Stimuli reduce the dimensionality of cortical activity

Luca Mazzucato[1] **Alfredo Fontanini**[1,2] **and Giancarlo La Camera**[1,2]

[1]*Department of Neurobiology and Behavior and*

[2]*Graduate Program in Neuroscience, State University of New York at Stony Brook, Stony Brook, NY 11794*

ABSTRACT: The activity of ensembles of simultaneously recorded neurons can be represented as a set of points in the space of firing rates. Even though the dimension of this space is equal to the ensemble size, neural activity can be effectively localized on smaller subspaces. The dimensionality of the neural space is an important determinant of the computational tasks supported by the neural activity. Here, we investigate the dimensionality of neural ensembles from the sensory cortex of alert rats during period of ongoing (inter-trial) and stimulus-evoked activity. We find that dimensionality grows linearly with ensemble size, and grows significantly faster during ongoing activity compared to evoked activity. We explain these results using a spiking network model based on a clustered architecture. The model captures the difference in growth rate between ongoing and evoked activity and predicts a characteristic scaling with ensemble size that could be tested in high-density multi-electrode recordings. Moreover, the model predicts the existence of an upper bound on dimensionality. This upper bound is inversely proportional to the amount of pair-wise correlations and, compared to a homogeneous network without clusters, it is larger by a factor equal to the number of clusters. The empirical estimation of such bounds depends on the number and duration of trials. Together, these results provide a framework to analyze neural dimensionality in alert animals, its behavior under stimulus presentation, and its theoretical dependence on ensemble size, number of clusters, and pair-wise correlations in spiking network models.

Emails: luca dot mazzucato, alfredo dot fontanini, giancarlo dot lacamera at stonybrook.edu

KEYWORDS: gustatory cortex; dimensionality; hidden Markov models; network dynamics; ongoing activity; spiking network models

Contents

1 Introduction

Understanding the dynamics of neural activity and how it is generated in cortical circuits is a fundamental question in Neuroscience. The spiking activity of ensembles of simultaneously recorded neurons can be represented in terms of sequences of firing rate vectors, as shown e.g. in frontal [1–3], gustatory [4, 5], motor [6], premotor and somatosensory cortex [7]. The dimension of each firing rate vector is equal to the number of ensemble neurons N and the collection of rate vectors across trials takes the form of a set of points in the N-dimensional space of firing rates. Such points may not fill the whole space, but be restricted to lie inside a lower-dimensional subspace (see e.g. [8]). Roughly, dimensionality is the minimal number of dimensions necessary to provide an accurate description of the neural dynamics. If ensemble neurons are independent of each other, neural activities at different times will scatter around in the space of firing rate, filling a large portion of the space. In

this case, dimensionality will be maximal and equal to the size of the ensemble N. At the other extreme, if all neurons have the same time-dependent firing rate, ensemble activity localizes along a line. In this case, dimensionality is minimal and equal to one. These simple examples suggest that dimensionality captures information about the structure of a cortical circuit and the functional relations among the simultaneously recorded neurons, such as their pair-wise correlation [9].

Different definitions of dimensionality have been introduced for different tasks and across neural systems [8–13]. Such measures of dimensionality can shed light on the underlying neural computation; for example, they can predict the onset of an error trial in a recall task [13], or can allow the comparison of classification accuracy between different brain areas (e.g., IT vs. V4) and synthetic algorithms [12]. Here, we investigate a measure of dimensionality closely related to the pair-wise correlations of simultaneously recorded neurons [9]. We elucidate the dependence of dimensionality on experimental parameters, such as ensemble size and interval length, and we show that it varies across experimental conditions. We address these issues by comparing recordings of ensembles of neurons from the gustatory cortex (GC) of alerts rats to a biologically plausible network model based on neural clusters with recurrent connectivity. This model captures neural activity in GC during periods of ongoing and stimulus-evoked activity, explaining how the spatiotemporal dynamics of ensemble activity is organized in sequences of metastable states and how single-neuron firing rate distributions are modulated by stimulus presentation [5]. Here, we show that the same model expounds the observed dependence of dimensionality on ensemble size and how such dependence is reduced by the presentation of a stimulus. By comparing the clustered network model with a homogeneous network without clusters, we find that the clustered network has a larger dimensionality that depends on the number of clusters and the pair-wise correlations among ensemble neurons. A simple theory explains these results and allows extrapolating the scaling of dimensionality to very large ensembles. Our theory shows that recurrent networks with clustered connectivity provide a substrate for high-dimensional neural representations, which may lead to computational advantages.

2 Results

2.1 Dimensionality of the neural activity

We investigate the dimensionality of sequences of firing rate vectors generated in the GC of alert rats during periods of ongoing or evoked activity (see Appendix A.1). To provide an intuitive picture of the meaning of dimensionality adopted in this paper, consider the firing rate vectors from N simultaneously recorded neurons. These vectors can occupy, a priori, the entire N-dimensional vector space minimally required to describe the population activity of N independent neurons. However, the sequence of firing rate vectors generated by the neural dynamics may occupy a subspace that is spanned by a smaller number $m < N$ of coordinate axes. For example, the data obtained by the ensemble of three simulated spike counts in Fig. 1 mostly lie on a 2D space, the plane shaded in gray. Although 3 coordinates are still required to specify all data points, a reduced representation of the data, such as that obtained from PCA, would quantify the dimension of the relevant subspace as being close to 2. To quantify this fact we use the following definition of dimensionality [9]

$$d = \left(\sum_{i=1}^{N} \tilde{\lambda}_i^2 \right)^{-1} , \tag{2.1}$$

where N is the ensemble size and $\tilde{\lambda}_i$ are the normalized eigenvalues of the covariance matrix, each expressing the fraction of the variance explained by the corresponding principal component (see Appendix A.5 for details). According to this formula, if the first n eigenvalues express each a fraction $1/n$ of the variance while the remaining eigenvalues vanish, the dimensionality is $d = n$. In less symmetric situations, d reflects roughly the

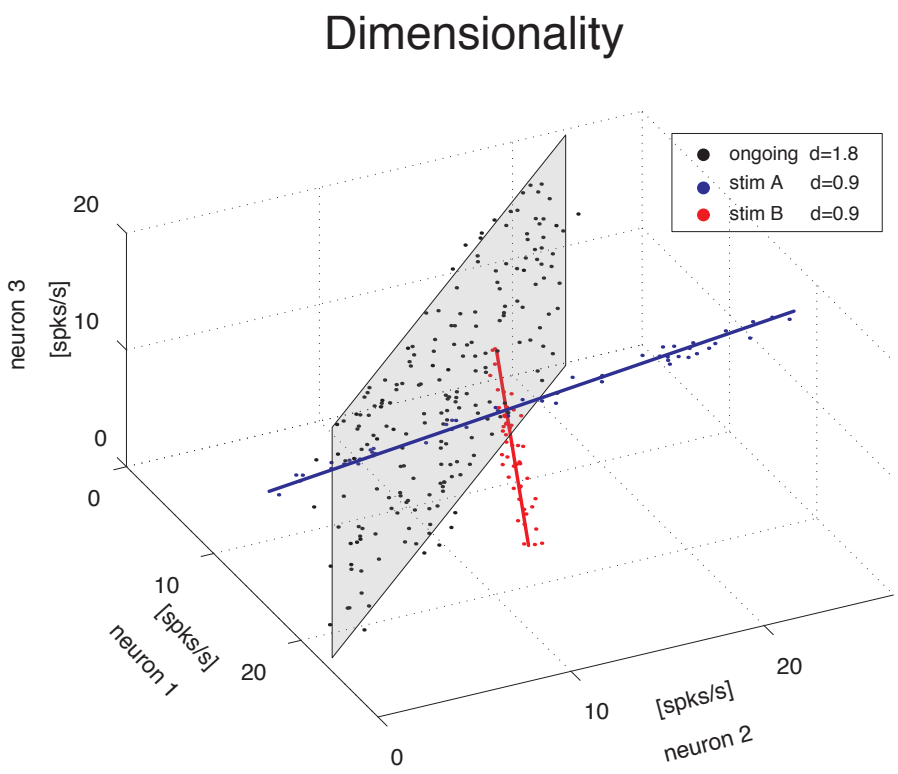

Figure 1. *Dimensionality of the neural representation.* Pictorial representation of the firing rate activity of an ensemble of $N = 3$ neurons. Each dot represents a three-dimensional vector of ensemble firing rates in one trial. Ensemble ongoing activity localizes around a plane (black dots cloud surrounding the shaded black plane), yielding a dimensionality of $d = 1.8$. Activity evoked by each of two different stimuli localizes around a line (red and blue dots clouds and lines), yielding a dimensionality of $d = 0.9$ in both cases.

dimension of the linear subspace explaining most variance about all data points. In the example of the data on the gray plane of Fig. 1, $d = 1.8$, which is close to 2, as expected. Similarly, data points lying mostly along the blue and red straight lines in Fig. 1 have a dimensionality of 0.9, close to 1. In all cases, $d > 0$ and $d \leq N$, where N is the ensemble size. The blue and red data points in Fig. 1 were obtained from a fictitious scenario where neuron 1 and neuron 2 were selective to surrogate stimuli A and B, respectively, and are meant to mimic two possible evoked responses. The subspace containing responses to both stimuli A and B would have a dimensionality $d_{A+B} = 1.7$, similar to the dimensionality of the data points distributed on the grey plane (meant instead to represent spike counts during ongoing activity in the same fictitious scenario). Thus, a dimensionality close to 2 could originate from different patterns of activity, such as occupying a plane or two straight lines. Other and more complex scenarios are, of course, possible. In general, the dimensionality will reflect existing functional relationships among ensemble neurons (such as pair-wise correlations) as well as the response properties of the same neurons to external stimuli. The pictorial example of Fig. 1 caricatures a stimulus-induced reduction of dimensionality, as found in the activity of simultaneously recorded neurons from the GC of alert rats, as we show next.

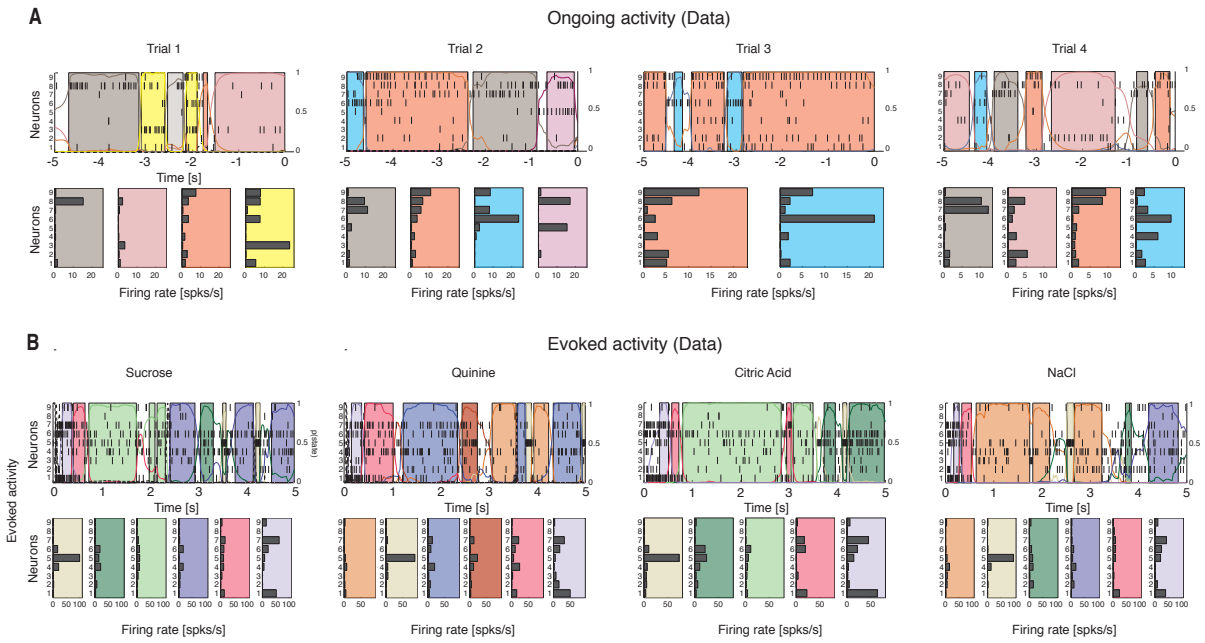

Figure 2. *Ensemble neural activity is characterized by sequences of states*. A: Upper panels: Representative trials from one ensemble of nine simultaneously recorded neurons during ongoing activity, segmented according to their ensemble states (HMM analysis, thin black vertical lines are action potentials; states are color-coded; smooth colored lines represent the probability for each state; shaded colored areas indicate intervals where the probability of a state exceeds 80%). Lower panels: Average firing rates across simultaneously recorded neurons (states are color-coded as in the upper panels). X-axis for population rasters: time preceding the next event at ($0 =$ stimulus delivery); Y-axis for population rasters: left, ensemble neuron index, right, probability of HMM states; X-axis for average firing rates panels: firing rates (spks/s); Y-axis for firing rate panels: ensemble neuron index. B: Ensemble rasters and firing rates during evoked activity for four different tastes delivered at $t = 0$: sucrose, sodium chloride, citric acid and quinine (notations as in panel A).

2.2 Dimensionality is proportional to ensemble size

We computed the dimensionality of the neural activity of ensembles of 3 to 9 simultaneously recorded neurons in the gustatory cortex of alert rats during the 5 s inter-trial period preceding (ongoing activity) and following (evoked activity) the delivery of a taste stimulus (see Appendix A.3 for details). Ensemble activity in single trials during both ongoing (Fig. 2A) and evoked activity (Fig. 2B) could be characterized in terms of sequences of metastable states, where each state is defined as a collection of firing rates across simultaneously recorded neurons [4, 5]. Transitions between consecutive states were detected via a Hidden Markov Model (HMM) analysis, which provides the probability that the network is in a certain state at every 1 ms bin (Fig. 2, color-coded lines superimposed to raster plots). The ensemble of spike trains was considered to be in a given state if the posterior probability of being in that state exceeded 80% in at least 50 consecutive 1-ms bins (Fig. 2, color-coded shaded areas). Transitions among states were triggered by the co-modulation of a variable number of ensemble neurons and occurred at seemingly random times [5]. For this reason, the dimensionality of the neural activity was computed based on the firing rate vectors in each HMM state (one firing rate vector per state per trial; see Appendix A.3 for details). The average dimensionality of ongoing activity across sessions was $d_{ongoing} = 2.6 \pm 1.2$ (mean±SD; range: $[1.2, 5.0]$; 27 sessions). An example of the eigenvalues for the representative ensemble of Fig. 2A is shown in Fig. 3A, where $d = 4.7$. The dimensionality of ongoing activity

was approximately linearly related to ensemble size (Fig. 3B, linear regression, $r = 0.4$, slope 0.26 ± 0.11, $p = 0.04$). During evoked activity dimensionality did not differ across stimuli (one-way ANOVA, no significant difference across tastants, $p > 0.8$), hence all evoked data points were combined for further analysis. An example of the eigenvalue distribution of the ensemble in Fig. 2B (the same ensemble as in Fig. 2A and 3A) is shown in Fig. 3C where $d_{evoked} = 1.3 \sim 1.7$ across 4 different taste stimuli. Across all sessions, dimensionality was overall smaller ($d_{evoked} = 2.0 \pm 0.6$, mean$\pm$SD, range: $[1.1, 3.9]$) and had a reduced slope as a function of N compared to ongoing activity (Fig. 3D, linear regression, $r = 0.39$, slope 0.13 ± 0.03, $p < 10^{-4}$). However, since dimensionality depends on the number and duration of the trials used for its estimation (Fig. 3E), a proper comparison requires matching trial number and duration for each data point, as described next.

2.3 Stimulus-induced reduction of dimensionality

We matched the number and duration of the trials for each data point and ran a two-way ANOVA with condition (ongoing vs. evoked) and ensemble size as factors. Both the main dimensionality ($F_{1,202} = 11.93$, $p < 0.001$) and the slope were significantly smaller during evoked activity (test of interaction, $F_{6,202} = 5.09$, $p < 10^{-4}$). There was also a significant effect of ensemble size ($F_{6,202} = 18.72$, $p < 10^{-14}$), confirming the results obtained with the separate regression analyses. These results suggest that stimuli induce a reduction of the effective space visited by the firing rate vector during evoked activity. Unlike ongoing activity, the dependence of dimensionality on ensemble size was modulated during different epochs of the post-stimulus period (two-way ANOVA with condition and time course as factors; main effect of time course $F_{4,495} = 3.80$, $p < 0.005$; interaction term: $F_{4,495} = 4.76$, $p < 0.001$). In particular, the dependence of d on the ensemble size N disappeared immediately after stimulus presentation ($d_{evoked} = 0.04 \pm 0.03$, compared to a trial- and bin length-matched average of $d_{ongoing} = 0.10 \pm 0.01$ during ongoing activity) and returned to the ongoing activity level after approximately 1 second (Fig. 3F).

In summary, we found that dimensionality depends on ensemble size during both ongoing and evoked periods, but such dependence is significantly reduced in the post-stimulus period; in particular, it disappears in the first second after taste delivery. In other words, while state sequences during ongoing activity explore a large portion of the available firing rate space, the presentation of a stimulus initially collapses the state sequence along a stereotyped and lower-dimensional response [14, 15].

2.4 Clustered spiking network model of dimensionality

To gain a mechanistic understanding of the different dimensionality of ongoing and evoked activity we have analyzed a spiking network model with clustered connectivity which has been shown to capture many essential features of the data [5]. In particular, the model reproduces the transitions among latent states in both ongoing and evoked activity. The network (see Appendix A.7 for details) comprises Q clusters of excitatory neurons characterized by stronger synaptic connections within each cluster and weaker connections between neurons in different clusters. All neurons receive recurrent input from a pool of inhibitory neurons that keeps the network in a balanced regime of excitation and inhibition in the absence of external stimulation (Fig. 4A). In very large networks (technically, in networks with an infinite number of neurons), the stable configurations of the neural activity are characterized by a finite number of active clusters whose firing rates depend on the number clusters active at any given moment, as shown in Fig. 4B (where $Q = 30$). In a finite network, however, finite size effects ignite transitions among these configurations, inducing network states (firing rate vectors) on randomly chosen subsets of neurons that resemble the HMM states found in the data (Fig. 5; see [5] for details). The dimensionality of the simulated sequences during ongoing and evoked activity was computed as done for the data, finding similar results. For the examples in Fig. 5, we found $d_{ongoing} = 4.0$ for ongoing activity (Fig. 6A) between $d_{evoked} = 2.2$ and $d_{evoked} = 3.2$ across tastes during evoked activity (Fig. 6C). Across all simulated

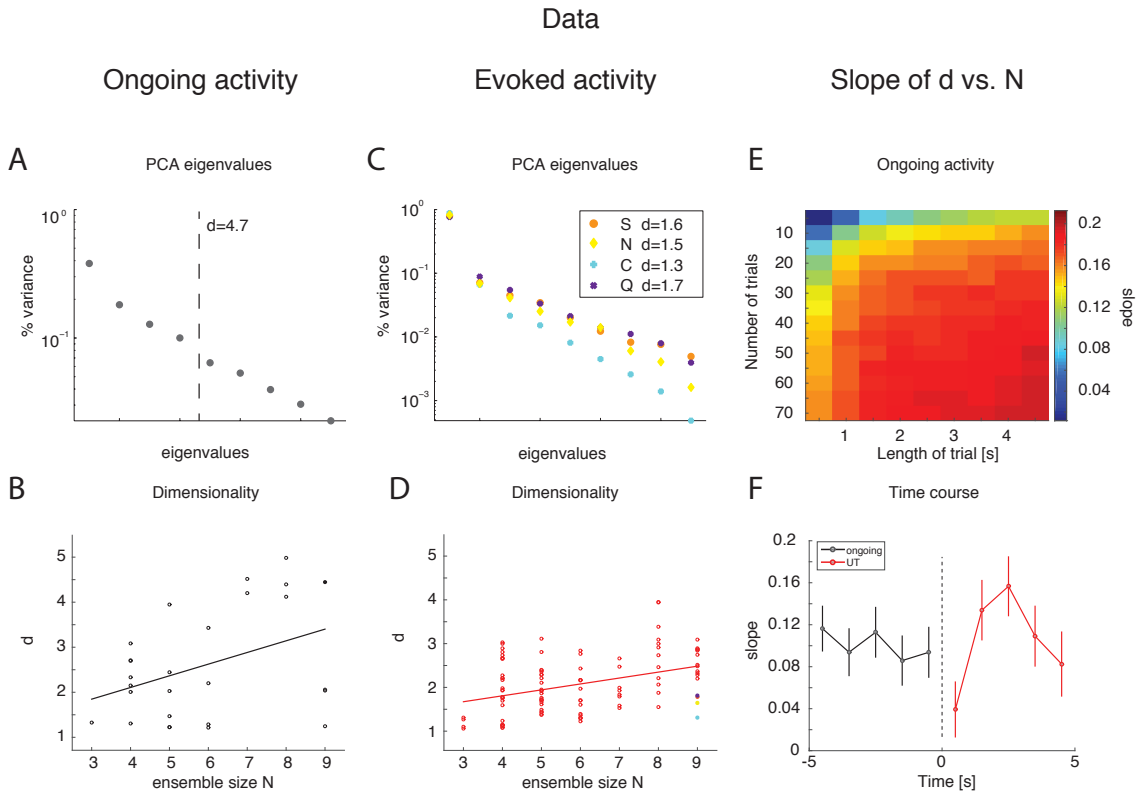

Figure 3. *Dependence of dimensionality on ensemble size (data).* A: Principal eigenvalue distribution for an ensemble of nine neurons during ongoing activity (corresponding to the filled dot in panel B) in the empirical dataset (Fig. 2). The dashed vertical line represents the value of the dimensionality for this ensemble ($d = 4.7$). X-axis: eigenvalue number; Y-axis: fraction of variance explained by each eigenvalue. B: Dimensionality of neural activity across all ensembles in the empirical dataset during ongoing activity (circles, linear regression fit overlaid). X-axis: ensemble size; Y-axis: dimensionality. C: Principal eigenvalue distribution for the ensemble in panel A during evoked activity. The eigenvalue distributions for sucrose (S, orange), sodium chloride (N, yellow), citric acid (C, cyan), and quinine (Q, blue) are presented (corresponding to the color-coded dots in panel D). X-axis: eigenvalue number; Y-axis: percentage of variance explained by each eigenvalue. D: Dimensionality of neural activity across all ensembles in the empirical dataset during evoked activity (notations as in panel B). E: The slope of the linear regression of dimensionality (d) vs. ensemble size (N) as a function of the length of the trial interval and the number of trials used to estimate the dimensionality. X-axis: length of trial interval [s]; Y-axis: number of trials. F: Time course of the slope of d vs. N, evaluated with HMM states appearing in a 1 second interval (error bars represent SD). A significant time course is triggered by stimulus presentation (see Section 2 for details).

sessions, we found an average $d_{ongoing} = 2.9 \pm 0.9$ (mean±SD) for ongoing activity and $d_{evoked} = 2.4 \pm 0.7$ for evoked activity. The model captured the essential properties of dimensionality observed in the data: the dimensionality did not differ across different tastes (one-way ANOVA, $p > 0.2$) and depended on ensemble size during both ongoing (Fig. 6B; slope $= 0.36 \pm 0.07$, $r = 0.77$, $p < 10^{-4}$) and evoked periods (Fig. 6D; slope $= 0.12 \pm 0.04$, $r = 0.29$, $p = 0.01$). As for the data, the dependency on ensemble size was smaller for evoked compared to ongoing activity. We performed a trial-matched two-way ANOVA as done on the data and found, also in the model, a main effect of condition (ongoing vs. evoked: $F_{1,159} = 20.92$, $p < 10^{-4}$), a main effect of ensemble size ($F_{6,159} = 3.58$, $p = 0.002$), and a significant interaction ($F_{6,159} = 3.58$, $p = 0.002$). Since the model was not fine-tuned to find these results, the different dimensionalities of ongoing and evoked

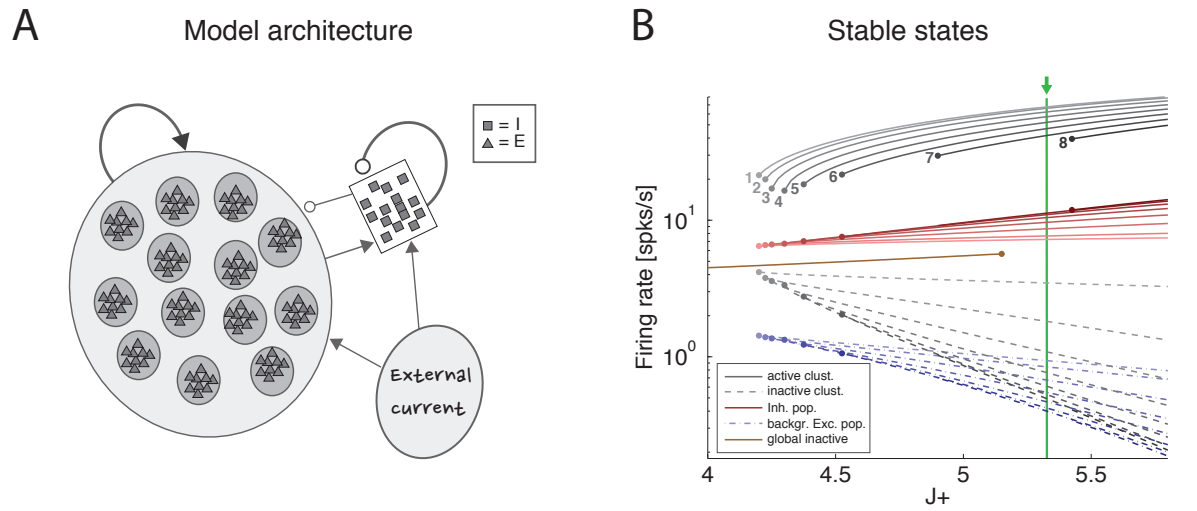

Figure 4. *Recurrent network model*. A: Schematic recurrent network architecture. Triangles and squares represent excitatory and inhibitory LIF neurons respectively. Darker disks indicate excitatory clusters with potentiated intra-cluster synaptic weights. B: Mean field solution of the recurrent network. Firing rates of the stable states for each subpopulation are shown as function of the intra-cluster synaptic potentiation parameter J_+: firing rate activity in the active clusters (solid grey lines), firing rate in the inactive clusters (dashed grey lines), activity of the background excitatory population (dashed blue lines), activity of the inhibitory population (solid red lines). In each case, darker colors represent configurations with larger number of active clusters. Numbers denote the number of active clusters in each stable configuration. Configurations with 1 to 8 active clusters are stable in the limit of of infinite network size. A global configuration where all clusters are inactive (brown line) becomes unstable at the value $J_+ = 5.15$. The vertical green line represents the value of $J_+ = 5.3$ chosen for the simulations. X-axis: intra-cluster potentiation parameter J_+ in units of J_{EE}; Y-axis: Firing rate (spks/s).

activity are likely the consequence of the organization in clusters and of the ensuing dynamics during ongoing and evoked activity.

2.5 Scaling of dimensionality with ensemble size and pair-wise correlations

The dependence of dimensionality on ensemble size observed in the data (Fig. 3B) and in the model (Fig. 6B) raises the question of whether or not the dimensionality would converge to an upper bound as one increases the number of simultaneously recorded neurons. In general, this question is important in a number of settings, related e.g. to coding in motor cortex [8, 10], performance in a discrimination task [13], or coding of visual stimuli [12]. In this section we provide an estimate of the dependence of dimensionality of ensembles of gustatory neurons on a number of relevant factors, both during ongoing and evoked activity. By definition, the dimensionality depends on the number of neurons N, the trial-to-trial variability of their firing rates, and their mutual dependencies (i.e., their pair-wise correlations). To gain insight into this dependency we computed the dimensionality in an asynchronous homogeneous network (i.e., having no clusters and low pair-wise correlations), and in the clustered network of Fig. 4. We consider first the case of a homogeneous network of neurons having the same firing rates distributions (which were approximately lognormal, Fig. 7A) and the same mean pair-wise spike count correlations as found in the data ($\rho \sim 0.01 - 0.2$). This would require solving a homogeneous recurrent network self-consistently for the desired firing rates and correlations. As a proxy for this scenario, we generated 20 sessions of 40 Poisson spike trains having exactly the desired properties (including the case of independent neurons for which $\rho = 0$). Two examples with $\rho = 0$ and $\rho = 0.1$, respectively, are

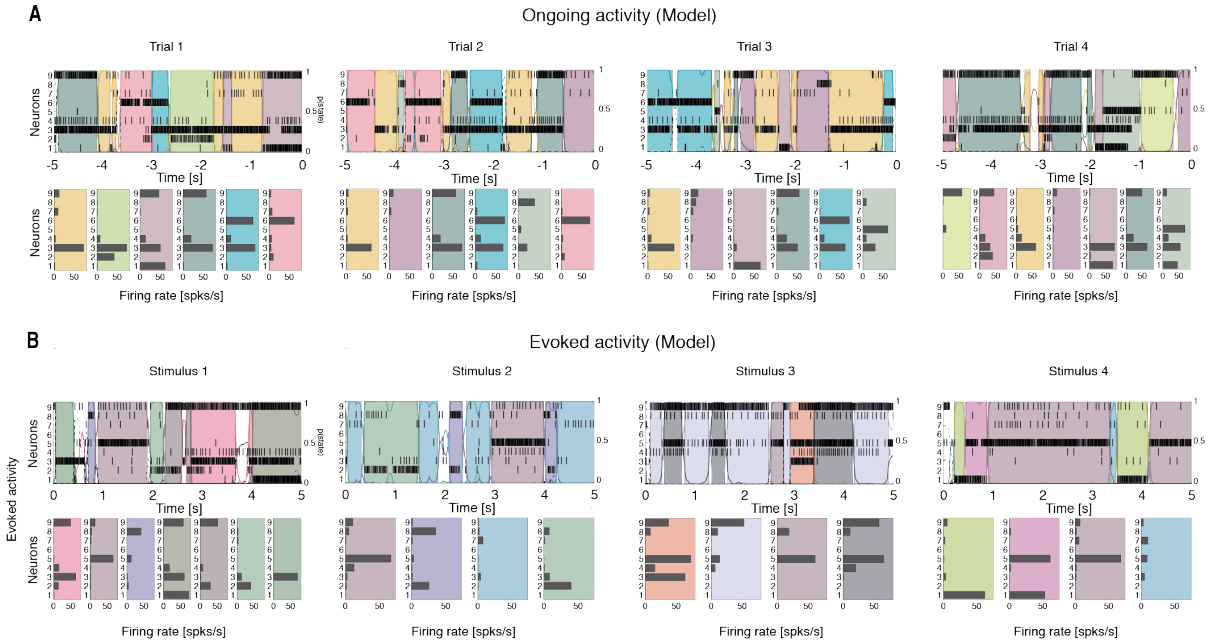

Figure 5. *Ensemble activity in the recurrent network model is characterized by sequences of states*. Representative trials from one ensemble of nine simultaneously recorded neurons sampled from the recurrent network, segmented according to their ensemble states (notations as in Fig. 1). A: ongoing activity. B: Ensemble activity evoked by four different stimuli, modeled as an increase in the external current to selected clusters (see Methods for details).

shown in Fig. 7B-C. Since in the asynchronous homogeneous network there are no transitions and hence no hidden states, the dimensionality was estimated based on the rate vectors in bins of 200 ms duration (using bin widths of 50 to 500 ms did not change the results; see Methods for details). We found that the dimensionality grows linearly with ensemble size in the absence of correlations, but is a concave function of N in the presence of spike count correlations (circles in Fig. 7D). Thus, as expected, the presence of correlations reduces the dimensionality. A simple theoretical calculation mimicking this scenario shows that d in this case converges slowly to an upper bound that depends on the inverse of the square of the pair-wise correlations. For example, in the case of uniform correlations (ρ) and equal variances of the spike counts

$$d(N,\rho) = \frac{1}{\rho^2 + (1-\rho^2)/N}\ , \tag{2.2}$$

(see Appendix A.5 for a derivation) shows that $d = N$ in the absence of correlations, and $d \leq 1/\rho^2$ for large networks in the presence of correlations. These properties remain approximately true in the case where the variances σ_i^2 of the spike counts are drawn from a distribution with mean $E[\sigma_i^2] = \sigma^2$ and variance $\mathrm{var}[\sigma_i^2] = \delta\sigma^4$. Dimensionality is reduced compared to the case of equal variances, in particular, in the absence of correlations we find $d \simeq \frac{\sigma^4}{\sigma^4+\delta\sigma^4} N$ for large N (see Appendix A.5 for details). The analytical results are shown in Fig. 7E (full lines for theoretical formula), together with their estimates based on 1000 data points (same number of data points as in Fig. 7D; see Appendix A), with either equal (dashed lines) or lognormal-distributed variances ('+') to mimic the empirical distribution in GC (not shown). Note the qualitative agreement with the surrogate data representing the homogeneous network in Fig. 7D; indeed, the latter can be well fit by Eq. (A.12), a version of Eq. (A.5) corrected to account for a finite number of trials and unequal variances).

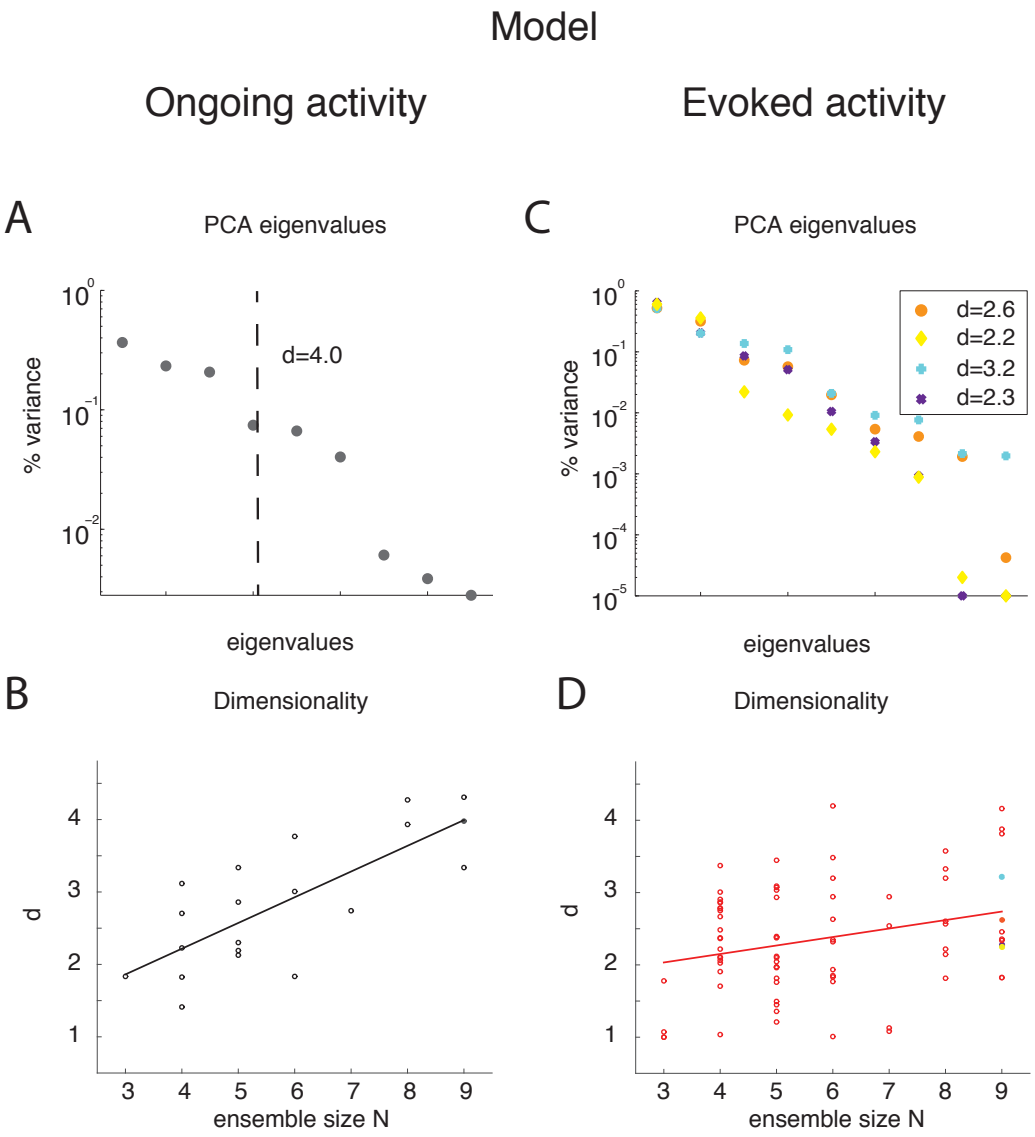

Figure 6. *Dependence of dimensionality on ensemble size (model).* A: Principal eigenvalue distribution for an ensemble of nine neurons during ongoing activity (corresponding to the filled dot in panel B) in the model network of Fig. 5 (notations as in Fig. 3A). B: Dimensionality of neural activity across all ensembles in the model during ongoing activity (linear regression fit overlaid). X-axis: ensemble size; Y-axis: dimensionality. C: Principal eigenvalue distribution for the ensemble in panel A during evoked activity. The eigenvalue distributions for four are presented (corresponding to the color-coded dots in panel D). X-axis: eigenvalue number; Y-axis: percentage of variance explained by each eigenvalue. D: Dimensionality of neural activity across all ensembles in the model during evoked activity (notations as in panel B).

2.6 Scaling of dimensionality in the presence of clusters

We next compared the dimensionality of the homogeneous networks activity to that predicted by the clustered network model of Fig. 4. To ease comparison with the homogeneous network, dimensionality was computed based on the spike counts in 200 ms bins rather than the HMMs firing rate vectors as in Fig. 6 (see Appendix A.5 for details). We found that the dependence of d on N in the clustered network depends on how the neurons are

sampled. If the sampling is completely random, so that any neuron has the same probability of being added to the ensemble regardless of cluster membership, a concave dependence on N will appear, much like the case of the homogeneous network (Fig. 8A, dashed lines). However, if neurons are selected one from each cluster until all clusters have been sampled once, then one neuron from each cluster until all clusters have been sampled twice, and so on, until all the neurons in the network have been sampled, then the dependence of d on N shows an abrupt transition when $N = Q$, i.e., when the number of sampled neurons reaches the number of clusters in the network (Fig. 8A, full lines; see Fig 8B for raster plots with $Q = 30$ and $N = 50$). In the following, we refer to this sampling procedure as "ordered sampling", as a reminder that neurons are selected randomly from each cluster, but the clusters are selected in serial order. For $N \leq Q$, the dimensionality grows linearly with ensemble size in both ongoing (slope 0.24 ± 0.01, $r = 0.79$, $p < 10^{-10}$, black line) and evoked periods (slope 0.19 ± 0.01, $r = 0.84$, $p < 10^{-10}$; red line), and was larger during ongoing than evoked activity (trial-matched two-way ANOVA, main effect: $F_{1,948} = 168$, $p < 10^{-30}$; interaction: $F_{5,948} = 4.1$, $p < 0.001$). These results are in keeping with the empirical and model results based on the HMM analysis (Fig. 3 and 6). The new finding is that, in the case of ordered sampling, the dependence of dimensionality on ensemble size disappears for $N \geq Q$ both during ongoing (slope 0.010 ± 0.003, $r = 0.1$, $p < 0.001$) and evoked periods (slope 0.009 ± 0.002, $r = 0.13$, $p < 10^{-4}$; Fig. 8A, full lines). The average dimensionality over the range $30 \leq N \leq 100$ was significantly larger for ongoing, $d_{ongoing} = 8.74 \pm 0.06$, than for evoked activity, $d_{evoked} = 7.15 \pm 0.04$ (trial-matched two-way ANOVA, main effect: $F_{1,2212} = 488$, $p < 10^{-30}$), confirming that dimensionality is larger during ongoing than evoked activity also in this case. The difference in dimensionality between ongoing and evoked activity also holds in the case of random sampling on the entire range of N values (Fig. 8A, dashed lines), confirming the generality of this finding.

2.7 Dimensionality is larger in the presence of clusters

Intuitively, in the clustered network the dimensionality saturates at $N = Q$ because additional neurons will be highly correlated with already sampled ones. For $N \leq Q$, each new neurons activity adds an independent degree of freedom to the neural dynamics and thus increases its dimensionality. For $Q > N$, additional neurons are highly correlated with an existing neuron, adding little or no additional contribution to d. Indeed, compared to the low overall correlations found across all neuron pairs in the data (and used as desiderata for the homogeneous network), neurons belonging to the same model cluster had a much higher spike count correlation of $\rho = 0.92$ $[0.56, 0.96]$ (median and $[25, 75]$-percentile), while neurons belonging to different clusters had a much lower correlation of $\rho = -0.02$ $[-0.10, 0.06]$. Overall, the median spike count correlation in the model, regardless of cluster membership, was $\rho = -0.029$ $[-0.109, 0.083]$. For comparison, the empirical correlations during ongoing activity was $\rho = -0.002$ $[-0.047, 0.051]$, with maximal values of $\rho \sim 0.5$; during evoked activity $\rho = 0.001$ $[-0.085, 0.113]$ with maximal values of $\rho \sim 0.9$. Plugging these values into a correlation matrix reflecting the clustered architecture and the sampling procedure used in Fig. 8B, we obtained the matrix shown in Fig. 8C, where pairwise correlations depend on whether or not the neurons belong to the same cluster (for the first 40 neurons, adjacent pairs belong to the same cluster; the last 10 neurons belong to the remaining clusters). It is natural to interpret such correlation matrix as the noisy observation of a block-diagonal matrix such that neurons in the same cluster have uniform correlation while neurons from different clusters are uncorrelated. For such a correlation matrix the dimensionality can be evaluated exactly (see Eq. (A.10)). In the approximation where all neurons have the same variance, this reduces to

$$d = \begin{cases} N\,, & N \leq Q \\ \frac{N}{1+m\rho^2[1-(Q-p)/N]}\,, & N > Q \end{cases} \tag{2.3}$$

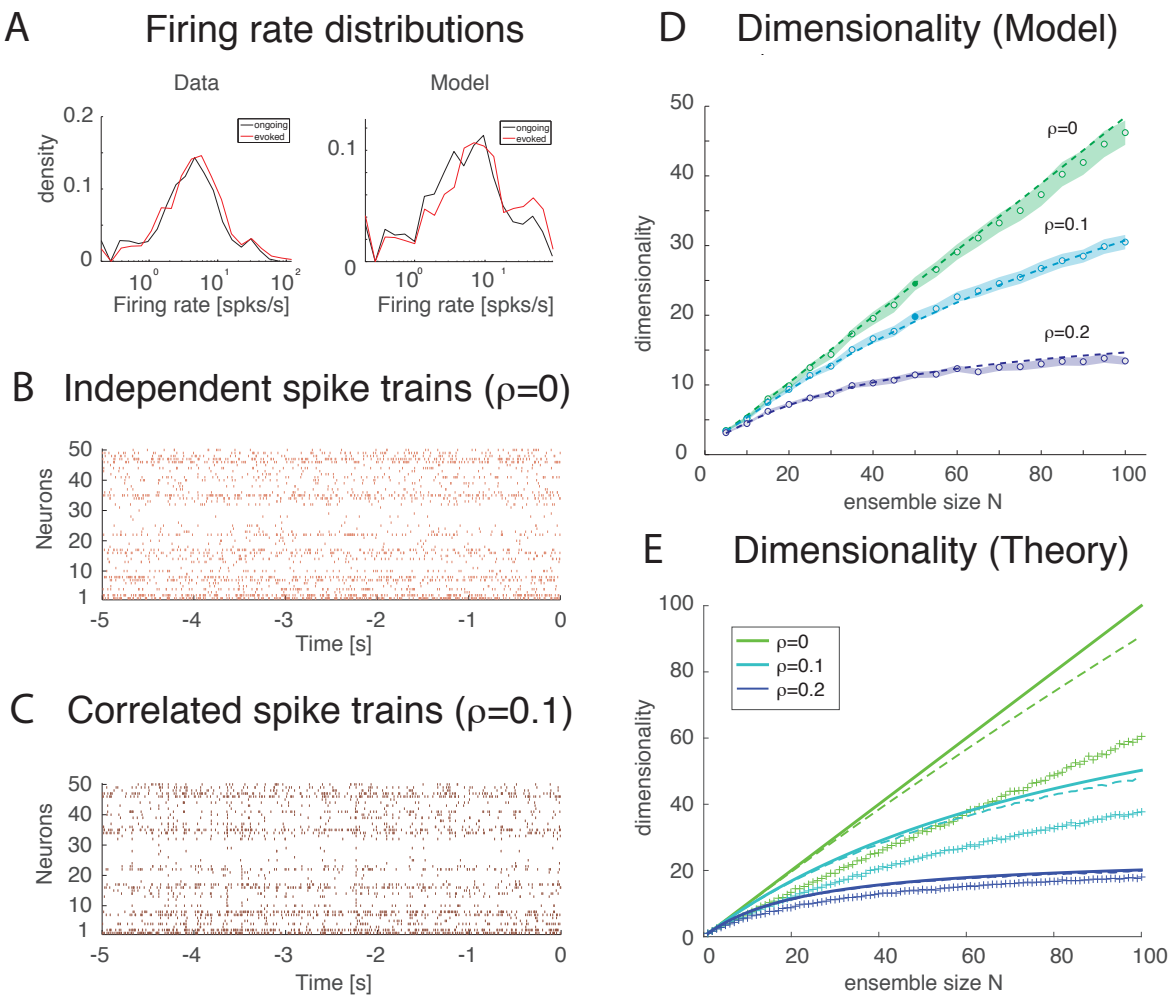

Figure 7. *Dimensionality and correlation*. A: Empirical single neuron firing rate distributions in the data (left) and in the model (right), for ongoing (black) and evoked activity (red). The distributions are approximately lognormal. X-axis: Firing rate (spks/s); Y-axis: density. B: Example of independent Poisson spike trains with firing rates matched to the firing rates obtained in simulations of the spiking network model. C: Example of correlated Poisson spike trains with firing rates matched to the firing rates obtained in simulations of the spiking network model. Pair-wise correlations of $\rho = 0.1$ were used (see Methods). X-axis: time [s]; Y-axis: neuron index. D: Dimensionality as a function of ensemble size N in the recurrent network model (ongoing activity, black) and in an ensemble of Poisson spike trains with spike count correlations $\rho = 0, 0.1, 0.2$ and firing rates matched to the model simulations. Dashed lines represent the fit of Eq. (6) to the data (see Methods for details), with best-fit parameters (mean with 95% confidence interval) $a = 1.10(1.09, 1.14)$, $b = 2.09(2.06, 2.12)$, $c = -3.16(-4.51, -1.80)$. Filled circles (from top to bottom): dimensionality of the data (raster plots) shown in panel B, C and Fig. 8A (top panel). X-axis: ensemble size; Y-axis, dimensionality. Inset: enlargement of the plot in the region $N \leq Q$. E: Theoretical prediction for the dependence of dimensionality on ensemble size N and spike count correlation ρ for the case of uniform correlation and equal variances, Eq. (7) (full lines; darker shades represent increasing correlations). Dashed and '+': dimensionality were estimated from $1,000$ "nominal" firing rates for each N built so as to converge to Eq. (5) for an infinite number of trials, in the case of equal (dashed) and lognormally distributed ('+') spike count variance. X-axis: ensemble size N; Y-axis: dimensionality.

where $N = mQ + p$. This formula is plotted in Fig. 8D for relevant values of ρ and N and it explains the origin of the abrupt transition in dimensionality at $Q = N$. (The reasons for a dimensionality lower than Q for $N \leq Q$ in the data - see Fig. 8A - are i) the finite number of data points (250) used for its estimation and ii) the non-uniform distribution of variances; see e.g. '+' in Fig. 7E). Note that the formula also predicts cusps in dimensionality (which become local maxima for large ρ) whenever the ensemble size is an exact multiple of the number of clusters. This is also visible in the simulated data of Fig. 8A, where local maxima seem to appear at $N = 30, 60, 90$ with $Q = 30$ clusters. It is also worth mentioning that for low intra-cluster correlations the dependence on N predicted by Eq. (2.3) becomes smoother and the cusps harder to detect (not shown), suggesting that the behavior of a clustered network with weak clusters tends to converge to the behavior of a homogeneous asynchronous network - therefore lacking sequences of hidden states. Thus, the complexity of the network dynamics is reflected in how its dimensionality scales with N, assuming that one may sample one neuron per cluster (i.e., via "ordered sampling"). Even though it is not clear how to perform ordered sampling empirically (see Discussion), this result is nevertheless useful since it represents an upper bound also in the case of random sampling (see Fig. 8A, dashed lines). Eq. (2.3) predicts that $d \leq Q/\rho^2$, with this value reached asymptotically for large N. In the case of random sampling, growth to this bound is even slower (Fig. 8A). For comparison, in a homogeneous network $d \leq 1/\rho^2$ from Eq. (2.2), a smaller bound by a factor of Q. Homogeneous dimensionality is dominated by clustered dimensionality also in the more realistic case of non-uniform variances, where similar bounds are found in both cases (see Appendix A.5 for details). For the values of synaptic strength used in the model to capture the dynamics of multistable states [5], correlations within clusters are large (> 0.8), resulting in bounds close to Q (empirically, finite size effects result in a much lower estimate of this bound, e.g., ~ 9 in Fig. 8A). However, the central result here is that, a parity of pair-wise correlations, the dimensionality in a clustered network is larger than in a homogeneous network by a factor equal to the number of clusters Q.

3 Discussion

In this paper we have investigated the dimensionality of the neural activity in the gustatory cortex of alert rats. Dimensionality was defined as a collective property of ensembles of simultaneously recorded neurons that reflects the effective space occupied by the ensemble activity during either ongoing or evoked activity. If one represents ensemble activity in terms of firing rate vectors, whose dimension is the number of ensemble neurons N, then the collection of rate vectors across trials takes the form of a set of points in the N-dimensional space of firing rates. Roughly, dimensionality is the minimal number of dimensions necessary to provide an accurate description of such set of points, which may be localized on a lower-dimensional subspace inside the whole firing rate space.

One of the main results of this paper is that the dimensionality of evoked activity is smaller than that of ongoing activity, i.e., stimulus presentation quenches dimensionality. More specifically, the dimensionality is linearly related to the ensemble size, with a significantly larger slope during ongoing activity compared to evoked activity (compare Fig. 3B and 3D). We explained this phenomenon using a biologically plausible, mechanistic spiking network model based on recurrent connectivity with clustered architecture. The model was recently introduced in [5] to account for the observed dynamics of ensembles of GC neurons as sequences of metastable states, where each state is defined as a vector of firing rates across simultaneously recorded neurons. The model captures the reduction in trial-to-trial variability and the multiple firing rates attained by single neurons across different states observed in GC upon stimulus presentation. Here, the same model was found to capture also the stimulus-induced reduction of dimensionality. While the set of active clusters during ongoing activity varies randomly, allowing the ensemble dynamics to explore a large portion of firing rate space, the evoked set of active clusters is limited mostly to the stimulus-selective clusters only (see [5] for a detailed

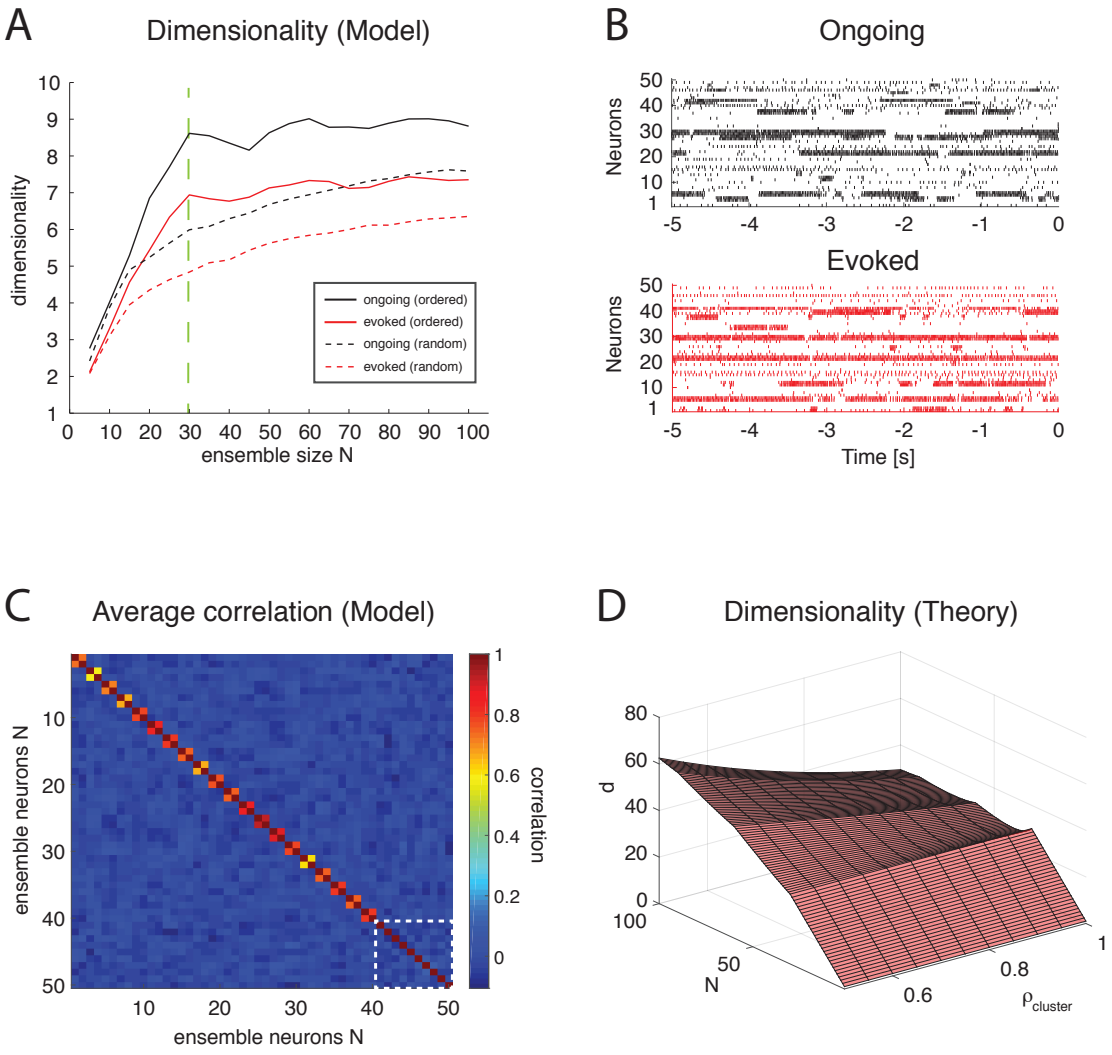

Figure 8. *Dimensionality in a clustered network.* A: Trial-matched dimensionality as a function of ensemble size in the recurrent network model (ongoing and evoked activity in black and red, respectively). Filled lines represent ordered sampling, where ensembles to the left of the green vertical line ($N = Q = 30$) contain at most one neuron per cluster, while to the right of the line they contain one or more neurons from all clusters. Dashed lines represent random sampling of neurons, regardless of cluster membership. X-axis: ensemble size; Y-axis, dimensionality. B: Representative trial of an ensemble of 50 neurons sampled from the recurrent network in Fig. 4 during ongoing activity (upper plot, in black) or evoked activity (lower plot, in red) for the case of ordered sampling (full lines). Neurons are sorted according to their cluster membership (adjacent neuron pairs with similar activity belong to the same cluster, for neurons #1 up to #40; the last ten neurons are sampled from the remaining clusters). X-axis: time to stimulus presentation at $t = 0$ (s); Y-axis: neuron index. C: Average correlation matrix for twenty ensembles of $N = 50$ neurons from the clustered network model with $Q = 30$. For the first 40 neurons, adjacent pairs belong to the same cluster; the last 10 neurons (delimited by a dashed white square) belong to the remaining clusters (neurons are ordered as in panel B). Thus, neurons $1, 3, 5, \ldots, 39$ (20 neurons) belong to the first 20 clusters; neurons $2, 4, 6, \ldots, 40$ (20 neurons) belong also the first 20 clusters; and neurons $41, 42, 43, \ldots, 50$ (10 neurons) belong to the remaining 10 clusters. X-axis, Y-axis: neuron index. D: Plot of Eq. A.12 giving d vs. N and ρ (uniform within-cluster correlations) for the sampling procedure of Fig. 8B. X-axis: ensemble size N; Y-axis: dimensionality.

analysis). The dynamics of cluster activation in the model thus explains the more pronounced dependence of dimensionality on ensemble size found during ongoing compared to evoked activity.

We presented a simple theory of how dimensionality depends on the number of simultaneously recorded neurons N, their pair-wise spike count correlations, their variance, and the number and duration of recording trials (see also [8]). We found that dimensionality increases with N and decreases with the amount of pair-wise correlations among the neurons. When neurons have the same (uniform) correlations, dimensionality is maximal when all neurons have the same spike count variance, and it decreases as the distribution of single neurons spike count variances become more heterogeneous (Fig. 7E, '+'). Finally, introducing clustered correlations in the theory, and sampling one neuron per cluster as in Fig. 8A (full lines), results in cusps at values of N that are multiples of the number of clusters (Fig. 8D), in keeping with the predictions of the spiking network model (Fig. 8A, full lines).

3.1 Dimensionality scaling with ensemble size

The increased dimensionality with sample size, especially during ongoing activity, was found empirically in datasets with 3 to 9 neurons per ensemble, but could be extrapolated for larger N in a spiking network model with homogeneous or clustered architecture. In homogeneous networks with finite correlations the dimensionality is predicted to increase sub-linearly with N (Eq. (2.2)), whereas in the clustered network it may exhibit cusps at multiple values of the number of clusters (Fig. 8A), and would saturate abruptly (for uniform firing rate variances) to a value that depends on the ratio of the number of clusters Q and the amount of pair-wise correlations, $d \leq Q/\rho^2$. Testing this prediction requires the ability to sample neurons one from each cluster, until all clusters are sampled, and seems beyond the current recording techniques. However, looking for natural groupings of neurons based on response similarities could uncover spatial segregation of clusters [16] and allow sampling neurons according to this procedure. Moreover, the model predicts a slower approach to a similar bound also in the case of random sampling. Dimensionality in a homogeneous network is instead bounded by $1/\rho^2$, and hence it is a factor Q smaller than in the clustered network. Dimensionality is maximal in a population of independent neurons ($\rho = 0$), where it grows linearly with N; however neurons of recurrent networks have finite amounts of correlations which can be rather large inside clusters (in our data, we have found neuron pairs with ρ 0.9). Since the presence of even low correlations can dramatically reduce the dimensionality (see e.g. Fig. 7), the neural activity in a clustered architecture can reach values much higher values at parity of correlations, representing an intermediate case between a homogeneous network and a population of independent neurons.

3.2 Dimensionality and trial-to-trial variability

Cortical recordings from alert animals show that neurons produce irregular spike trains with variable spike counts across trials [17–19]. Despite many efforts, it remains a key issue to establish whether variability is detrimental [20, 21] or useful [22] for neural computation. trial-to-trial variability is reduced by the presentation of a stimulus [23], a phenomenon that would not occur in a population of independent or homogeneously connected neurons [24]. Recent work has shown that the stimulus-induced reduction of trial-to-trial variability can be explained by the slow dynamics generated in a recurrent spiking network model with clustered connectivity [9, 24, 25]. Slow fluctuations in firing rates across neurons can ignite metastable sequences of neural activity, closely resembling metastable sequences observed experimentally [1–7]. The slow, metastable dynamics of cluster activation produces high variability in the spike count during ongoing activity. While cluster activations occur at random times during ongoing activity periods, stimulus presentation locks cluster activation at its onset, leading to a decrease in trial-to-trial variability. Similarly, a stimulus-induced reduction of dimensionality is obtained in the same model. In this case, preferred cluster activation due to stimulus onset generates an increase

in pair-wise correlations that reduce dimensionality. Note that the two properties (trial-to-trial variability and dimensionality) are conceptually distinct. An ensemble of Poisson spike trains can be highly correlated (hence have low dimensionality), yet the Fano Factor of each spike train will still be 1 (hence high), independently of the correlations among neurons. In a recurrent network, however, dimensionality and trial-to-trial variability may become intertwined and exhibit similar properties, such as the stimulus-induced reduction observed in a model with clustered connectivity. A deeper investigation of the link between dimensionality and trial-to-trial variability in recurrent networks is left for future studies.

3.3 Alternative definitions of dimensionality

Following [9] we have defined dimensionality (Eq. (2.1)) as the dimension of an effective linear subspace of firing rate vectors containing the most variance of the neural activity. It differs from the typical dimensionality reduction based on PCA in that the latter retains only the number of eigenvectors explaining a predefined amount of variance (see e.g. [26, 27]), whereas Eq. (2.1) includes contribution from all eigenvalues. Other definitions of neural dimensionality have been proposed in the literature, which aim at capturing different properties of the neural activity, typically during stimulus-evoked activity. A measure of dimensionality related to ours, and referred to as "complexity," was introduced in [12]. According to their definition, population firing rate vectors from all evoked conditions were first decomposed along their kernel Principal Components [28]. A linear classifier was then trained on an increasing number of leading PCs in order to perform a discrimination task, where the number of PCs used was defined as the complexity of the representation. In general, the classification accuracy improves with increasing complexity, and it may saturate when all PCs containing relevant features are used - with the remaining PCs representing noise or information irrelevant to the task. Reaching high accuracy at low complexity implies good generalization performance, i.e., the ability to classify novel variations of a stimulus in the correct category. Neural representations in monkey inferotemporal cortex (IT) were found to require lower complexity than in area V4, confirming ITs premier role in classifying visual objects despite large variations in shape, orientation and background [12]. Complexity relies on a supervised algorithm and is an efficient tool to capture the generalization properties of evoked representations (see e.g. [29]) for its relevance to visual object recognition). A second definition of dimensionality, sometimes referred to as "shattering dimensionality" in the Machine Learning literature, was used to assess the discrimination properties of the neural representation [13]. Given a set of firing rate vectors in an N-dimensional space, one can split them into two classes (e.g. white and black colorings) in 2^N different ways, and train a classifier to learn as many of those binary classification labels as possible. The shattering dimensionality is then defined as (the logarithm of) the largest number of binary classifications that can be implemented. This measure of dimensionality was found to drop significantly in monkey prefrontal cortex during error trials in a recall task and thus predicts the ability of the monkey to correctly perform the task [13]. A flexible and informative neural representation is one that achieves a large shattering dimensionality (good discrimination) while keeping a low complexity (good generalization). On the other hand, both complexity and shattering dimensionality represent measures of classification performance in task-related paradigms, and their definition requires a set of evoked conditions to be classified via a supervised learning algorithm. While both definitions could be applied to neural activity in our stimulus-evoked data, their interpretation cannot be extended to periods of ongoing activity, as the latter is not associated to desired targets in a way that can be learned by a classification algorithm. Since our main aim was to compare the dimensionality of ongoing and evoked activity, the unsupervised approach of [9] and their notion of "effective" dimensionality was better suited for our analysis. A related definition of dimensionality has been used by [8] to investigate neural representations of movements in motor cortex.

3.4 Ongoing activity and task complexity

Our results show that neural activity during ongoing periods occupies a space of larger dimensionality compared to evoked activity. Although based on a different measure of dimensionality, recent results on the relation between the dimensionality of evoked activity and task complexity suggest that evoked dimensionality is roughly equal to the number of task conditions [13]. It is natural to ask whether the dimensionality of ongoing activity provides an estimate of the complexity of the hardest task that can be supported by the neural activity. Moreover, based on the clustered network model, the presence of clusters imposes an upper value $d \leq Q/\rho^2$ during ongoing activity, suggesting that a discrimination task with up to $\sim Q$ different conditions may be supported. Applied to our data, this implies that GC can support tasks of larger complexity, such as discrimination and choice tasks with a larger number of tastants and conditions. The experience of taste consumption is by itself multidimensional, including both chemo- and oro-sensory aspects (i.e., taste identity [15] and concentration [30], texture, temperature [31, 32], taste and odor mixtures [33]) and psychological aspects (hedonic value [14, 15, 34]), anticipation [35], novelty [36] and satiety effects [37]. It is tempting to speculate that neural activity during ongoing periods explores these different dimensions, while evoked activity is confined to the features of the particular taste being delivered in a specific context. According to our network model interpretation, the ability to solve more complex tasks would be related to the number of clusters in GC, but also on the contribution of other areas involved in gustatory processing. Establishing a precise experimental and theoretical link between the number of clusters and task complexity supported by the neural activity is an important question left for future studies.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

This work was supported by a National Institute of Deafness and Other Communication Disorders Grant K25-DC013557 (LM), by the Swartz Foundation Award 66438 (LM), by National Institute of Deafness and Other Communication Disorders Grant R01-DC010389 (AF), by a Klingenstein Foundation Fellowship (AF), and partly by a National Science Foundation Grant IIS1161852 (GLC). We thank Dr. Stefano Fusi for useful discussions and David Ecker at the Research Technologies DoIT of Stony Brook University for access to its computational resources.

A Methods

A.1 Experimental procedures

Adult female Long Evans rats were used for this study [5, 35]. Animals received ad lib. access to food and water, unless otherwise mentioned. Movable bundles of sixteen microwires attached to a "mini-microdrive" [35, 38] were implanted in GC (AP 1.4, ML $\pm$ 5 from bregma, DV -4.5 from dura). After electrode implantation, intra-oral cannulae (IOC) were inserted bilaterally [39, 40]. At the end of the surgery a positioning bolt for restraint was cemented in the acrylic cap. Rats were given at least 7 days for recovery before starting the behavioral procedures outlined below. All experimental procedures were approved by the Institutional Animal Care and Use Committee of Stony Brook University and complied with University, state, and federal regulations on the care and use of laboratory animals. More details can be found in [35]. Rats were habituated to being restrained

and to receiving fluids through IOCs, and then trained to self-deliver water by pressing a lever following a 75 dB auditory cue at a frequency of 4 KHz. The interval at which lever-pressing delivered water was progressively increased to 40 ± 3 s (ITI). During training and experimental sessions additional tastants were automatically delivered at random times near the middle of the ITI, at random trials and in the absence of the anticipatory cue. Upon termination of each recording session the electrodes were lowered by at least 150 μm so that a new ensemble could be recorded. A computer-controlled, pressurized, solenoid-based system delivered $\sim 40\mu$l of fluids (opening time $\sim$ 40 ms) directly into the mouth through a manifold of 4 polymide tubes slid into the IOC. The following tastants were delivered: 100 mM NaCl, 100 mM sucrose, 100 mM citric acid, and 1 mM quinine HCl. Water ($\sim 50\mu$l) was delivered to rinse the mouth clean through a second IOC five seconds after the delivery of each tastant. Each tastant was delivered for at least 6 trials in each condition.

A.2 Data analysis

Single neuron action potentials were amplified, bandpass filtered (at $300 - 8$ KHz), digitized and recorded to a computer (Plexon, Dallas, TX). Single units of at least $3 : 1$ signal-to-noise ratio were isolated using a template algorithm, cluster cutting techniques and examination of inter-spike interval plots (Offline Sorter, Plexon, Dallas, TX). All data analyses and model simulations were performed using custom software written in Matlab (Mathworks, Natick, MA, USA), Mathematica (Wolfram Research, Champaign, IL), and C. Starting from a pool of 299 single neurons in 37 sessions, neurons with peak firing rate lower than 1 Hz (defined as silent) were excluded from further analysis, as well as neurons with a large peak around the $6 - 10$ Hz in the spike power spectrum, which were considered somatosensory [14, 35, 41]. Only ensembles with 3 or more simultaneously recorded neurons were further analyzed (167 non-silent, non-somatosensory neurons from 27 ensembles). We analyzed ongoing activity in the 5 seconds interval preceding either the auditory cue or taste delivery, and evoked activity in the 5 seconds interval following taste delivery in trials without anticipatory cue, wherein significant taste-related information is present [15].

A.3 Hidden Markov Model (HMM) analysis

Here we briefly outline the procedure used in [5], see this reference and [4, 7, 42] for further details. Under the HMM, a system of N recorded neurons is assumed to be in one of a predetermined number of hidden (or latent) states [43, 43]. Each state m is defined as a vector of N firing rates $\nu_i(m)$, $i = 1, \ldots, N$, one for each simultaneously recorded neuron. In each state, the neurons were assumed to discharge as stationary Poisson processes (Poisson-HMM). We matched the model to the data segmented in 1-ms bins (see below). We assumed that in such short bins either zero or one spikes can be emitted by each neuron; otherwise, one spike was randomly assigned to one of the neurons, effectively reducing ensemble activity from a Poisson to a Bernoulli process. We denote by $y_i(t)$ the spiking activity of the i-th neuron in the interval $[t, t + dt]$, $y_i(t) = 1$ if the neuron emitted a spike and $y_i(t) = 0$ otherwise. Denoting with S_t the hidden state of the ensemble at time t, the probability of having a spikes from neuron i in a given state m in the interval $[t, t + dt]$ is given by $p(y_i(t) = 1|S_t = m) = 1 - e^{\nu_i(m)dt}$.

The firing rates $\nu_i(m)$ completely define the states and are also called "emission probabilities" in HMM parlance. The emission and transition probabilities were found by maximization of the log-likelihood of the data given the model via the expectation-maximization (EM), or Baum-Welch, algorithm [43], a procedure known as "training the HMM". For each session and type of activity (ongoing vs. evoked), ensemble spiking activity from all trials was binned at 1 ms intervals prior to training assuming a fixed number of hidden states M [4, 42]. For each given number of states M, the Baum-Welch algorithm was run 5 times, each time with random initial conditions for the transition and emission probabilities. The range of hidden states M for the HMM analyses were $M_{min} = 10$ and $M_{max} = 20$ for spontaneous activity, and $M_{min} = 10$ and $M_{max} = 40$

for evoked activity. Such numbers were based on extensive exploration of the parameter space and previous studies [4, 5, 7, 42, 44]. For evoked activity, each HMM was trained on all four tastes simultaneously. Of the models thus obtained, the one with largest total likelihood M^* was taken as the best HMM match to the data, and then used to estimate the probability of the states given the model and the observations in each bin of each trial (a procedure known as "decoding"). During decoding, only those hidden states with probability exceeding 80% in at least 50 consecutive bins were retained (henceforth denoted simply as "states"). The firing rate fits $\nu_i(m)$ in each trial were obtained from the analytical solution of the maximization step of the Baum-Welch algorithm,

$$\nu_i(m) = -\frac{1}{dt} \ln \left(1 - \frac{\sum_{t=1}^{T} r_m(t) y_i(t)}{\sum_{t=1}^{T} r_m(t)} \right) . \tag{A.1}$$

Here, $[y_i(1), \ldots, y_i(T)]$ is the spike train of the i-th neuron in the current trial, and T is the total duration of the trial. $r_m(t) = P(S_t = m | y(1), \ldots, y(T))$ is the probability that the hidden state S_t at time t is m, given the observations.

A.4 Spike count correlations

Given neuron i and neuron j's spike trains, we computed the spike count correlation coefficient r_{ij},

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} ,$$

where C is the covariance matrix of the spike counts estimated as

$$C_{ij} = \frac{1}{n-1} \sum_{b=1}^{n} \left(n_i(b) - \frac{1}{n} \sum_{b'} n_i(b') \right) \left(n_j(b) - \frac{1}{n} \sum_{b'} n_j(b') \right) , \tag{A.2}$$

where $n_i(b)$ is the spike count of neuron i in bin b (of 200 ms duration), and the sums run over all 200 ms bins of each trial and over all trials in each session (a total of n temporal windows across bins and trials). Varying the bin size from 200 ms to 5 seconds did not affect these estimates appreciably. Significance of the correlation was estimated as follows [45]: $N_{shuffle} = 200$ trial-shuffled correlation coefficients r'_{ij} were computed, then a p-value was determined as the fraction of shuffled coefficients r'_{ij} whose absolute value exceeded the absolute value of the experimental correlation, $p = (\#(|r'_{ij}| > |r_{ij}|))/N_{shuffle}$. For example, a correlation r was significant at $p = 0.05$ confidence level if no more than 10 shuffled correlation coefficients out of 200 exceeded r.

A.5 Dimensionality measure

We defined the dimensionality of the neural activity as [9]

$$d = \frac{1}{\sum_{i=1}^{N} \tilde{\lambda}_i^2} , \tag{A.3}$$

where the $\tilde{\lambda}_i$ are the principal eigenvalues expressed as fractions of the total amount of variance explained, i.e. $\tilde{\lambda}_i = \lambda_i / (\sum_j \lambda_j)$, where λ_j are the eigenvalues of the covariance matrix of the firing rates (see below). The eigenvalues λ_j were found with a standard Principal Component Analysis (PCA) of the set of all firing rate vectors. The latter were indexed by neuron number and either HMM state or bin, depending on the analysis. For the analysis of Fig. 3 and 6, the firing rate vectors were the $\nu_i(m)$ given by Eq. (A.1), where $i = 1, \ldots, N$ spans the ensemble neurons and m the HMM states; all data from either ongoing or evoked activity were used.

For the analysis of Fig. 3E, where the duration and number of trials were varied, only the firing rate vectors of the HMM states present in the given trial snippet were used (even if present for only a few ms). For the analysis of Fig. 7D and 8B, the firing rate vectors were given by the spike counts in $T = 200$ ms bins divided by T, $n_i(b)/T$, where $n_i(b)$ is as defined in Eq. (A.2). Dimensionality values were averaged across 20 simulated sessions for each ensemble size N; in each session, 40 trials of 5 s duration, resulting in $N_T = 1000$ bins, were used (using bin widths of 50 to 500 ms did not change the results). Note that since bins of equal length T were used in this case, the dimensionality d is the same as found by performing PCA on the C matrix of spike counts, Eq. (A.2).

In our data, d roughly corresponded to the number of principal components explaining between 80 to 90% of the variance. However, note that all eigenvalues are retained in our definition of dimensionality given in Eq. (A.3) above.

The dimensionality can be computed exactly in some relevant special cases. The calculation is simplified by the observation that Eq. (A.3) is equivalent to

$$d = \frac{[\mathrm{Tr}(C_f)]^2}{\mathrm{Tr}(C_f^2)} ,$$

where C_f is the covariance matrix of the firing rate vectors, $Tr(A) = \sum_{i=1}^{N} A_{ii}$ is the trace of matrix A, and $\mathrm{Tr}(A^2) = \sum_{i,j=1}^{N} A_{ij}A_{ji}$. We consider in the following only the case of firing rates in equal bins, hence we can replace C_f with the covariance matrix of the spike count C in the definition of d:

$$d = \frac{[\mathrm{Tr}(C)]^2}{\mathrm{Tr}(C^2)} , \tag{A.4}$$

Dimensionality in the case of uniform pair-wise correlations. When all the pair-wise correlations r_{ij} are identical, $r_{ij} = \rho$ for all $i \neq j$,

$$r_{ij} = \begin{bmatrix} 1 & \rho & \cdot & \rho \\ \rho & 1 & & \cdot \\ \cdot & & \cdot & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} ,$$

the dimensionality (Eq. (A.4)) is given by

$$d = \frac{(\sum_{i=1}^{N} \sigma_i^2)^2}{\rho^2 \sum_{i \neq j}^{N} \sigma_i^2 \sigma_j^2 + \sum_{i=1}^{N} \sigma_i^4},$$

where $\sigma_i^2 = C_{ii}$ is the variance of the spike count and $C_{ij} = \rho\sqrt{\sigma_i^2 \sigma_j^2}$ for $i \neq j$. Note from this formula that d does not depend on the distribution of the firing rates, but only on their variances (modulo a common scaling factor). From the above expression we obtain

$$d = \frac{1}{\rho^2 + (1 - \rho^2) g(N)} , \tag{A.5}$$

where

$$g_N = \frac{a_N}{b_N^2} = \frac{\sum_{i=1}^{N} \sigma_i^4}{(\sum_{i=1}^{N} \sigma_i^2)^2} ,$$

having defined

$$a_N = \sum_{i=1}^{N} \sigma_i^4 \,, \qquad b_N = \sum_{i=1}^{N} \sigma_i^2 \,. \tag{A.6}$$

Note that since both a_N and b_N scale as N when N is large, in general $g_N \sim 1/N$ for large N. If all spike counts have equal variance, $\sigma_i = \sigma$, we find exactly $g_N = 1/N$:

$$d = \frac{1}{\rho^2 + (1-\rho^2)/N} = \frac{N}{N\rho^2 + 1 - \rho^2} \,, \tag{A.7}$$

and the dependence of d on the variance drops out. Note that for uncorrelated spike counts ($\rho = 0$) this formula gives $d = N$, whereas for any finite correlation we find the upper bound $d = 1/\rho^2$. For $N > 1$, the dimensionality is inversely related to the amount of pair-wise correlation ρ.

In the case where spike counts have variances σ_i^2 drawn from a probability distribution with mean $E[\sigma_i^2] = \sigma^2$ and variance $\mathrm{var}[\sigma_i^2] = \delta\sigma^4$, one can evaluate Eq. (A.5) approximately by its Taylor expansion around the expectation values

$$E[a_N] = N(\sigma^4 + \delta\sigma^4) \,, \qquad E[b_N^2] = N^2\sigma^4 + N\delta\sigma^4 \,. \tag{A.8}$$

Here we used the fact that, given a random vector X_i with mean μ_i and covariance Σ_{ij}, and a constant symmetric matrix A_{ij}, the expectation value of the quadratic form

$$E[\sum_{i,j} X_i A_{ij} X_j] = \sum_{i,j} (A_{ij}\Sigma_{ji} + \mu_i A_{ij} \mu_j) \,.$$

At the leading order in the Taylor expansion one finds

$$d = \frac{1}{\rho^2 + (1-\rho^2)\frac{\sigma^4+\delta\sigma^4}{N\sigma^4+\delta\sigma^4}} \,. \tag{A.9}$$

In the case of uncorrelated spike counts ($\rho = 0$), dimensionality still depends linearly on the ensemble size N, but with a smaller slope $\frac{\sigma^4}{\sigma^4+\delta\sigma^4} < 1$ compared to the case of equal variance in Eq. (A.7).

Dimensionality in the case of neural clusters. Given an ensemble of N neurons arranged in Q clusters (motivated by the model network described later in Sec. A.7), we created ensembles of uncorrelated spike trains for $N \leq Q$ and correlated within each cluster for $N > Q$. Thus, if $N \leq Q$ the correlation matrix is the $N \times N$ identity matrix. If $N > Q$, the $(Q+1)$-th neuron was added to the first cluster, with correlation ρ with the other neuron of the cluster, and uncorrelated to the neurons in the remaining clusters. The $(Q+2)$-th neuron was added to the second cluster, with correlation ρ with the other neuron of the second cluster, and uncorrelated to the neurons in the remaining clusters, and so on. Similarly, the $(2Q+p)$-th neuron $(p \leq Q)$ was added to the p-th cluster, with pair-wise correlation ρ with the other neurons of the same cluster, but no correlation with the neurons in the remaining clusters; and so on. In general, for $N = mQ + p$ neurons (where $m = [N/Q]_- \geq 1$ is the largest integer smaller than N/Q), the procedure picked $m+1$ neurons per cluster for the first p cluster and m neurons per cluster for the remaining $Q - p$ clusters, with uniform pair-wise correlations ρ in the same cluster while neurons from different clusters were uncorrelated. The resulting correlation matrix r was block diagonal

$$r = \mathrm{diag}(R_1, \ldots, R_Q) \,,$$

where each of the Q blocks contains the correlations of neurons from the same cluster. Inside each block R_i,

the off-diagonal terms are equal to the uniform within-cluster correlation ρ:

$$R_i = \begin{bmatrix} 1 & \rho & \cdot & \rho \\ \rho & 1 & & \cdot \\ \cdot & & \cdot & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} ,$$

The first p blocks have size $(m+1)\times(m+1)$ and the last $Q-p$ blocks have size $m\times m$, so that $(m+1)p + m(Q-p) = N$. The remaining elements of matrix r (representing pair-wise correlations of neurons belonging to different clusters) were all zero. Recalling that $C_{ij} = r_{ij}\sigma_i\sigma_j$, one finds $\mathrm{Tr}(C) = pb_{m+1} + (Q-p)b_m$ and $\mathrm{Tr}(C^2) = \rho^2[pb_{m+1}^2 + (Q-p)b_m^2] + (1-\rho^2)[pa_{m+1} + (Q-p)a_m]$, where a_n and b_n are defined in Eq. (A.6), from which one obtains

$$d = \begin{cases} b_N^2/a_N \ , & N \le Q \\ \frac{[pb_{m+1}+(Q-p)b_m]^2}{\rho^2[pb_{m+1}^2+(Q-p)b_m^2]+(1-\rho^2)[pa_{m+1}+(Q-p)a_m]} \ , & N > Q \end{cases} \tag{A.10}$$

In the approximation where all neurons have the same variance this simplifies to

$$d = \begin{cases} N \ , & N \le Q \\ \frac{N}{1+m\rho^2[1-(Q-p)/N]} \ , & N > Q \end{cases} \tag{A.11}$$

Recall that in the formulae above m and p depend on N. For finite ρ, Eq. (A.11) predicts the bound $d \le Q/\rho^2$ for any N, with this value reached asymptotically for large N. When single neuron variances σ_i^2 are drawn from a distribution with mean $E[\sigma_i^2] = \sigma^2$ and variance $\mathrm{var}[\sigma_i^2] = \delta\sigma^4$, an expression for the dimensionality can be easily obtained from Eq. (A.10) at leading order around the expectation values in Eq. A.8 (not shown).

Dependence on the number of trials (Fig. 7E). The estimate of d from data depends on the number and duration of the trials (Fig. 3E). To understand this phenomenon in a simple theoretical setting we generated $N \times N_T$ "nominal" firing rates, thought of as originating from N neurons, each sampled N_T times (trials). The single neuron firing rates were sampled according to a log-normal distribution with equal means and covariance leading to Eq. (A.5), i.e., $C_{ij} = \rho\sigma_i\sigma_j(1-\delta_{ij}) + \sigma_i^2\delta_{ij}$, with $\delta_{ij} = 1$ if $i = j$, and zero otherwise (note that the actual distribution used is immaterial since the dimensionality only depends on the covariance matrix, see Eq. (A.4)). We considered the two cases of equal variance for all ensemble neurons, $\sigma_i = \sigma$ for all i (corresponding to Eq. (A.7); dashed lines in Fig. 7E) or variances σ_i sampled from a lognormal distribution (corresponding to Eq. (A.5); "+" in Fig. 7E). The same N and N_T as used for the analysis of the Poisson spike trains and model simulations in Fig. 7D were used (where the "trials" were N_T bins of 200 ms duration for each ensemble size N). The covariance of the data thus generated was estimated according to Eq. (A.2), based on which the dimensionality Eq. (A.4) was computed (Fig. 7E). The estimated dimensionality depends on N and N_T and was averaged across 100 values of d, each obtained as explained above. Note that in this simplified setting increasing the duration of each trial is equivalent to adding more trials, i.e., the effect of having a trial 400 ms long producing 2 firing rates (one for each 200 ms bin) is equivalent to having two trials of 200 ms duration.

Model fitting. The dependence of the data"s dimensionality on ensemble size N was fitted by a straight line via standard least-squares,

$$d = \beta_1 \cdot N + \beta_0 \ ,$$

separately for ongoing and evoked activity (Fig. 3 and 6). Comparison between the dimensionality of evoked and ongoing activity was carried out with a 2-way ANOVA with condition (evoked vs. ongoing) and ensemble

size (N) as factors. Since d depends on the number and duration of the trials used to estimate the covariance matrix (Fig. 3E), we matched both the number of trials and trial length in comparisons of ongoing and evoked dimensionality. If multiple tastes were used, the evoked trials were each matched to a random subset of an equal number of ongoing trials. The dependence of dimensionality d on ensemble size N in a surrogate dataset of Poisson spike trains with mean pairwise correlation ρ (generated according to the algorithm described in the next section) was modeled as

$$d(\rho, N) = \frac{1}{(a\rho)^2 + (1-(a\rho)^2)g(N)} \ , \tag{A.12}$$

(Fig. 7D, dashed lines) with $g(N) = b/N + c/N^2$ and a, b and c constant parameters to be tuned to fit the data. Eq. (A.12) can be understood as the theoretical formula Eq. (A.5) corrected to account for a finite number of trials via an effective correlation parameter $a\rho$, and with only the first two terms of the asymptotic expansion of $g(N)$ taken into account. We fitted Eq. (A.12) to the data via non-linear least-squares [46] simultaneously on datasets with $N = 5, 10, \ldots, 100$ and $\rho = 0, 0.01, 0.05, 0.1, 0.2$, with 20 ensembles for each value (Fig. 7D; only the fits for $\rho = 0, 0.1, 0.2$ are shown).

A.6 Generation of correlated Poisson spike trains

Ensembles of independent and correlated Poisson spike trains were generated for the analysis of Fig. (7). Ensembles of independent stationary Poisson spike trains with given firing rates ν_i were generated by producing their interspike intervals according to an exponential distribution with parameter ν_i. Stationary Poisson spike trains with fixed pairwise correlations (but no temporal correlations) were generated according to the method reported in [47], that we briefly outline below. We split each trial into 1 ms bins and consider the associated binary random variable $X_i(t) = 1$ if the i-th neuron emitted a spike in the t-th bin, and $X_i(t) = 0$ if no spike was emitted. These samples were obtained by first drawing a sample from an auxiliary N-dimensional Gaussian random variable $U \sim N(\gamma\, \lambda)$ and then thresholding it into 0 and 1: $X_i = 1$ if $U_i > 0$, and $X_i = 0$ otherwise. Here, $\gamma = \gamma_1, \gamma_2, \ldots, \gamma_N$ is the mean vector and $\lambda = \lambda_{ij}$ is the covariance matrix of the N-dimensional Gaussian variable U. For appropriately chosen parameters γ_i and λ_{ij} the method generates correlated spike trains with the desired firing rates ν_i and pairwise spike count correlation coefficients r_{ij}. The prescription for γ_i and λ_{ij} is most easily expressed as a function of the desired probabilities μ_i of having a spike in a bin of width dt, $\mu_i = P(X_i(t) = 1)$, and the pairwise covariance S_{ij} of the random binary vectors $X_i(t)$ and $X_j(t)$, from which γ_i and λ_{ij} can be obtained by inverting the following relationships:

$$\begin{aligned} \mu_i &= \Phi(\gamma_i) \ , \\ S_{ii} &= \Phi(\gamma_i)\Phi(-\gamma_i) \ , \\ S_{ij} &= \Phi_2(\gamma_i, \gamma_j, \lambda_{ij}) - \Phi(\gamma_i)\Phi(\gamma_j) \ , \quad i \neq j \ . \end{aligned}$$

Here, $\Phi(x)$ is the cumulative distribution of a univariate Gaussian with mean 0 and variance 1 evaluated at x, and $\Phi_2(x, y, \lambda)$ is the cumulative distribution of a bivariate Gaussian with means 0, variances 1 and covariance λ evaluated at (x, y) (note that the distributions Φ and Φ_2 are unrelated to the N-dimensional Gaussian $U \sim N(\gamma, \lambda)$). Without loss of generality we imposed unit variances for U_i, i.e. $\lambda_{ii} = 1$. We related the spike probabilities μ_i to the firing rates ν_i as $\mu_i = 1 - e^{-\nu_i dt}$, with $1 - \mu_i$ being the probability of no spikes in the same bin. When dt approaches zero, $\mu_i \simeq \nu_i dt$ and the spike trains generated as vectors of binary random variables by sampling $U \sim N(\gamma, \lambda)$ will approximate Poisson spike trains. We used $dt = 1$ ms since in this case it was extremely unlikely to observe two or more spikes in bins of this duration, as required by the Poisson hypothesis. In order to have a fair comparison with the data generated by the spiking network model (described in the next section), the mean firing rates of the Poisson spike trains were matched to the average firing rates

obtained from the simulated data. Since γ and λ were the same in all bins, values of $X_i(t)$ and $X_i(s)$ were independent for $t \neq s$ (i.e., the spike trains had no temporal correlations). As a consequence, the random binary vectors have the same pair-wise correlations as the spike counts, and the S_{ij} are related to the desired r_{ij} by $S_{ij} = r_{ij}\sqrt{\mu_i(1-\mu_i)\mu_j(1-\mu_j)}$, where $\mu_i(1-\mu_i)$ is the variance of X_i. See [47] for further details.

A.7 Spiking network model

We modeled the data with a recurrent spiking network of $N = 5000$ randomly connected leaky integrate-and-fire (LIF) neurons, of which 4000 excitatory (E) and 1000 inhibitory (I). Connection probability $p_{\beta\alpha}$ from neurons in population $\alpha \in E, I$ to neurons in population $\beta \in E, I$ were $p_{EE} = 0.2$ and $p_{EI} = p_{IE} = p_{ii} = 0.5$; a fraction $f = 0.9$ of excitatory neurons were arranged into Q different clusters, with the remaining neurons belonging to an unstructured ("background") population [48]. Synaptic weights $J_{\beta\alpha}$ from neurons in population $\alpha \in E, I$ to neurons in population $\beta \in E, I$ scaled with N as $J_{\beta\alpha} = j_{\beta\alpha}/\sqrt{N}$, with $j_{\beta\alpha}$ constants having the following values (units of mV): $j_{EI} = 3.18, j_{IE} = 1.06, j_{II} = 4.24, j_{EE} = 1.77$. Within an excitatory cluster synaptic weights were potentiated, i.e. they took average values of $\langle J\rangle_+ = J_+ j_{EE}$ with $J_+ > 1$, while synaptic weights between units belonging to different clusters were depressed to average values $\langle J\rangle_- = J_- j_{EE}$, with $J_- = 1 - \gamma f(J_+ - 1) < 1$, with $\gamma = 0.5$. The latter relationship between J_+ and J_- helps to maintain balance between overall potentiation and depression in the network [48]. Below spike threshold, the membrane potential V of each LIF neuron evolved according to

$$\tau_m \frac{dV}{dt} = -V + \tau_m(I_{\text{rec}} + I_{\text{ext}} + I_{\text{stim}}) \ ,$$

with a membrane time constant $\tau_m = 20$ ms for excitatory and 10 ms for inhibitory units. The input current was the sum of a recurrent input I_{rec}, an external current I_{ext} representing an ongoing afferent input from other areas, and an external stimulus I_{stim} representing e.g. a delivered taste during evoked activity only. In our units, a membrane capacitance of 1 nF is set to 1. A spike was said to be emitted when V crossed a threshold V_{thr}, after which V was reset to a potential $V_{reset} = 0$ for a refractory period of $\tau_{ref} = 5$ ms. Spike thresholds were chosen so that, in the unstructured network (i.e., with $J_+ = J_- = 1$), the E and I populations had average firing rates of 3 and 5 spikes/s, respectively [48]. The recurrent synaptic input I^i_{rec} to unit i evolved according to the dynamical equation

$$\tau_s \frac{dI^i_{\text{rec}}}{dt} = -I^i_{\text{rec}} + \sum_{j=1}^{N} J_{ij} \sum_k \delta(t - t^j_k) \ ,$$

where t^j_k was the arrival time of k-th spike from the j-th pre-synaptic unit, and τ_s was the synaptic time constant (3 and 2 ms for E and I units, respectively), resulting in an exponential post-synaptic current in response to a single spike, $\frac{J_{ij}}{\tau_s} e^{-t/\tau_s}\Theta(t)$, where $\Theta(t) = 1$ for $t \geq 0$, and $\Theta(t) = 0$ otherwise. The ongoing external current to a neuron in population α was constant and given by

$$I_{\text{ext}} = N_{\text{ext}} p_{\alpha 0} J_{\alpha 0} \nu_{\text{ext}} \ ,$$

where $N_{ext} = n_E N$, $p_{\alpha 0} = p_{EE}$, $J_{\alpha 0} = j_{\alpha 0}/\sqrt{N}$ with $j_{E0} = 0.3$, $j_{I0} = 0.1$, and $\nu_{ext} = 7$ spikes/s. During evoked activity, stimulus-selective units received an additional input representing one of the four incoming stimuli. The stimuli targeted combinations of neurons as observed in the data. Specifically, the fractions of neurons responsive to $n = 1, 2, 3$ or all 4 stimuli were $17\%(27/162)$, $22\%(36/162)$, $26\%(42/162)$, and $35\%(57/162)$ [5, 15]. Each stimulus had constant amplitude ν_{stim} ranging from 0 to $0.5\nu_{ext}$. In the following we measure the stimulus amplitude as percentage of ν_{ext} (e.g., "10%" corresponds to $\nu_{stim} = 0.1\nu_{ext}$). The

onset of each stimulus was always $t = 0$, the time of taste delivery. The stimulus current to a unit in population α was constant and given by $I_{stim} = N_{ext} p_\alpha 0 J_\alpha 0 \nu_{stim}$.

A.8 Mean field analysis of the model

The stationary states of the spiking network model in the limit of large N were found with a mean field analysis [5, 48–50]. Under typical conditions, each neuron of the network receives a large number of small post-synaptic currents (PSCs) per integration time constant. In such a case, the dynamics of the network can be analyzed under the diffusion approximation within the population density approach. The network has $\alpha = 1, \ldots, Q+2$ sub-populations, where the first Q indices label the Q excitatory clusters, $\alpha = Q+1$ labels the "background" units, and $\alpha = Q+2$ labels the homogeneous inhibitory population. In the diffusion approximation [51–53], the input to each neuron is completely characterized by the infinitesimal mean μ_α and variance σ_α^2 of the post-synaptic potential (see [5] for the expressions of the infinitesimal mean and variance for all subpopulations). Parameters were chosen so that the network with $J_+ = J_- = 1$ (where all $E \to E$ synaptic weights are equal) would operate in the balanced asynchronous regime [45, 54, 55], where incoming contributions from excitatory and inhibitory inputs balance out, neurons fire irregular spike trains with weak pair-wise correlations. The unstructured network has only one dynamical state, i.e., a stationary point of activity where all E and I neurons have constant firing rate ν_E and ν_I, respectively. In the structured network (where $J_+ > 1$), the network undergoes continuous transitions among a repertoire of states, as shown in the main text. To avoid confusion between network activity states and HMM states, we refer to the former as network "configurations" instead of states. Admissible networks configurations must satisfy the $Q+2$ self-consistent mean field equations [48]

$$\nu_\alpha = F_\alpha\left(\mu_\alpha(\overrightarrow{\nu}), \sigma_\alpha^2(\overrightarrow{\nu})\right) \ ,$$

where $\overrightarrow{\nu} = [\nu_1, \ldots, \nu_Q, \nu_E^{(bg)}, \nu_I]$ is the firing rate vector and $F_\alpha(\mu_\alpha, \sigma_\alpha^2)$ is the current-to-rate response function of the LIF neurons. For fast synaptic times, i.e. $\tau_s/\tau_m << 1$, $F_\alpha(\mu_\alpha, \sigma_\alpha^2)$ is well approximated by [56, 57]

$$F_\alpha(\mu_\alpha, \sigma_\alpha) = \left(\tau_{\text{ref}} + \tau_{m,\alpha}\sqrt{\pi}\int_{H_{\text{eff},\alpha}}^{\Theta_{\text{eff},\alpha}} e^{u^2}\left[1 + \text{erf}(u)\right]\right)^{-1} ,$$

where

$$\begin{aligned} \Theta_{\text{eff},\alpha} &= \frac{V_{\text{thr},\alpha} - \mu_\alpha}{\sigma_\alpha} + a k_\alpha \ , \\ H_{\text{eff},\alpha} &= \frac{V_{\text{reset},\alpha} - \mu_\alpha}{\sigma_\alpha} + a k_\alpha, \end{aligned}$$

where $k_\alpha = \sqrt{\tau_{s,\alpha}/\tau_{m,\alpha}}$ is the square root of the ratio of synaptic time constant to membrane time constant, and $a = |\zeta(1/2)|/\sqrt{2} \sim 1.03$. This theoretical response function has been fitted successfully to the firing rate of neocortical neurons in the presence of in vivo-like fluctuations [58–60]. The fixed points $\overrightarrow{\nu}^*$ of the mean field equations were found with Newtons method [61]. The fixed points can be either stable (attractors) or unstable depending on the eigenvalues λ_α of the stability matrix

$$S_{\alpha\beta} = \frac{1}{\tau_{s,\alpha}}\left(\frac{\partial F_\alpha\left(\mu_\alpha(\overrightarrow{\nu}), \sigma_\alpha^2(\overrightarrow{\nu})\right)}{\partial \nu_\beta} - \frac{\partial F_\alpha\left(\mu_\alpha(\overrightarrow{\nu}), \sigma_\alpha^2(\overrightarrow{\nu})\right)}{\partial \sigma_\alpha^2}\frac{\partial \sigma_\alpha^2}{\partial \nu_\beta} - \delta_{\alpha\beta}\right) \ ,$$

evaluated at the fixed point $\overrightarrow{\nu}^*$ [62]. If all eigenvalues have negative real part, the fixed point is stable (attractor). If at least one eigenvalue has positive real part, the fixed point is unstable. Stability is meant with respect to an

approximate linearized dynamics of the mean and variance of the input current:

$$\begin{aligned}\tau_{s,\alpha}\frac{dm_\alpha}{dt} &= -m_\alpha + \mu_\alpha(\overrightarrow{\nu}) \,, \\ \frac{\tau_{s,\alpha}}{2}\frac{ds^2_\alpha}{dt} &= -s^2_\alpha + \sigma^2_\alpha(\overrightarrow{\nu}) \,, \\ \nu_\alpha(t) &= F_\alpha\left(m_\alpha(\overrightarrow{\nu}), s^2_\alpha(\overrightarrow{\nu})\right) \,,\end{aligned}$$

where μ_α and σ^2_α are the stationary values for fixed $\overrightarrow{nu}$ given earlier. For fast synaptic dynamics in the asynchronous balanced regime, these rate dynamics are in very good agreement with simulations ([63] see [64, 65] for more detailed discussions).

A.9 Metastable configurations in the network model

The stable configurations of a network with an infinite number of neurons were obtained in the mean field approximation of the previous section and are shown in Fig. 4B for $Q = 30$ and a range of values of the relative potentiation parameter J_+. Above the critical point $J_+ = 4.2$, stable configurations characterized by a finite number of active clusters emerge (grey lines; the number of active clusters is reported next to each line). For a given J_+, the firing rate is the same in all active clusters and is inversely proportional to the total number of active clusters. Stable patterns of firing rates are also found in the inhibitory population (red lines), in the inactive clusters (having low firing rates; grey dashed lines), and in the unstructured excitatory population (dashed blue lines). For a fixed value of J_+, multiple stable configurations coexist with different numbers of active clusters. For example, for $J_+ = 5.3$, stable configurations with up to 7 active clusters are stable, each configuration with different firing rates. This generates multistable firing rates in single neurons, i.e., the property, also observed in the data, that single neurons can attain more than 2 firing rates across states [5]. Note that if $J_+ \leq 5.15$ an alternative stable configuration of the network with all clusters inactive (firing rates < 10 spikes/s) is also possible (single brown line). Strictly speaking, the configurations in Fig. 4B are stable only in a network containing an infinite number of uncorrelated neurons. In a finite network (or when neurons are strongly correlated) these configurations can lose stability due to strong fluctuations which ignite transitions among the different configurations. Full details are reported in [5].

A.10 Model simulations and analysis of simulated data

The dynamical equations of the LIF neurons were integrated with the Euler algorithm with a time step of $dt = 0.1$ ms. We simulated 20 different networks (referred to as "sessions" in the following) during both ongoing and evoked activity. We chose four different stimuli per session during evoked activity (to mimic taste delivery). Trials were 5 seconds long. The HMM analyses for Figs. 2 and 5 were performed on ensembles of randomly selected excitatory neurons with the same procedure used for the data (see previous section "Hidden Markov Model (HMM) analysis"). The ensemble sizes were chosen so as to match the empirical ensemble sizes (3 to 9 randomly selected neurons). For the analysis of Fig. 8B, random ensembles of increasing size (from 5 to 100 neurons) were used from simulations with $Q = 30$ clusters. When the ensemble size was less than the number of clusters ($N \leq Q$), each neuron was selected randomly from a different cluster; when ensemble size was larger than the number of clusters, one neuron was added to each cluster until all clusters were represented, and so on until all N neurons had been chosen. To allow comparison with surrogate Poisson spike trains, the dimensionality of the simulated data was computed from the firing rate vectors in $T = 200$ ms bins as explained in Sec. A.5. For control, the dimensionality was also computed from the firing rate vectors in hidden states obtained from an HMM analysis, obtaining qualitatively similar results.

References

[1] M. Abeles, H. Bergman, I. Gat, I. Meilijson, E. Seidemann, N. Tishby, and E. Vaadia, *Cortical activity flips among quasi-stationary states*, *Proc Natl Acad Sci USA* **92** (1995) 8616–8620.

[2] E. Seidemann, I. Meilijson, M. Abeles, H. Bergman, and E. Vaadia, *Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task*, *J Neurosci* **16** (1996), no. 2 752–68.

[3] D. Durstewitz, N. M. Vittoz, S. B. Floresco, and J. K. Seamans, *Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning*, *Neuron* **66** (2010), no. 3 438–48.

[4] L. M. Jones, A. Fontanini, B. F. Sadacca, P. Miller, and D. B. Katz, *Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles*, *Proc Natl Acad Sci U S A* **104** (2007), no. 47 18772–7.

[5] L. Mazzucato, A. Fontanini, and G. La Camera, *Dynamics of multistable states during ongoing and evoked cortical activity*, *The Journal of Neuroscience* **35** (2015), no. 21 8214–8231.

[6] C. Kemere, G. Santhanam, B. M. Yu, A. Afshar, S. I. Ryu, T. H. Meng, and K. V. Shenoy, *Detecting neural-state transitions using hidden markov models for motor cortical prostheses*, *J Neurophysiol* **100** (2008), no. 4 2441–52.

[7] A. Ponce-Alvarez, V. Nacher, R. Luna, A. Riehle, and R. Romo, *Dynamics of cortical neuronal ensembles transit from decision making to storage for later report*, *J Neurosci* **32** (2012), no. 35 11956–69.

[8] P. Gao and S. Ganguli, *On simplicity and complexity in the brave new world of large-scale neuroscience*, *Current opinion in neurobiology* **32** (2015) 148–155.

[9] L. F. Abbott, K. Rajan, and H. Sompolinsky, *Interactions between Intrinsic and Stimulus-Evoked Activity in Recurrent Neural Networks*, ch. 4. Oxford University Press, 2011.

[10] S. Ganguli, J. W. Bisley, J. D. Roitman, M. N. Shadlen, M. E. Goldberg, and K. D. Miller, *One-dimensional dynamics of attention and decision making in lip*, *Neuron* **58** (2008), no. 1 15–25.

[11] S. Ganguli and H. Sompolinsky, *Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis*, *Annual review of neuroscience* **35** (2012) 485–508.

[12] C. F. Cadieu, H. Hong, D. Yamins, N. Pinto, N. J. Majaj, and J. J. DiCarlo, *The neural representation benchmark and its evaluation on brain and machine*, *arXiv preprint arXiv:1301.3530* (2013).

[13] M. Rigotti, O. Barak, M. R. Warden, X. J. Wang, N. D. Daw, E. K. Miller, and S. Fusi, *The importance of mixed selectivity in complex cognitive tasks*, *Nature* **497** (2013), no. 7451 585–90.

[14] D. B. Katz, S. A. Simon, and M. A. Nicolelis, *Dynamic and multimodal responses of gustatory cortical neurons in awake rats*, *J Neurosci* **21** (2001), no. 12 4478–89.

[15] A. Jezzini, L. Mazzucato, G. La Camera, and A. Fontanini, *Processing of hedonic and chemosensory features of taste in medial prefrontal and insular networks*, *The Journal of Neuroscience* **33** (2013), no. 48 18966–18978.

[16] R. Kiani, C. J. Cueva, J. B. Reppas, D. Peixoto, S. I. Ryu, and W. T. Newsome, *Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex*, *Neuron* **85** (2015), no. 6 1359–1373.

[17] M. N. Shadlen and W. T. Newsome, *Noise, neural codes and cortical organization*, *Curr Opin Neurobiol* **4** (1994), no. 4 569–79.

[18] A. Fontanini and D. B. Katz, *Behavioral states, network states, and sensory response variability*, *J Neurophysiol* **100** (2008), no. 3 1160–8.

[19] R. Moreno-Bote, *Poisson-like spiking in circuits with probabilistic synapses.*, *PLoS computational biology* **10** (2014), no. 7 e1003522–e1003522.

[20] M. Gur, A. Beylin, and D. M. Snodderly, *Response variability of neurons in primary visual cortex (v1) of alert monkeys*, *The Journal of neuroscience* **17** (1997), no. 8 2914–2920.

[21] B. White, L. F. Abbott, and J. Fiser, *Suppression of cortical neural variability is stimulus-and state-dependent*, *Journal of neurophysiology* **108** (2012), no. 9 2383–2392.

[22] M. D. McDonnell and L. M. Ward, *The benefits of noise in neural systems: bridging theory and experiment*, *Nature Reviews Neuroscience* **12** (2011), no. 7 415–426.

[23] M. M. Churchland, B. M. Yu, J. P. Cunningham, L. P. Sugrue, M. R. Cohen, G. S. Corrado, W. T. Newsome, A. M. Clark, P. Hosseini, B. B. Scott, D. C. Bradley, M. A. Smith, A. Kohn, J. A. Movshon, K. M. Armstrong, T. Moore, S. W. Chang, L. H. Snyder, S. G. Lisberger, N. J. Priebe, I. M. Finn, D. Ferster, S. I. Ryu, G. Santhanam, M. Sahani, and K. V. Shenoy, *Stimulus onset quenches neural variability: a widespread cortical phenomenon*, *Nat Neurosci* **13** (2010), no. 3 369–78.

[24] A. Litwin-Kumar and B. Doiron, *Slow dynamics and high variability in balanced cortical networks with clustered connections*, *Nat Neurosci* **15** (2012), no. 11 1498–505.

[25] G. Deco and E. Hugues, *Neural network mechanisms underlying stimulus driven variability reduction*, *PLoS Comput Biol* **8** (2012), no. 3 e1002395.

[26] B. M. Broome, V. Jayaraman, and G. Laurent, *Encoding and decoding of overlapping odor sequences*, *Neuron* **51** (2006), no. 4 467–482.

[27] M. N. Geffen, B. M. Broome, G. Laurent, and M. Meister, *Neural encoding of rapidly fluctuating odors*, *Neuron* **61** (2009), no. 4 570–586.

[28] G. Montavon, M. L. Braun, and K.-R. Müller, *Kernel analysis of deep networks*, *The Journal of Machine Learning Research* **12** (2011) 2563–2581.

[29] J. J. DiCarlo, D. Zoccolan, and N. C. Rust, *How does the brain solve visual object recognition?*, *Neuron* **73** (2012), no. 3 415–434.

[30] B. F. Sadacca, J. T. Rothwax, and D. B. Katz, *Sodium concentration coding gives way to evaluative coding in cortex and amygdala*, *The Journal of Neuroscience* **32** (2012), no. 29 9999–10011.

[31] T. Yamamoto, N. Yuyama, and Y. Kawamura, *Cortical neurons responding to tactile, thermal and taste stimulations of the rat's tongue*, *Brain research* **221** (1981), no. 1 202–206.

[32] T. Yamamoto, R. Matsuo, Y. Kiyomitsu, and R. Kitamura, *Sensory inputs from the oral region to the cerebral cortex in behaving rats: an analysis of unit responses in cortical somatosensory and taste areas during ingestive behavior*, *Journal of neurophysiology* **60** (1988), no. 4 1303–1321.

[33] J. X. Maier and D. B. Katz, *Neural dynamics in response to binary taste mixtures*, *Journal of neurophysiology* **109** (2013), no. 8 2108–2117.

[34] S. E. Grossman, A. Fontanini, J. S. Wieskopf, and D. B. Katz, *Learning-related plasticity of temporal coding in simultaneously recorded amygdala–cortical ensembles*, *The Journal of neuroscience* **28** (2008), no. 11 2864–2873.

[35] C. L. Samuelsen, M. P. Gardner, and A. Fontanini, *Effects of cue-triggered expectation on cortical processing of taste*, *Neuron* **74** (2012), no. 2 410–422.

[36] F. Bermudez-Rattoni, *The forgotten insular cortex: its role on recognition memory formation*, *Neurobiology of learning and memory* **109** (2014) 207–216.

[37] I. E. de Araujo, R. Gutierrez, A. J. Oliveira-Maia, A. Pereira, M. A. Nicolelis, and S. A. Simon, *Neural ensemble coding of satiety states*, *Neuron* **51** (2006), no. 4 483–494.

[38] A. Fontanini and D. B. Katz, *State-dependent modulation of time-varying gustatory responses*, *J Neurophysiol* **96** (2006), no. 6 3183–93.

[39] A. Fontanini and D. B. Katz, *7 to 12 hz activity in rat gustatory cortex reflects disengagement from a fluid self-administration task*, *J Neurophysiol* **93** (2005), no. 5 2832–40.

[40] M. Phillips and R. Norgren, *A rapid method for permanent implantation of an intraoral fistula in rats*, *Behavioral Research, Methods, and Instrumentation* (1970), no. 2 124.

[41] N. K. Horst and M. Laubach, *Reward-related activity in the medial prefrontal cortex is driven by consumption*, *Front Neurosci* **7** (2013) 56.

[42] S. Escola, A. Fontanini, D. Katz, and L. Paninski, *Hidden markov models for the stimulus-response relationships of multistate neural systems*, *Neural Comput* **23** (2011), no. 5 1071–132.

[43] L. R. Rabiner, *A tutorial on hidden markov models and selected applications in speech recognition*, *Proceedings of the IEEE* **77** (1989) 257–286.

[44] P. Miller and D. B. Katz, *Stochastic transitions between neural states in taste processing and decision-making*, *J Neurosci* **30** (2010), no. 7 2559–70.

[45] A. Renart, J. de la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris, *The asynchronous state in cortical circuits*, *Science* **327** (2010), no. 5965 587–90.

[46] P. W. Holland and R. E. Welsch, *Robust regression using iteratively reweighted least-squares*, *Communications in Statistics-theory and Methods* **6** (1977), no. 9 813–827.

[47] J. H. Macke, P. Berens, A. S. Ecker, A. S. Tolias, and M. Bethge, *Generating spike trains with specified correlation coefficients*, *Neural Computation* **21** (2009), no. 2 397–423.

[48] D. J. Amit and N. Brunel, *Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex*, *Cereb Cortex* **7** (1997), no. 3 237–52.

[49] N. Brunel and V. Hakim, *Fast global oscillations in networks of integrate-and-fire neurons with low firing rates*, *Neural Comput* **11** (1999), no. 7 1621–71.

[50] S. Fusi and M. Mattia, *Collective behavior of networks with linear (vlsi) integrate-and-fire neurons*, *Neural Comput* **11** (1999), no. 3 633–52.

[51] H. Tuckwell, *Introduction to theoretical neurobiology*, vol. 2. Cambridge University Press, 1988.

[52] P. Lansky and S. Sato, *The stochastic diffusion models of nerve membrane depolarization and interspike interval generation*, *J Peripher Nerv Syst* **4** (1999), no. 1 27–42.

[53] M. J. Richardson, *The effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons.*, *Phys Rev E Stat Nonlin Soft Matter Phys* **69** (2004) 051,918.

[54] C. van Vreeswijk and H. Sompolinsky, *Chaos in neuronal networks with balanced excitatory and inhibitory activity*, *Science* **274** (1996), no. 5293 1724–1726.

[55] C. v. Vreeswijk and H. Sompolinsky, *Chaotic balanced state in a model of cortical circuits*, *Neural computation* **10** (1998), no. 6 1321–1371.

[56] N. Brunel and S. Sergi, *Firing frequency of leaky intergrate-and-fire neurons with synaptic current dynamics*, *J Theor Biol* **195** (1998), no. 1 87–95.

[57] N. Fourcaud and N. Brunel, *Dynamics of the firing probability of noisy integrate-and-fire neurons*, *Neural Comput* **14** (2002), no. 9 2057–110.

[58] A. Rauch, G. La Camera, H. R. Luscher, W. Senn, and S. Fusi, *Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents*, *J Neurophysiol* **90** (2003), no. 3 1598–612.

[59] G. La Camera, A. Rauch, D. Thurbon, H. R. Luscher, W. Senn, and S. Fusi, *Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons*, *J Neurophysiol* **96** (2006), no. 6 3448–64.

[60] G. La Camera, M. Giugliano, W. Senn, and S. Fusi, *The response of cortical neurons to in vivo-like input current: theory and experiment : I. noisy inputs with stationary statistics*, *Biol Cybern* **99** (2008), no. 4-5 279–301.

[61] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical recipes the art of scientific computing*, 2007.

[62] M. Mascaro and D. J. Amit, *Effective neural response function for collective population states*, *Network* **10** (1999), no. 4 351–73.

[63] G. La Camera, A. Rauch, H. R. Luscher, W. Senn, and S. Fusi, *Minimal models of adapted neuronal response to in vivo-like input currents*, *Neural Comput* **16** (2004), no. 10 2101–24.

[64] A. Renart, N. Brunel, and X.-J. Wang, *Mean-field theory of recurrent cortical networks: from irregularly spiking neurons to working memory*, pp. 431–490. Boca Raton, FL: CRC, 2004.

[65] M. Giugliano, G. La Camera, S. Fusi, and W. Senn, *The response of cortical neurons to in vivo-like input current: theory and experiment: Ii. time-varying and spatially distributed inputs*, *Biol Cybern* **99** (2008), no. 4-5 303–18.