

State-Dependent Population Coding in Primary Auditory Cortex

Marius Pachitariu,¹ Dmitry R. Lyamzin,^{2,3} Maneesh Sahani,¹ and Nicholas A. Lesica²

¹Gatsby Computational Neuroscience Unit, and ²Ear Institute, University College London, London WC1E 6BT, United Kingdom, and ³Graduate School of Systemic Neurosciences, Ludwig-Maximilians-University Munich, 80802 Munich, Germany

Sensory function is mediated by interactions between external stimuli and intrinsic cortical dynamics that are evident in the modulation of evoked responses by cortical state. A number of recent studies across different modalities have demonstrated that the patterns of activity in neuronal populations can vary strongly between synchronized and desynchronized cortical states, i.e., in the presence or absence of intrinsically generated up and down states. Here we investigated the impact of cortical state on the population coding of tones and speech in the primary auditory cortex (A1) of gerbils, and found that responses were qualitatively different in synchronized and desynchronized cortical states. Activity in synchronized A1 was only weakly modulated by sensory input, and the spike patterns evoked by tones and speech were unreliable and constrained to a small range of patterns. In contrast, responses to tones and speech in desynchronized A1 were temporally precise and reliable across trials, and different speech tokens evoked diverse spike patterns with extremely weak noise correlations, allowing responses to be decoded with nearly perfect accuracy. Restricting the analysis of synchronized A1 to activity within up states yielded similar results, suggesting that up states are not equivalent to brief periods of desynchronization. These findings demonstrate that the representational capacity of A1 depends strongly on cortical state, and suggest that cortical state should be considered as an explicit variable in all studies of sensory processing.

Key words: brain state; cortex; noise correlations; population code

Introduction

The representation of sensory inputs in the activity of primary cortical areas provides the basis for higher-level processing. Characterizing this primary representation is critical for understanding sensory function, as its nature determines the suitability of different strategies for subsequent computations, and its fidelity constrains behavioral performance. The study of sensory representations is complicated by the fact that neuronal activity is determined not only by external inputs, but also by other sources that are internal to the brain. In cortex, the processing of incoming stimuli can depend strongly on brain state (Steriade et al., 2001; Castro-Alamancos, 2004a; Haider and McCormick, 2009; Harris and Thiele, 2011). In asleep, anesthetized, and awake animals, the state of the cortex can vary along a continuum of synchronized and desynchronized states with different population dynamics. When the cortex is in a synchronized state (also known as an inactivated state), activity is characterized by slow fluctuations

between intrinsically generated up and down states, corresponding to periods of concerted spiking and silence across large areas, and these up and down states play a major role in shaping activity patterns (Marguet and Harris, 2011; Okun et al., 2012). Synchronized states are commonly observed during slow-wave sleep and under certain anesthetics, but recent studies have shown that the cortex can also be in a synchronized state when animals are awake (Crochet and Petersen, 2006; Greenberg et al., 2008; Poulet and Petersen, 2008; Xu et al., 2012; Luczak et al., 2013; Polack et al., 2013; Sachidhanandam et al., 2013; Tan et al., 2014; Zhou et al., 2014).

During active sensory processing in awake animals, the cortex often transitions to a desynchronized (or activated) state in which up and down states are suppressed and activity is strongly modulated by sensory inputs. Studies in the visual and somatosensory systems have observed dramatic differences between responses in synchronized and desynchronized states (Castro-Alamancos, 2004b; Hasenstaub et al., 2007; Goard and Dan, 2009; Hirata and Castro-Alamancos, 2011), and there are indications that such differences may also be present in the primary auditory cortex (A1; Ter-Mikaelian et al., 2007; Curto et al., 2009; Otazu et al., 2009; Marguet and Harris, 2011; Guo et al., 2012; Zhou et al., 2014). In this study, we measured the activity of populations of single units in gerbil A1 in synchronized and desynchronized states under different anesthetics and observed strong effects that were evident at both the single cell and population level. We found that cortical state modulated the selectivity, reliability, and diversity of spike patterns, as well as the strength of noise correlations, in a manner that greatly impacted the fidelity of the population code.

Received Aug. 11, 2014; revised Oct. 17, 2014; accepted Nov. 18, 2014.

Author contributions: M.P., M.S., and N.A.L. designed research; D.R.L. and N.A.L. performed research; M.P. and N.A.L. analyzed data; M.P., M.S., and N.A.L. wrote the paper.

This work was supported by the Gatsby Charitable Foundation and the Wellcome Trust. We thank K. Harris and J. Bizley for helpful discussions.

The authors declare no competing financial interests.

This article is freely available online through the *JNeurosci* Author Open Choice option.

Correspondence should be addressed to Dr Nicholas Lesica, Ear Institute, University College London, 332 Gray's Inn Road, London WC1X 8EE, UK. E-mail: n.lesica@ucl.ac.uk.

Copyright © 2015 Pachitariu et al.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

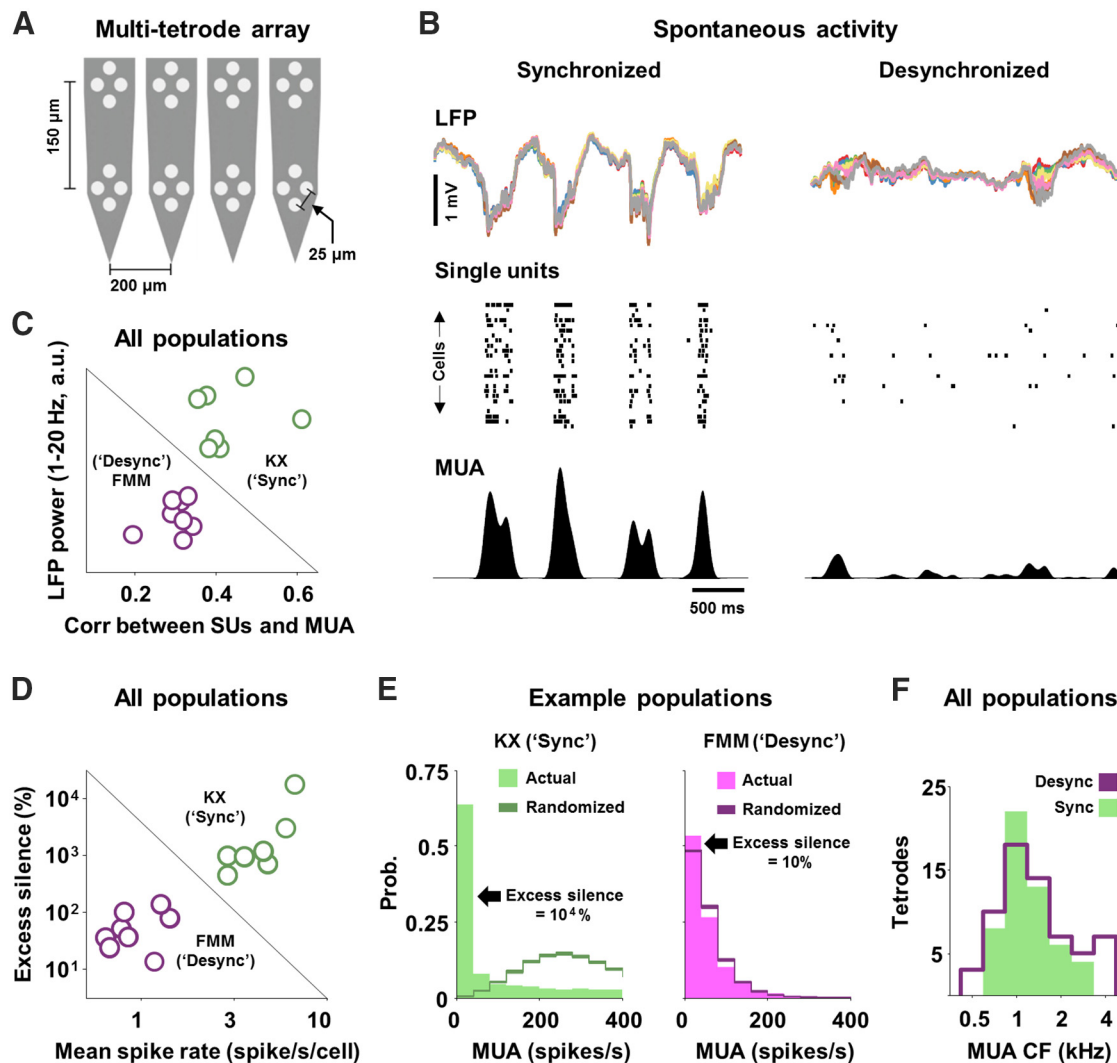


Figure 1. Synchronized and desynchronized states in A1. **A**, A schematic diagram of the multitetrode array used to record A1 activity. **B**, Examples of a short segment of spontaneous activity recorded in synchronized (under KX) and desynchronized (under FMM) A1. Top row, The LFP (0.1–100 Hz) recorded on each of the eight tetrodes with the signal for each tetrode shown in a different color. Each tetrode signal is the sum across its four electrodes. Middle row, A raster plot of the spiking of all of the single units in the population. Each row shows the spike times for one cell. Bottom row, MUA; the sum of the activity of all isolated single units after smoothing with a Gaussian window with a width of 50 ms. **C**, A scatter plot showing the low-frequency LFP power (1–20 Hz) and average correlation between the MUA and spiking of each single unit for spontaneous activity in all of the synchronized (green) and desynchronized (purple) populations that were analyzed. **D**, A scatter plot showing the excess silence and average mean spike rates for spontaneous activity in all of the synchronized (green) and desynchronized (purple) populations that were analyzed. **E**, The distributions of MUA spike rates during spontaneous activity before and after randomizing the spike times of each cell for example synchronized and desynchronized populations. The filled distributions correspond to the actual activity and the lines correspond to the distributions obtained from fifty different randomizations. The excess silence, i.e., the probability of complete silence in the actual activity of the population relative to that in the randomized activity, is indicated. **F**, The distributions of CFs for the MUA on each tetrode for all of the synchronized (green) and desynchronized (purple) populations that were analyzed. The CF was the frequency at which the MUA was most sensitive, i.e., the frequency for which the MUA was significantly larger than spontaneous activity at the lowest intensity.

Materials and Methods

In vivo recordings

Adult male gerbils (70–90 g, P60–P120) were anesthetized for surgery with one of three different anesthetics: ketamine/xylazine (KX), fentanyl/medetomidine/midazolam (FMM), or urethane. For KX, an initial injection of 1 ml per 100 g body weight was given of ketamine (100 mg/ml), xylazine (2% w/v), and saline in a ratio of 5:1:19, and the same solution was infused continuously during recording at a rate $\sim 2.5 \mu\text{l}/\text{min}$. For FMM, an initial injection of 0.2 ml per 100 g body weight was given with fentanyl (0.05 mg/ml), medetomidine (1 mg/ml), and midazolam (5 mg/ml) in a ratio of 4:1:10, and the same solution was infused continuously during recording at a rate of $\sim 0.08 \mu\text{l}/\text{min}$. For urethane, an initial injection of urethane and saline containing 0.15 g of urethane per 100 g body weight was given. Internal temperature was monitored and maintained at 38.7°C and heart rate was consistently ~ 300 bpm under all anesthetics. A small metal rod was mounted on

the skull and used to secure the head of the animal in a stereotaxic device in a sound-attenuated chamber. A craniotomy was made over A1, an incision was made in the dura mater, and a multitetrode array (Fig. 1A; Neuronexus) was inserted into the brain. Only recordings from A1, determined by the direction of the tonotopic gradient (Thomas et al., 1993), were analyzed. Recordings were made between 1 and 1.5 mm from the cortical surface (most likely in layer V; Happel et al. (2010)).

Sound delivery

Sounds were generated with a 48 kHz sampling rate, attenuated, and delivered to speakers coupled to tubes inserted into both ear canals for diotic sound presentation along with microphones for calibration. The frequency response of these speakers measured at the entrance of the ear canal was flat (± 5 dB SPL) between 0.2 and 5 kHz. The properties of each sound are given below.

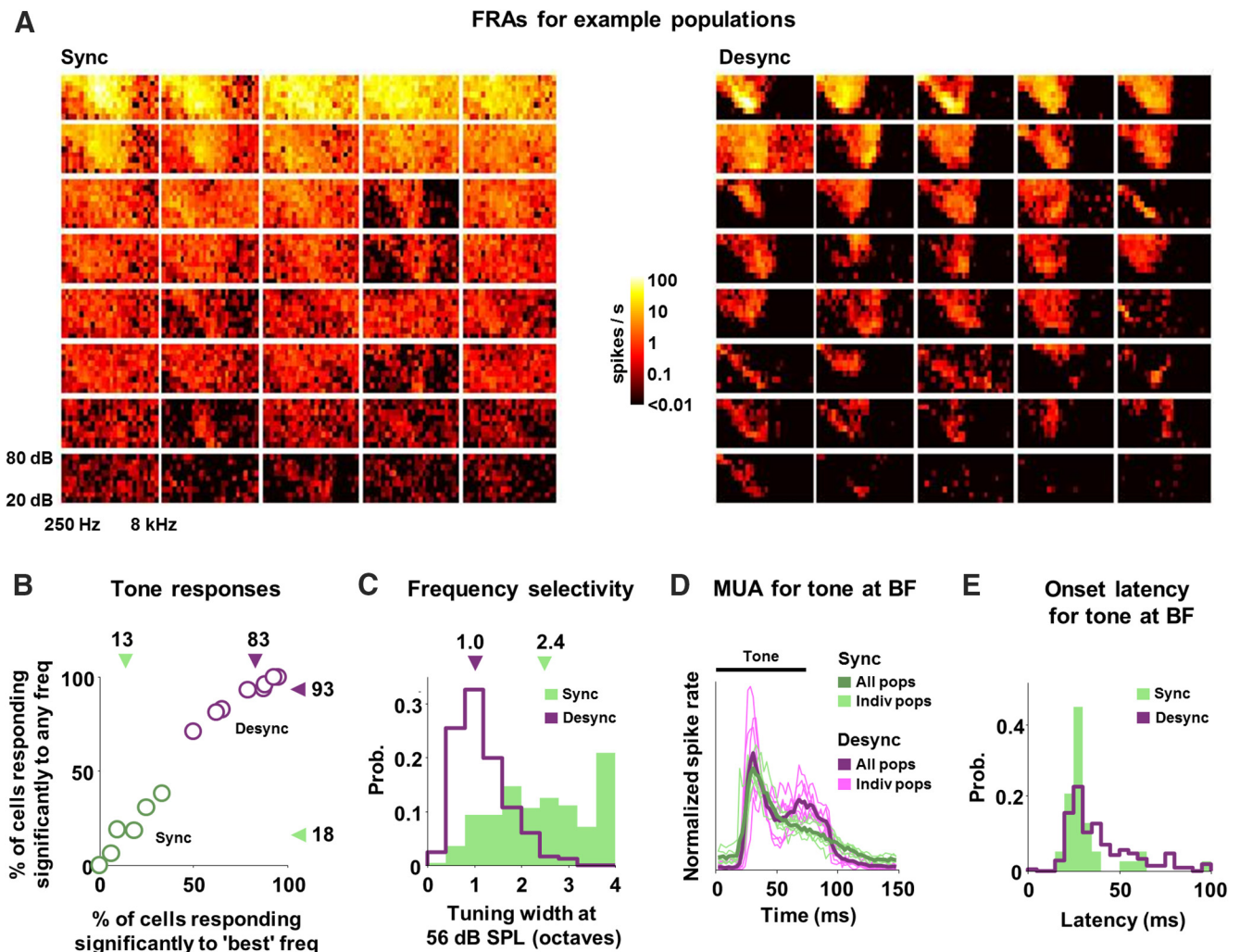


Figure 2. The impact of cortical state on responses to tones. **A**, FRAs for example populations in synchronized and desynchronized A1. Each image shows the average spike rate of responses to tones of different frequencies and intensities for one cell. Cells were ordered according to how strongly their activity was modulated by the tones as measured by the variance in their average spike rates across all frequencies and intensities. The two cells that were most weakly modulated in the synchronized population and the one cell that was most weakly modulated in the desynchronized population are not shown. **B**, A scatter plot showing the percentage of cells in each synchronized (green) and desynchronized (purple) population that responded to the best frequency for that population (i.e., the frequency that evoked a significant response from the largest fraction of cells) and the fraction of cells that responded significantly to at least one of the frequencies tested. A response was considered significant if the average spike rate was >2 SDs above the average spontaneous rate. The median values are indicated by the arrows. **C**, The distribution of the frequency tuning widths for cells in synchronized (green) and desynchronized (purple) A1. Tuning width was measured as the range of frequencies for which the average spike rate was at least half of its maximum value for tones at 56 dB SPL. The median values are indicated by the arrows. **D**, The tone-evoked MUA for populations in synchronized (green) and desynchronized (purple) A1. The thin lines show the MUA for each population in response to its best frequency, and the thick lines show the medians of the thin lines. The MUA for all populations were normalized to have the same sum. The timing of the tone is indicated by the horizontal bar. **E**, The distributions of onset latencies for responses to tones at best frequency for all cells that responded significantly to tones in synchronized (green) and desynchronized (purple) A1.

- (1) *Silence*. Ten minutes without the presentation of any sound. The spontaneous activity recorded during this period was used to measure strength of up and down states based on the low-frequency power in the local field potentials (LFPs), the correlation between single-unit spiking and the multiunit activity (MUA), and excess silence as described below.
- (2) *Tone set 1*. Seventy-five millisecond tones with frequencies ranging from 256 to 8192 Hz in 0.2 octave steps and intensities ranging from 16 to 80 dB SPL in 8 dB steps with 5 ms cosine on and off ramps and a 75 ms pause between tones. Tones were presented 10 times each in random order. Responses to these sounds were used to measure frequency response areas (FRAs) and tuning width as described below, as well as center frequencies for the MUA on each tetrode.
- (3) *Tone set 2*. Seventy-five millisecond tones with frequencies ranging from 256 Hz to 3104 Hz in 0.6 octave steps at 56 dB SPL with 5 ms cosine on and off ramps and 75 ms pause between sounds.

Tones were presented 200 times each in random order. These sounds were used to measure first spike latencies and tone responsiveness as described below.

- (4) *Frequency-modulated tones*. Chirps in which the frequency either increased from 64 to 8192 Hz or decreased from 8192 to 64 Hz at speeds of 16, 32, 64, 128, 256, or 512 octaves/s with 2 ms cosine on and off ramps and a 250 ms pause between chirps. Chirps were presented 128 times each in the sequential order shown in Figure 2D. Responses to these sounds were used to measure direction and speed selectivity, temporal precision, reliability, and information as described below.
- (5) *Speech*. One to three 2.5 s segments of female speech from the UCL SCRIBE database (<http://www.phon.ucl.ac.uk/resource/scribe>) at a peak intensity of 75 dB SPL. Each segment was presented between 256 and 1024 times in sequential order. Responses to these sounds were used to measure temporal precision, reliability, and information as described below. For decoding and analyses of

spike pattern similarity, responses to seven 0.25 s tokens of speech were extracted from the responses to each 2.5 ms segment. When separating trials in which the response to a token occurred during an ongoing up state from those in which the token triggered an up state, the responses to one token from each set were removed from the analysis because those tokens did not reliably evoke a response. For urethane experiments, 10 s of silence were inserted between every 16 trials of speech.

Spike sorting

The procedure for the isolation of single-unit spikes consisted of (1) bandpass filtering each channel between 500 and 5000 Hz, (2) whitening each tetrode, i.e., projecting the signals from the four channels into a space in which they are uncorrelated, (3) identifying potential spikes as snippets with energy (Choi et al., 2006) that exceeded a threshold (with a minimum of 0.7 ms between potential spikes), (4) projecting each of the snippets into the space defined by the first three principal components for each channel, (5) identifying clusters of snippets within this space using KlustaKwik (<http://klustakwik.sourceforge.net>) and Klusters (Hazan et al., 2006), and (6) quantifying the likelihood that each cluster represented a single unit using isolation distance (Schmitzer-Torbert et al., 2005). Isolation distance assumes that each cluster forms a multidimensional Gaussian cloud in feature space and measures, in terms of the SD of the original cluster, the increase in the size of the cluster required to double the number of snippets within it. The number of snippets in the “noise” cluster (nonisolated multiunit activity) for each tetrode was always at least as large as the number of spikes in any single-unit cluster. Only single-unit clusters with an isolation distance >20 were analyzed. The average number of single-units per tetrode was similar in recordings from synchronized (4.43) and desynchronized A1 (4.37).

Data analysis

Low-frequency LFP power. The low-frequency power in the LFP for each population was measured from spontaneous activity (sound 1 described above). For each tetrode on the array, the voltage signals were averaged across the four channels. For each of these tetrode signals, the power spectrum was computed using Welch’s averaged, modified periodogram method for 6 s segments with 50% overlap. The low-frequency power was measured as the sum of the power between 1 and 20 Hz. The values reported for each population are the average across the eight tetrodes on the array. The units associated with the reported values are arbitrary, but are the same for all populations.

Correlation between single-unit spiking and multiunit activity in spontaneous activity. The degree of concerted spiking in each population was measured from spontaneous activity (sound 1 described above) as the average value of the correlation between spiking of each cell and the MUA. The activity of each cell was represented as a spike-count vector with 50 ms bins. The MUA for each population was defined as the sum of the activity of all of the individual cells in the population. The correlation between the single-unit spiking and MUA in spontaneous activity was used to classify the cortical state as synchronized or desynchronized for urethane experiments: during periods when the value was <0.2 , the cortex was classified as desynchronized, and during periods when the value was >0.35 , the cortex was classified as synchronized.

Excess silence. The degree of concerted spiking in each population was measured from spontaneous activity (sound 1 described above) as excess silence, defined as the fraction of time during which the population was silent relative to that expected for a population of cells with the same mean rates that were spiking independently. For this analysis, the activity of each cell was represented as a spike count vector with 25 ms bins. The fraction of time bins in which there were no spikes across the entire population were compared before and after randomizing the spike times of each cell.

Tone responsiveness. Responses to tone set 2 (sound 3 described above) were evaluated in two ways: (1) the fraction of cells in each population that responded significantly (average spike rate >2 SDs above average spontaneous rate) to the best frequency for that population (i.e., the frequency that evoked a significant response from the largest fraction of cells), and (2) the fraction of cells that responded significantly to at least one of the frequencies tested.

Frequency tuning width. The width of the frequency tuning curve for each cell was measured from responses to tone set 1 (sound 2 described above) at 56 dB SPL as the range of frequencies for which the spike rate averaged over all trials was at least half of its maximum value. Spontaneous spike rates were not subtracted before measurement.

Direction selectivity. The direction selectivity index (DSI) for each cell was measured from responses to frequency-modulated (FM) tones (sound 4 described above). For each of the six FM speeds, the direction selectivity index was measured from the average spike rate of responses to the two directions as $(\text{higher rate} - \text{lower rate}) / (\text{higher rate} + \text{lower rate})$. The DSI reported for each cell is the highest of the values measured for the six speeds. Spontaneous spike rates were not subtracted before measurement.

Speed selectivity. The speed selectivity index (SSI) for each cell was measured from responses to FM tones (sound 4 described above). For each of the two FM directions, the speed selectivity index was measured from the average spike rate of responses to the six speeds as $(\text{highest rate} - \text{lowest rate}) / (\text{highest rate} + \text{lowest rate})$. The SSI reported for each cell is the higher of the two measured for the two directions. Spontaneous spike rates were not subtracted before measurement.

Temporal precision. The critical level of spike timing precision for each cell was measured from responses to speech (sound 5 described above) using a method that we have described previously (Garcia-Lazaro et al., 2013). The responses for each cell were represented as binary vectors with 2 ms bins and the single-spike information (Brenner et al., 2000) was measured as described below. The original spike times were then jittered by adding noise drawn from a uniform distribution and the information was recomputed. The critical level of precision was defined as the amount of jitter (i.e., the width of the noise distribution) that reduced the information in the responses to 95% of its original value.

Reliability. The reliability of responses across trials for each cell was measured from responses to speech (sound 5 described above) using a method that we have described previously (Sahani and Linden, 2003). To quantify reliability, we measured the signal-to-noise ratio (SNR) defined as the ratio of unbiased estimates of the signal (repeatable) and noise (not repeatable) response power with responses represented as binary vectors with 2 ms bins.

Information throughput and efficiency. The mutual information between the stimulus and the responses of each cell was measured from responses to speech (sound 5 described above). The mutual information between two variables measures how much the uncertainty about the value of one variable is reduced by knowing the value of the other. The mutual information between a sensory stimulus and a neural response can be computed as the difference between the entropy of the response before and after conditioning on the stimulus:

$$I(r; s) = H(r) - H(r|s) = - \sum_r p(r) \log_2 p(r) + \sum_s p(s) \sum_r p(r|s) \log_2 p(r|s)$$

To measure the information that is carried by spike trains about speech without having to specify which features of the speech were relevant, we used the approach pioneered by Strong et al. (1998) of discretizing a continuous stimulus into separate “stimuli” in time. To measure information, the total entropy of the response is compared with the average entropy of the response in each time bin (the noise entropy):

$$I(r; s) = H(r) - H(r|t) = - \sum_r p(r) \log_2 p(r) + \left\langle \sum_r p(r(t)) \log_2 p(r(t)) \right\rangle_t$$

We measured the single-spike information for each cell, which is equivalent to the information in the peristimulus time histogram (PSTH; Brenner et al., 2000), by representing responses as binary vectors with 2 ms bins and computing the information in single bin “words”. All infor-

mation calculations were performed using the Direct Method via info-Toolbox for MATLAB (Magri et al., 2009) with bias correction via the shuffling method and quadratic extrapolation (Panzeri et al., 2007). The stability of all calculations was verified by ensuring that the values obtained using only half of the recorded trials differed from those obtained using all trials by <5%.

Spike pattern similarity. The similarity of the spike patterns evoked by different speech tokens for each population was measured from responses to speech (sound 5 described above). From each 2.5 s segment of speech, responses to seven 0.25 s tokens were extracted. The responses of each population to each trial of each token were represented as binary matrices with rows corresponding to cells and columns corresponding to 10 ms time bins (see Fig. 5A). The similarity of trial-averaged spike patterns was measured as the average value of the correlation between the average responses across all pairs of tokens. The similarity of single-trial spike patterns was measured as the fractional increase in the average value of the Euclidean distance between the responses across all pairs of tokens relative to the average value of the Euclidean distance between spike patterns evoked by the same token.

The similarity of the spatial structure of the spike patterns was measured following the approach of Luczak et al. (2009). The spatial structure of spiking for each token was measured as the set of correlations between the responses of each pair of cells (i.e., the correlations between the rows of the binary response matrices). The similarity of the spatial structure across tokens was measured as the average value of the correlation between the set of pairwise correlations for all pairs of tokens.

The similarity of the temporal order of the spike patterns was measured following the approach of Luczak et al. (2009). The responses of each cell to each trial of each token were represented as binary vectors with 1 ms bins. The MUA for each population was defined as the sum of the activity of all of the individual cells in the population. The temporal order of spiking for each token was measured as the set of latencies obtained by taking the center of mass of the correlation function between each cell and the MUA (after smoothing with a Gaussian window with a width of 8 ms). The similarity of the temporal order was measured as the average value of the correlation between the sets of latencies for all pairs of tokens.

Signal and noise correlations. The signal and noise correlations between each pair of cells in each population were measured from responses to speech (sound 5 described above). The response of each cell to each trial was represented as a binary vector with 10 ms time bins. The total correlation for each pair of cells was obtained by computing the correlation coefficient between the actual responses. The signal correlation was computed after shuffling the order of repeated trials for each time bin. The noise correlation was obtained by subtracting the signal correlation from the total correlation.

Population decoding. A support vector machine was trained (using the LIBSVM package from <http://www.csie.ntu.edu.tw/~cjlin/libsvm> with default parameters) to decode the single-trial responses of each population to speech (sound 5 described above). From each 2.5 s segment of speech, responses to seven 0.25 s tokens were extracted. The responses of each population to each trial of each token were represented as binary matrices with rows corresponding to cells and columns corresponding to 10 ms time bins (see Fig. 5A). The classifier was trained on responses to 75% of trials and used to predict which token evoked the responses on other 25% of trials. The values reported for each population are the average performance obtained using 10 different subsets of trials for training and prediction. To test the effects of noise correlations on decoding, the order of repeated trials for each cell for each time bin were shuffled before training and prediction.

Classification of up and down states. To classify up and down states in synchronized A1, the MUA was computed as described above and represented as a spike count vector with 10 ms time bins. The MUA was filtered with a 10 bin median filter and the population was considered to be in an up state in any bin in which the filtered MUA was greater than zero.

Separation of trials in which the response to a speech token occurred during an ongoing up state from those in which the token triggered an up state. For responses to speech in synchronized A1, the MUA was com-

puted as described above and represented as a spike count vector with 5 ms time bins. The MUA was filtered with a 3 bin median filter and, for each token, the time of the first peak in the mean MUA across trials that was a least 75% as large as the maximum overall value was determined. Trials in which there was no activity within ± 25 ms of this peak were ignored. For the remaining trials, if there was any activity in the period from 75 to 25 ms before this peak, the response was classified as having occurred during an ongoing up state; otherwise, the response was classified as having triggered an up state.

Results

Synchronized and desynchronized states in A1

There are many aspects of neural activity that have been used to define cortical states. Recent studies comparing membrane potentials, single-unit spiking, MUA, and LFPs under different experimental conditions have demonstrated that cortical states are not discrete, but rather form a continuum with dynamics that are observable across different intracellular and extracellular properties (Harris and Thiele, 2011). At one end of this continuum are synchronized states in which spontaneous activity is dominated by slow fluctuations between up and down states that are concerted across a population. These fluctuations between up and down states are evident in intracellular measurements as transitions between depolarized and hyperpolarized membrane potentials, and in extracellular measurements as transitions between periods of vigorous population-wide spiking and silence, or strong low-frequency LFP fluctuations. At the other end of the continuum are desynchronized states in which the concerted fluctuations between up and down states are suppressed and neighboring cells spike independently.

To study population coding in synchronized and desynchronized cortical states, we compared activity recorded with a multielectrode array in gerbil A1 (Fig. 1A) under several different anesthetics. The cortical states imposed by anesthesia may, of course, differ from those that occur naturally. However, comparisons of spontaneous and evoked activity in rodent A1 have revealed similar dynamical properties in the synchronized and desynchronized states observed under anesthesia and those in awake animals (Luczak et al., 2007, 2013; Bermudez Contreras et al., 2013). Furthermore, the use of anesthesia enabled us to control synchronization and desynchronization without additional influences related to the particular task in which an animal is engaged, thus allowing us to perform a general comparison of A1 responses in the presence or absence of intrinsically generated up and down states.

To achieve a stable and consistent synchronized or desynchronized state throughout an entire experiment, we recorded activity under either KX or FMM. The up and down states that are typical of a synchronized cortical state were always evident in the populations recorded under KX, but were largely absent in those recorded under FMM. Short segments of the spontaneous LFP, single-unit spiking, and multiunit activity (defined as the summed spiking of all of the individual cells in the population) for two example populations are shown in Figure 1B. To assess the cortical state for each population, we measured the strength of up and down states based on: (1) the low-frequency power in the LFP, (2) the degree to which the spiking of individual cells was similar to the MUA, and (3) the excess silence in the population spiking, i.e., the fraction of time during which the population was silent relative to that expected for a population of cells with the same mean rates that were spiking independently.

The up and down state dynamics that are indicative of a synchronized cortical state were strong under KX and weak under FMM. As shown in Figure 1C, populations recorded under KX

($n = 7$) had more low-frequency LFP power and more strongly correlated spiking than those recorded under FMM ($n = 8$). Populations recorded under KX also exhibited more excess silence than those recorded under FMM, as shown in Figure 1D. For populations recorded under KX, the distribution of MUA spike rates during spontaneous activity changed dramatically after randomizing the spike times of each cell, indicating that periods of spiking and silence were concerted across the populations, whereas the same manipulation had a much weaker impact on the distributions for populations recorded under FMM (the distributions for two example populations before and after randomizing the spike times for each cell are shown in Fig. 1E). The suppression of up and down states under FMM was accompanied by an overall decrease in the level of spontaneous spiking (mean rates: 4.3 spikes/s for KX, $n = 284$ cells, 1.3 spikes/s for FMM, $n = 245$ cells).

The majority of our analysis (all figures but the last) is based on populations recorded in the low-frequency region of A1 under KX and FMM that exhibited stable synchronized and desynchronized states, respectively. These populations were well matched in their preferred frequencies; Figure 1F shows the distribution of center frequencies (CFs) for the MUA on each tetrode under KX ($n = 56$, 7 populations each with 8 tetrodes) and FMM ($n = 64$, 8 populations each with 8 tetrodes). To confirm that the state-dependent effects that we observed when comparing different populations were also evident when comparing synchronized and desynchronized states within the same population, we also recorded from three populations (131 cells in total) under urethane in which A1 exhibited spontaneous fluctuations between synchronized and desynchronized states (Curto et al., 2009; Marguet and Harris, 2011; Okun et al., 2012; Bermudez Contreras et al., 2013). Our analysis of these populations is summarized in the final figure.

The impact of cortical state on responses to pure tones

We began by examining A1 responses to tones. Although, on average, the spike rates evoked by tones were higher than spontaneous rates in both states (median increase: 0.68 spikes/s for synchronized, $n = 251$, 1.24 spikes/s for desynchronized, $n = 224$), the relative increase was much higher in the desynchronized state, as illustrated in the FRAs for two example populations shown in Figure 2A. For tones presented at 56 dB SPL, we measured the fraction of cells in each population that responded significantly above their spontaneous rate to the best frequency for that population (i.e., the frequency that evoked a significant response from the largest fraction of cells), as well as the fraction of cells that responded significantly to at least one of the frequencies tested. As shown in Figure 2B, only a small fraction of cells in synchronized A1 responded significantly above their spontaneous rate (median values: 13% for best tone, 18% for any tone, $n = 6$ populations), but in desynchronized A1, nearly all cells responded significantly in some populations (median values: 83% for best tone, 93% for any tone, $n = 8$ populations). These differences in population medians between synchronized and desynchronized A1, as well as all of the other differences in population medians between synchronized and desynchronized A1 reported in Figures 1 through 6, were significant with $p < 0.001$ (Wilcoxon rank-sum test).

It is possible that increased responsiveness in desynchronized A1 could be accompanied by a loss of selectivity, but this was not the case. As shown in Figure 2C, frequency selectivity (width of spike rate tuning at half max for tones at 56 dB SPL) was much sharper in desynchronized A1 (median value: 1 octave, $n = 224$

cells) than in synchronized A1 (median value: 2.4 octaves, $n = 251$ cells). There were also state-dependent differences in the temporal profiles of the responses to tones. As shown in Figure 2D, the MUA for populations in both synchronized and desynchronized A1 reached a peak ~ 30 ms after tone onset. However, whereas the MUA decreased gradually after this initial peak in synchronized A1, the MUA in desynchronized A1 reached a second peak with a latency of ~ 70 ms (note that this second peak does not correspond to an offset response, as it precedes the end of the tone). The second peak in the spike rates of desynchronized A1 populations was not caused by a subset of cells with long latencies; as shown in Figure 2E, the distributions of the onset latencies for all cells that responded significantly to tones in synchronized A1 ($n = 41$) and desynchronized A1 ($n = 223$) had a single dominant mode at ~ 30 ms (median values: 26 ms for synchronized, 33 ms for desynchronized).

The impact of cortical state on responses to frequency-modulated tones

For some populations, we also examined the effects of cortical state on responses to FM tones ($n = 3$ populations for a total of 108 cells in synchronized A1, $n = 5$ populations for a total of 175 cells in desynchronized A1). The responses of example cells from synchronized and desynchronized A1 to FM tones are shown in Figure 3A. We began by measuring the selectivity of each cell for the direction and speed of FMs. We quantified selectivity for direction (or speed) based on the maximum and minimum spike rates observed across all directions (or speeds) as (max rate – min rate)/(max rate + min rate). Cells in synchronized A1 were generally either nonresponsive or weakly selective (median selectivity index: 0.14 for direction, 0.36 for speed), whereas cells in desynchronized A1 were highly selective for both speed and direction (median selectivity index: 0.7 for direction, 0.91 for speed), as shown in Figure 3B.

We also assessed the fidelity of each cell's response to FMs by measuring the precision and reliability of spiking across repeated trials. We found that responses in synchronized A1 were highly variable, whereas responses in desynchronized A1 contained temporally precise firing events that were reliable across trials. To quantify the temporal precision of the responses, we measured the timescale at which spike timing needs to be considered to capture the information in single spikes (i.e., the information in the PSTH) from each cell. We defined the precision for each cell by jittering the spike times with successively larger amounts of noise until the information in the responses decreased to 95% of its original value (Garcia-Lazaro et al., 2013). As shown in Figure 3C, the median precision was 63 ms in synchronized A1 and 24 ms in desynchronized A1.

To quantify the reliability of the responses across trials, we measured the SNR defined as the ratio of unbiased estimates of the signal (repeatable) and noise (not repeatable) response power (Sahani and Linden, 2003), with responses represented as binary vectors with 2 ms bins. As shown in Figure 3D, cells in desynchronized A1 were, on average, nearly an order of magnitude more reliable than those in synchronized A1 (median SNR: 0.004 for synchronized, 0.03 for desynchronized). Finally, to quantify the overall fidelity of A1 responses in a manner that combines precision and reliability, we measured the throughput and the efficiency of the single-spike information for each cell. The information throughput (bits/s) in desynchronized A1 cells was ~ 2.5 times higher than that in synchronized A1 cells (median values: 1.3 bits/s for synchronized, 3.3 bits/s for desynchronized) and the information efficiency (bits/spike) in desynchronized

A1 cells was eight times higher than that in synchronized A1 cells (median values: 0.4 bits/spike for synchronized, 3.2 bits/spike for desynchronized), as shown in Figure 3E.

The impact of cortical state on the temporal precision and reliability of responses to speech

We next examined the fidelity of responses to speech in synchronized and desynchronized A1. The responses of two example cells from synchronized and desynchronized A1 to a short segment of speech are shown in Figure 4A. As with responses to FMs, we found that responses to speech in synchronized A1 were highly variable, while responses to speech in desynchronized A1 were precise and reliable. As shown in Figure 4B, the median precision in responses to speech was 31 ms in synchronized A1 ($n = 245$ cells) and 13 ms in desynchronized A1 ($n = 284$ cells). Note that, for both synchronized and desynchronized A1, these values are approximately half as large as those measured for responses to FMs. Responses to speech for cells in desynchronized A1 were, on average, six times more reliable than those in synchronized A1 (median SNR: 0.005 for synchronized, 0.029 for desynchronized), with the SNR of the most reliable cells in desynchronized A1 approaching values typically observed for responses to speech in sub-cortical areas (Horvath and Lesica, 2011), as shown in Figure 4C. There were also strong state dependencies in the throughput and the efficiency of the single-spike information in responses to speech: the information throughput (bits/s) in desynchronized A1 cells was three times higher than that in synchronized A1 cells (median values: 1.2 bits/s for synchronized, 3.8 bits/s for desynchronized) and the information efficiency (bits/spike) in desynchronized A1 cells was five times higher than that in synchronized A1 cells (median values: 0.5 bits/spike for synchronized, 2.6 bits/spike for desynchronized), as shown in Figure 4D.

The impact of cortical state on the similarity of spike patterns evoked by different speech tokens

The above results demonstrate that individual cells in desynchronized A1 respond reliably to repeated presentations of the same sound. However, the representation in A1 depends not only on the fidelity of individual cells, but also on the extent to which different sounds evoke different spike patterns across the population. Pre-

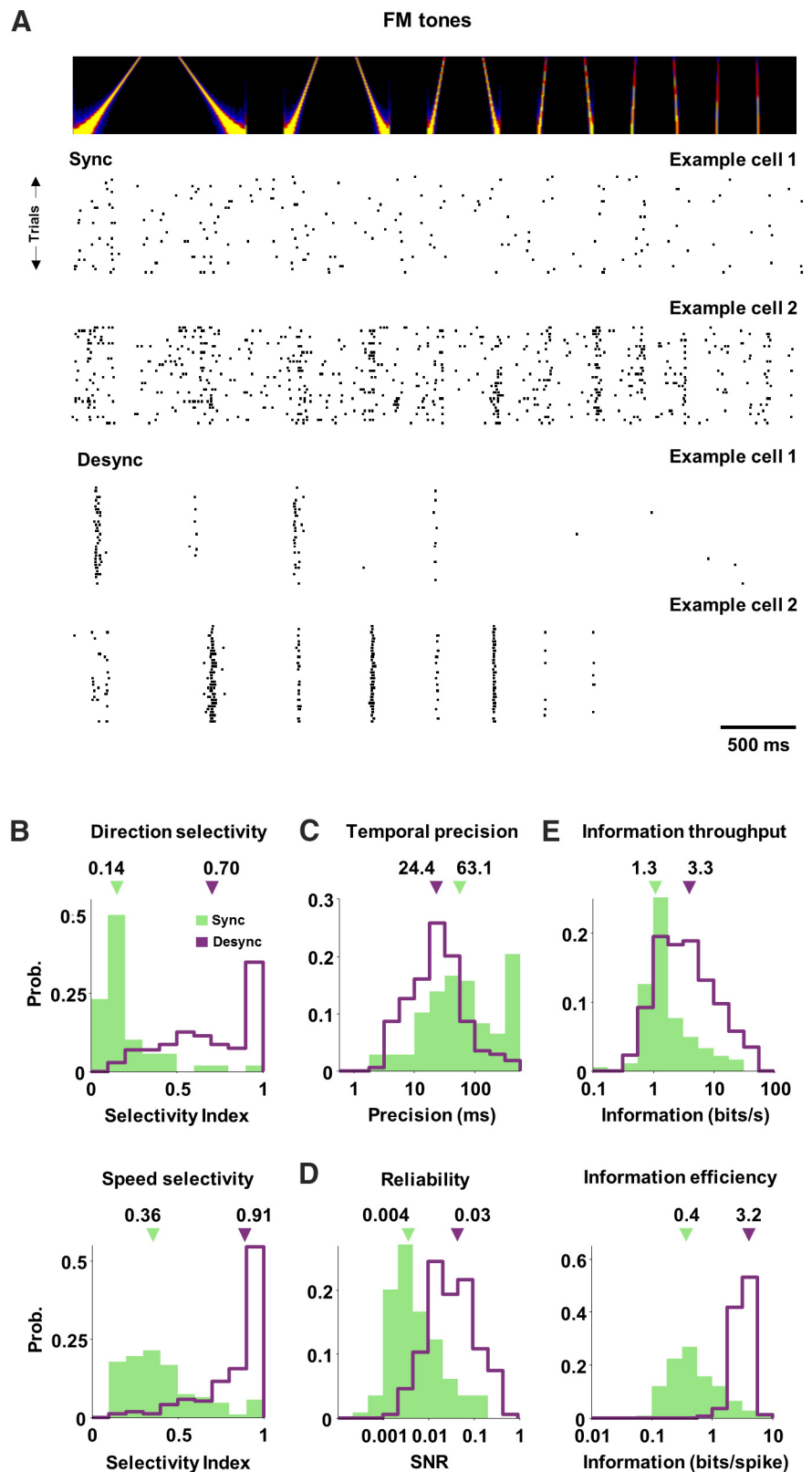


Figure 3. The impact of cortical state on responses to frequency-modulated tones. **A**, Responses of example cells from synchronized and desynchronized A1 to repeated presentations of FM tones. Top row, The spectrogram of the sounds; bottom rows, raster plots for individual cells. Each row in the raster plots shows the spike times for one trial. **B–E**, Distributions of the direction selectivity index, speed selectivity index, temporal precision, reliability, information throughput, and information efficiency of responses of individual cells in synchronized (green) and desynchronized (purple) A1 to FM tones, plotted as in Figure 2C.

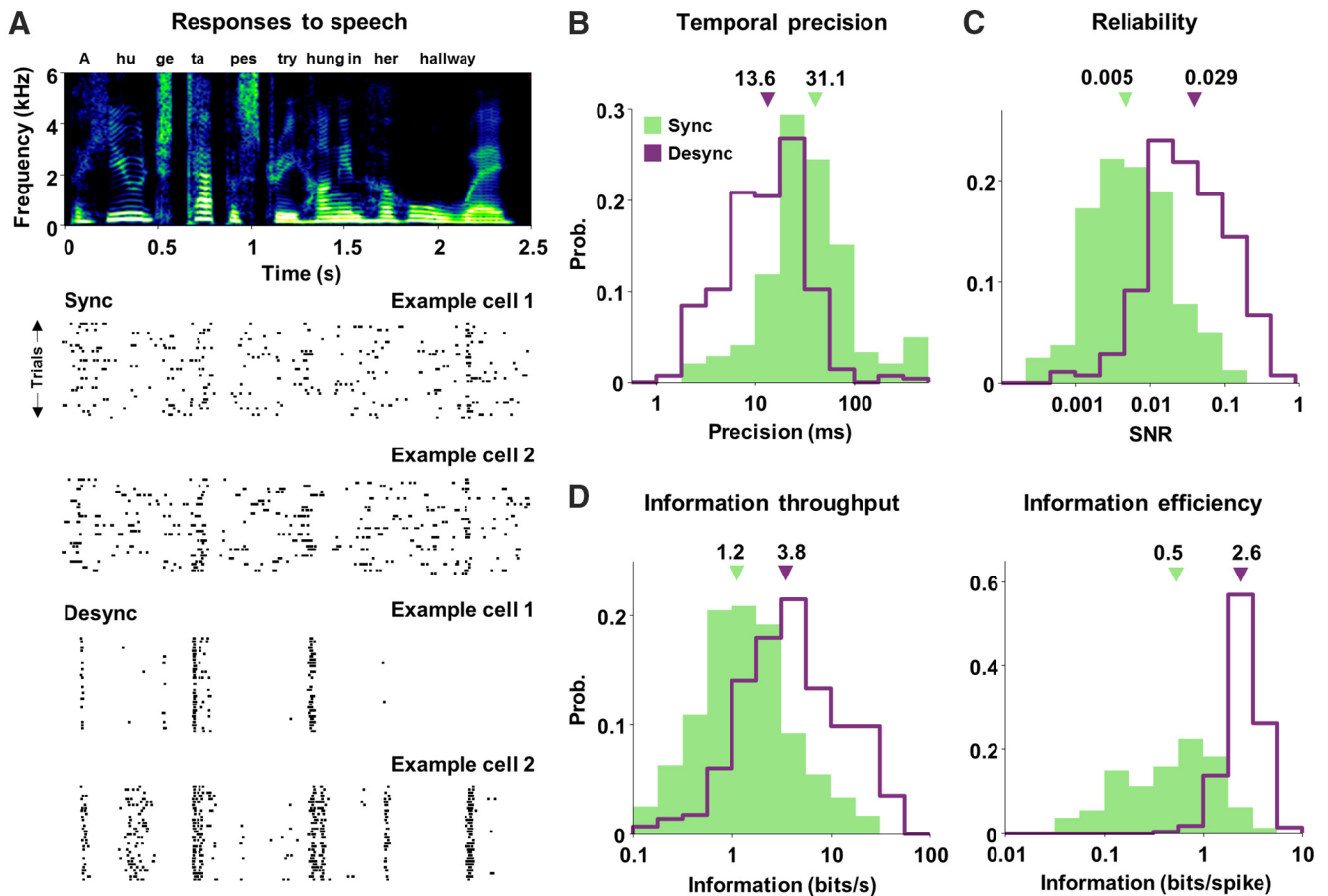


Figure 4. The impact of cortical state on the temporal precision and reliability of responses to speech. **A**, Responses of example cells from synchronized and desynchronized A1 to repeated presentations of speech. Top row, The spectrogram of the sound, the bottom rows show raster plots for individual cells. Each row in the raster plots shows the spike times for one trial. **B–D**, Distributions of the temporal precision, reliability, information throughput, and information efficiency of responses of individual cells in synchronized (green) and desynchronized (purple) A1 to speech, plotted as in Figure 2C.

vious studies in rodent A1 have shown that responses can be highly constrained, with different sounds evoking spike patterns that are remarkably similar (Luczak et al., 2009, 2013; Bathellier et al., 2012). We examined the similarity of responses evoked by different segments of speech and found that, although there was a high degree of similarity between responses in synchronized A1, responses in desynchronized A1 were much more diverse.

We represented population spike patterns as binary matrices (Fig. 5A) and measured the average similarity between both the single-trial and trial-averaged patterns evoked by different speech tokens. The spike patterns in synchronized A1 were much more similar across tokens than those in desynchronized A1, both for the average patterns evoked by each token across trials and for the patterns evoked on single trials. As shown in Figure 5B, the median correlation between average patterns for each pair of tokens was 0.51 for synchronized A1 (7 populations each with between 1 and 3 sets of 7 different tokens for total $n = 12$) and 0.19 for desynchronized A1 (8 populations for total $n = 14$). This result indicates a qualitative difference between synchronized and desynchronized A1: if the intrinsic dynamics in synchronized A1 simply added noise to the responses observed in desynchronized A1, the similarity between the trial-averaged patterns in the two states would be the same. The difference between synchronized and desynchronized A1 was also evident when comparing the spike patterns evoked on single trials. As shown in Figure 5B, the median fractional increase in the average distance between

single-trial patterns for each pair of tokens relative to the average distance between patterns for the same token was 4% for synchronized A1 and 20% for desynchronized A1 (note that although the distances may seem small even for desynchronized A1, they are sufficient to support nearly perfect classification in the high dimensional response space, as shown below).

To examine the similarity of spike patterns in more detail, we followed the approaches of previous studies for comparing patterns based on their spatial and temporal structure (Luczak et al., 2009, 2013). We represented the spatial structure of spiking for each token by the set of correlations between the spike patterns of each pair of cells in the population (i.e., the correlations between the rows of the binary spike pattern matrices). Figure 5C shows the set of pairwise correlations for two example populations for two different speech tokens (each square in each image shows the correlation between one pair of cells for a given token). In synchronized A1, the spatial structure of spiking was largely preserved across tokens, while in desynchronized A1, the spatial structure varied from token to token. To quantify the degree to which the spatial structure of spiking for each population was similar across tokens, we measured the correlation between the spatial structures for each pair of tokens and averaged across all pairs of tokens. As shown in Figure 5D, the spatial structure of spiking in synchronized A1 was twice as similar across tokens as that in desynchronized A1 (median values: 0.83 for synchronized, 0.42 for desynchronized).

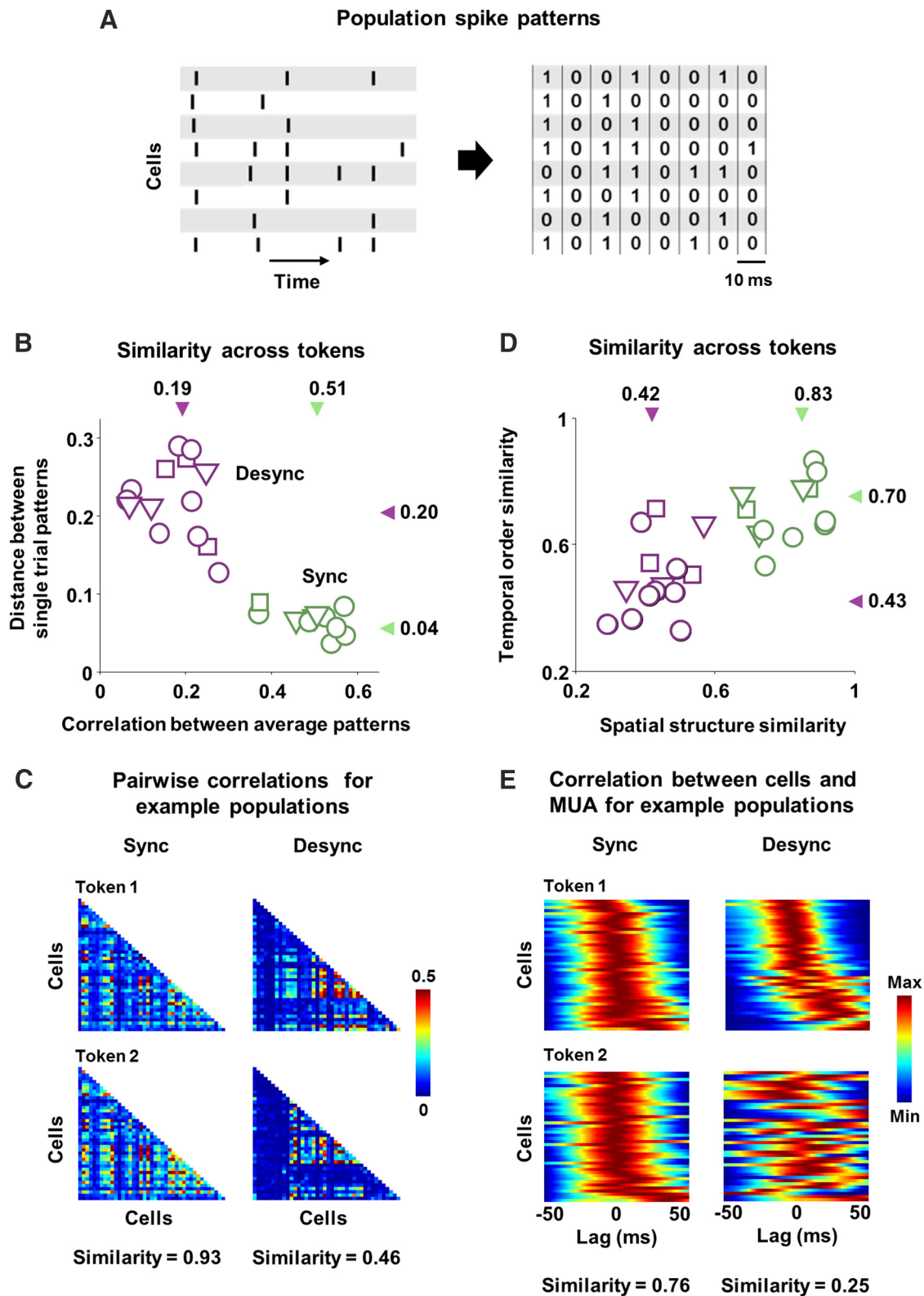


Figure 5. The impact of cortical state on the similarity of spike patterns evoked by different speech tokens. **A**, The responses of each population to each trial of each token were represented as binary matrices with rows corresponding to cells and columns corresponding to 10 ms time bins. **B**, A scatter plot showing the similarity of the spike patterns across speech tokens for each synchronized (green) and desynchronized (purple) population for both single-trial responses and responses averaged across trials. For trial average similarity, values are the average correlation between the average spike patterns evoked by each pair of tokens. For single-trial similarity, values are the average fractional increase in the distance between spike patterns evoked by each pair of tokens relative to the average distance between patterns evoked by the same token. For those populations for which responses were recorded for more than one set of tokens, multiple symbols are shown (circles for token set 1, squares for token set 2, and triangles for token set 3). The median values (with each token set for each population treated as a separate measurement) are indicated by the arrows. **C**, The pairwise correlations for the responses of example synchronized and desynchronized A1 populations to different speech tokens. Each square in each image shows the correlation for one pair of cells. The images in the top row show the correlations for the first token and the images in the bottom row show the correlations for the second token. (Figure legend continues.)

We also examined the degree to which the temporal order of spiking for each population was similar across tokens. We represented the temporal order of spiking for each token by the set of latencies measured from the center of mass of the correlation function between the spiking of each cell in the population and the MUA (i.e., the correlation function between each row of the binary spike pattern matrices and the sum of all rows). Figure 5E shows the set of correlation functions for two example populations for two different speech tokens (each row in each image shows the correlation function between one cell and the MUA). In synchronized A1, the temporal order of spiking was largely preserved across tokens, while in desynchronized A1, the temporal order varied from token to token (for the images in Fig. 5E, the cells in each population were ordered according to their latency for the first token and plotted in the same order for the second token). To quantify the degree to which the temporal order of spiking for each population was similar across tokens, we measured the correlation between the latencies for each pair of tokens and averaged across all pairs of tokens. As shown in Figure 5D, the temporal order of spiking was much more similar across tokens in synchronized A1 than in desynchronized A1 (median values: 0.7 for synchronized, 0.43 for desynchronized).

The impact of cortical state on signal correlations, noise correlations, and population decoding

The above results demonstrate that the degree of similarity in the spike patterns evoked by different sounds differs strongly between synchronized and desynchronized A1. However, the extent to which A1 can support discrimination of different sounds depends not only on the range of evoked patterns, but also on the structure of the trial-to-trial variability in these patterns across the population. For each population, we separated the correlations in responses to speech into signal correlations, the correlations in the fraction of the response that was repeatable across trials, and noise correlations, the correlations in the trial-to-trial variability. Figure 6A shows the distributions of pairwise correlations in responses to speech for each population. Although there was a significant difference in the signal correlations in synchronized and desynchronized A1 (median values: 0.012 for synchronized, $n = 6451$ pairs, and 0.017 for desynchronized, $n = 6101$ pairs), the dependence of noise correlations on cortical state was much more striking; whereas noise correlations in synchronized A1 were strong (median value: 0.07), those in desynchronized A1 were extremely weak (median value: 0.002). These results were consistent across a wide range of time scales (Fig. 6B). As shown in Figure 6C, there was also a positive dependency between signal and noise correlations in both states (though this relationship was much stronger in synchronized A1), indicating that cells that

preferred similar acoustic features also tended to have a higher degree of shared variability.

To quantify how the differences between spike patterns in synchronized and desynchronized A1 impact the representation of speech, we trained a support vector machine to predict which speech token evoked a given single-trial response. As shown in Figure 6D, decoding of population spike patterns from desynchronized A1 was highly accurate (median performance: 99% correct), while decoding of patterns from synchronized A1 was substantially worse (median performance: 62% correct). Decoding of synchronized A1 responses was also impacted by noise correlations; when noise correlations were removed by shuffling the trial order before training the classifier and decoding, median performance increased from 62% correct to 82% correct ($p < 0.001$, Wilcoxon signed rank test).

Spike patterns evoked by different speech tokens in synchronized A1 are similar and have strong noise correlations even within up states

It has been hypothesized that up states in synchronized cortex may be equivalent to brief periods of desynchronization (Destexhe et al., 2007; Castro-Alamancos, 2009). This implies that the differences in the spike patterns in synchronized and desynchronized A1 that we have observed can be accounted for by the global dynamics of up and down states in synchronized A1, and that if only the activity within up states is considered, the differences between synchronized and desynchronized A1 should be small. We found, however, that restricting the analysis of synchronized A1 to activity within up states had little impact on our results.

Figure 7A shows the probability of being in an up state for an example population from synchronized A1 during repeated presentations of a short segment of speech. The timing of up and down states in this population was strongly modulated by the sound, and this effect was consistent across all of the populations that we studied in synchronized A1; the reliability of the timing of up and down states across trials measured as the SNR for binary vectors specifying whether the population was in an up or down state in 10 ms time bins was 0.17 ± 0.09 (7 populations each between 1 and 3 different speech segments for total $n = 12$). Figure 7B shows the MUA for an example population from synchronized A1 across repeated presentations of two different speech tokens. Each row of the image shows the MUA for one trial, and the trials are ordered by the time of the earliest activity. There were very few trials in which the tokens evoked no response (median value: 4% of trials across 7 populations each with between 12 and 18 different tokens for total $n = 96$). In most trials, either the response to the onset of the token occurred during an ongoing up state (median value: 43% of trials) or the onset of the token triggered an up state (median value: 50% of trials).

We repeated the analyses of population spike patterns described in the previous section after separating trials in which the response to a token occurred during an ongoing up state from those in which the token triggered an up state (see Materials and Methods for a description of how trials were classified). Whether considering the similarity in the spike patterns evoked by different sounds (Fig. 7C), noise correlations (Fig. 7D), or decoding performance (Fig. 7E), the differences between different classes of responses in synchronized A1 were small, and the differences between synchronized and desynchronized A1 were large. Surprisingly, although the differences between the different classes of responses in synchronized A1 were small, the responses on trials in which an up state was triggered were more like desynchronized

←

(Figure legend continued.) The similarity of the correlations for token 1 and token 2 are shown. Similarity was measured as the correlation between the set of pairwise correlations for each token. **D**, A scatter plot showing the similarity of the spatial pattern and temporal order of spiking across speech tokens for each synchronized (green) and desynchronized (purple) population, plotted as in **B**. **E**, The correlation function between the spiking of individual cells and the multiunit activity for the responses of example synchronized and desynchronized A1 populations to different speech tokens. Each row in each image shows the correlation function for one cell. For plotting, the correlation functions for all cells were scaled to have the same maximum and minimum values, and the cells were ordered according to their latency with respect to the MUA for the first token. The latency was measured as the center of mass of the correlation function. The ordering of the images was the same for the first and second tokens. The similarity of the latencies for token 1 and token 2 is shown. Similarity was measured as the correlation between the set of latencies for each token.

responses (i.e., had more diverse spike patterns, weaker noise correlations, and allowed for better decoding performance) than those that occurred during ongoing up states (see Fig. 7 for population medians and significance for Wilcoxon signed rank tests).

Differences between synchronized and desynchronized states in the same population

All of the above results are based on comparing synchronized and desynchronized states in different populations. To confirm that the same state-dependent effects on population coding were also evident when comparing synchronized and desynchronized states within the same population, we recorded from three populations under urethane in which A1 exhibited spontaneous fluctuations between synchronized and desynchronized states (Curto et al., 2009; Marguet and Harris, 2011; Okun et al., 2012; Bermudez Contreras et al., 2013). Figure 8A shows the spontaneous LFP and MUA for an example population over a period of ~1 h, along with the responses to speech for an example cell recorded during the same period (each 10 s period of silence for measurement of spontaneous activity was followed by 40 s of speech). Measuring the strength of up and down states in populations recorded under urethane based on the same measures used to assess the cortical states observed under KX and FMM in Figure 1 (low-frequency LFP power, correlation between the spiking of individual cells and the MUA, and excess silence) revealed clear transitions between synchronized and desynchronized states. As illustrated in Figure 8B for the same example population, relatively strong low-frequency LFPs were accompanied by highly correlated spiking and a large degree of excess silence, whereas periods of relatively weak low-frequency LFPs were accompanied by weakly correlated spiking and less excess silence.

For all populations recorded under urethane, we classified the cortical state based on the average correlation between the spiking of individual cells and the MUA in spontaneous activity; periods during which this value was >0.35 were classified as synchronized, whereas periods during which this value was <0.2 were classified as desynchronized. As shown in Figures 8C,D, the state-dependent differences that we observed in the LFP, MUA, and single-unit spiking properties under urethane were consistent across the three populations that we studied. As with the

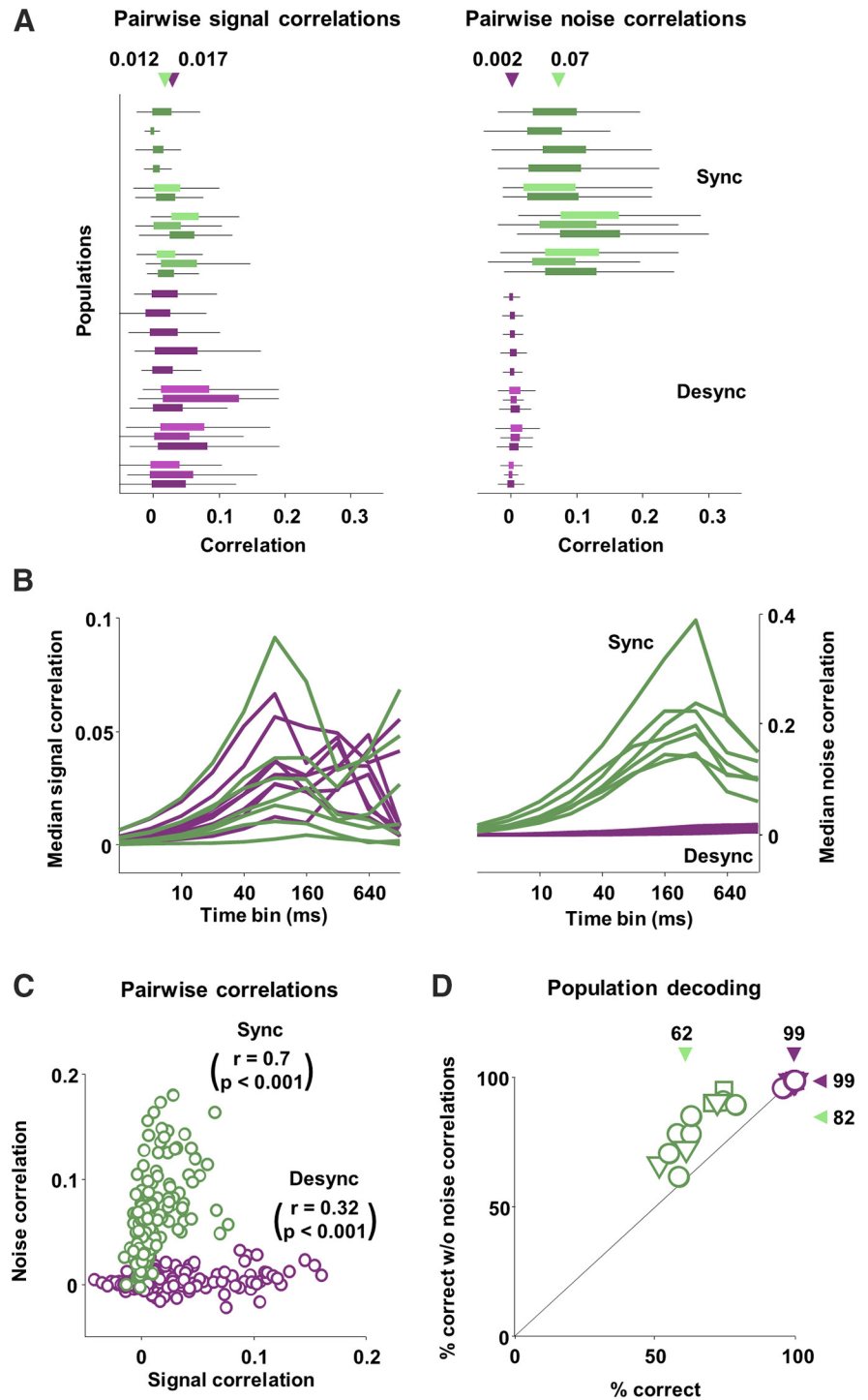


Figure 6. The impact of cortical state on signal correlations, noise correlations and population decoding. **A**, Box-and-whisker plots showing the distributions of pairwise signal and noise correlations for each synchronized (green) and desynchronized (purple) population. The box spans the 25th to 75th percentiles, and the whiskers span the 5th to 95th percentiles. For those populations for which responses were recorded for more than one segment of speech, multiple distributions are shown (darkest color for token set 1, middle color for token set 2, and lightest color for token set 3). The median values for all pairs across all populations are indicated by the arrows. **B**, The median pairwise signal and noise correlations in responses to speech across different time scales for each synchronized (green) and desynchronized (purple) population. The response of each cell to each trial was represented as a binary vector with a range of time bins as indicated on the horizontal axis. **C**, A scatter plot showing the signal and noise correlations for all pairs of cells that we studied in synchronized (green) and desynchronized (purple) A1. The correlation coefficients between the signal and noise correlations in each state are indicated. **D**, A scatter plot showing the performance of a support vector machine in decoding the responses of each synchronized (green) and desynchronized (purple) population to different speech tokens with and without noise correlations, plotted as in Figure 5B.

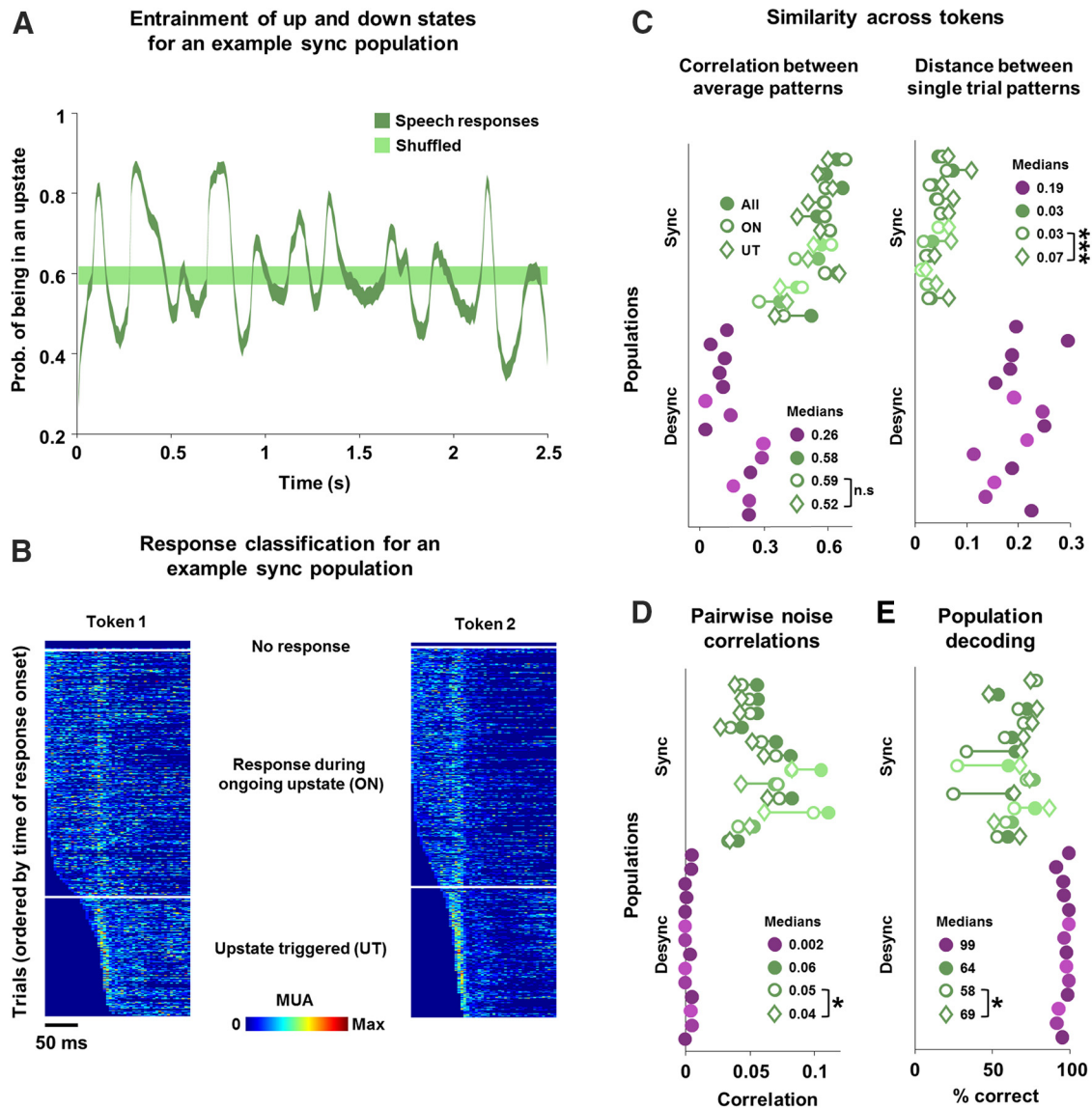


Figure 7. Spike patterns evoked by different speech tokens in synchronized A1 are similar and have strong noise correlations even within up states. **A**, The probability of being in an up state in 10 ms time bins for an example population from synchronized A1 during repeated presentations of a short segment of speech. The dark line shows the probabilities for the actual responses. The light line shows the probability computed after shuffling the order of time bins on each trial and indicates the overall probability of being in an up state. The thickness of lines indicates 95% confidence intervals. **B**, The MUA for an example population from synchronized A1 across repeated presentations of two different speech tokens. Each row of the image shows the MUA for one trial, and the trials are ordered by the time of the earliest activity. Trials were separated into those with no response, those in which the response occurred during an ongoing up state, and those in which the token triggered an up state. **C**, Plots showing the similarity of the spike patterns across speech tokens for each synchronized (green) and desynchronized (purple) population for both responses averaged across trials (left) and single-trial responses (right). For trial average similarity, values are the average correlation between the average spike patterns evoked by each pair of tokens. For single-trial similarity, values are the average fractional increase in the distance between spike patterns evoked by each pair of tokens relative to the average distance between patterns evoked by the same token. For those populations for which responses were recorded for more than one set of tokens, multiple symbols are shown (darkest color for token set 1, middle color for token set 2, and lightest color for token set 3). The symbols indicate the median value for each population across all pairs of tokens. The median values across all populations are noted on the figure (with each token set for each population treated as a separate measurement). Responses from synchronized A1 were analyzed for all trials (All), trials in which the response occurred during an ongoing up state (ON), and those in which the token triggered an up state (UT). The *p* value for a Wilcoxon signed rank test comparing the ON and UT medians is indicated. **D, E**, Plots showing the pairwise noise correlations and the performance of a support vector machine in decoding responses, plotted as in **C**.

populations recorded under KX and FMM, there were also state-dependent differences in spontaneous spike rates under urethane; however, whereas spike rates in desynchronized states under FMM were lower than those in synchronized states under KX (Fig. 1D), spike rates in desynchronized states were higher than those in synchronized states under urethane (Fig. 8D).

To examine the effects of cortical state on the population coding of speech under urethane, we grouped responses to speech according to whether the cortex was classified as synchronized or

desynchronized during the preceding period of spontaneous activity and repeated the analyses of population spike patterns described above. With respect to the similarity in the spike patterns evoked by different sounds (Fig. 8E), noise correlations (Fig. 8F), and decoding performance (Fig. 8G), the differences between responses in synchronized and desynchronized states for the three populations recorded under urethane mirrored those that we observed when comparing states across different populations under KX and FMM above.

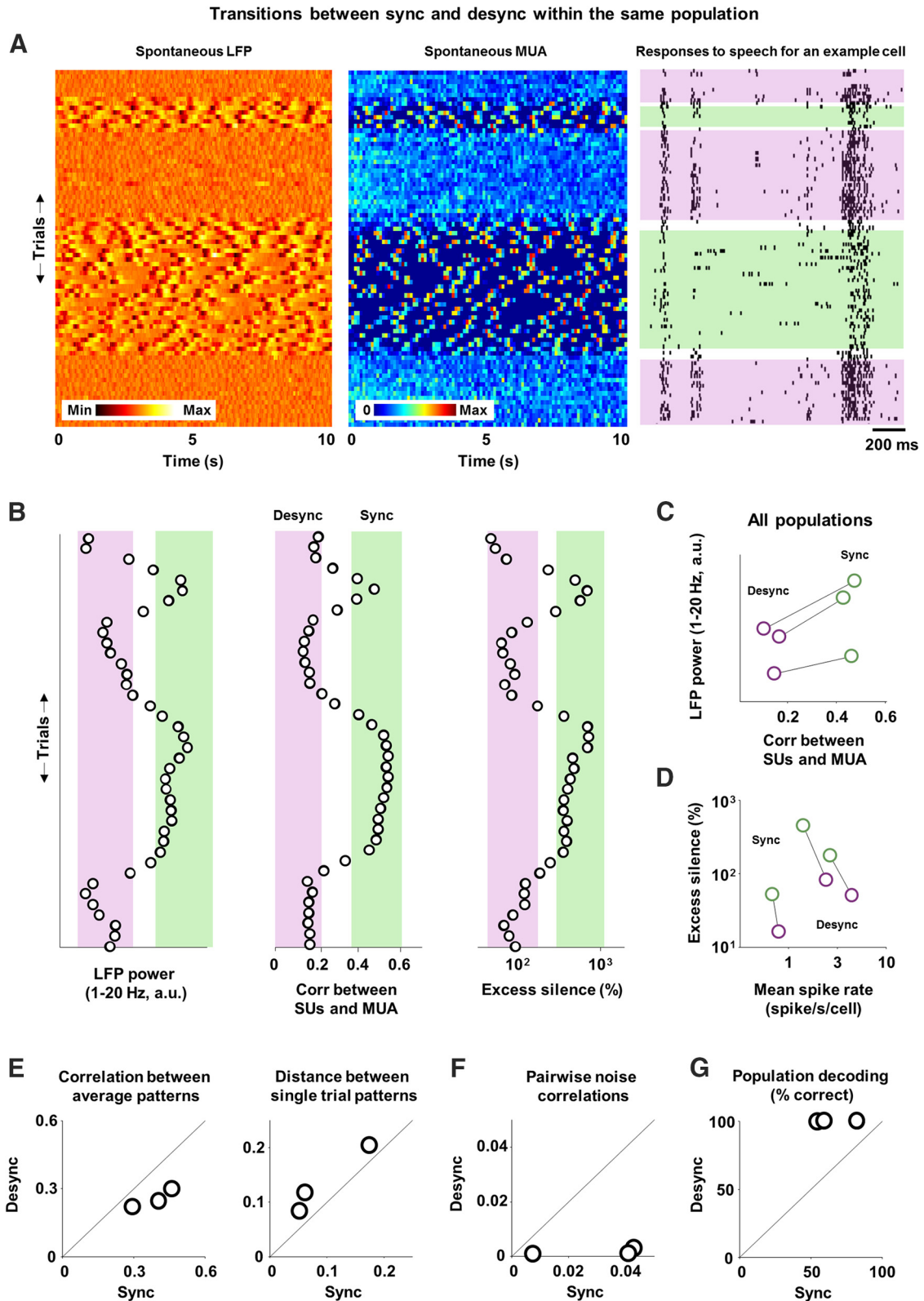


Figure 8. Differences between synchronized and desynchronized states in the same population. **A**, Left and middle, The spontaneous LFP and MUA for an example A1 population under urethane over a period of ~1 h. Each 10 s period of silence for measuring spontaneous activity was followed by 40 s of speech. The MUA was defined as the sum of the activity of all of the individual cells in the population. Right, The responses of an example cell to repeated presentations of speech in synchronized and desynchronized states. Each row in the raster plots shows the spike times for one trial. The periods during which the cortex was classified as synchronized and desynchronized are indicated by the shading. **B**, Middle, The median value of correlation between the spiking of each cell in the population and the MUA for each 10 s trial of spontaneous activity shown in **A**. The activity of each cell was represented as a spike count vector with 50 ms bins. During periods when the value was <0.2, the cortex was classified as desynchronized, and during periods when the value was >0.35, the cortex was classified as synchronized. Left and right, The low-frequency LFP power (1–20 Hz) and the excess silence in the spontaneous activity during the same period. **C**, A scatter plot showing the low-frequency LFP power (1–20 Hz) and average correlation between the MUA and spiking of each single unit for three populations recorded under urethane during periods in which the (*Figure legend continues.*)

Discussion

We have shown that responses to tones and speech in A1 depend strongly on cortical state. We found that responses to FM tones and speech in desynchronized A1 were temporally precise and reliable across trials, with median precision that was several times higher than in synchronized A1. Whereas different speech tokens evoked similar spike patterns in synchronized A1, we found that responses in desynchronized A1 were much more diverse, with similarity in both the spatial structure and the temporal order of spiking across tokens that was approximately half that in synchronized A1. This diversity of spike patterns, together with extremely weak noise correlations, allowed us to decode responses to different speech tokens from desynchronized A1 with nearly perfect performance. These state-dependent differences in the population coding of speech were evident in comparisons both across different populations, as well between synchronized and desynchronized states within the same populations.

Our finding that gerbil A1 has the capacity to represent sounds with high fidelity in the desynchronized state is consistent with behavioral studies in rodents that have demonstrated the essential role of A1 in auditory processing (Wetzel et al., 1998; Rybalco et al., 2006; Cooke et al., 2007; Porter et al., 2011) and learning (Bao et al., 2004; Reed et al., 2011; Aizenberg and Geffen, 2013; Banerjee and Liu, 2013). Several previous studies of synchronized and desynchronized rodent A1 have reported differences that are qualitatively consistent with our results. In rats anesthetized with urethane, the change from synchronized to desynchronized states was accompanied by a decrease in the trial-to-trial variability of A1 responses to clicks (Curto et al., 2009) and amplitude-modulated noise (Marguet and Harris, 2011), as well as a decrease in noise correlations (Renart et al., 2010). A study in awake rats found that the temporal order of population spiking was conserved across synchronized and desynchronized states (Luczak et al., 2013), which may seem inconsistent with our finding that the temporal order of spiking was similar across different sounds in synchronized A1, but not in desynchronized A1. However, the comparison by Luczak et al. (2013) was based on the average temporal order across all sounds tested in the two states, rather than on the order for individual sounds as in our study. We also observed a consistent temporal order in desynchronized A1 when averaging across two separate 5 min segments on ongoing speech (data not shown), but our results show that the intrinsic factors that impose this consistency across sounds provide only a weak constraint on the temporal order in responses to any particular sound.

Another study in awake rats found that A1 responses in engaged animals were suppressed relative to those in passive animals (Otazu et al., 2009). Although this study did not explicitly measure cortical state, the results of previous studies suggest that

engaged and passive behavioral conditions in rodents are typically associated with desynchronized and synchronized states, respectively (Harris and Thiele, 2011). Our data are consistent with the results of Otazu et al. (2009); the average spike rates in responses to speech were lower in desynchronized A1 than in synchronized A1 (median values: 5.2 spikes/s for synchronized, $n = 245$, 3.3 spikes/s for desynchronized, $n = 284$).

Our results differ from those of previous studies with respect to differences between activity in desynchronized cortex and activity during up states in synchronized cortex. Previous studies have shown that membrane potential dynamics during up states in anesthetized animals are similar to those during prolonged periods of desynchronization in awake animals, suggesting that up states may be equivalent to brief periods of desynchronization (Destexhe et al., 2007; Castro-Alamancos, 2009). Our results argue against this hypothesis, at least at the level of population spike patterns, as restricting our analysis of synchronized A1 to activity within up states had little impact on our results. Our finding that noise correlations in synchronized A1 persist even when only up states are considered also differ from those of recent studies that have shown that noise correlations within up states in synchronized cortex are weak (Renart et al., 2010).

Such discrepancies suggest that there may be important differences in the synchronized and desynchronized states observed under different anesthetics and in different behavioral states. Although many aspects of the synchronized and desynchronized states that we observed under different anesthetics were similar, there were also noticeable differences. For example, whereas spontaneous spike rates in desynchronized states under FMM were much lower than those in synchronized states under KX, the opposite was true under urethane, where spike rates in desynchronized states were higher than those in synchronized states for 82% of cells. Understanding the relationships between the different synchronized and desynchronized states that have been observed in studies of behaving animals is also difficult. For example, although studies across different sensory modalities are generally consistent in suggesting that transitions from passive to active behavioral states are associated with a suppression of up and down states, this desynchronization can be accompanied by either an increase or decrease in overall activity depending on context (Castro-Alamancos, 2004b; Niell and Stryker, 2010; Schneider et al., 2014; Zhou et al., 2014). One recent study linking different forms of desynchronization to the action of different neuromodulator pathways may provide a potential explanation for these results: Castro-Alamancos and Gulati (2014) found that cholinergic stimulation in S1 produced desynchronization with increased activity, whereas noradrenergic stimulation produced desynchronization with decreased activity. However, even the notion that active behavioral states are associated with desynchronized cortical states requires further refinement; recent studies in S1 and V1 (Sachidhanandam et al., 2013; Tan et al., 2014) found that the cortex could be in a synchronized state even when animals were performing a task (and, in the case of V1, switched to a desynchronized state only after the onset of visual stimulation). To determine how cortical states observed under different anesthetics and in different behavioral states are related, further studies involving direct comparisons of population activity under different conditions are needed.

Our results add to a growing body of evidence demonstrating the importance of cortical state for sensory processing (Harris and Thiele, 2011). Early evidence suggested that the interactions between spontaneous and evoked activity were additive (Arieli et al., 1996; Azouz and Gray, 2003; Ringach, 2009), but recent stud-

←

(Figure legend continued.) cortex was classified as synchronized (green) or desynchronized (purple). The corresponding values for each population are connected by lines. **D**, A scatter plot showing the excess silence and average mean spike rates for spontaneous activity for three populations recorded under urethane, plotted as in **C**. **E**, Plots showing the similarity of the spike patterns across speech tokens for each of the three populations recorded under urethane in synchronized and desynchronized states. For trial average similarity, values are the average correlation between the average spike patterns evoked by each pair of tokens. For single-trial similarity, values are the average fractional increase in the distance between spike patterns evoked by each pair of tokens relative to the average distance between patterns evoked by the same token. The median values for each population across all pairs of tokens are shown. **F**, **G**, Plots showing the pairwise noise correlations and the performance of a support vector machine in decoding responses, plotted as in **E**.

ies have shown that these interactions can be much more complex, with sensory inputs causing transitions between up and down states and intrinsic dynamics placing strong constraints on activity patterns (MacLean et al., 2005; Hasenstaub et al., 2007; Rigas and Castro-Alamancos, 2007; Curto et al., 2009; Luczak et al., 2009, 2013; Bathellier et al., 2012). The ability of stimuli to trigger an up state may facilitate the detection of stimulus onsets; indeed, in our sample of populations in synchronized A1, trials in which the onset of a speech token triggered an up state contained an average of 18% more spikes than those in which responses occurred during an ongoing upstate ($p < 0.001$, Wilcoxon signed rank test). Recent studies have also provided evidence that network dynamics can aid in the processing of ongoing stimuli. For example, the entrainment of slow rhythms in A1 has been shown to facilitate the processing of complex sound streams (Kayser et al., 2009; Giraud and Poeppel, 2012; Lakatos et al., 2013; Zion Golumbic et al., 2013) and our finding that the dynamics of up and down states can be entrained by speech are consistent with these results. Thus, rather than simply reflecting a general suppression of network dynamics, the high fidelity representation of sounds that we observed in desynchronized A1 may result from network dynamics being strongly driven by sound rather than by intrinsic sources. Elucidating the role of network dynamics in desynchronized cortex and characterizing how they interact with sensory inputs are challenges for future studies.

References

- Aizenberg M, Geffen MN (2013) Bidirectional effects of aversive learning on perceptual acuity are mediated by the sensory cortex. *Nat Neurosci* 16:994–996. [CrossRef Medline](#)
- Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* 273:1868–1871. [CrossRef Medline](#)
- Azouz R, Gray CM (2003) Adaptive coincidence detection and dynamic gain control in visual cortical neurons in vivo. *Neuron* 37:513–523. [CrossRef Medline](#)
- Banerjee SB, Liu RC (2013) Storing maternal memories: hypothesizing an interaction of experience and estrogen on sensory cortical plasticity to learn infant cues. *Front Neuroendocrinol* 34:300–314. [CrossRef Medline](#)
- Bao S, Chang EF, Woods J, Merzenich MM (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat Neurosci* 7:974–981. [CrossRef Medline](#)
- Bathellier B, Ushakova L, Rumpel S (2012) Discrete neocortical dynamics predict behavioral categorization of sounds. *Neuron* 76:435–449. [CrossRef Medline](#)
- Bermudez Contreras EJ, Schjetnan AG, Muhammad A, Bartho P, McNaughton BL, Kolb B, Gruber AJ, Luczak A (2013) Formation and reverberation of sequential neural activity patterns evoked by sensory stimulation are enhanced during cortical desynchronization. *Neuron* 79:555–566. [CrossRef Medline](#)
- Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck RR (2000) Synergy in a neural code. *Neural Comput* 12:1531–1552. [CrossRef Medline](#)
- Castro-Alamancos MA (2004a) Dynamics of sensory thalamocortical synaptic networks during information processing states. *Prog Neurobiol* 74:213–247. [CrossRef Medline](#)
- Castro-Alamancos MA (2004b) Absence of rapid sensory adaptation in neocortex during information processing states. *Neuron* 41:455–464. [CrossRef Medline](#)
- Castro-Alamancos MA (2009) Cortical up and activated states: implications for sensory information processing. *Neurosci*:625–634.
- Castro-Alamancos MA, Gulati T (2014) Neuromodulators produce distinct activated states in neocortex. *J Neurosci* 34:12353–12367. [CrossRef Medline](#)
- Choi JH, Jung HK, Kim T (2006) A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios. *IEEE Trans Biomed Eng* 53:738–746. [CrossRef Medline](#)
- Cooke JE, Zhang H, Kelly JB (2007) Detection of sinusoidal amplitude modulated sounds: deficits after bilateral lesions of auditory cortex in the rat. *Hear Res* 231:90–99. [CrossRef Medline](#)
- Crochet S, Petersen CC (2006) Correlating whisker behavior with membrane potential in barrel cortex of awake mice. *Nat Neurosci* 9:608–610. [CrossRef Medline](#)
- Curto C, Sakata S, Marguet S, Itskov V, Harris KD (2009) A simple model of cortical dynamics explains variability and state dependence of sensory responses in urethane-anesthetized auditory cortex. *J Neurosci* 29:10600–10612. [CrossRef Medline](#)
- Destexhe A, Hughes SW, Rudolph M, Crunelli V (2007) Are corticothalamic “up” states fragments of wakefulness? *Trends Neurosci* 30:334–342. [CrossRef Medline](#)
- Garcia-Lazaro JA, Belliveau LA, Lesica NA (2013) Independent population coding of speech with sub-millisecond precision. *J Neurosci* 33:19362–19372. [CrossRef Medline](#)
- Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci* 15:511–517. [CrossRef Medline](#)
- Goard M, Dan Y (2009) Basal forebrain activation enhances cortical coding of natural scenes. *Nat Neurosci* 12:1444–1449. [CrossRef Medline](#)
- Greenberg DS, Houweling AR, Kerr JN (2008) Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nat Neurosci* 11:749–751. [CrossRef Medline](#)
- Guo W, Chambers AR, Darrow KN, Hancock KE, Shinn-Cunningham BG, Polley DB (2012) Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. *J Neurosci* 32:9159–9172. [CrossRef Medline](#)
- Haider B, McCormick DA (2009) Rapid neocortical dynamics: cellular and network mechanisms. *Neuron* 62:171–189. [CrossRef Medline](#)
- Happel MF, Jeschke M, Ohl FW (2010) Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. *J Neurosci* 30:11114–11127. [CrossRef Medline](#)
- Harris KD, Thiele A (2011) Cortical state and attention. *Nat Rev Neurosci* 12:509–523. [CrossRef Medline](#)
- Hasenstaub A, Sachdev RN, McCormick DA (2007) State changes rapidly modulate cortical neuronal responsiveness. *J Neurosci* 27:9607–9622. [CrossRef Medline](#)
- Hazan L, Zugaro M, Buzsáki G (2006) Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. *J Neurosci Methods* 155:207–216. [CrossRef Medline](#)
- Hirata A, Castro-Alamancos MA (2011) Effects of cortical activation on sensory responses in barrel cortex. *J Neurophysiol* 105:1495–1505. [CrossRef Medline](#)
- Horvath D, Lesica NA (2011) The effects of interaural time difference and intensity on the coding of low-frequency sounds in the mammalian mid-brain. *J Neurosci* 31:3821–3827. [CrossRef Medline](#)
- Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61:597–608. [CrossRef Medline](#)
- Lakatos P, Musacchia G, O’Connell MN, Falchier AY, Javitt DC, Schroeder CE (2013) The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77:750–761. [CrossRef Medline](#)
- Luczak A, Barthó P, Marguet SL, Buzsáki G, Harris KD (2007) Sequential structure of neocortical spontaneous activity in vivo. *Proc Natl Acad Sci U S A* 104:347–352. [CrossRef Medline](#)
- Luczak A, Barthó P, Harris KD (2009) Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* 62:413–425. [CrossRef Medline](#)
- Luczak A, Bartho P, Harris KD (2013) Gating of sensory input by spontaneous cortical activity. *J Neurosci* 33:1684–1695. [CrossRef Medline](#)
- MacLean JN, Watson BO, Aaron GB, Yuste R (2005) Internal dynamics determine the cortical response to thalamic stimulation. *Neuron* 48:811–823. [CrossRef Medline](#)
- Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S (2009) A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci* 10:81. [CrossRef Medline](#)
- Marguet SL, Harris KD (2011) State-dependent representation of amplitude-modulated noise stimuli in rat auditory cortex. *J Neurosci* 31:6414–6420. [CrossRef Medline](#)
- Niell CM, Stryker MP (2010) Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* 65:472–479. [CrossRef Medline](#)

- Okun M, Yger P, Marguet SL, Gerard-Mercier F, Benucci A, Katzner S, Busse L, Carandini M, Harris KD (2012) Population rate dynamics and multi-neuron firing patterns in sensory cortex. *J Neurosci* 32:17108–17119. [CrossRef Medline](#)
- Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. *Nat Neurosci* 12:646–654. [CrossRef Medline](#)
- Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol* 98:1064–1072. [CrossRef Medline](#)
- Polack PO, Friedman J, Golshani P (2013) Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nat Neurosci* 16:1331–1339. [CrossRef Medline](#)
- Porter BA, Rosenthal TR, Ranasinghe KG, Kilgard MP (2011) Discrimination of brief speech sounds is impaired in rats with auditory cortex lesions. *Behav Brain Res* 219:68–74. [CrossRef Medline](#)
- Poulet JFA, Petersen CC (2008) Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature* 454:881–885. [CrossRef Medline](#)
- Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, Kilgard MP (2011) Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron* 70:121–131. [CrossRef Medline](#)
- Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD (2010) The asynchronous state in cortical circuits. *Science* 327:587–590. [CrossRef Medline](#)
- Rigas P, Castro-Alamancos MA (2007) Thalamocortical up states: differential effects of intrinsic and extrinsic cortical inputs on persistent activity. *J Neurosci* 27:4261–4272. [CrossRef Medline](#)
- Ringach DL (2009) Spontaneous and driven cortical activity: implications for computation. *Curr Opin Neurobiol* 19:439–444. [CrossRef Medline](#)
- Rybalko N, Suta D, Nwabueze-Ogbo F, Syka J (2006) Effect of auditory cortex lesions on the discrimination of frequency-modulated tones in rats. *Eur J Neurosci* 23:1614–1622. [CrossRef Medline](#)
- Sachidhanandam S, Sreenivasan V, Kyriakatos A, Kremer Y, Petersen CC (2013) Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat Neurosci* 16:1671–1677. [CrossRef Medline](#)
- Sahani M, Linden J (2003) How linear are auditory cortical responses? In: *Advances in neural information processing systems* 15 (Becker S, Thrun S, Obermayer K, eds), pp 301–308. Cambridge, MA: MIT.
- Schmitzer-Torbert N, Jackson J, Henze D, Harris K, Redish AD (2005) Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* 131:1–11. [CrossRef Medline](#)
- Schneider DM, Nelson A, Mooney R (2014) A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513:189–194. [CrossRef Medline](#)
- Steriade M, Timofeev I, Grenier F (2001) Natural waking and sleep states: a view from inside neocortical neurons. *J Neurophysiol* 85:1969–1985. [CrossRef Medline](#)
- Strong S, Koberle R, de Ruyter van Steveninck R, Bialek W (1998) Entropy and information in neural spike trains. *Phys Rev Lett* 80:197–200. [CrossRef](#)
- Tan AY, Chen Y, Scholl B, Seidemann E, Priebe NJ (2014) Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature* 509:226–229. [CrossRef Medline](#)
- Ter-Mikaelian M, Sanes DH, Semple MN (2007) Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. *J Neurosci* 27:6091–6102. [CrossRef Medline](#)
- Thomas H, Tillein J, Heil P, Scheich H (1993) Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*): I. Electrophysiological mapping of frequency representation and distinction of fields. *Eur J Neurosci* 5:882–897. [CrossRef Medline](#)
- Wetzel W, Ohl FW, Wagner T, Scheich H (1998) Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. *Neurosci Lett* 252:115–118. [CrossRef Medline](#)
- Xu S, Jiang W, Poo MM, Dan Y (2012) Activity recall in a visual cortical ensemble. *Nat Neurosci* 15:449–455:S1–2. [CrossRef Medline](#)
- Zhou M, Liang F, Xiong XR, Li L, Li H, Xiao Z, Tao HW, Zhang LI (2014) Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat Neurosci* 17:841–850. [CrossRef Medline](#)
- Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party.” *Neuron* 77:980–991. [CrossRef Medline](#)