bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Between-area communication through the lens of within-area neuronal dynamics

Olivia Gozel[1,2,*] and Brent Doiron[1,2,*]

[1]Departments of Neurobiology and Statistics, University of Chicago, Chicago, IL 60637, USA
[2]Grossman Center for Quantitative Biology and Human Behavior, University of Chicago, Chicago, IL 60637, USA
[*]gozel@uchicago.edu bdoiron@uchicago.edu

Abstract

A core problem in systems neuroscience is deciphering the origin of shared fluctuations in neuronal activity: does it emerge through local network interactions, or is it inherited from external sources? We explore this issue with large-scale networks of spatially ordered spiking neuron models where a downstream network receives input from an upstream sender network. We show that linear measures of communication between sender and receiver networks can discriminate between emergent or inherited population dynamics. Faithful communication requires a match in the dimensionality of the sender and receiver population activities, along with an alignment of their shared fluctuations. Nonetheless, even in scenarios with seemingly weak linear communication, the timing of spike responses in the receiver network remains sensitive to those in the sender network. Our work exposes the benefits and limitations of linear measures when analyzing between-area communication in circuits with diverse neuronal dynamics.

Introduction

The brain is composed of a multitude of distributed areas which interact to support the complex computations needed for cognition. While past experimental investigations were typically limited to single neuron recordings, recent technological advances allow for sampling from large populations of neurons simultaneously (Urai et al., 2022). This newfound ability was initially used to characterize the local dynamics of a population of neurons from the same brain area (Churchland et al., 2010; Schölvinck et al., 2015; Murray et al., 2017; Williamson et al., 2019; Jiang et al., 2020). Presently, many studies measure distributed population activity, giving a more holistic view of neuronal processing (Urai et al., 2022). However, despite these much broader datasets, the

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

science of the mechanics by which different brain areas communicate with one another is still in its infancy.

An often used measure of neuron-to-neuron interaction is the joint trial-to-trial covariability, or noise correlation, of their spike train responses (Cohen and Kohn, 2011; Doiron et al., 2016). The idea is that neuron pairs that have high correlations are likely members of the same putative neuronal circuit (Abeles, 1991; Shadlen and Newsome, 1998; Ocker et al., 2017). While pairwise correlations can be informative (Doiron et al., 2016), the large-scale nature of population recordings presents a challenge when attempting to expose the salient aspects of population-wide interactions simply from an analysis of neuron pairs. Dimensionality reduction techniques have been developed to frame population activity within a space of the appropriate size: large enough to capture the core shared variability across a population, yet small enough to be tractable (Cunningham and Yu, 2014). Recently, a dimensionality reduction analysis method has been developed to assess the quality of between-area communication by computing how well activity in an upstream area can linearly predict activity in a downstream area (Semedo et al., 2019). It has been applied to simultaneous recordings along the primate visual pathways, both during anesthesia and a cognitive task, and suggests that areas interact through a communication subspace (Semedo et al., 2019; Srinath et al., 2021). These novel analysis techniques organize and quantify the magnitude and effective structure of between-area communication, offering important tools to help expose how the circuit structure of cortical networks shape distributed processing. Yet these techniques do not on their own provide insight into the circuit mechanisms that support or impede communication.

The propagation of brain activity has been the focus of extensive circuit modeling attempts. Feedforward networks are the base structure of many contemporary models of object classification, and have been used with great success to model the performance of visual system hierarchy (Yamins and DiCarlo, 2016). However, networks of spiking neuron models with random, sparse feedforward connectivity produce propagation that leads to excessive, often rhythmic, synchronization (Abeles, 1991; Diesmann et al., 1999; Reyes, 2003; Kumar et al., 2010; Rosenbaum et al., 2011). By contrast, a single population of spiking neuron models with sparse, yet strong, excitatory (E) and inhibitory (I) recurrent connections can show temporally irregular, roughly asynchronous spiking dynamics (Van Vreeswijk and Sompolinsky, 1998; Amit and Brunel, 1997; Renart et al., 2010; Rosenbaum et al., 2017), mimicking what is often considered the default state of cortical networks (Shadlen and Newsome, 1998; Renart et al., 2010). However, neurophysiological recordings over a range of sensory and cognitive states show a wide distribution of spike count correlations whose average is low, but positive and significantly different from zero (Cohen and Kohn, 2011; Doiron et al., 2016). Recent modeling work shows how population dynamics with stable firing rates yet moderate population-wide noise correlations can be produced when structured synaptic wiring is considered, such as discrete block structure (Darshan et al., 2018), low-rank recurrent components (Landau and Sompolinsky, 2018; Mastrogiuseppe and Ostojic, 2018), or distance dependent connection probability (Keane and Gong, 2015; Rosenbaum et al., 2017; Huang et al., 2019). While these results provide new insights into how circuit structure determines shared variability, they have been restricted to within-population dynamics. On the other hand, recent modeling efforts have shed some light on the interaction between brain areas (Chaudhuri et al., 2015; Muller et al., 2018; Hahn et al., 2019), yet often without a consideration of response fluctuations. Thus,

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

there remains a gap in understanding how the circuit-based theories of shared variability within a population extend to the distribution (or propagation) of variability between populations.

In this work, we investigate how complex within-area neuronal dynamics affect interactions between distinct brain areas using a network of model spiking neurons with biologically plausible synaptic dynamics, as well as spatially-dependent synaptic architecture. We determine conditions when communication between an upstream sender network and a downstream receiver network, as assessed by a linear measure, is disrupted. Indeed, the emergence of complex spatio-temporal dynamics within-area may or may not lead to poor communication depending on its origin. Specifically, a substantial decrease of dimensionality or a misalignment of shared fluctuations in the receiver with respect to the sender disrupts communication. Interestingly, however, we show that the receiver is nevertheless effectively driven by the sender using a perturbation experiment. These results expose the limitations of linear measures when deciphering between-area communication in the presence of complex spatio-temporal neuronal dynamics.

Results

Sender and receiver networks exhibit rich population-wide dynamics

To investigate if and how complex within-area neuronal dynamics affect communication with other brain areas, we explore the responses of a three-layer network of spiking neuron models (Fig. 1a). Neurons in the input layer are modeled as homogeneous Poisson processes with a uniform rate. The sender (second layer) and receiver (third layer) populations each consist of excitatory (E) and inhibitory (I) neurons which are spatially organized on a square grid. It is well known that brain connectivity is spatially structured with connection probability falling off with the distance between pre- and post-synaptic neurons (Holmgren et al., 2003; Levy and Reyes, 2012; Rossi et al., 2020). Accordingly, following past work (Rosenbaum and Doiron, 2014; Rosenbaum et al., 2017; Huang et al., 2019) we model within- and between-layer connectivity as obeying a two-dimensional Gaussian (whose spatial widths are denoted by σ; see Methods), and we assume periodic boundary conditions on our domain. In both sender and receiver layers, we set a larger width for the recurrent than the feedforward connections ($\sigma_{\text{rec}} > \sigma_{\text{ffwd}}$), because such networks exhibit correlated neuronal spiking dynamics (Rosenbaum et al., 2017).

Unless otherwise specified, our network is set with the parameters reported in Table 1. We call these the standard parameters because they yield stable spiking dynamics in both the sender and receiver networks, as reflected by temporally irregular spiking activity (Fig. 1b) and an average pairwise spike-count correlation over all spatial scales close to zero (Rosenbaum et al., 2017). It is known from previous work that when the E/I balance is destabilized, spatio-temporal patterns of spiking activity intrinsically emerge within the network (Rosenbaum et al., 2017; Huang et al., 2019; Keane and Gong, 2015; Muller et al., 2018). The E/I balance can either be destabilized in space by increasing the width of recurrent inhibition σ_I (Fig. 1c), or in time by increasing the inhibitory neuron time constant τ_I^{decay} (Fig. 1d). The exact spatio-temporal characteristics of the emerging

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

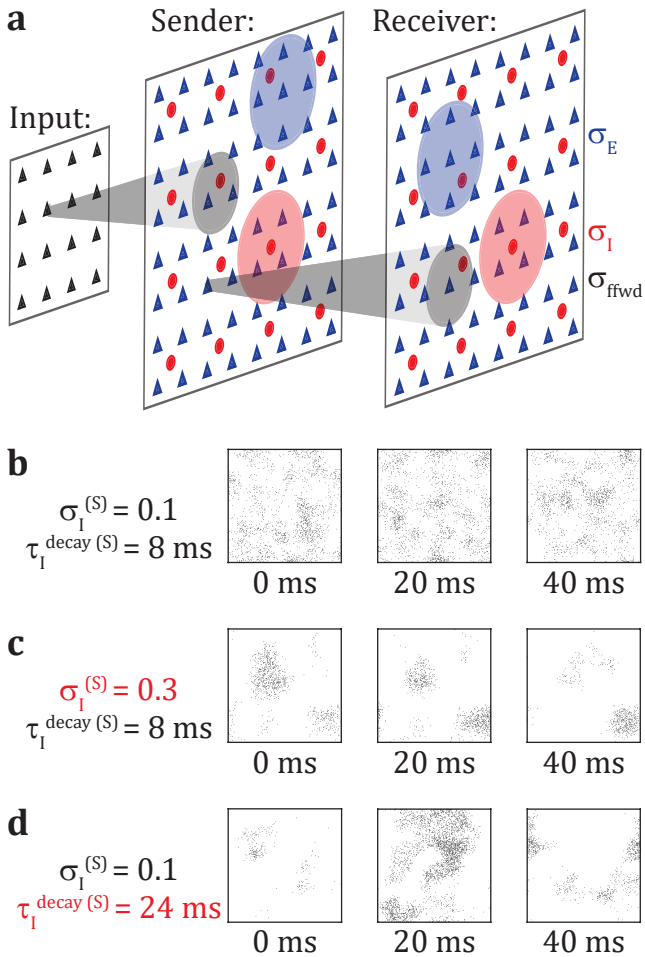

Figure 1: A spiking network with spatially-organized connectivity yields diverse neuronal dynamics.
(**a**) The input layer produces homogeneous Poisson spike trains (black triangles) and connects to the Sender layer (S), which itself connects to the Receiver layer (R). S and R consist of excitatory (E, blue triangles) and inhibitory (I, red circles) conductance-based neurons which are recurrently connected. All neurons are organized on two-dimensional $[0,1] \times [0,1]$ spatial grids. All connections are spatially-organized according to a wrapped Gaussian (periodic boundary conditions) with widths σ_{ffwd}, σ_E and σ_I for the feedforward and recurrent connections respectively (in all panels $\sigma_{\text{ffwd}} = 0.05$ and $\sigma_E = 0.01$). (**b,c,d**) Raster plot snapshots of the spiking activity in S over 2 ms time-windows separated by 20 ms: (b) standard network in the stable regime where $\sigma_I = 0.1$ and $\tau_I^{\text{decay}} = 8$ ms, (c) S is destabilized with a spatial failure of I to balance E ($\sigma_I^{(\text{S})} = 0.3$), and (d) S is destabilized with a temporal failure of I to balance E ($\tau_I^{\text{decay (S)}} = 24$ ms).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

patterns of activity are different. However, their global effect on within-area neuronal dynamics, and therefore on the interactions between the sender and receiver populations, is expected to be qualitatively similar. For simplicity, in the following sections we mainly report the results when destabilizing the network spatially through an increase of σ_I. This modeling framework gives us control of the emergence of complex spiking dynamics without the need for an external source of noise.

The emergence of spatio-temporal patterns yields an increase in spike count correlations

Spike count correlations are a common experimental measure of underlying neuronal dynamics (Cohen and Kohn, 2011; Ocker et al., 2017). Their distribution over all pairs shows high heterogeneity with a positive mean close to – but significantly different from – zero (Shadlen and Newsome, 1998; Renart et al., 2010). Our model can readily capture those features (Fig. 2). With the standard parameters the network produces spiking activity that is irregular (Fig. 2a, left), with a near symmetric distribution of pairwise correlations having a mean close to zero (Fig. 2b, left). Emergence of spatio-temporal patterns in the receiving area through a spatial destabilization of the E/I balance (Fig. 2a, right) induces a broader distribution of spike count correlations with a heavy positive tail (Fig. 2b, right). Yet the distribution of pairwise correlations between neurons from the sending and receiving populations remains relatively narrow with a mean close to zero (Fig. 2c).

By contrast, when the E/I balance is instead destabilized in the sender population, it yields the emergence of spatio-temporal patterns (Fig. 2d, left), which are subsequently propagated to the receiving population (Fig. 2d, right). Consequently the distribution of spike count correlations is heavily-tailed in both the sender and receiver networks (Fig. 2e). Furthermore, we observe a heavy tail in the distribution of between-area spike count correlations as well (Fig. 2f).

While pairwise correlations can be a signature of the dynamical regime of a network, we take advantage of our access to large amounts of synthetic data to move beyond statistics on pairwise neuronal activity. In the next section, we investigate the full structure of the shared neuronal variability within each population.

Within-area shared fluctuations depend on the spatial arrangement of sampled neurons

We randomly select 50 neurons from a local portion of the grid delineated by a disc (Fig. 3a, left). By changing the radius of the disc we can explore how shared variability depends on the spatial scale of the sampled population. We only select neurons whose average firing rate is sufficiently responsive (above 2 Hz) and compute their full covariance matrix C. Through Factor Analysis (FA) (Everitt, 1984; Yu et al., 2009), we separate C into a shared component, C_{shared}, and a private

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

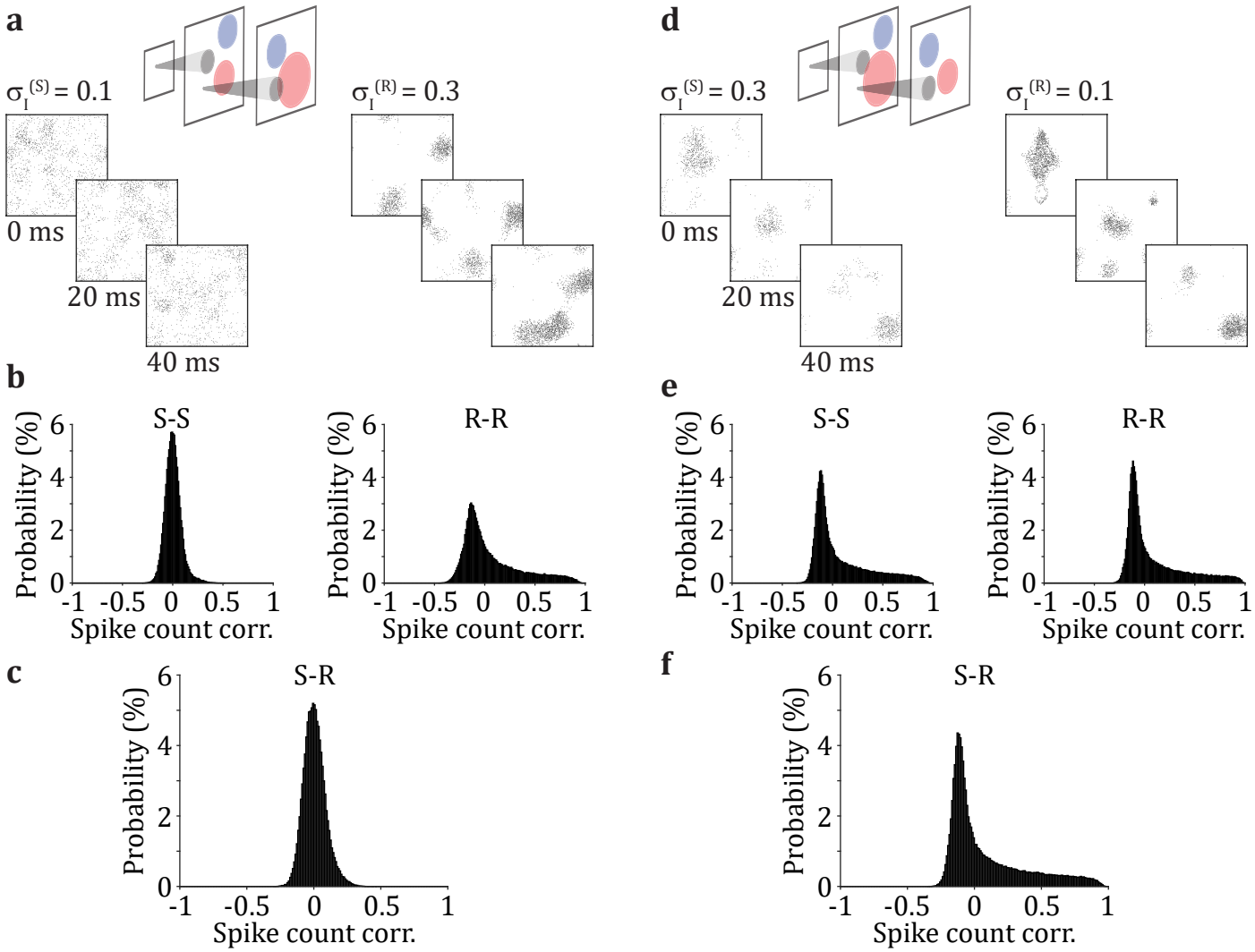

Figure 2: The emergence of spatio-temporal patterns yields an increase in spike count correlations.
(**a**) Simultaneous rasters of activity in Sender (S, left) and Receiver (R, right) layers when the width of recurrent inhibitory connections is larger in R than in S. (**b**) Corresponding distributions of the pairwise spike count correlations for neuron pairs within-area S (left, mean of the distribution: $\mu = 0.0045$) and within-area R (right, $\mu = 0.0658$). (**c**) Distribution of the pairwise spike count correlations between a neuron in S and a neuron in R ($\mu = 0.0061$). (**d**) Simultaneous rasters of activity in Sender (S, left) and Receiver (R, right) layers when the width of recurrent inhibitory connections is larger in S than in R. (**e**) Corresponding distributions of the pairwise spike count correlations for neuron pairs within-area S (left, $\mu = 0.0839$) and within-area R (right, $\mu = 0.0866$). (**f**) Distribution of the pairwise spike count correlations between a neuron in S and a neuron in R ($\mu = 0.0846$).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

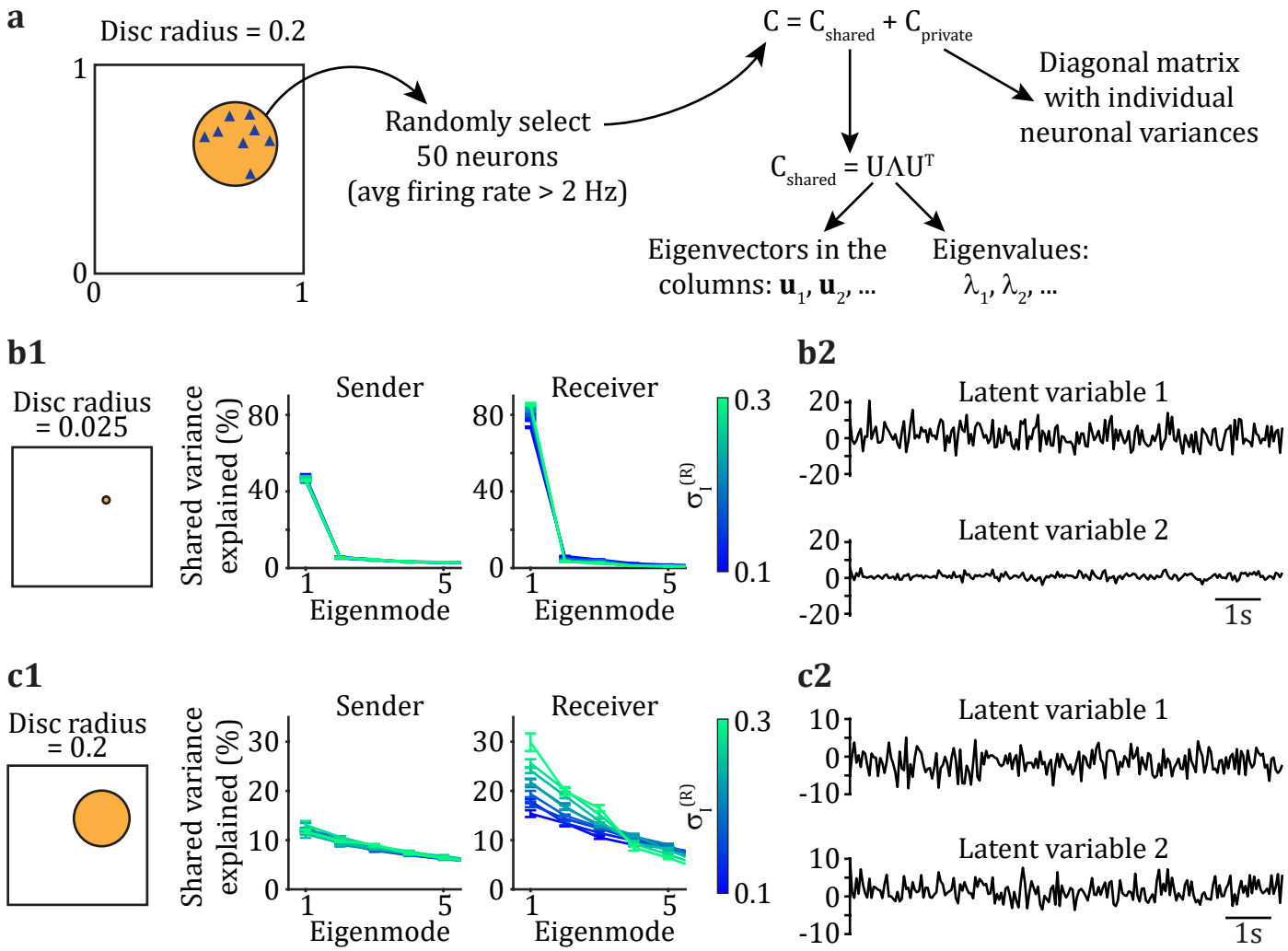

Figure 3: The structure of within-area shared fluctuations is spatial-scale dependent. (**a**) Fifty neurons, each with average firing rates larger than 2 Hz, are selected from the neuronal grid by random sampling from discs with different radii. The shared covariance matrix, C_{shared}, is obtained from the full covariance matrix, C, through Factor Analysis. The corresponding shared eigenvectors and eigenvalues are then obtained through Singular Value Decomposition. (**b1**) Shared variance explained by the first five eigenmodes within the sender network (S) and the receiver network (R) when modifying σ_I in R. Neurons are randomly sampled from a small disc (disc radius = 0.025). (**b2**) Projection of 10 s of the receiving population activity on the first two shared eigenvectors to obtain the first two shared latent variables when neurons are sampled from a small disc. The network is implemented with the standard parameters ($\sigma_I^{(S)} = \sigma_I^{(R)} = 0.1$). (**c1,c2**) Same as (b1,b2), except that neurons are sampled from a large disc (disc radius = 0.2). See Fig. S1 for the case when modifying σ_I in S.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

component, C_{private} (Fig. 3a, right). FA, in contrast to Probabilistic Principal Component Analysis (PPCA), does not assume isotropic noise. Hence the elements of the diagonal matrix C_{private} are not constrained to be identical. FA thus determines the directions of highest covariance, and not largest individual variance as done in PPCA. Singular Value Decomposition is then applied to C_{shared} to obtain the shared eigenvectors and associated eigenvalues which characterize the structure of the shared fluctuations (Fig. 3a).

When neurons are sampled from a small disc, the first eigenmode explains most of the shared variance, in both the sender and receiver areas (Fig. 3b1). It is reflected by a high variability of the first latent variable, while the variability of the second latent variable is relatively small (Fig. 3b2). A destabilization of the E/I balance in the receiving population does not change the amount of shared variance explained by the first eigenmodes (Fig. 3b1, right). When neurons are sampled from a large disc, however, the structure of shared variability is substantially different. Indeed, the first eigenmode no longer explains an outsized amount of variance compared to the other eigenmodes (Fig. 3c1), as reflected by similar variance of the first two latent variables (Fig. 3c2). As the breadth of the inhibitory recurrent connections in the receiving population increases, the percentage of shared variance explained by the first eigenmodes grows (Fig. 3c1, right). Similar results were observed when the E/I balance was destabilized in the sending instead of the receiving population (Fig. S1).

These results emphasize a spatial scale dependence of the structure of shared within-area fluctuations in networks with a spatially-organized architecture. While current experimental recordings are constrained to a fixed spatial scale, we take advantage of the flexibility of our model to get a broad understanding of the effect of within-area neuronal dynamics on interaction between brain regions using a wide range of spatial scales.

Spatio-temporal pattern formation generates low-dimensional shared variability

To better quantify the structure of within-area shared variability we consider the distribution of the eigenvalues $\{\lambda_i\}$ of the shared covariance matrix C_{shared}. In particular, we measure the dimension of shared variability with the participation ratio of the eigenspectrum: $\text{PR} = \left(\sum \lambda_i\right)^2 / \sum \lambda_i^2$ (Mazzucato et al., 2016; Litwin-Kumar et al., 2017). Contrary to other measures of dimensionality (Mante et al., 2013; Kaufman et al., 2014; Williamson et al., 2016; Semedo et al., 2019), PR does not require an arbitrary threshold to give an integer value of dimension. Rather, PR is the squared first moment of the eigenspectrum $\{\lambda_i\}$ normalized by the second moment. If the shared fluctuations preferentially take place over a few dimensions, reflected by a few eigenvalues, λ_i, which are much larger than the others, it yields a low PR (Fig. 4a, left). On the other hand, if the shared fluctuations are broadly distributed over the whole eigenspace, as reflected by a uniform distribution of the eigenvalues, the resulting PR is high (Fig. 4a, right).

Using within-area shared dimensionality we investigate how the emergence of spatio-temporal patterns affects the structure of shared variability, both within the area and in connected areas.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

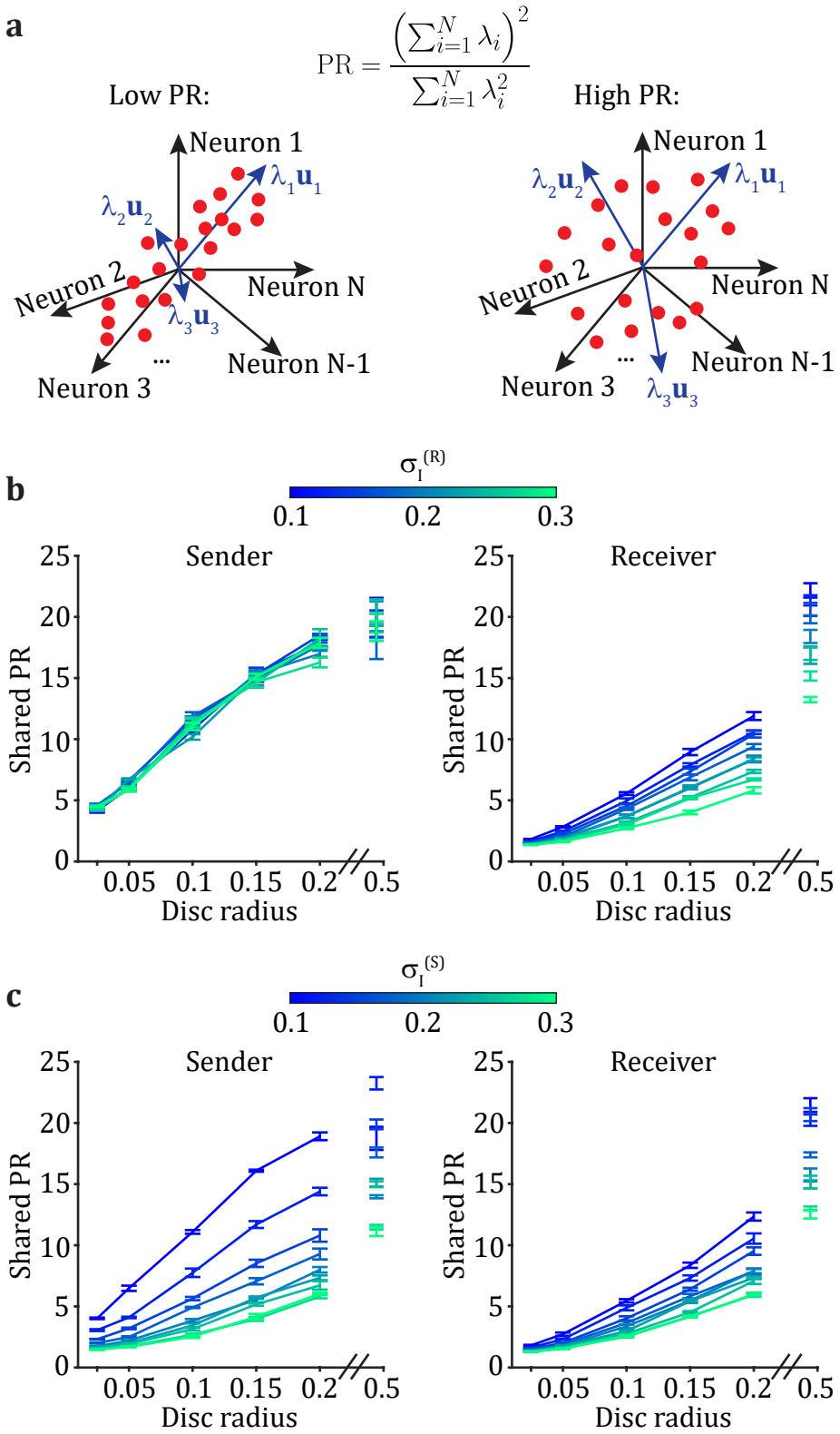

Figure 4: Caption is on the next page

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4: Spatio-temporal pattern formation generates low-dimensional shared variability.
(**a**) Shared dimensionality is measured by the participation ratio: $\mathrm{PR} = \left(\sum \lambda_i\right)^2 / \sum \lambda_i^2$, where $\{\lambda_i\}$ is the eigenspectrum of the shared covariance matrix. When the eigenvalues are heterogeneous in magnitude, dimensionality is low (left). Whereas when eigenvalues are more uniform in magnitude, dimensionality is high (right). (**b,c**) Shared PR as a function of the radius of the disc from which neurons are sampled, either in the sender (S, left) or in the receiver (R, right) population. We destabilize network activity by modifying σ_I in R (b) or in S (c).

When neurons are sampled from a larger disc the estimated shared PR increases (Fig. 4b, left). This is expected given how the dominant eigenmode changes with disc radius (Fig. 3b1,c1). However, when sampling neurons from the whole grid (disc radius = 0.5), the shared PR stays much lower than the theoretical upper bound of 50 (Fig. 4b, left). Because PR depends strongly on the spatial arrangement of the sampled neurons, we cannot simply associate a unique PR value to the network dynamics as a whole. Instead, we will focus on how changes in network dynamics owing to E/I destabilization are reflected in changes in PR.

As activity propagates from the sender to the receiver network we observe a decrease in PR (at fixed disc radius), even in the stable network with standard parameters (Fig. 4b). The more the E/I balance is destabilized in the receiving population, the larger the decrease in PR (Fig. 4b) when compared to that of the sending population. Strikingly, shared dimensionality in the receiver network is similarly low irrespective if spatio-temporal patterns emerge locally or if they are inherited from the sender network (compare Fig. 4b with Fig. 4c). This raises an important dilemma for the interpretation of changes in the structure of population-wide shared fluctuations. Namely, that a change in the dimension of population activity can be due to either a shift of the internal dynamics within a population or be inherited from shifts in upstream areas. This ambiguity prompts us to next consider how the sender and receiver networks directly communicate their shared fluctuations.

Inter-area communication is oppositely affected if patterns emerge within the receiver network or if they are inherited from the sender network

In the previous section we used the participation ratio PR as the metric of within-area neuronal dynamics. However, as pointed out, its value is dependent on the spatial scale from which neurons are sampled. To give a distance-independent metric of the structure of population-wide activity we apply spatio-temporal Fourier analysis to the full receiver population activity (see Methods; Huang et al. (2022)). This provides the power of the network activity as a function of the temporal frequency ω and the spatial wavenumber k (Fig. S2). While the spatio-temporal structure is rich, in what follows we use only the peak power of the receiver population activity (over all ω and k) as a proxy for the characterization of within-area neuronal dynamics. The larger the peak power, the more structured are the spatio-temporal dynamics of the population-wide activity.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

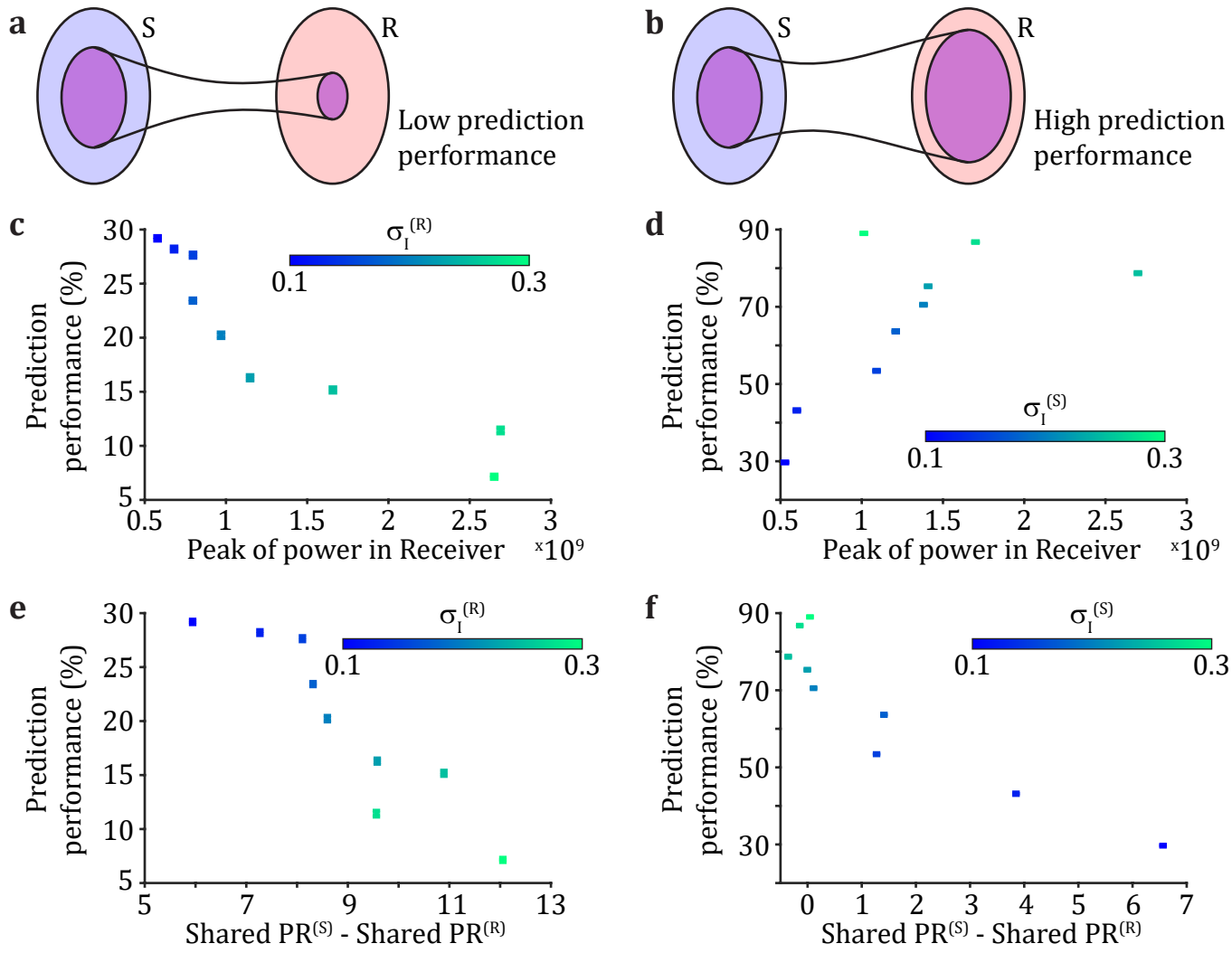

Figure 5: Inter-area communication is oppositely affected if patterns emerge within the receiving population or if they are inherited from the sending population. (**a,b**) Schematic for inter-area communication. Population activity in the sender network (S) is depicted in light blue and population activity in the receiver network (R) is depicted in light red. The communication subspace is depicted in purple and quantifies how well population activity in S can linearly predict population activity in R. A low prediction performance of the communication subspace indicates that little of the activity fluctuations in R can be explained by S (a). In contrast, if most of the activity fluctuations in R can be predicted by S activity, prediction performance of the communication subspace is high (b). (**c,d**) Prediction performance of the communication subspace between S and R when σ_I is modified in R (c) or in S (d). (**e,f**) Prediction performance of the communication subspace is higher when shared PR in S and R is similarly low, while it is lower when shared PR in S is much higher than in R, no matter whether spatio-temporal patterns emerge in R (e) or in S (f). Neurons are sampled from a disc with radius 0.2; errorbars are mean ± SEM.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To assess the interaction between sending and receiving populations we use a recently developed measure of area-to-area communication based on reduced-rank regression (Semedo et al., 2019). Briefly, the activity in the receiver network, R, is predicted from the activity in the sender network, S, through a linear model: $\hat{R} = SB_{\mathrm{RRR}}$. The rank of the regression matrix B_{RRR} is constrained to be a low value m (see Methods). Prediction performance is given by the comparison between R and $\hat{R}$, quantifying the ability of the sender population activity in linearly predicting the receiver population activity through a low-dimensional communication subspace (Fig. 5a,b).

As the peak power of spiking activity in the receiver network increases (by increasing σ_I locally), there is a decrease in prediction performance of the communication subspace (Fig. 5c). We note that the exact value of prediction performance depends on the spatial scale from which neurons are sampled (Fig. S3): the larger the spatial scale, the lower the prediction performance. When the E/I balance is instead destabilized in the sender network, the peak power in the receiver network also increases due to the inheritance of spatio-temporal patterns from the sending area. Indeed, as mentioned earlier, the structure of the receiver population activity shows the same signatures, in terms of shared dimensionality (Fig. 4b,c) and peak power (Fig. 5c,d), irrespective of whether the population-wide activity is generated locally or inherited from the sender network.

Despite similar dynamics in the receiver population regardless of the origin of the E/I destabilization, prediction performance of the communication subspace is unambiguous. On one hand, if complex spatio-temporal patterns emerge in the receiver network, they disrupt communication between the sending and receiving areas (Fig. 5c). On the other hand, communication is improved if spatio-temporal patterns emerge in the sender network. In total, the characteristics of neuronal dynamics in a single area are not sufficient to predict the quality of communication to that area.

Interestingly, the difference in shared dimensionality in the sender and receiver areas appears to be a better predictor of the quality of communication between the two. Specifically, a mismatch in the within-area dimensionality leads to poor communication between sender and receiver networks, no matter the origin of spatio-temporal patterns of activity (Fig. 5e,f). We note that if the E/I balance is destabilized temporally by increasing the time constant of inhibitory neurons, τ_I^{decay}, similar observations are made (Fig. S4). This last result motivates a more thorough analysis of how the relation between the dimensionality of sender and receiver population dynamics affects communication.

A misalignment of shared variability is associated with poor communication between connected areas

As exposed above, destabilization of the E/I balance within a population leads to low-dimensional shared variability. If the match of the dimension of shared variability is all that is required for good communication then a combined E/I destabilization in the sender and receiver networks (so that both have low dimension) should yield a high prediction performance of the communication subspace. If instead prediction performance is unexpectedly low, this would suggest that the matching in shared dimensionality is not the only mechanism at play for faithful communication.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

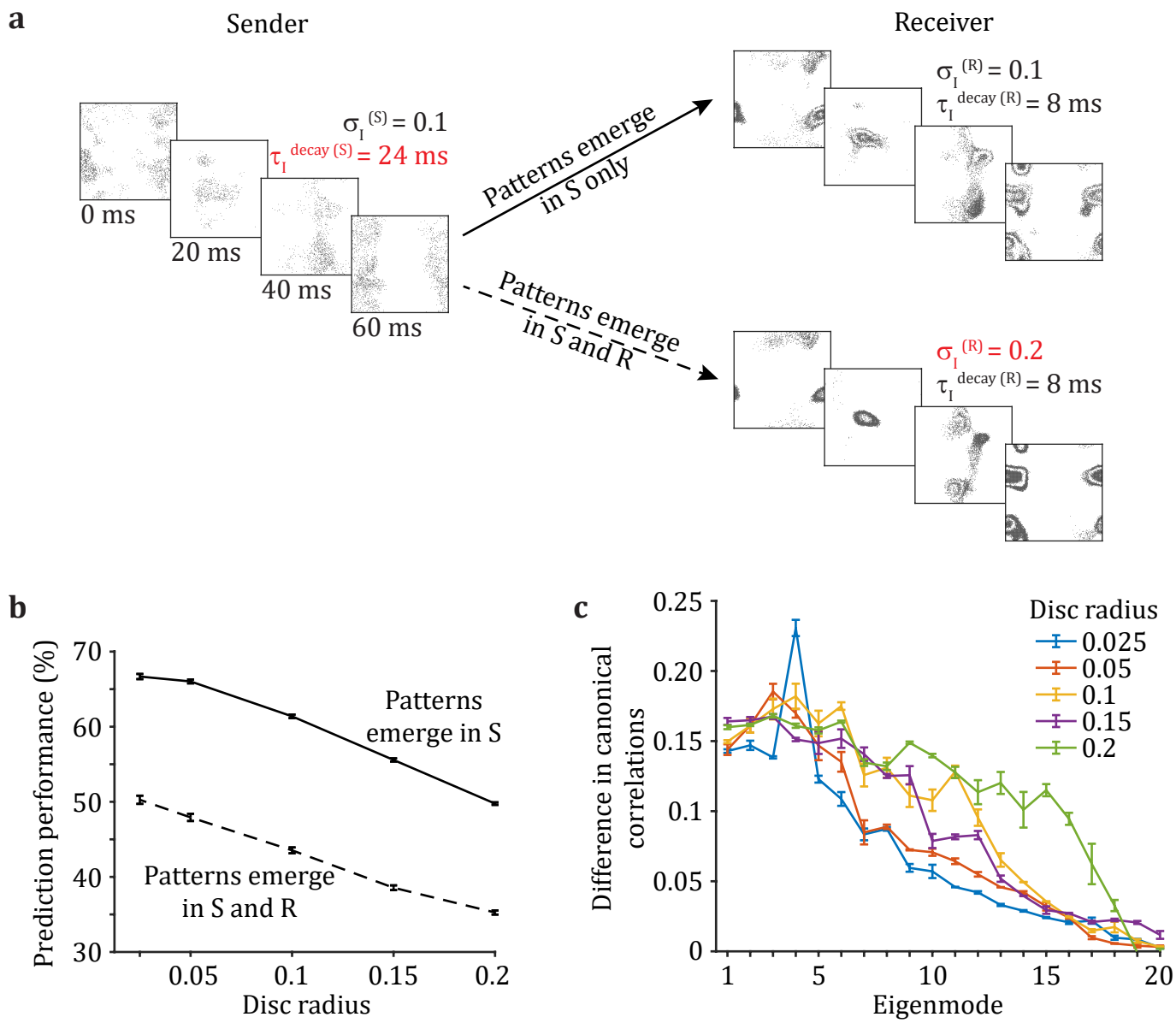

Figure 6: A misalignment of shared variability is associated with poor communication between connected areas.
(**a**) The sender layer (S) is destabilized through a slower inhibitory timescale. The receiver layer (R) is either in the stable regime (top, patterns emerge in S only), or destabilized through broader recurrent inhibition (bottom, patterns emerge in S and R). (**b**) The prediction performance of the communication subspace between S and R is higher when only S is destabilized compared to the case when both S and R are destabilized. (**c**) Difference in canonical correlations between the case when only S is destabilized and the case when S and R are destabilized (see Fig. S5 for the canonical correlations in each case individually).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To test this hypothesis we compare the communication across two distinct sender-receiver networks. Both networks have the sender E/I balance destabilized through a longer inhibitory time constant. In the first network the receiving population is kept in an intrinsically stable regime (Fig. 6a, top), while in the second network the receiver E/I balance is destabilized through a broadening of recurrent inhibition (Fig. 6a, bottom). As expected, both networks have a good match in the dimensionality of the sender and receiver populations (Fig. S5). However, prediction performance of the communication subspace is substantially lower when different spatio-temporal patterns emerge in the sender and receiver networks individually, as compared to the case where the receiving area inherits its low dimensional nature from the sending area (Fig. 6b). Therefore, dimensionality matching is not a sufficient condition for good communication.

We then hypothesize that a misalignment of the low-dimensional manifolds of shared variability in the sender and receiver populations is the cause of disruption of their communication. We measure the alignment of shared variability in sender and receiver areas by computing their aligned canonical correlations (Gallego et al., 2020). We conservatively select a dimensionality of the manifold of 20, which is higher than what was used in previous studies (Gallego et al., 2020). In our case, the value of the canonical correlation for each eigenmode indicates to what extent the given latent dimension can be well aligned between the sender and receiver areas through a linear transformation. We observe that the first eigenmodes can be well aligned and that the larger the spatial scale from which neurons are sampled, the better the alignment of their eigenmodes (Fig. S5). Strikingly, we observe larger canonical correlations between the sending and receiving populations when only the sender network is destabilized, compared to the case when both the sender and the receiver networks are destabilized (Fig. 6c). Hence a misalignment of the within-area shared fluctuations between sender and receiver populations causes poor communication.

To summarize, a mismatch in either the dimensionality or the alignment of the manifolds of shared variability yields poor communication between connected brain areas (Fig. 7). Specifically, faithful communication, as assessed by a linear measure, is possible if shared dimensionality is similarly low in sender and receiver networks, and if the manifold of the shared variability of the receiving population can be well aligned to the one in the sending area (Fig. 7a). If a good alignment of the shared variability in sending and receiving populations cannot be achieved through a linear transformation, for example in the case where different spatio-temporal patterns emerge in both areas independently, communication is disrupted (Fig. 7b). Finally, if the sender population displays high-dimensional shared variability while the receiver population exhibits low-dimensional shared variability, such as when spatio-temporal patterns of activity only emerge in the receiving area, the resulting communication is also poor (Fig. 7c).

Activity in the receiving population is effectively driven by the sending population activity even when communication is poor

We have shown that communication between sender and receiver networks can be disrupted in different ways. But what are the implications of a poor communication for the propagation of

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

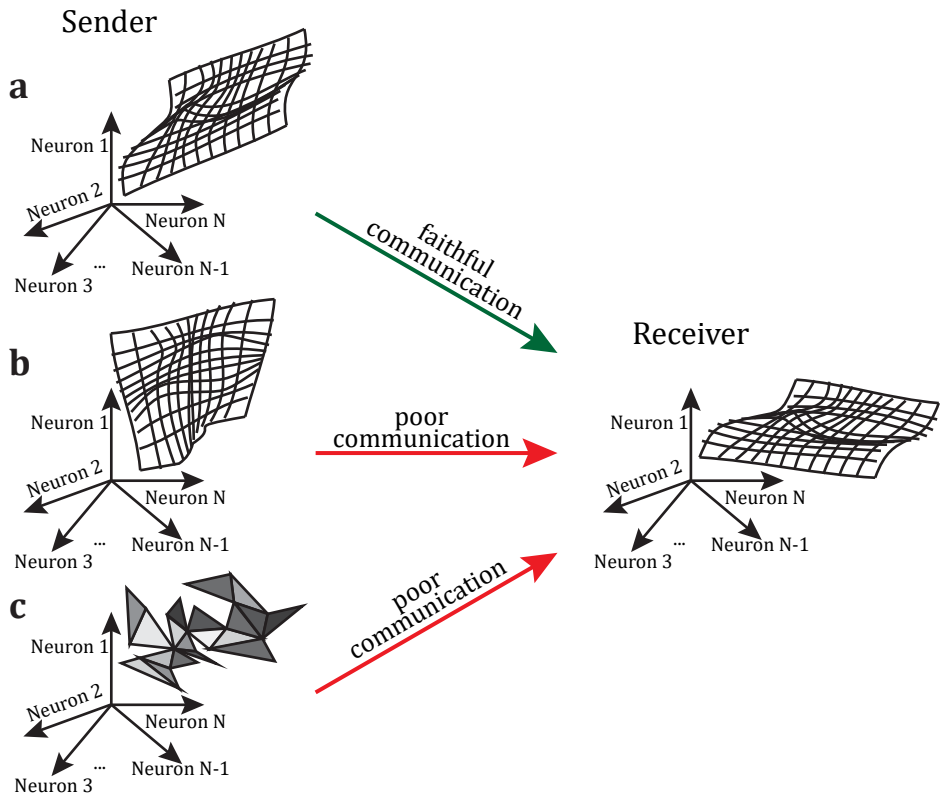

Figure 7: Communication is disrupted if there is a mismatch in the dimensionality or the alignment of within-area shared fluctuations in the sender and receiver networks. Schematic of the manifolds of shared fluctuations in the sending and receiving areas. (**a**) When shared dimensionality is similarly low-dimensional in the sender and receiver layers and the shared manifolds can be aligned, communication is faithful. (**b**) When shared dimensionality is similarly low-dimensional in the sender and receiver layers, but the manifolds cannot be properly aligned through a linear transformation, communication is poor. (**c**) When shared dimensionality is higher in the sender than in the receiver layer, communication is poor.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

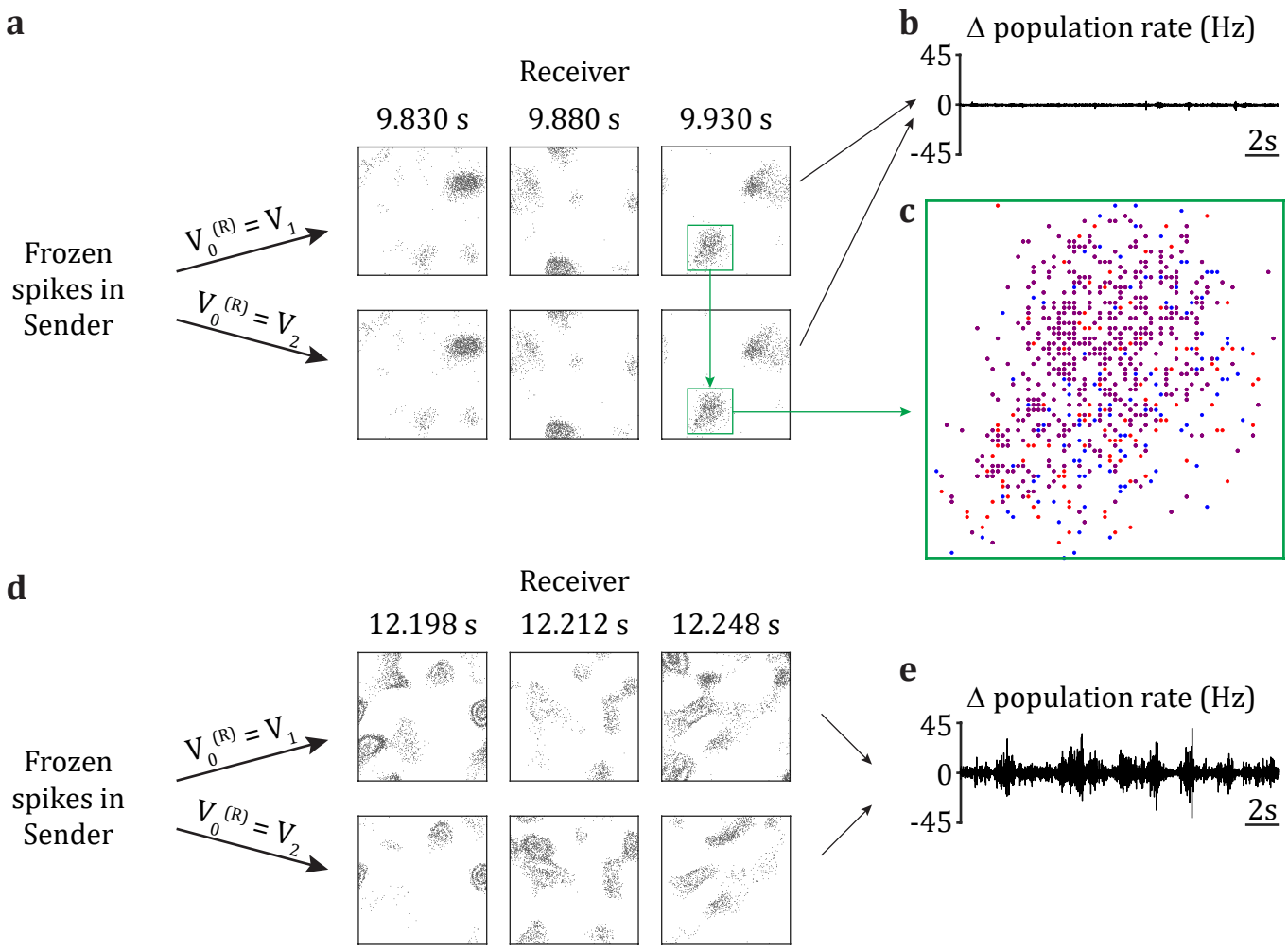

Figure 8: Receiver is driven by Sender, even when prediction performance of the communication subspace is low.
(**a**) Networks with $\sigma_I^{(S)} = 0.1$ and $\sigma_I^{(R)} = 0.3$ yield poor prediction performance of the communication subspace (Fig. 5c). Spikes are generated in the input layer and drive the Sender layer, where spikes are frozen across trials. The membrane potential of neurons in the Receiver (R) is set to two different initial conditions: V_1 or V_2. Raster plots at three time points are presented ($\Delta t = 2$ ms). (**b**) Difference in the Receivers' excitatory population firing rate (spikes per second) over time between the two trials with different membrane potential initial conditions. (**c**) Zoom in the raster plot of (a) at time 9.930 s. Blue dots are spikes for the first trial, red dots are spikes for the second trial and purple dots are overlapping spikes of the first and second trials. (**d**) Same as (a), for the network where $\tau_I^{\text{decay}(S)} = 8$ ms and $\tau_I^{\text{decay}(R)} = 24$ ms, which yields poor prediction performance of the communication subspace (Fig. S4). (**e**) Corresponding difference in the Receiver's excitatory population firing rate.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

activity from sending to receiving populations? By design, the architecture of our network involves strong excitatory synaptic connections between the sender and receiver layers. Further, the input from the sender area is the only one that drives the neurons in the receiver layer. Therefore, disrupted communication between the sender and receiver networks is not due to a simple weakening of interactions, or competition from other sources. Rather, the linear regression framework required to measure the communication subspace may be blind to nonlinear sender – receiver network interactions.

We test this hypothesis by setting the network parameters in a regime with low prediction performance of the communication subspace due to emergence of low-dimensional shared fluctuations in the receiver network (Fig. 5c). We generate an instance of spikes in the input layer and record the resulting spike times in the sender layer. We then compare two response trials in the receiving layer, where the only difference between them is the initial membrane voltage of the excitatory neurons in the receiving area (Fig. 8a). Despite the receiver network being in a regime where internally generated patterns occur, the difference in initial states does not affect the macroscopic pattern of spiking activity across trials. Indeed, the two trials exhibit the same statistical properties: the difference in population firing rate between the two trials remains low throughout the network simulations (Fig. 8b). Overall, this shows that the macroscopic population activity in the receiving area is entrained to the activity in the sending layer. Thus, despite poor communication between the sender and receiver networks, as measured through linear analysis, there remains a very strong interaction that dictates receiver population activity.

While the macroscopic network activity is reliable across trials, an analysis of the microscopic activity shows that it is nevertheless unreliable. The exact spike train sequences across the network differ significantly between the two trials (Fig. 8c), so that any observer of a single neuron would easily distinguish the trials. These results agree with previous studies where a weak perturbation to network activity yields only a transient change in firing rate but a long-lasting decorrelation of spike sequences in balanced networks, often termed microscopic chaos (London et al., 2010; Monteforte and Wolf, 2012). Our network architecture involves spatially-organized recurrent and feedforward connections. Such networks can either be set in an asynchronous regime or a correlated balanced state, where balance is achieved over the whole spatial domain rather than locally (Rosenbaum and Doiron, 2014). Similar to disordered recurrent networks, spatially-organized networks in the asynchronous regime yield microscopic chaos when weakly perturbed (Rosenbaum and Doiron, 2014). Here we show that when the network is set in a correlated balanced state, where macroscopic spatio-temporal patterns emerge, a weak perturbation still only yields microscopic chaos. Therefore, our results extend our understanding of neuronal dynamics in weakly perturbed networks with biologically-constrained architecture.

We observe similar results when the E/I balance is destabilized temporally in the receiver network (Fig. 8d,e), a scenario which also leads to poor communication (Fig. S4). A difference in the initial membrane voltage of the excitatory neurons in the receiver layer does not affect the global structure of spatio-temporal patterns of activity (Fig. 8d). The difference in population firing rate is larger than for the spatial destabilization (compare Fig. 8e to Fig. 8b). Yet it remains within a relatively narrow range and does not diverge with time.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In summary, we have shown evidence that the receiver layer is always effectively driven by the sender layer in our network, even in scenarios with apparently disrupted communication. Because our communication measure is linear, these results highlight a nonlinear mapping of activity between sender and receiver layers. Therefore, it points for the need of nonlinear measures when deciphering between-area communication in neuronal networks with complex spatio-temporal dynamics.

Discussion

The mechanisms that produce the low-dimensional shared variability across a neuronal population can be organized into two broad categories. First, shared variability of the neurons in a brain region may be inherited (in part) from connecting brain areas (Wimmer et al., 2015; Gómez-Laberge et al., 2016; Semedo et al., 2019). Second, shared variability may be an emergent property of a brain area, owing to local, complex recurrent interactions between neurons (Darshan et al., 2018; Landau and Sompolinsky, 2018; Mastrogiuseppe and Ostojic, 2018; Huang et al., 2019). Population recordings restricted to a single brain area cannot easily disentangle the contributions of each mechanic to the total shared variability. Rather, multi-area brain recordings will be needed to expose these separate mechanisms (Urai et al., 2022). In our study, we used computational modeling to explore signatures of inherited or emergent shared variability within a brain-area by measuring the (linear) communication between distinct, but connected, brain areas.

A suitable modeling framework to investigate the interplay between inheritance and emergence of neuronal variability in brain circuits has only recently become available. At one extreme the inheritance of neuronal fluctuations has been extensively studied through the analysis of activity propagation in feedforward networks (Abeles, 1991; Diesmann et al., 1999; Reyes, 2003; Rosenbaum et al., 2011). However, since those circuits explicitly lacked within layer recurrent connections, they could not model the emergence of rich population dynamics. At the other extreme, within-population recurrent, yet unstructured, excitatory and inhibitory connections were introduced to model asynchronous activity reminiscent of the baseline state of cortical variability (Van Vreeswijk and Sompolinsky, 1998; Amit and Brunel, 1997; Renart et al., 2010; Shadlen and Newsome, 1998). However, these networks did not produce rich, low dimensional, fluctuations shared across the population. Over the past few years novel modeling frameworks have included structure to the within-population recurrent wiring that permits low-dimensional fluctuations to intrinsically emerge within the network. Different approaches have been taken. On the one hand, forcing a low-rank structure of the recurrent connectivity matrix yields low-dimensional activity, revealing a strong relationship between structure and dynamics (Mastrogiuseppe and Ostojic, 2018; Landau and Sompolinsky, 2018). On the other hand, low-dimensional shared fluctuations can emerge due to a destabilization of the E/I balance despite a connectivity matrix with high rank (Darshan et al., 2018; Huang et al., 2019). In our work, we leverage those modeling frameworks to investigate how the emergence and inheritance of low-dimensional neuronal fluctuations affect between-area interactions in well-controlled settings.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We assess the interaction between a sender network and a receiver network through a linear communication subspace measure. Brain recordings are believed to frequently operate in a linear regime, in particular in early sensory areas (Stringer et al., 2021). Besides, linear methods are routinely used for the analysis of neuronal dynamics and in many cases have proven successful. At the single population level, they have unraveled stable low-dimensional manifolds in working memory networks (Murray et al., 2017) as well as in motor areas (Jiang et al., 2020). In addition, the suppression of rich spatio-temporal dynamics has been shown to increase the amount of linear Fisher information that is propagated down layers (Huang et al., 2022). However, cognition arises through the interaction of distributed brain regions. Recent technological advances have allowed a glimpse into distributed processing through simultaneous recordings from distinct neuronal populations, allowing us to answer the question of how much of the dynamics in a population can be explained by the activity of another recorded brain area. At this multiple populations level, linear methods have also been critical. They have exhibited selectivity (Kaufman et al., 2014) and a low-dimensional subspace (Semedo et al., 2019; Srinath et al., 2021) in the communication between brain areas. Our work shows that those methods provide a useful lens to investigate the inheritance of neuronal fluctuations by a receiver displaying linear mapping, in particular in low-dimensional inheritance settings.

However, when the receiving area is in a nonlinear regime, where complex spatio-temporal dynamics intrinsically emerge, we show that between-area communication cannot be properly assessed through linear measures. Our results support a recent study revealing that when a network exhibits strong pairwise correlations, reminiscent of low-dimensional pattern formation, connectivity inference is biased towards an excess of connectivity between highly correlated neurons (Das and Fiete, 2020). Therefore, linear inference methods are only appropriate when neuronal dynamics operate in a linear regime, where activity is high-dimensional and unstructured (Das and Fiete, 2020). Even though early sensory areas are believed to mostly operate in a linear regime to faithfully encode sensory inputs, brain areas involved in higher-level cognition exhibit low-dimensional spatio-temporal dynamics (Wang, 2002; Mante et al., 2013; Lara et al., 2018; Chen et al., 2021). Besides, it has recently been shown that microscopic irregularity can subsist even in the presence of macroscopic fluctuations (Pyle and Rosenbaum, 2017; Darshan et al., 2018). Accordingly, our results indicate that even when the receiver network is in the pattern-forming regime it is nevertheless effectively driven by the sender network, as reflected by microscopic chaos upon a weak perturbation. It is simply that the linear methods of communication that we have used (Semedo et al., 2019) are blind to this interaction. Therefore, novel nonlinear methods will have to be developed to accurately assess communication involving brain areas with more complex dynamics, such as those involved in higher-level cognition.

Finally, the simplicity of our modeling framework is critical to thoroughly study the interplay between inheritance and emergence of neuronal fluctuations in tightly controlled settings. Indeed, the use of feedforwardly connected distinct sender and receiver networks, each only involving local recurrent connections, allows us to observe the differential effects of low-dimensional shared fluctuations on between-area communication depending on their well-defined origin. However, brain circuits show recurrent architecture spanning a wide range of spatial scales. Furthermore, cognition is believed to arise from distributed computational processes. Even in early sensory areas

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

historically thought of as mostly feedforward, the importance of feedback interactions in sensory processing has started to be exposed (Semedo et al., 2022). Therefore, a natural extension of our work will be the implementation of more complex circuit architectures, starting by the introduction of feedback connections, to provide novel insights into the mechanistic interplay of inheritance and emergence of shared fluctuation across spatial scales.

Methods

Network structure

A three-layer spiking network model is implemented similar to previous work (Huang et al., 2019). The input layer consists of 2,500 excitatory neurons whose spikes are taken from independent homogeneous (space and time) Poisson processes with a uniform rate of 10 Hz. The sender and receiver layers each consist of 40,000 excitatory (E) and 10,000 inhibitory (I) neurons which are arranged on a unit square domain $\Gamma = [0,1] \times [0,1]$ with periodic boundary conditions. The probability of connection between a presynaptic neuron belonging to class $\beta \in \{\text{E,I}\}$ located at position $\vec{y} = (y_1, y_2)$ and a postsynaptic neuron belonging to class $\alpha \in \{\text{E,I}\}$ located at position $\vec{x} = (x_1, x_2)$ depends on their pairwise distance measured periodically on Γ:

$$p_{\alpha\beta}\left(\vec{x}, \vec{y}\right) = \frac{K_{\alpha\beta}^{\text{out}}}{N_\alpha} g(x_1 - y_1, \sigma_\beta) g(x_2 - y_2, \sigma_\beta) \tag{1}$$

where $K_{\alpha\beta}^{\text{out}}$ is the out-degree, so $\bar{p}_{\alpha\beta} = \frac{K_{\alpha\beta}^{\text{out}}}{N_\alpha}$ is the mean connection probability, and $g(u, \sigma)$ is a wrapped Gaussian distribution:

$$g(u, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{+\infty} e^{(-u+k)^2/(2\sigma^2)} \tag{2}$$

Excitatory feedforward connections between layers and recurrent excitatory and inhibitory connections within layers are spatially distributed according to a Gaussian with width σ_{ffwd}, σ_E and σ_I respectively.

Neuronal dynamics

Excitatory and inhibitory neurons in sender and receiver networks are modeled as conductance-based exponential integrate-and-fire neurons:

$$C_m \frac{\mathrm{d}V_j^\alpha}{\mathrm{d}t} = -g_L^\alpha \left(V_j^\alpha - V_L\right) + g_L^\alpha \Delta_T^\alpha \mathrm{e}^{(V_j^\alpha - V_T)/\Delta_T^\alpha} + I_j^\alpha \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Network connectivity	$\sigma_{\text{ffwd}} = 0.05$	$\sigma_{\text{rec}} = 0.1$	$\sigma_E = 0.1$	$\sigma_I = 0.1$
	$\bar{p}_E^{\text{ffwd}} = 0.1$	$\bar{p}_I^{\text{ffwd}} = 0.05$	$j_E^{\text{ffwd}} = 240$	$j_I^{\text{ffwd}} = 400$
	$\bar{p}_{EE}^{\text{rec}} = 0.01$	$\bar{p}_{EI}^{\text{rec}} = 0.04$	$\bar{p}_{IE}^{\text{rec}} = 0.03$	$\bar{p}_{II}^{\text{rec}} = 0.04$
	$j_{EE}^{\text{rec}} = 80$	$j_{EI}^{\text{rec}} = -240$	$j_{IE}^{\text{rec}} = 40$	$j_{II}^{\text{rec}} = -300$
Neuronal dynamics	$C_m = 1$ ms	$g_L^E = 1/15$	$g_L^I = 1/10$	$V_L = -60$ mV
	$V_{\text{th}} = -10$ mV	$V_{\text{reset}} = -65$ mV	$t_{\text{ref}}^E = 1.5$ ms	$t_{\text{ref}}^I = 0.5$ ms
	$V_T = -50$ mV	$\Delta_T^E = 2$ mV	$\Delta_T^I = 0.5$ mV	
	$\tau_E^{\text{rise}} = 1$ ms	$\tau_E^{\text{decay}} = 5$ ms	$\tau_I^{\text{rise}} = 1$ ms	$\tau_I^{\text{decay}} = 8$ ms

Table 1: Standard parameters for the simulations

with $\alpha \in \{\text{E}, \text{I}\}$. A neuron spikes when its membrane voltage V_j^α reaches the spiking threshold V_{th}. Then its membrane voltage is reset at V_{reset} for a refractory period t_{ref}^α. The total current received by neuron j belonging to class α, $I_j^\alpha(t)$, is given by the summation of feedforward and recurrent input:

$$\frac{I_j^\alpha(t)}{C_m} = \sum_{k=1}^{N_{\text{E}}^{\text{ffwd}}} \frac{j_{jk}^{\text{ffwd}}}{\sqrt{N}} \sum_n \eta_{\text{E}}(t - t_n^k) + \sum_{\beta \in \{\text{E},\text{I}\}} \sum_{k=1}^{N_\beta^{\text{rec}}} \frac{j_{jk}^{\text{rec}}}{\sqrt{N}} \sum_n \eta_\beta(t - t_n^k) \tag{4}$$

with $N = N_{\text{E}} + N_{\text{I}}$ the total number of neurons within the layer of interest. The postsynaptic current, $\eta_\beta(t)$, is induced by presynaptic spiking and depends on the class ($\beta \in \{\text{E}, \text{I}\}$) of the presynaptic neuron. Assuming a single presynaptic spike at time $t = 0$, it is given by the difference of two exponentials with rise timescale τ_β^{rise} and decay timescale $\tau_\beta^{\text{decay}}$:

$$\eta_\beta(t) = \begin{cases} \frac{\mathrm{e}^{-t/\tau_\beta^{\text{decay}}} - \mathrm{e}^{-t/\tau_\beta^{\text{rise}}}}{\tau_\beta^{\text{decay}} - \tau_\beta^{\text{rise}}}, & t \geqslant 0 \\ 0, & t < 0 \end{cases} \tag{5}$$

Equations are numerically integrated using the forward Euler method with a timestep of 0.05 ms. Unless specified otherwise, all neuronal and connectivity parameters are identical in the sender and the receiver layers (Table 1).

Shared covariance matrix and within-area shared dimensionality

Shared covariance of within-area neuronal activity $\vec{x}$ is assessed through factor analysis (FA):

$$p(\vec{x}) = \mathcal{N}(\vec{\mu}, C_{\text{shared}} + C_{\text{private}}) \tag{6}$$

where C_{private} is a diagonal matrix whose elements represent the individual neuronal variances and C_{shared} represents the shared component of the full covariance matrix (Everitt, 1984; Yu et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2009). Singular value decomposition is applied to C_{shared}:

$$C_{\text{shared}} = U \Lambda U^T \tag{7}$$

where the columns of U are the eigenvectors and the elements of the diagonal matrix Λ, λ_i, are the associated eigenvalues ordered from larger to smaller.

The dimensionality of the shared covariance matrix is estimated using participation ratio (PR) (Mazzucato et al., 2016; Litwin-Kumar et al., 2017):

$$\text{PR} \equiv \dim(C_{\text{shared}}) = \frac{\left(\sum_{i=1}^{N} \lambda_i\right)^2}{\sum_{i=1}^{N} \lambda_i^2} \tag{8}$$

Power spectrum

We estimate the Fourier transform of the excitatory spiking activity (1 ms bins) in space and time using the Welch's method with 200 ms timewindows overlapping by 10 ms, as done previously (Huang et al., 2022). We then compute the corresponding power by taking the average of the squared absolute value of the single-sided Fourier transform. As order parameter of the degree of pattern-formation, we use the peak of the power over all temporal frequencies ω and spatial wavenumbers k (excluding the null frequencies).

Between-area communication

To assess communication between sender and receiver networks, we use a recently developed communication subspace measure based on reduced-rank regression (Semedo et al., 2019):

$$\hat{R} = S B_{\text{RRR}}, \qquad \text{with} \quad \text{rank}(B_{\text{RRR}}) = m \tag{9}$$

with S the activity in the sender network and $\hat{R}$ the estimated activity in the receiver network. It can be shown that B_{RRR} is a low-rank approximation of the ordinary least-squares solution (Semedo et al., 2019):

$$B_{\text{RRR}} = B_{\text{OLS}} \Gamma_m \Gamma_m^T \tag{10}$$

with $B_{\text{OLS}} = \left(S^T S\right)^{-1} S^T R$. We define the optimal rank m as the smallest number of dimensions for which the predictive performance is within one SEM of the peak performance over 20 cross-validation folds. The dimensions in activity S that are most predictive of activity R according to the communication subspace measure are called "predictive dimensions". It can be shown that they are the m columns of $B_{\text{OLS}} \Gamma_m$. The predictive performance of the communication subspace is defined as the average Normalized Squared Error over 20 cross-validation folds (Semedo et al., 2019).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Alignment of shared variability

To align the simultaneously recorded latent dynamics in the sending and receiving populations, we use a method based on canonical correlation analysis which has recently been used to align within-area latent dynamics over several days where a different number of neurons with different identities were recorded (Gallego et al., 2020). More specifically, we know from equation (7) that the singular value decomposition of the shared covariance matrix is given by: $C_{\text{shared}} = U\Lambda U^T$. Hence the columns of U are the eigenvectors of C_{shared}, ordered from the one which explains the most shared variance to the one that explains the least. We keep $m = 20$ shared latent variables, as they are sufficient to explain most of the shared variance. For the sender and receiver networks separately, we project the activity of the 50 neurons that were sampled onto the 20 shared latent variables to obtain an $m \times T$ matrix of latent dynamics L_k, where T is the number of timepoints and $k \in \{S, R\}$. Then, we compute the QR decomposition of their transpose: $L_k^T = Q_k R_k$, where Q_k is a $T \times m$ matrix and R_k an $m \times m$ matrix. The singular value decomposition of the covariance of Q_S with Q_R is given by: $Q_S^T Q_R = \tilde{U}\tilde{S}\tilde{V}^T$. Canonical correlation analysis finds new latent directions to maximize the pairwise correlations between the sending and receiving populations. The projection of the latent dynamics onto these new latent directions is implemented through the corresponding $m \times m$ matrices: $M_S = R_S^{-1}\tilde{U}$ and $M_R = R_R^{-1}\tilde{V}$. We can then project the latent dynamics onto those new latent directions: $\tilde{L}_k^T = L_k^T M_k$. Finally, the aligned canonical correlations are given by the pairwise correlations between the aligned latent dynamics: $\tilde{L}_S \tilde{L}_R^T = \tilde{U}^T Q_S^T Q_R \tilde{V} = \tilde{S}$. They are the elements of the diagonal matrix $\tilde{S}$, which are ordered from larger to lower value (Gallego et al., 2020).

Acknowledgements

This work was supported by grants from the Simons Foundation Collaboration on the Global Brain, the National Institutes of Health Grants (1U19NS107613-01 and R01 EB026953), and the Vannevar Bush Faculty Fellowship (N00014-18-1-2002).

References

Abeles, M., Corticonics: Neural circuits of the cerebral cortex (Cambridge University Press, 1991).

Amit, D.J. and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cerebral cortex (New York, NY: 1991) 7, 237–252.

Chaudhuri, R., Knoblauch, K., Gariel, M.A., Kennedy, H., and Wang, X.J. (2015). A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. Neuron 88, 419–431.

Chen, G., Kang, B., Lindsey, J., Druckmann, S., and Li, N. (2021). Modularity and robustness of frontal cortical networks. Cell 184, 3717–3730.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Churchland, M.M., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nature neuroscience 13, 369–378.

Cohen, M.R. and Kohn, A. (2011). Measuring and interpreting neuronal correlations. Nature neuroscience 14, 811.

Cunningham, J.P. and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. Nature neuroscience 17, 1500–1509.

Darshan, R., Van Vreeswijk, C., and Hansel, D. (2018). Strength of correlations in strongly recurrent neuronal networks. Physical Review X 8, 031072.

Das, A. and Fiete, I.R. (2020). Systematic errors in connectivity inferred from activity in strongly recurrent networks. Nature Neuroscience 23, 1286–1296.

Diesmann, M., Gewaltig, M.O., and Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. Nature 402, 529–533.

Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G.K., and Josić, K. (2016). The mechanics of state-dependent neural correlations. Nature neuroscience 19, 383–393.

Everitt, B., An introduction to latent variable models (London: Chapman & Hall, 1984).

Gallego, J.A., Perich, M.G., Chowdhury, R.H., Solla, S.A., and Miller, L.E. (2020). Long-term stability of cortical population dynamics underlying consistent behavior. Nature neuroscience 23, 260–270.

Gómez-Laberge, C., Smolyanskaya, A., Nassi, J.J., Kreiman, G., and Born, R.T. (2016). Bottom-up and top-down input augment the variability of cortical neurons. Neuron 91, 540–547.

Hahn, G., Ponce-Alvarez, A., Deco, G., Aertsen, A., and Kumar, A. (2019). Portraits of communication in neuronal networks. Nature Reviews Neuroscience 20, 117–127.

Holmgren, C., Harkany, T., Svennenfors, B., and Zilberter, Y. (2003). Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. The Journal of physiology 551, 139–153.

Huang, C., Pouget, A., and Doiron, B. (2022). Internally generated population activity in cortical networks hinders information transmission. Science Advances In Press, https://www.biorxiv.org/content/early/2020/02/04/2020.02.03.932723.

Huang, C., et al. (2019). Circuit models of low-dimensional shared variability in cortical networks. Neuron 101, 337–348.

Jiang, X., Saggar, H., Ryu, S.I., Shenoy, K.V., and Kao, J.C. (2020). Structure in neural activity during observed and executed movements is shared at the neural population level, not in single neurons. Cell reports 32, 108006.

Kaufman, M.T., Churchland, M.M., Ryu, S.I., and Shenoy, K.V. (2014). Cortical activity in the null space: permitting preparation without movement. Nature neuroscience 17, 440–448.

Keane, A. and Gong, P. (2015). Propagating waves can explain irregular neural dynamics. Journal of Neuroscience 35, 1591–1605.

Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nature reviews neuroscience 11, 615–627.

Landau, I.D. and Sompolinsky, H. (2018). Coherent chaos in a recurrent neural network with structured connectivity. PLoS computational biology 14, e1006309.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Lara, A.H., Cunningham, J.P., and Churchland, M.M. (2018). Different population dynamics in the supplementary motor area and motor cortex during reaching. Nature communications 9, 1–16.

Levy, R.B. and Reyes, A.D. (2012). Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. Journal of Neuroscience 32, 5609–5619.

Litwin-Kumar, A., Harris, K.D., Axel, R., Sompolinsky, H., and Abbott, L. (2017). Optimal degrees of synaptic connectivity. Neuron 93, 1153–1164.

London, M., Roth, A., Beeren, L., Häusser, M., and Latham, P.E. (2010). Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. Nature 466, 123–127.

Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. nature 503, 78–84.

Mastrogiuseppe, F. and Ostojic, S. (2018). Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. Neuron 99, 609–623.

Mazzucato, L., Fontanini, A., and La Camera, G. (2016). Stimuli reduce the dimensionality of cortical activity. Frontiers in systems neuroscience 10, 11.

Monteforte, M. and Wolf, F. (2012). Dynamic flux tubes form reservoirs of stability in neuronal circuits. Physical Review X 2, 041007.

Muller, L., Chavane, F., Reynolds, J., and Sejnowski, T.J. (2018). Cortical travelling waves: mechanisms and computational principles. Nature Reviews Neuroscience 19, 255–268.

Murray, J.D., et al. (2017). Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. Proceedings of the National Academy of Sciences 114, 394–399.

Ocker, G.K., et al. (2017). From the statistics of connectivity to the statistics of spike times in neuronal networks. Current opinion in neurobiology 46, 109–119.

Pyle, R. and Rosenbaum, R. (2017). Spatiotemporal dynamics and reliable computations in recurrent spiking neural networks. Physical review letters 118, 018103.

Renart, A., et al. (2010). The asynchronous state in cortical circuits. Science 327, 587–590.

Reyes, A.D. (2003). Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. Nature neuroscience 6, 593–599.

Rosenbaum, R. and Doiron, B. (2014). Balanced networks of spiking neurons with spatially dependent recurrent connections. Physical Review X 4, 021039.

Rosenbaum, R., Smith, M.A., Kohn, A., Rubin, J.E., and Doiron, B. (2017). The spatial structure of correlated neuronal variability. Nature neuroscience 20, 107–114.

Rosenbaum, R., Trousdale, J., and Josic, K. (2011). The effects of pooling on spike train correlations. Frontiers in neuroscience 5, 58.

Rossi, L.F., Harris, K.D., and Carandini, M. (2020). Spatial connectivity matches direction selectivity in visual cortex. Nature 588, 648–652.

Schölvinck, M.L., Saleem, A.B., Benucci, A., Harris, K.D., and Carandini, M. (2015). Cortical state determines global variability and correlations in visual cortex. Journal of Neuroscience 35, 170–178.

Semedo, J.D., Zandvakili, A., Machens, C.K., Byron, M.Y., and Kohn, A. (2019). Cortical areas interact through a communication subspace. Neuron 102, 249–259.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Semedo, J.D., et al. (2022). Feedforward and feedback interactions between visual cortical areas use different population activity patterns. Nature communications 13, 1–14.

Shadlen, M.N. and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. Journal of neuroscience 18, 3870–3896.

Srinath, R., Ruff, D.A., and Cohen, M.R. (2021). Attention improves information flow between neuronal populations without changing the communication subspace. Current Biology 31, 5299–5313.

Stringer, C., Michaelos, M., Tsyboulski, D., Lindo, S.E., and Pachitariu, M. (2021). High-precision coding in visual cortex. Cell 184, 2767–2778.

Urai, A.E., Doiron, B., Leifer, A.M., and Churchland, A.K. (2022). Large-scale neural recordings call for new insights to link brain and behavior. Nature neuroscience 25, 1–9.

Van Vreeswijk, C. and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. Neural computation 10, 1321–1371.

Wang, X.J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36, 955–968.

Williamson, R.C., Doiron, B., Smith, M.A., and Byron, M.Y. (2019). Bridging large-scale neuronal recordings and large-scale network models using dimensionality reduction. Current opinion in neurobiology 55, 40–47.

Williamson, R.C., et al. (2016). Scaling properties of dimensionality reduction for neural populations and network models. PLoS computational biology 12, e1005141.

Wimmer, K., et al. (2015). Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area mt. Nature communications 6, 1–13.

Yamins, D.L. and DiCarlo, J.J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nature neuroscience 19, 356–365.

Yu, B.M., et al. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. Journal of neurophysiology 102, 614–635.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

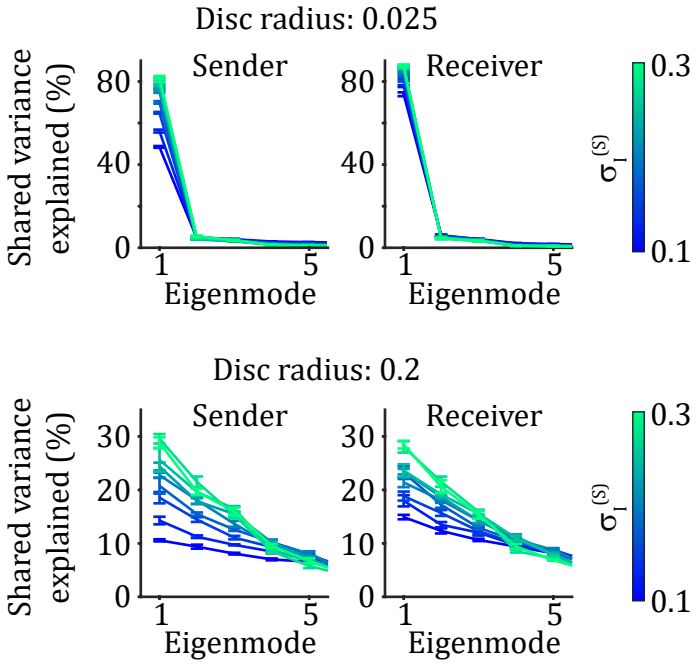

Figure S1: The structure of within-area shared fluctuations is spatial-scale dependent. Shared variance explained by the first five eigenmodes within sender network (S, left) and within receiver network (R, right) when modifying σ_I in S. Neurons are randomly sampled from a small disc (top) or from a large disc (bottom).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

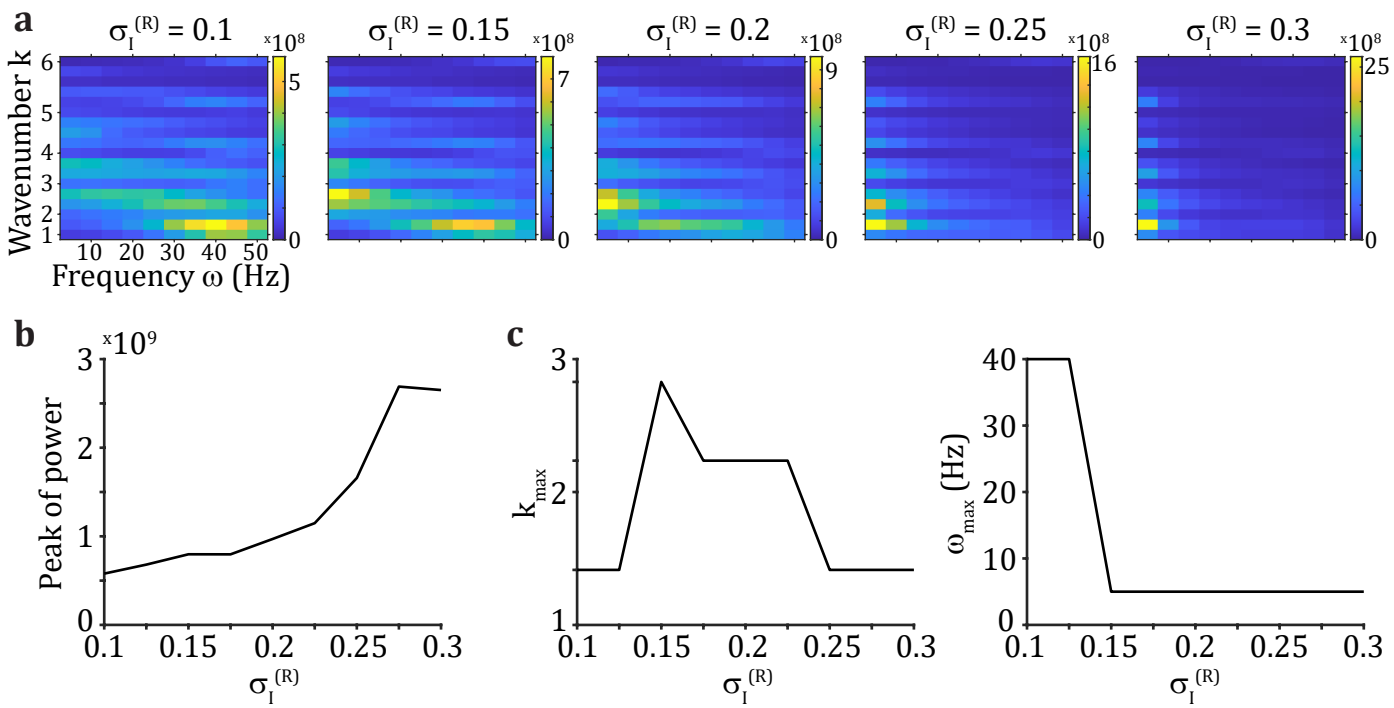

Figure S2: Spatio-temporal power spectrum as the network is spatially destabilized. (**a**) Power as a function of temporal frequency ω and spatial wavenumber k when modifying σ_I in the receiver network (R). The wavenumber is given by $k = \sqrt{x^2 + y^2}$, where x and y are the coordinates in the 2-dimensional grid space. (**b**) Peak of the power over all temporal frequencies ω and spatial wavenumbers k. (**c**) Spatial wavenumber $k_{\max}$ (left) and temporal frequency $\omega_{\max}$ (right) at the location of the peak of power.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

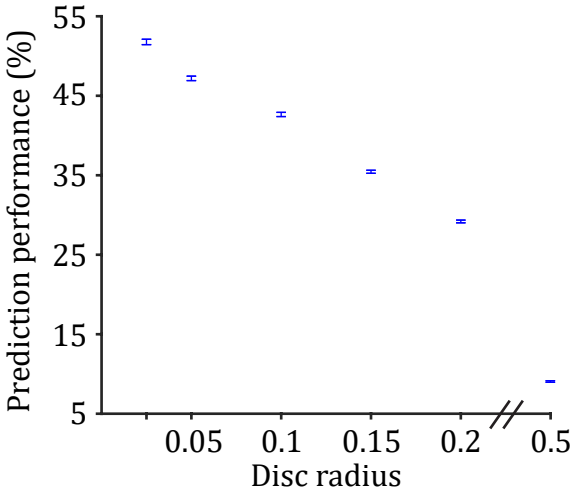

Figure S3: Dependence on the size of the spatial domain from which neurons are sampled on prediction performance of the communication subspace.
Scenario with the standard network parameters ($\sigma_I^{(S)} = \sigma_I^{(R)} = 0.1$).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

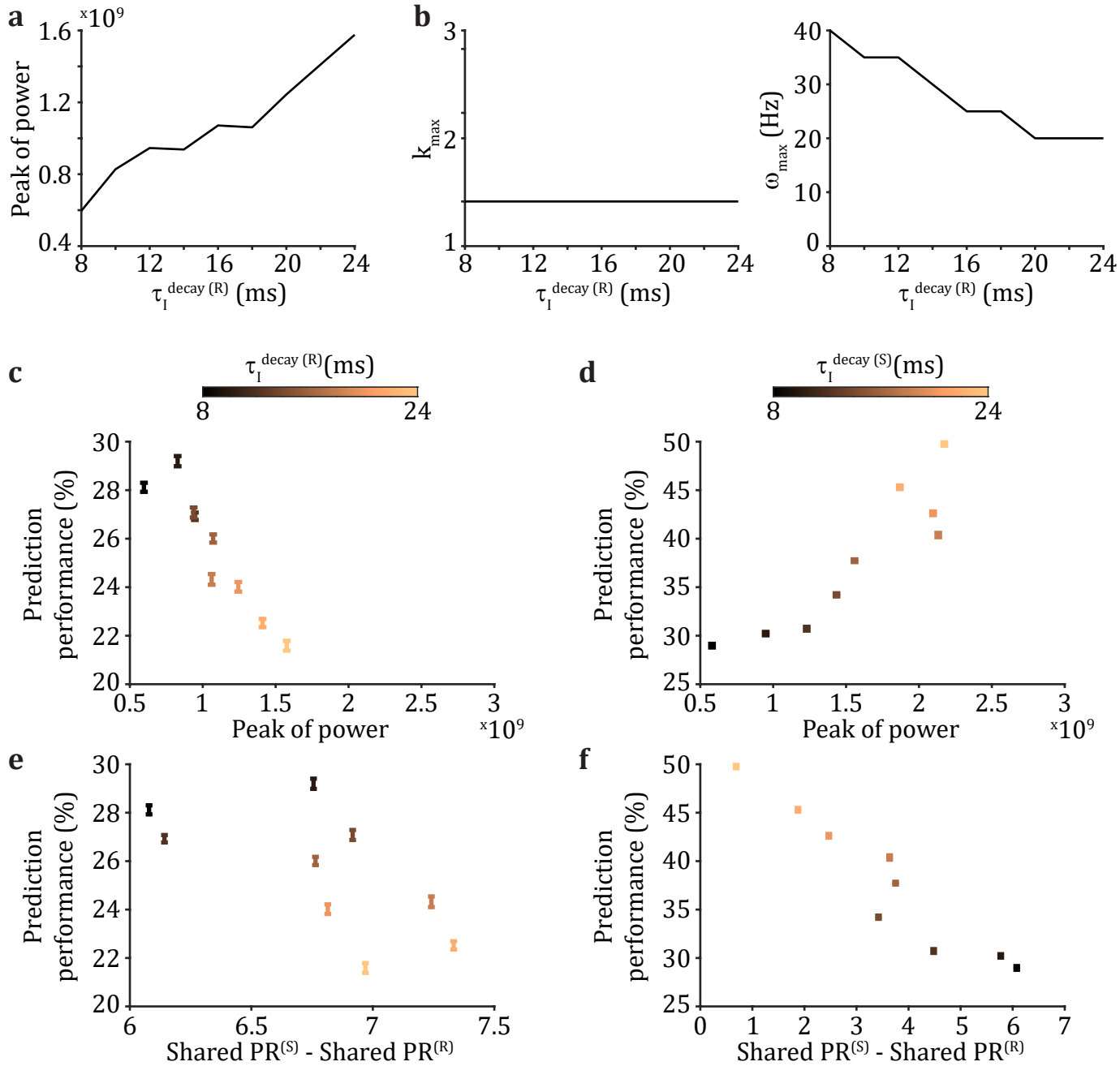

Figure S4: Inter-area communication is oppositely affected if temporal patterns emerge within the receiver network or if they are inherited from the sender network.

(**a**) Peak of the power over all temporal frequencies ω and spatial wavenumbers k as a function of τ_I^{decay} in the receiver network (R). (**b**) Spatial wavenumber $k_{\max}$ (left) and temporal frequency $\omega_{\max}$ (right) at the location of the peak of power. (**c,d**) Prediction performance of the communication subspace between S and R when τ_I^{decay} is modified in R (c) or in S (d). (**e,f**) Prediction performance of the communication subspace is higher when shared dimensionality in S and R is similarly low, while it is lower when shared dimensionality in S is much higher than in R, no matter if spatio-temporal patterns emerge in R (e) or in S (f). Neurons are sampled from a disc with radius 0.2; errorbars are mean ± SEM.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.11.487906; this version posted May 11, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

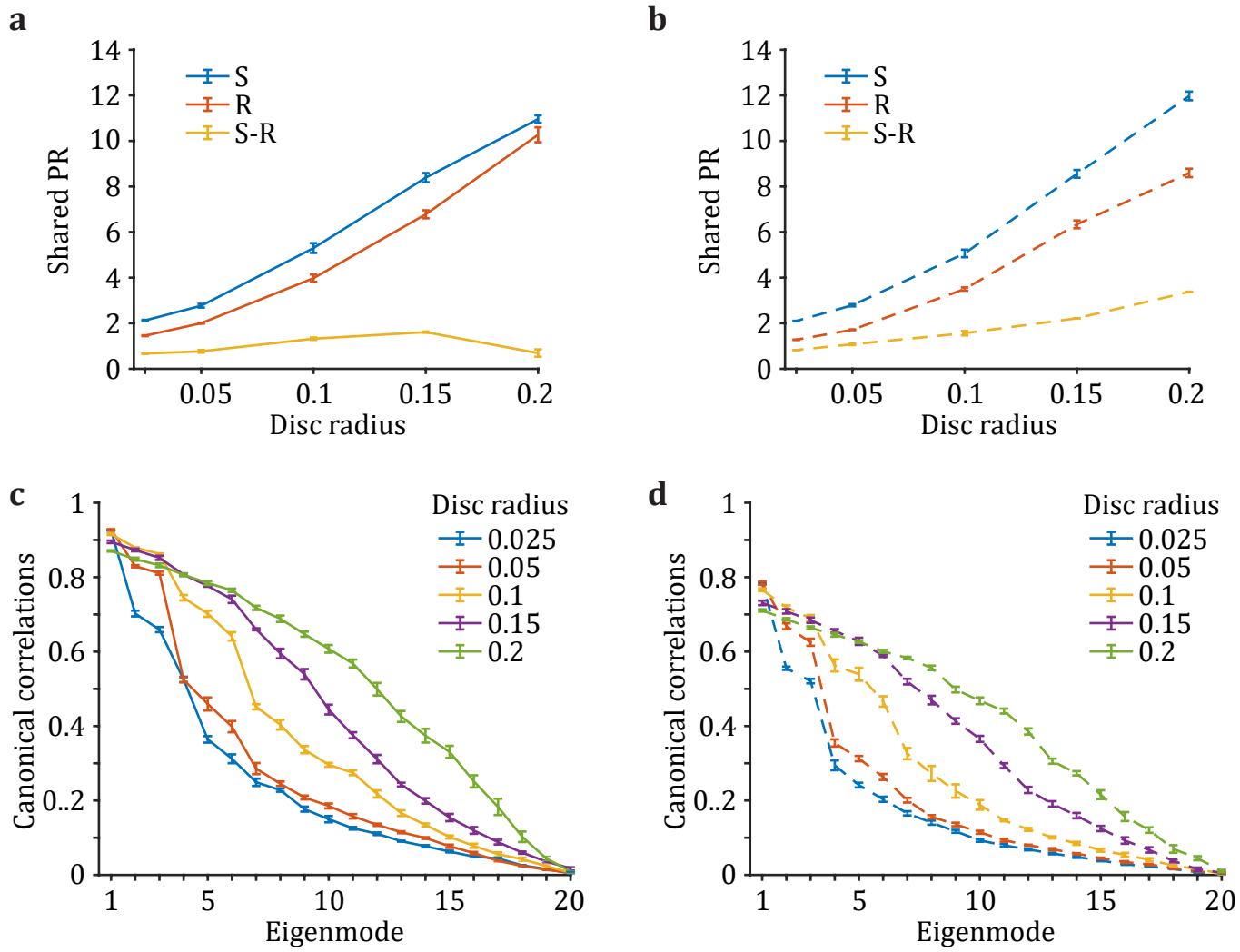

Figure S5: Emergence of novel spatio-temporal patterns in the receiver network yields a misalignment of the shared variability in sending and receiving populations. (**a**) Shared dimensionality in the sender network (S), the receiver network (R), and the difference between the two (S-R) in the case where spatio-temporal patterns emerge in S only (Fig. 6a, top). (**b**) Same as (a), except that different spatio-temporal patterns emerge in S and R (Fig. 6a, bottom). (**c,d**) Canonical correlations between the shared fluctuations in S and R when spatio-temporal patterns emerge in S only (c), and when different spatio-temporal patterns emerge in S and R (d).