task states, the slope of the neural transfer function decreases (e.g., via feedback inhibition or neuronal adaptation), reducing the dynamic range of input-output responses. This results in reduced overall output variability, as well as reduced shared variability (e.g., correlations) from connected neural populations. The collective empirical and theoretical results provide strong evidence that observed neural variability and correlations are state-dependent, and can be attributed to interpretable changes in the activity dynamics governed by the transfer functions of neural populations.

Previous studies have provided anecdotal evidence to characterize neural variability and correlations during rest and task states in both large-scale brain systems and local neural populations (3, 10, 11, 15, 55). Studies investigating state-dependent changes in neural variability have found consistent results, demonstrating that task states quench neural variability in animal spiking data (3), rodent calcium imaging data (47), and human fMRI (1, 2) as well as intracranial field potentials recorded in humans (56). However, studies characterizing task-evoked neural correlation changes have found conflicting effects across the human and animal literature. Studies in the human fMRI literature have reported correlation increases across networks during tasks, with those studies suggesting that increased/decreased correlations facilitate better/worse communication, respectively (17, 19, 20). In contrast, studies in local spiking populations have found mostly correlation decreases in spike rates during task states, consistent with the hypothesis that correlation decreases support more robust information coding (10, 11). To adjudicate between these two perspectives, we used theoretical modeling and empirical data analysis from two distinct data modalities to clarify the interpretation of correlation changes from rest to task states in large-scale neural systems. Across different data sets and different task states, we found mostly task-related correlation decreases that coincided with an increase in the neural dimensionality. Moreover, our results converged using both correlation and covariance, suggesting that these decreases were due to reductions in shared signal dynamics rather than changes in unshared signal variance (or noise) (39, 40, 43). Thus, our results support the latter hypothesis, consistent with the notion that during task states neural systems reduce their shared signaling to represent more information more robustly as has been previously proposed (1, 11).

Empirical studies in large-scale functional networks using fMRI have reported FC increases during task states (20). More broadly, results from the fMRI literature suggest that the global network architecture transitions from a state of segregation to integration during task

states characterized by within-network FC decreases and/or cross-network FC increases (15–17). However, we recently found that task-evoked activity inappropriately inflates FC estimates if the mean-evoked activity response is not properly removed (24). Indeed, when properly accounting for the mean-evoked response, we found that FC changes from rest to task states were dominated by FC decreases (see Fig. 3I). This finding is consistent with a recent study measuring task-evoked and resting-state FC differences during a movie watching paradigm, where FC was measured during long continuous epochs (which remove the transient onset of the BOLD response which artificially inflates correlations) (57). The correction of the mean-evoked response in our paradigm (24) brought the empirical results in line with our modeling results, suggesting a counterintuitive interpretation of FC changes during tasks: co-activation in the presence of FC decreases between pairs of regions is consistent with high (or saturating levels of) interaction between those nodes. This can be understood in the context of the two-unit neural mass model: If two regions are connected at rest and both are highly active on task, the high levels of input into sigmoid transfer function of both units reduces the overall shared signal variability of the two units. In turn, this reduces the inter-region correlation and covariance, suggesting that observed decreases are a result of the quenching of shared signal dynamics rather than increases in noise variance (39, 40, 43). This interpretation is consistent with the notion that resting-state functional networks, which typically co-activate, decrease their within-network FC during tasks (15–17, 58). However, future work will need to directly evaluate this hypothesis.

Though we largely observed FC decreases during task states in human fMRI data, we identified a small number of correlations that increased during task (see Fig. 3F). We also identified a few brain regions that increased variability during task (see Fig. 3C). This is likely due to the fact that regions that are inactive during rest have low shared and overall variability. Thus, any increase in task-evoked activity will push such a population's input-output responses from a subthreshold to a linear response range (see Fig. 1C), supporting some increases in shared and overall variability (3, 41). Similar findings have also been reported in the NHP spiking literature, where spike count correlations between regions with low baseline correlations have been found to increase across regions with attentional modulation (31, 42). Our results corroborate some of these findings, demonstrating that there are few pairs of regions that increase their correlations during tasks. However, we were unable to find any increases in spike count covariance (rather than correlations) in our NHP data, suggesting that some of the identified correlation increases are due to decreases in unshared variance, rather than increases in shared variance (39, 40).

A recent computational study investigating the role of neurotransmitter modulation (e.g., norepinephrine) has demonstrated that increases in neural gain can facilitate the transition between states of segregation and integration (59). Though our model was not designed to to account for complex network topology, our model suggests that regions that are both co-active during task and strongly connected at rest would decrease their correlations, consistent with FC decreases within network. This is consistent with results from Shine et al., given that strong within-network FC decreases corresponds to increases in global participation coefficients and network integration. Further, increases in neural gain (as implemented by Shine and colleagues) is dynamically consistent with FC decreases: increasing neural gain shrinks the dynamic range

of a population's activation function, which results in a reduction of output variability relative to the input (see Fig. 1C). Thus, while we expect similar outcomes with our particular model formulation and previous work, it will be important for future work to determine how the dynamical mechanisms investigated here extend to networks with complex and heterogeneous connectivity topologies.

The present results may appear to contradict some reports that task engagement increases (rather than decreases) overall neural communication (18, 60). Yet there are several key differences between those previous findings and the present results. First, many of the previous results focused on communication through coherence, which often involves frequency-specific coupling of neural signals. This involves the phase-alignment of neural activity on faster timescales, which relates only indirectly to the slower correlation measures of spike rate activity and metabolic demand focused on here (61, 62). Thus, it is possible that phase-based measures of coupling (also using magnetoencephalography and electroencephalography) (63–65) can indicate that large-scale neural populations increase their communication during tasks relative to rest, without contradicting the present results. A second key difference between the present and most previous results is our emphasis on the absolute amount of correlation change from rest to task, rather than changes in network organization. Focusing on changes in network organization effectively baselines to the overall level of FC in each brain state. Several recent studies focusing on network reorganization from rest to task found that inter-network-community FC increased (18). While counterintuitive, these results are consistent with the present results, given that the global average FC can decrease while increasing inter-network-community FC in a relative way. This is possible because within-network-community FC can decrease more than inter-network-community FC, yielding an overall FC decrease. Importantly, the present results suggest a reinterpretation of those previous findings, in which apparent inter-network-community FC increases coincide with overall FC reductions during tasks, which increases network dimensionality and overall capacity for information processing.

In conclusion, we constructed a mechanistic framework for interpreting changes in neural variability and correlations by investigating the effects of evoked activity on the underlying neural attractor dynamics. Using theoretical modeling and empirical data analysis across two highly distinct neural data sets, we demonstrated convergent evidence suggesting that task states quench neural variability and correlations likely due to strengthening neural attractor dynamics across large-scale neural systems. Our work extends previous research establishing similar attractor mechanisms in sensory cortex (4) to characterize the role of attractor dynamics across large-scale neural systems. We expect these findings to spur new investigations to better understand how we can best leverage neural correlation analyses to infer processes governed by neural attractor dynamics, providing a deeper understanding of the causal processes in the brain.

Methods

fMRI: Data and paradigm

The present study was approved by the Rutgers University institutional review board. Data were collected as part of the Washington University-Minnesota Consortium of the Human Connectome Project (HCP) (35). A subset of data (*n* = 352) from the HCP 1200 release was used for empirical analyses. Specific details and procedures of subject recruitment can be found in (35). The subset of 352 participants was selected based on: quality control assessments (i.e., any participants with any quality control flags were excluded, including 1) focal anatomical anomaly found in T1w and/or T2w scans, 2) focal segmentation or surface errors, as output from the HCP structural pipeline, 3) data collected during periods of known problems with the head coil, 4) data in which some of the FIX-ICA components were manually reclassified; low-motion participants (i.e., exclusion of participants that had any fMRI run in which more than 50% of TRs had greater than 0.25mm framewise displacement); removal according to family relations (unrelated participants were selected only, and those with no genotype testing were excluded). A full list of the 352 participants used in this study will be included as part of the code release.

All participants were recruited from Washington University in St. Louis and the surrounding area. We split the 352 subjects into two cohorts of 176 subjects: an exploratory cohort (99 women) and a replication cohort (84 women). The exploratory cohort had a mean age of 29 years of age (range=22-36 years of age), and the replication cohort had a mean age of 28 years of age (range=22-36 years of age). All subjects gave signed, informed consent in accordance with the protocol approved by the Washington University institutional review board. Whole-brain multiband echo-planar imaging acquisitions were collected on a 32-channel head coil on a modified 3T Siemens Skyra with TR=720 ms, TE=33.1 ms, flip angle=52º, Bandwidth=2,290 Hz/Px, in-plane FOV=208x180 mm, 72 slices, 2.0 mm isotropic voxels, with a multiband acceleration factor of 8. Data for each subject were collected over the span of two days. On the first day, anatomical scans were collected (including T1-weighted and T2-weighted images acquired at 0.7 mm isotropic voxels) followed by two resting-state fMRI scans (each lasting 14.4 minutes), and ending with a task fMRI component. The second day consisted with first collecting a diffusion imaging scan, followed by a second set of two resting-state fMRI scans (each lasting 14.4 minutes), and again ending with a task fMRI session.

Each of the seven tasks was collected over two consecutive fMRI runs. The seven tasks consisted of an emotion cognition task, a gambling reward task, a language task, a motor task, a relational reasoning task, a social cognition task, and a working memory task. Briefly, the emotion cognition task required making valence judgements on negative (fearful and angry) and neutral faces. The gambling reward task consisted of a card guessing game, where subjects were asked to guess the number on the card to win or lose money. The language processing task consisted of interleaving a language condition, which involved answering questions related to a story presented aurally, and a math condition, which involved basic arithmetic questions presented aurally. The motor task involved asking subjects to either tap their left/right fingers,

squeeze their left/right toes, or move their tongue. The reasoning task involved asking subjects to determine whether two sets of objects differed from each other in the same dimension (e.g., shape or texture). The social cognition task was a theory of mind task, where objects (squares, circles, triangles) interacted with each other in a video clip, and subjects were subsequently asked whether the objects interacted in a social manner. Lastly, the working memory task was a variant of the N-back task.

Further details on the resting-state fMRI portion can be found in (66), and additional details on the task fMRI components can be found in (32). All fMRI results reported in the main article reflect results found with the first cohort of subjects. Independent replication of these effects are reported in Fig. S2 with the replication cohort.

fMRI: Preprocessing

Minimally preprocessed data for both resting-state and task fMRI were obtained from the publicly available HCP data. Minimally preprocessed surface data was then parcellated into 360 brain regions using the (67) atlas. We performed additional preprocessing steps on the parcellated data for resting-state fMRI and task state fMRI to conduct neural variability and FC analyses. This included removing the first five frames of each run, de-meaning and de-trending the time series, and performing nuisance regression on the minimally preprocessed data (68). Nuisance regression removed motion parameters and physiological noise. Specifically, six primary motion parameters were removed, along with their derivatives, and the quadratics of all regressors (24 motion regressors in total). Physiological noise was modeled using aCompCor on time series extracted from the white matter and ventricles (69). For aCompCor, the first 5 principal components from the white matter and ventricles were extracted separately and included in the nuisance regression. In addition, we included the derivatives of each of those components, and the quadratics of all physiological noise regressors (40 physiological noise regressors in total). The nuisance regression model contained a total of 64 nuisance parameters. This was a variant of previously benchmarked nuisance regression models reported in (68).

We excluded global signal regression (GSR), given that GSR artificially induces negative correlations (70, 71), which could bias analyses of whether global correlations decrease during task performance. We included aCompCor as a preprocessing step here given that aCompCor does not include the circularity of GSR (regressing out some global gray matter signal of interest) while including some of the benefits of GSR (some extracted components are highly similar to the global signal) (72). This logic is similar to a recently-developed temporal-ICA-based artifact removal procedure that seeks to remove global artifact without removing global neural signals, which contains behaviorally relevant information such as vigilance (73, 74). We extended aCompCor to include the derivatives and quadratics of each of the component time series to further reduce artifacts. Code to perform this regression is publicly available online using python code (version 2.7.15) (https://github.com/ito-takuya/fmriNuisanceRegression). Following nuisance regression, the time

series for each run (task-state and rest-state) were z-normalized such that variances across runs would be on the same scale (i.e., unit variance).

Task data for task FC analyses were additionally preprocessed using a standard general linear model (GLM) for fMRI analysis. For each task paradigm, we removed the mean evoked task-related activity for each task condition by fitting the task timing (block design) for each condition using a finite impulse response (FIR) model (24). (There were 24 task conditions across seven cognitive tasks.) We used an FIR model instead of a canonical hemodynamic response function given recent evidence suggesting that the FIR model reduces both false positives and false negatives in the identification of FC estimates (24). This is due to the FIR model's ability to flexibly fit the mean-evoked response across all blocks. Removing the mean-evoked response of a task condition is critical to isolate second order effects such as correlations, and is standard when performing in spike count correlations (10, 24). Analogous statistical preprocessing steps were critical when comparing neural correlation measures across human fMRI data and NHP spiking data.

FIR modeled task blocks were modeled separately for task conditions within each of the seven tasks. In particular, two conditions were fit for the emotion cognition task, where coefficients were fit to either the face condition or shape condition. For the gambling reward task, one condition was fit to trials with the punishment condition, and the other condition was fit to trials with the reward condition. For the language task, one condition was fit for the story condition, and the other condition was fit to the math condition. For the motor task, six conditions were fit: (1) cue; (2) right hand trials; (3) left hand trials; (4) right foot trials; (5) left foot trials; (6) tongue trials. For the relational reasoning task, one condition was fit to trials when the sets of objects were matched, and the other condition was fit to trials when the objects were not matched. For the social cognition task, one condition was fit if the objects were interacting socially (theory of mind), and the other condition was fit to trials where objects were moving randomly. Lastly, for the working memory task, 8 conditions were fit: (1) 2-back body trials; (2) 2-back face trials; (3) 2-back tool trials; (4) 2-back place trials; (5) 0-back body trials; (6) 0-back face trials; (7) 0-back tool trials; (8) 0-back place trials. Since all tasks were block designs, each time point for each block was modeled separately for each task condition (i.e., FIR model), with a lag extending up to 25 TRs after task block offset.

fMRI: Task state activation analysis

We performed a task GLM analysis on fMRI task data to evaluate the evoked global brain activity during task states. The task timing for each of the 24 task conditions was convolved with the SPM canonical hemodynamic response function to obtain task-evoked activity estimates (75). FIR modeling was not used when modeling task-evoked activity. Coefficients were obtained for each parcel in the Glasser et al. (2016) cortical atlas for each of the 24 task conditions.

fMRI: Task state versus resting-state variability analysis

To compare task state versus resting-state variability, we regressed out the exact same FIR design matrix used on task-state regression on resting-state data. This was possible given that the number of timepoints of the combined resting-state scans in the HCP data set exceeded the number of timepoints of the combined task-state scans (4800 resting-state TRs > 3880 task-state TRs). This step was to ensure that any spurious change to the residual task time series through the removal of the mean-evoked response through FIR regression would also induce spurious changes in the resting-state data. However, results were qualitatively identical without the regression of the task design matrix on resting-state data.

Post-regression, we obtained equivalently lengthed residual time series for both resting-state and task state fMRI data. To calculate the average variability during task blocks, we first z-normalized each task run with zero-mean and unit variance such that we could appropriately compare the neural variability of task blocks across different runs. Given the hemodynamic response is temporally lagged from stimulus onset, we convolved the task timing time series (i.e., boxcar time series with amplitude=1) with a canonical hemodynamic response function and identified time points exceeding 0.5. This ensured that our measure of variability did not include the lagged and transient onset of the BOLD response. We then extracted those time windows, computed the standard deviation (SD) for each task condition, and then averaged across all task conditions to obtain our statistic of task-evoked neural variability across all task conditions (separately for each region, for each subject). To identify the resting-state neural variability, we applied the same exact procedure to resting-state time series using the task state time series. A sanity check for our analysis was that the 'intrinsic-state' neural variability is close to 1 (given that the time series was normalized to have unit variance), while the task state neural variability is significantly less than 1 (Fig. 3A).

We compared the neural variability of the entire brain during task state periods versus resting-state periods. For each subject, we computed the SD during task and resting state separately, and then averaged across all brain regions. This resulted in two values per subject, representing task state and resting-state variability. We then performed a two-way group paired t-test across subjects to assess statistical significance (Fig. 3A). We also computed the task state versus resting-state difference in neural variability for each brain region separately (Fig. 3B). We corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$ (Fig. 3B,C). Cortical surface visualizations were constructed using Connectome Workbench (version 1.2.3) (35).

fMRI: Task state versus resting-state correlation analysis

We compared task state versus resting-state FC (i.e., neural correlations), after performing the exact same preprocessing steps as mentioned above. This included the regressing out of the FIR design matrix on task state data for each task condition, z-normalization of runs to zero-mean and unit variance, and the selection of task timed windows for each task condition using a temporally lagged window which accounts for the temporally

lagged hemodynamic response function. The same exact post-processing procedure was performed on resting-state data.

We computed the correlation between all pairs of brain regions for each task condition using the isolated task blocks. We then averaged the Fisher's *z*-transformed correlation values across all task conditions to obtain a general task state FC matrix (Fig. 3E). We repeated the same procedure (i.e., using the same task-timed blocks) on resting-state FC to obtain an equivalent resting-state FC matrix for each subject (Fig. 3D). We directly compared task-state FC to resting-state FC by performing two-way group paired t-tests for every pair of brain regions using the Fisher's *z*-transformed correlation values. Statistical significance was assessed using an FDR-corrected threshold of $p<0.05$ (Fig. 3F). To compare the average global correlation during task state and resting-state, we computed the average correlation between all pairs of brain regions during task and resting-state, performing a group paired t-test (Fig. 3G). To compare the average global connectivity profile of every brain region (Cole et al., 2010), we computed the average Fisher *z*-transformed correlation of a single region to all other brain regions during task and rest and performed a two-way group paired t-test between task and rest (Fig. 3H,I). Statistical significance was assessed using an FDR-corrected threshold of $p<0.05$.

Spiking data: Data collection

The behavioral paradigm for each monkey was a motion-color categorization task (Fig. 2D). Experimental methods for electrophysiology data collected for NHP was previously reported in (33) and (34). Data was collected *in vivo* from two (one female) behaving adult rhesus macaques (*Macaca mulatta*) across 55 sessions. Data from six distinct cortical regions were recorded simultaneously from acutely inserted electrodes. Cortical regions included: MT, V4, PIT, LIP, FEF, and LPFC (Fig. 2C). Spikes from each region were sorted offline into isolated neurons. However, given our interest in inter-region neural correlations across large scale neural systems, we pooled spikes from each functional area into a single spike rate time series. For each trial, spikes were sorted for a 5s period, beginning 2.5s prior to stimulus onset, and until 3.5s after stimulus onset. Further details regarding electrophysiological data collection can be found here:
http://www.sciencemag.org/content/348/6241/1352/suppl/DC1 and here:
http://www.pnas.org/content/pnas/suppl/2018/07/09/1717075115.DCSupplemental/pnas.1717075115.sapp.pdf

All statistical analyses in the main article (detailed below) were performed on a single monkey. Independent replication was performed on the second monkey, and is reported in Fig. S3.

Spiking data: Task versus rest variability analysis

Neural variability analysis was analyzed using an analogous approach to both the computational model and fMRI data. However, since we had no true 'resting-state' activity for the monkey data set, we use the inter-trial interval (ITI; 0.25s - 1s variable duration, see Fig. 2D) as "resting-state activity". We used the 0.25s - 1s interval immediately preceding the trial's

fixation period to avoid any reward/feedback signals from the previous trial. (Reward/feedback from the previous trial was provided more than 1.5s prior to the fixation period.) Spike counts were calculated by taking a 50ms sliding window with 10ms increments, consistent with previous studies (3). The spike-triggered average across all trials for a given task rule (e.g., motion rule versus color rule) was calculated and removed from each trial, as is common in the spike count literature (10). Spike counts were then z-normalized across time for each trial with zero-mean and unit variance for the entire 5s trial period. Trials in which there was less than 250ms (or 25 time points) worth of spiking data for both the ITI and/or task cue presentation were excluded. This was done to reduce variability of the estimated spike count correlations, since correlations with few observations are highly variable.

We computed the SD during the ITI and task cue period across 25 consecutive trials using the average spike rate from each cortical recording. This was repeated for all trials for each subject. We used SD to calculate variability rather than Fano factor (3). This choice was due to the insight from our model illustrating that the mean-evoked activity and the corresponding SD interact in a nonlinear manner, and that the Fano factor is computed as the variance over the mean.

The statistical difference in task versus rest neural variability was computed by using a two-way, paired t-test across all bins of consecutive trials. The global neural variability change was computed by averaging the SD across all recording areas for each bin of 25 trials (Fig. 4C). Statistics for the regional neural variability change were corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$.

In addition, we computed the SD during the ITI and task cue period separately within trial (across time points) (Fig. S8). This analysis demonstrated that moment-to-moment variability (rather than event-to-event variability) was also quenched from rest to task periods, suggesting that variability quenching also occurs at faster timescales. The statistical difference in task versus rest neural variability was computed by using a two-way, paired t-test (paired by trial) across all trials for each monkey separately. The global neural variability change was computed by averaging the SD across all recording areas for each trial (Fig. S8A). Statistics for the regional neural variability change were corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$.

Spiking data: Task versus rest state correlation analysis

Neural correlations for spiking data were computed in an analogous approach to the fMRI FC approach, using the same preprocessing steps mentioned above for spike rate variability analysis. Specifically, the mean-evoked response across all trials for each task condition was removed from each trial, and normalized with zero-mean and unit-variance prior to computing spike count correlations. Spike count correlations were computed across trials, using groups of 25 trials as described above (Fig. 4E,F).

The difference in task versus rest intrinsic neural correlations was calculated using a two-way, paired t-test (paired by each bin of 25 trials) for each monkey separately using Fisher's *z*-transformed correlation values. The global neural correlation change was computed by

averaging the Fisher's *z*-transformed correlation values between all pairs of cortical regions, and comparing the averaged task versus rest correlation for each bin (Fig. 4D). Statistics for the pairwise neural correlation change (Fig. 4E,F,G) were corrected for multiple comparisons using an FDR-corrected threshold of $p<0.05$.

In addition, we computed the spike count correlation during the ITI and task cue period separately within trial (across time points) (Fig. S8). This analysis demonstrated that moment-to-moment correlations (rather than event-to-event correlations) were also quenched from rest to task periods, suggesting that correlation quenching also occurs at faster timescales. The statistical difference in task versus rest neural correlations was computed by using a two-way, paired t-test (paired by trial) across all trials for each monkey separately. The global neural correlation change was computed by averaging the correlation across all pairs of recording areas for each trial (Fig. S8B).

Dimensionality analysis

We evaluated the information-theoretic relevance of rest and task states by characterizing the dimensionality of neural activity. This was performed on both human fMRI and monkey spiking data. To estimate the statistical dimensionality of neural data, we used the formulation of dimensionality previously described in (45). We first obtain the covariance matrix W of activity for rest and task states separately. We then calculated

$$dim_W = \frac{(\sum_{i=1}^{m} \lambda_i)^2}{\sum_{i=1}^{m} \lambda_i^2} \qquad (1)$$

Where dim_W corresponds to the statistical dimensionality of W, and λ_i corresponds to the eigenvalues of the covariance matrix W. Intuitively, this is related to finding the number of components needed to explain variance greater than some fixed threshold, with more needed components reflecting a higher dimensionality of the data.

For human fMRI data, we estimated the task-state and resting-state dimensionality by calculating the whole-brain covariance matrix for each state. This was done by averaging across all 24 task conditions for task state, and by averaging across 24 equivalently lengthed resting-state periods. We applied equation 1 to the task-state and resting-state covariance matrix for each subject. Finally, we applied a two-way, group paired t-test comparing the dimensionality of task-state activity to resting-state activity (Fig. 5A). We replicated this effect in the replication cohort. In addition, we performed this analysis for each fMRI task separately (Fig. S11).

For NHP spiking data, we estimated the task (task cue period) and rest (ITI) dimensionality by calculating the covariance matrix across sets of 25 trials between all pairs of population recordings. We then applied equation 1 to task and rest periods for each covariance matrix. (Each covariance matrix was calculated using 25 consecutive trials.) Finally, we applied a two-way group paired t-test (across bins of 25 consecutive trials) comparing the dimensionality of task activity to rest activity (Fig. 5B). We replicated this effect in the held-out second monkey.

Model: Approximating a sigmoidal activation function with a loosely balanced excitatory-inhibitory circuit

Our modeling approach uses the simplest possible models to describe our phenomena of interest. We first characterize the relationship between activity and moment-to-moment variability, and subsequently characterize the relationship between activity and inter-region correlations. To study the former, we use a single neural mass model (one-dimensional model). To study the latter, we use a two-unit neural mass model (two-dimensional model). We then use the insights gained in these minimal models to provide a mechanistic grounding of empirical data.

Previous studies investigating the effects of evoked activity on neural variability have typically used either large-scale spiking or rate models (5–8, 38) or simplified E-I circuits where each node represented either E- or I-populations (4, 27, 76). Given our interest in looking at effects of evoked activity across large neural populations rather than within populations, we collapsed the neural population into a neural mass unit. Importantly, model simplification made it tractable to leverage dynamical systems tools to perform detailed phase plane analysis without loss of generality.

A loosely balanced E-I system prevents runaway excitation through inhibition while allowing for a dynamic range of excitatory activity (4, 6, 49). To model the excitatory activation subject to inhibitory regulation in a single neural mass, we combined the activation functions that the E and I populations would have if modeled separately by using the population activation function

$$F(x) = E(x) - I(x) \qquad (2)$$

where $E(x)$ reflects the E activation function and $I(x)$ reflects the I activation function. We modeled both the E and I activation functions using a soft plus function, such that

$$E(x) = log(1 + e^{x+b}) \qquad (3)$$

$$I(x) = log(1 + e^{x-b}) \qquad (4)$$

and b is defined as a bias. The mathematical form of F captures the assumption that the I-population activates with later (slower) onset than the E-population (consistent with loose E-I balance) (6). Moreover, F produces a nonlinear input-output relationship that is qualitatively identical to a sigmoid function (Fig. 1B). A sigmoidal activation function implies that for some optimal range of inputs, there is a dynamic input-output relationship (Fig. 1C). Any inputs exceeding this dynamic range produces a quenched output response consistent with the prevention of runaway excitation.

Model: One-dimensional minimal network model

We use the simplest model to mathematically characterize the relationship between evoked activity and neural variability: a one-dimensional neural population model. We used

Wilson-Cowan-type firing rate dynamics to simulate neural population activity (28). Specifically, our population's activity obeyed the equation

$$\tau_i \frac{dx_i}{dt} = -x_i + f(w_{ii}x_i + b_i + s_i + I) \qquad (5)$$

where x_i denotes the firing rate (or a measure of activity), $\tau_i = 0.1$ denotes the time constant, $w_{ii} = 1$ refers to local coupling (auto-correlation), $b_i = -0.5$ refers to the input threshold for optimal activity (or a bias term), s_i refers to the evoked stimulation ($s_i = 0$ for intrinsic activity), I refers to background spontaneous activity sampled from a Gaussian distribution with mean 0 and standard deviation 0.25, and f is a sigmoid input-output activation function, which is defined as

$$f(x) = \frac{1}{1+e^{-k*x}} \qquad (6)$$

where $k = 1$. Numerical simulations were computed using a Runge-Kutta second order method with a time step of dt=10ms (77). We simulated neural population activity injecting a fixed input (boxcar input) with amplitudes ranging from -5 to 5 in 0.01 increments (Fig. 6C). Neural variability for each input strength was calculated using the standard deviation of the time series following the input onset and preceding input offset. Each trial was run for 20 seconds. Fig. 6C was generated using input amplitudes of -3, 0, and 3.

To visualize the full dynamics of our single neural population, we visualized the one-dimensional phase space (i.e., flow field on a line) (46). In particular, we calculated the flow field by plotting $\dot{x}$ (i.e., $\frac{dx}{dt}$) as a function of x_i (equation 5). Notably, fixed point attractors (equilibrium states) are defined where $\dot{x} = 0$ (Fig. 6B).

Model: Two-dimensional minimal network model

To characterize the effects of evoked activity on neural correlations, we use a two-dimensional neural population model. We extended the one-dimensional network model to include two neural populations. The network dynamics obeyed the equations

$$\tau_1 \frac{dx_1}{dt} = -x_1 + f(w_{11}x_1 + w_{21}x_2 + b_1 + s_1 + I_1) \quad (7)$$

$$\tau_2 \frac{dx_2}{dt} = -x_2 + f(w_{22}x_2 + w_{12}x_1 + b_2 + s_2 + I_2) \quad (8)$$

where x_1 and x_2 describe the activity of each population, and all other variables are as described above. Inter-regional coupling was set to be greater than local coupling, given evidence from previous studies that global coupling is greater than local coupling (48, 78, 79). Specific network parameters for this network model were: $w_{11} = w_{22} = 2$, $w_{12} = w_{21} = 4$, $b_1 = b_2 = -3$, $\tau_1 = \tau_2 = 0.1$. I_1 and I_2 were sampled from a Gaussian distribution with mean 0 and standard deviation 1. For this network model, we decreased the slope of the sigmoid $k = 0.5$ to allow for a larger dynamic, linear response range. The full model code will be publicly available in the form of jupyter notebooks.

To quantify the relationship between evoked activity and neural correlations, we systematically simulated the network under different stimulation states (input strengths). Using the same methods as above, we simulated network activity for 50 seconds. We injected fixed input (uncorrelated inputs) into both neural populations with amplitudes from -5 to 5 in 0.01

increments (Fig. 7E). Notably, given that the injected stimulation is uncorrelated (due to 0-variance in a fixed input), it is non-trivial that the FC between two nodes would change in response to different inputs. Neural correlations were calculated using a Pearson correlation of the two time series following input onset and preceding input offset. (Note, Fig. 7A,B,C,D simulated activity for only 20 seconds, and are shown only for visualization purposes. Fig. 7B,D had a boxcar with an amplitude of 2 injected into both populations.)

The use of a minimal model constrained our network to two dimensions. This allowed us to leverage dynamical systems tools to visualize the flow field in the two-dimensional phase plane. To identify the fixed point attractors, we first calculated the nullclines for x_1 and x_2. Nullclines are defined as the values of x_1 and x_2 such that $\dot{x_2} = 0$ and $\dot{x_1} = 0$, respectively. The fixed points lie at the intersection of the nullclines. For our particular system of equations, the nullclines of x_1 and x_2 were defined, respectively, as

$$x_2 = \frac{f^{-1}(-x_1) - w_{11}x_1 - b_1 - s_1}{w_{21}} \qquad (9)$$

$$x_1 = \frac{f^{-1}(-x_2) - w_{22}x_2 - b_2 - s_2}{w_{12}} \qquad (10)$$

where parameters are identical to those used in equations 7 and 8. However, the background noise, parameter *I*, was removed when calculating the nullclines. Fixed point attractors (equilibrium states) are defined where $\dot{x} = 0$ (Fig. 7B). The full flow field was obtained by applying the system of equations (equations 7 and 8) to every point in the phase space (e.g., all values of x_1 and x_2).

Model: Evaluating fixed point attractor dynamics and the characteristic time scale

Our models accurately demonstrated that evoked activity decreased the neural variability and correlations from a stochastic dynamical network model. Since our network model was governed by firing rate equations which provided us full access to the system's dynamics, we sought to link dynamical mechanisms (in the absence of noise) with changes in the descriptive statistics. Such an analysis would provide us with a causal relationship between descriptive neural statistics used in empirical data analysis and the governing neural dynamics.

To understand how attractor dynamics influenced simulated activity in a network model, we characterized the dynamics around the network's fixed point attractor. Specifically, we performed a linear stability analysis around the fixed point (i.e., the equilibrium level of activity the system is drawn to during a particular state or input, e.g., Fig. 6B) in both the one-dimensional and two-dimensional network models. In the one-dimensional case, this analysis is equivalent to evaluating the first derivative of equation 5 at the fixed point (e.g., the slope of the line at the starred locations in Fig. 6B). We then calculate the *characteristic time scale* T at the fixed point x^* (in one-dimension) with the equation

$$T = \frac{1}{|f'(x^*)|} \qquad (11)$$

where f represents equation 5 (46). The characteristic time scale captures the speed with which the system approaches the fixed point attractor. We calculated the characteristic time scale across the same range of evoked stimulation strengths as in the neural variability

analysis. Fixed points were computed numerically by running the network model until it reached a steady state in the absence of noise or spontaneous activity.

The characteristic time scale is an established measure for one-dimensional systems. However, we sought to extend the characteristic time scale beyond a single dimension to evaluate shifting attractor dynamics in higher dimensions. We first performed a linear stability analysis in two dimensions by evaluating the Jacobian matrix for our two-dimensional system at the fixed point $(x^*{}_1, x^*{}_2)$

$$\mathbf{J}(x_1^*, x_2^*) = \begin{pmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} \end{pmatrix} \quad (12)$$

Where f_1 and f_2 refer to the equations governing neural populations 1 and 2 (equations 7 and 8, respectively). For our particular system of equations, the Jacobian was calculated as

$$\mathbf{J}(x_1^*, x_2^*) = \begin{pmatrix} \big(-1 + f'(w_{11}x_1 + w_{21}x_2 + b_1 + s_1)\big)\frac{1}{\tau_1} & \big(f'(w_{11}x_1 + w_{21}x_2 + b_1 + s_1)\big)\frac{1}{\tau_1} \\ \big(f'(w_{22}x_2 + w_{12}x_1 + b_2 + s_2)\big)\frac{1}{\tau_2} & \big(-1 + f'(w_{22}x_2 + w_{12}x_1 + b_2 + s_2)\big)\frac{1}{\tau_2} \end{pmatrix} \quad (13)$$

For each input strength (i.e., differing evoked states), we evaluated the Jacobian at the fixed point attractor. We then calculated the two eigenvalues (denoted λ_1 and λ_2) and eigenvectors (denoted v_1 and v_2) of the Jacobian using an eigendecomposition. To calculate the generalized characteristic time scale in two dimensions, we first calculated the linear combination of the eigenvectors weighted by the real eigenvalues, and computed the magnitude of the vector, such that

$$v_{sum}(x, y) = re(\lambda_1)v_1 + re(\lambda_2)v_2 \quad (14)$$

We then define the two dimensional characteristic time scale T as the reciprocal of the magnitude of $v_{sum}(x, y)$, such that

$$T = \frac{1}{|\sqrt{x^2+y^2}|} \quad (15)$$

We calculated T for a range of values $s_1, s_2 \in [-5, 5]$ in 0.01 increments, and correlated T across all values of s_1 and s_2 with the corresponding neural correlations (46).

Model: Simulating fMRI BOLD activity

We used the above model to simulate fMRI BOLD activity to demonstrate that changes in neural variability and correlations would extend to fMRI BOLD dynamics (Fig. S4,S5). Neural activity generated from our model simulations was transformed to fMRI BOLD activity using the Balloon-Windkessel model, a nonlinear transformation from neural activity to the fMRI BOLD signal (50, 80). Notably, the transformation assumes a nonlinear transformation of the normalized deoxyhemoglobin content, normalized blood inflow, resting oxygen extraction fraction, and the normalized blood volume. All state equations and biophysical parameters were taken directly from (50) (equations 3-4). The Balloon-Windkessel model was implemented in Python (version 2.7.13), and the implementation code has been made publicly available on GitHub (https://github.com/ito-takuya/HemodynamicResponseModeling).

Code and data availability

All code related to analyses and models in this study will be publicly released on GitHub. Code will provide supplementary details on data analysis, computational modeling, and figure generation. All fMRI human data is publicly available through the Human Connectome Project (http://www.humanconnectomeproject.org) (35). Inquiries related to monkey electrophysiology data can be made to authors S.L.B. and E.K.M.

References

1. He BJ (2011) Scale-Free Properties of the Functional Magnetic Resonance Imaging Signal during Rest and Task. *J Neurosci* 31(39):13786 LP–13795.

2. He BJ (2013) Spontaneous and task-evoked brain activity negatively interact. *J Neurosci* 33(11):4672–4682.

3. Churchland MM, et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13:369.

4. Hennequin G, Ahmadian Y, Rubin DB, Lengyel M, Miller KD (2018) The Dynamical Regime of Sensory Cortex: Stable Dynamics around a Single Stimulus-Tuned Attractor Account for Patterns of Noise Variability. *Neuron* 98(4):846–860.e5.

5. Litwin-Kumar A, Doiron B (2012) Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci* 15(11):1498–1505.

6. Huang C, et al. (2019) Circuit Models of Low-Dimensional Shared Variability in Cortical Networks. *Neuron* 101(2):337–348.e4.

7. Ponce-alvarez A, He BJ, Hagmann P, Deco G (2015) Task-Driven Activity Reduces the Cortical Activity Space of the Brain : Experiment and Whole-Brain Modeling. 1–26.

8. Deco G, Hugues E (2012) Neural Network Mechanisms Underlying Stimulus Driven Variability Reduction. *PLoS Comput Biol* 8(3):e1002395.

9. Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34(4):537–541.

10. Cohen MR, Kohn A (2011) Measuring and interpreting neuronal correlations. *Nat Neurosci* 14:811.

11. Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7:358.

12. Yeo BTT, et al. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165.

13. Power JD, et al. (2011) Functional Network Organization of the Human Brain. *Neuron*

72(4):665–678.

14. Ji JL, et al. (2018) Mapping the human brain's cortical-subcortical functional network organization. *Neuroimage*. doi:10.1016/j.neuroimage.2018.10.006.

15. Cole M, Bassett D, Power J, Braver T, Petersen S (2014) Intrinsic and task-evoked network architectures of the human brain. 83(1). doi:10.1016/j.neuron.2014.05.014.

16. Krienen FM, Yeo BTT, Buckner RL (2014) Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos Trans R Soc Lond B Biol Sci* 369(1653):20130526–20130526.

17. Gonzalez-Castillo J, Bandettini PA (2017) Task-based dynamic functional connectivity: Recent findings and open questions. *Neuroimage* (May):1–8.

18. Shine JM, et al. (2016) The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron* 92(2):544–554.

19. Tomasi D, Wang R, Wang G-J, Volkow ND (2014) Functional Connectivity and Brain Activation: A Synergistic Approach. *Cereb Cortex* 24(10):2619–2629.

20. Fiebach CJ, Rissman J, D'Esposito M (2006) Modulation of inferotemporal cortex activation during verbal working memory maintenance. *Neuron* 51(2):251–261.

21. Cohen MR, Maunsell JHR (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12(12):1594–1600.

22. Renart A, et al. (2010) The Asynchronous State in Cortical Circuits. *Science* 327(5965):587 LP–590.

23. Ecker AS, et al. (2010) Decorrelated neuronal firing in cortical microcircuits. *Science* 327(5965):584–587.

24. Cole MW, et al. (2018) Task activations produce spurious but systematic inflation of task functional connectivity estimates. *Neuroimage*. doi:10.1016/j.neuroimage.2018.12.054.

25. Cocchi L, Gollo LL, Zalesky A, Breakspear M (2017) Criticality in the brain: A synthesis of neurobiology, models and cognition. *Prog Neurobiol*. doi:10.1016/j.pneurobio.2017.07.002.

26. Breakspear M (2017) Dynamic models of large-scale brain activity. *Nat Neurosci* 20(3):340–352.

27. Rubin DB, Van Hooser SD, Miller KD (2015) The Stabilized Supralinear Network: A Unifying Circuit Motif Underlying Multi-Input Integration in Sensory Cortex. *Neuron* 85(2):402–417.

28. Wilson HR, Cowan JD (1972) Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophys J* 12(1):1–24.

29. Joglekar MR, Mejias JF, Yang GR, Wang X-J (2018) Inter-areal Balanced Amplification Enhances Signal Propagation in a Large-Scale Circuit Model of the Primate Cortex. *Neuron*

98(1):222–234.e8.

30. Ruff DA, Cohen MR (2014) Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci* 17(11):1591–1597.

31. Ruff DA, Cohen MR (2016) Stimulus Dependence of Correlated Variability across Cortical Areas. *J Neurosci* 36(28):7546–7556.

32. Barch DM, et al. (2013) Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* 80:169–189.

33. Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. *Science* 348(6241):1352–1355.

34. Brincat SL, Siegel M, von Nicolai C, Miller EK (2018) Gradual progression from sensory to task-related processing in cerebral cortex. *Proceedings of the National Academy of Sciences* 115(30):E7202 LP–E7211.

35. Van Essen DC, et al. (2013) The WU-Minn Human Connectome Project: an overview. *Neuroimage* 80:62–79.

36. Cole MW, Pathak S, Schneider W (2010) Identifying the brain's most globally connected regions. *Neuroimage* 49(4):3132–3148.

37. Gratton C, Laumann TO, Gordon EM, Adeyemo B, Petersen SE (2016) Evidence for Two Independent Factors that Modify Brain Networks to Meet Task Goals. *Cell Rep* 17(5):1276–1288.

38. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K (2016) The mechanics of state-dependent neural correlations. *Nat Neurosci* 19:383.

39. Cole MW, Yang GJ, Murray JD, Repovš G, Anticevic A (2016) Functional connectivity change as shared signal dynamics. *J Neurosci Methods* 259:22–39.

40. Duff EP, Makin T, Cottaar M, Smith SM, Woolrich MW (2018) Disambiguating brain functional connectivity. *Neuroimage* 173:540–550.

41. Al-Aidroos N, Said CP, Turk-Browne NB (2012) Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proc Natl Acad Sci U S A* 109(36):14675–14680.

42. Ruff DA, Cohen MR (2016) Attention Increases Spike Count Correlations between Visual Cortical Areas. *J Neurosci* 36(28):7523–7534.

43. Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. *Nat Rev Neurosci*. doi:10.1038/nrn3137.

44. Abbott LF, Rajan K, Sompolinsky H (2011) Interactions between Intrinsic and Stimulus-Evoked Activity in Recurrent Neural Networks. *The Dynamic Brain: An Exploration of Neuronal Variability and Its Functional Significance*:1–16.

45. Litwin-Kumar A, Harris KD, Axel R, Sompolinsky H, Abbott LF (2017) Optimal Degrees of Synaptic Connectivity. *Neuron* 0(0):1153–1164.e7.

46. Strogatz SH (1994) Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering (Cambridge, MA.

47. Jacobs EAK, Steinmetz NA, Carandini M, Harris KD (2018) Cortical state fluctuations during sensory decision making. *bioRxiv*:348193.

48. Deco G, et al. (2013) Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J Neurosci* 33(27):11239–11252.

49. Hennequin G, Agnes EJ, Vogels TP (2017) Inhibitory Plasticity: Balance, Control, and Codependence. *Annu Rev Neurosci* 40:557–579.

50. Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19(4):1273–1302.

51. Rosenbaum R, Rubin JE, Doiron B (2012) Short-term synaptic depression and stochastic vesicle dynamics reduce and shape neuronal correlations. *J Neurophysiol* 109(2):475–484.

52. Beer RD (1995) On the Dynamics of Small Continuous-Time Recurrent Neural Networks. *Adapt Behav* 3(4):469–509.

53. Ermentrout B, Terman D (2008) Foundations of Mathematical Neuroscience. 440.

54. MacKay DJC (2003) Information theory, inference and learning algorithms. Available at: https://books.google.ca/books?hl=en&lr=&id=AKuMj4PN_EMC&oi=fnd&pg=PR7&ots=EMlrca4vDj&sig=wBAwfSP_MiiGvnLPmvQ1zMRC2zM.

55. Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2010) Blood Oxygen Level-Dependent Signal Variability Is More than Just Noise. *Journal of Neuroscience* 30(14):4914–4921.

56. He BJ, Zempel JM (2013) Average is optimal: an inverted-U relationship between trial-to-trial brain activity and behavioral performance. *PLoS Comput Biol* 9(11):e1003348.

57. Lynch LK, et al. (2018) Task-evoked functional connectivity does not explain functional connectivity differences between rest and task conditions. *Hum Brain Mapp* 39(12):4939–4948.

58. Smith SM, et al. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A* 106(31):13040–13045.

59. Shine JM, Aburn MJ, Breakspear M, Poldrack RA (2018) The modulation of neural gain facilitates a transition between functional segregation and integration in the brain. *Elife* 7:e31130.

60. Fries P (2005) A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends Cogn Sci* 9(10):474–480.

61. Pesaran B, et al. (2018) Investigating large-scale brain dynamics using field potential

recordings: analysis and interpretation. *Nat Neurosci* 21(7):903–919.

62. Kahn I, et al. (2013) Optogenetic drive of neocortical pyramidal neurons generates fMRI signals that are correlated with spiking activity. *Brain Res* 1511(c):33–45.

63. Stam CJ, Nolte G, Daffertshofer A (2007) Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum Brain Mapp* 28(11):1178–1193.

64. Lin F-H, et al. (2004) Spectral spatiotemporal imaging of cortical oscillations and interactions in the human brain. *Neuroimage* 23(2):582–595.

65. Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*. Available at: https://www.sciencedirect.com/science/article/pii/S0896627310010755.

66. Smith SM, et al. (2013) Resting-state fMRI in the Human Connectome Project. *Neuroimage* 80:144–168.

67. Glasser MF, et al. (2016) A multi-modal parcellation of human cerebral cortex. *Nature* 536(7615):171–178.

68. Ciric R, et al. (2017) Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage* 154:174–187.

69. Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37(1):90–101.

70. Power JD, et al. (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* 84:320–341.

71. Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA (2009) The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *Neuroimage* 44(3):893–905.

72. Power JD, et al. (2018) Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. *Proceedings of the National Academy of Sciences*:201720985.

73. Glasser MF, et al. (2017) Using Temporal ICA to Selectively Remove Global Noise While Preserving Global Signal in Functional MRI Data. *bioRxiv*:193862.

74. Wong CW, Olafsson V, Tal O, Liu TT (2013) The amplitude of the resting-state fMRI global signal is related to EEG vigilance measures. *Neuroimage* 83:983–990.

75. Friston KJ, et al. (1994) Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2(4):189–210.

76. Ahmadian Y, Rubin DB, Miller KD (2013) Analysis of the stabilized supralinear network. *Neural Comput* 25(8):1994–2037.

77. Burden RL, Faires JD (2001) Numerical analysis. 2001. *Brooks/Cole, USA*.

78. Ito T, et al. (2017) Cognitive task information is transferred between brain regions via resting-state network topology. *Nat Commun* 8(1):1027.

79. Cole MW, Ito T, Bassett DS, Schultz DH (2016) Activity flow over resting-state networks shapes cognitive task activations. *Nat Neurosci* 19(12):1718–1726.

80. Buxton RB, Wong EC, Frank LR (1998) Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn Reson Med* 39(6):855–864.