UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Information processing by coordinated neuronal ensembles in the primary auditory cortex

Permalink

https://escholarship.org/uc/item/8hh4m98z

Author

See, Jermyn

Publication Date

2019

Peer reviewed|Thesis/dissertation

Information processing by coordinated neuronal ensembles in the primary auditory
cortex

by
Jermyn Zeming See

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
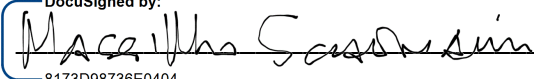DOCTOR OF PHILOSOPHY

in

Neuroscience

in the

GRADUATE DIVISION
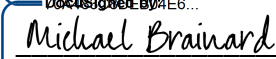of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

_____   Massimo Scanziani
8173D98736E0404...
                                                                                              Chair

DocuSigned by:

_____   Christoph Schreiner, MD, PhD
DocuSigned by:

_____   Vikaas Sohal
Vikaas Sohal
DocuSigned by:

_____   Michael Brainard
Michael Brainard
7ED6AD87FF16412...

_____
                                                                          Committee Members

Acknowledgements

The long and winding road towards a PhD is not embarked upon alone, and I am immensely grateful for the people that have supported me throughout this journey. First and foremost, I would like to thank my PhD advisors, Christoph Schreiner and Vikaas Sohal. Christoph welcomed me into his lab as a bright-eyed, bushy-tailed first-year graduate student with little computational experience and plans for a very ambitious project, but gave me the time, space and resources to hone my analytical skills and discover my passion for tackling big data problems. For this, I can only be eternally thankful. Vikaas took me on in spite of doubts that he could effectively mentor me based on my scientific and technical interests and I am indebted to him for accepting that risk, because I have learned so much from him, both scientifically and non-scientifically. Both of them have shown me what good mentorship looks like, and have provided me with a psychologically safe space to fail; that is, to not be afraid of taking risks, and to accept failures in my stride. I will take these values with me wherever I go, and hopefully provide my future mentees with a similar environment that I was so fortunate to have experienced.

Every Batman needs a Robin, and Craig Atencio was the very capable sidekick to my two PhD advisors. I remain hugely appreciative of the amount of time and patience that Craig spent mentoring me and teaching me the technical ropes, especially during my early days in lab, when I could not yet walk on my own two feet. He was also a great collaborator on my thesis work,

Contributions

Chapter 2 is reprinted largely as it appears in: See, J.Z., Atencio, C.A., Sohal, V.S., Schreiner, C.E., 2018. Coordinated neuronal ensembles in primary auditory cortical columns. Elife 7, 1–33. doi:10.7554/eLife.35587

Information processing by coordinated neuronal ensembles in the primary auditory cortex

Jermyn Z. See

Abstract

The primary auditory cortex (AI) is made up of highly interconnected populations of neurons that are responsible for integrating bottom-up auditory information from the lemniscal auditory pathway and top-down inputs from higher-order cortical areas. Despite that, most studies of information processing in AI focus on either single-unit spectro-temporal receptive field (STRF) estimation, or paired neuronal correlation analyses, and assume that AI neurons filter auditory information either as individual entities or pairs. Meanwhile, some recent studies have also shown how populations of AI neurons can also encode auditory behavior. Determining how AI encodes information will hence require an integrated approach that combines receptive field and multi-neuronal ensemble analyses.

In this dissertation, I show that I can accurately detect coordinated neuronal ensembles (cNEs), which we define as groups of neurons that have reliable synchronous activity, in AI. These cNEs are meaningful constructs that are active in both spontaneous and evoked activity, and their synchronous evoked activity cannot be trivially explained by receptive field overlap. cNEs also come in two flavors – one of them enhances stimulus representation over single neurons or simultaneously recorded random groups of neurons of the same size, while the other does not represent spectro-temporal features at all, and might reflect internally generated neuronal activity. Since single neurons can participate in multiple cNEs over the course of a

recording, I also show that neurons can multiplex information, and encode slightly different

spectro-temporal information, if they encode spectro-temporal information at all, when

associated with different cNEs. Meanwhile, the enhancement of information processing and the

reliability of representation of the stimulus by cNEs suggest that cNEs should be considered the

principal unit of information processing in AI.

Table of Contents

List of Figures

CHAPTER 2

CHAPTER 3

List of Tables

CHAPTER 2

CHAPTER 1

Introduction

Introduction

One of the biggest questions in systems neuroscience relates to how individual neurons work together to integrate sensory and cognitive information from various sources in the brain, and process that to influence behavior and motor output. The neocortex, evolutionarily the newest member of our brain, contains on the order of ~20 billion neurons (Pakkenberg and Gundersen, 1997) and is highly intra- and interconnected (Braitenberg and Schüz, 1991). On a microscopic scale, this means that each individual neuron in any particular area of the neocortex receives inputs from neurons in a large variety of subcortical and cortical areas, has the responsibility of integrating and processing this myriad of information, before transmitting the information downstream to the next brain area. However, each individual neuron does not perform this in a vacuum, since the information it receives is also received by thousands of its neighboring neurons, and would also be passed between its neighboring neurons, due to the high interconnectivity of neurons in the neocortex. As a result, the activity of each neuron is highly correlated with some of its neighbors, and this coordinated activity is thought to be important in information processing and transmission throughout the cortex (DeNardo et al., 2019; Reimann et al., 2017; Roland et al., 2017), in part due to the increased likelihood of information transmission downstream when ensembles of neurons are synchronously active (Matsumura et al., 1996; Stevens and Zador, 1998). This process repeats itself in each subsequent or downstream brain area on the order of milliseconds and eventually leads to perception and behavior.

To begin untangling the intricate puzzle of neuronal cooperativity, it was necessary to identify an area of the neocortex that has well-characterized and predictable responses to well-parameterized stimuli, such that it would be possible to probe, at high temporal resolutions, how

groups of neurons might process information differently from single neurons. The primary auditory cortex (AI) was a good candidate for this because AI neurons have been shown to have well-defined receptive fields (Depireux et al., 2001; Polley et al., 2007). AI is the first port of entry of auditory information into the neocortex via the auditory lemniscal pathway, and receives direct inputs from the medial geniculate body (MGB) in the thalamus in a tonotopic fashion (Lee and Winer, 2011), imparting to AI neurons their characteristic sharp tuning (Polley et al., 2007). Furthermore, AI neuronal responses to auditory stimuli are precise up to a few milliseconds (Kayser et al., 2010), making them ideal candidates for probing the importance of synchrony. However, tracing studies have shown that the thalamocortical inputs from MGB into AI only make up ~18% of all physical extrinsic input into AI, while a large majority of the extrinsic inputs originate from higher-order cortical areas (Lee and Winer, 2011). This structural convergence of bottom-up and top-down input makes AI a potentially good place to study the integration of different sources of information in populations of neurons.

Functionally, most of what we currently know about AI neurons is based on single-unit activity or paired neuronal correlation analyses. Single-unit studies in AI have been focused on characterizing receptive fields, assuming that AI neurons filter auditory information as individual and independent entities (Aertsen and Johannesma, 1981; Atencio et al., 2012; Calabrese et al., 2011). Meanwhile, correlational studies between pairs of neurons have shown that the synchronous activity of pairs of neurons can signal enhanced information encoding of the auditory stimulus (Atencio and Schreiner, 2013; Eggermont, 1992, 2006). More recently, population analyses, based on indiscriminate summation of single- or multi-unit activity can be used to decode the choice of an animal in an auditory discrimination task (Francis et al., 2018; Rodgers and DeWeese, 2014) or aspects of acoustic information (Abrams et al., 2017; David and

Shamma, 2013; Ince et al., 2013). Taken together, a necessary step to bridge the gap between single-neuron and population-level studies is to take correlational studies between pairs of neurons further, and investigate how ensembles of neurons might work together to enhance the encoding of the auditory stimulus. Only by doing so can we hope to reach a mechanistic understanding of how spectro-temporal filters can lead to auditory perception and behavior.

The idea that synchronous activity is important in information processing and transmission in the brain is not a new one (for review, see Averbeck et al., 2006). In Donald Hebb's seminal work, he postulated the existence of cell assemblies in the brain, hypothesizing that cell assemblies are distinct, fixed units of highly and strongly interconnected neurons whose activities would provide the basis of cognitive processes and large-scale behavioral phenomena (Hebb, 1949). Despite the age of the theory, it was not until the last decade or so did we have the wherewithal to scientifically investigate Hebb's cell assembly hypothesis. With the advent of technological advances in calcium imaging and high-density multi-channel electrodes (Hamel et al., 2015; Jun et al., 2017), neuroscientists have been able to monitor the simultaneous activity of large populations of neurons. Furthermore, advances in statistical analyses have also allowed us to detect groups of neurons with synchronous activity within simultaneously recorded populations of neurons, and identify coordinated activity events (Billeh et al., 2014; Gourévitch and Eggermont, 2010; Lopes-dos-Santos et al., 2013; Miller et al., 2014). In this dissertation, I made use of dense multi-channel electrodes to record the extracellular electrophysiology of large populations of neurons in AI, and used dimensionality reduction techniques (Lopes-dos-Santos et al., 2013) to extract and identify coordinated neuronal ensembles (cNEs), which were defined as groups of neurons that have reliable synchronous activity. This made cNEs distinct from Hebb's cell assemblies; while cell assemblies have been hypothesized to be groups of neurons with fixed

sequential firing patterns over time, akin to synfire chains (Ikegaya et al., 2004), cNEs are groups of neurons that fire synchronously within a predefined time window. I chose to define cNEs this way because AI responses to stimuli or thalamocortical excitation have been shown to occur within a narrow time window of <10ms (Atencio and Schreiner, 2010a; Kaur et al., 2005; Krause et al., 2014). Moreover, given the preciseness of AI responses to auditory stimuli (Kayser et al., 2010), it was necessary to follow cNE responses in AI at a very high temporal resolution, which becomes intractable if large time lags are considered, as in the case of synfire chains. Finally, previous analyses of zero or near-zero lag synchronous responses have provided rich insights into network function (Atencio and Schreiner, 2013; Harris, 2005; Nicolelis et al., 1995), and is more easily identified by statistical methods, and computationally less expensive to isolate, than when considering sequential firing patterns over time (Peyrache et al., 2009).

Chapter 2 elaborates upon these ideas that I have presented in this introduction/chapter. While presenting broadband stimuli to rats anesthetized with a cocktail mix of ketamine/xylazine, I performed high-density electrophysiological recordings across several laminae within AI columns of rats. From the recorded neuronal data, I isolated cNEs using the cNE detection algorithm (Lopes-dos-Santos et al., 2013) and validated their biological significance through various means. First, I validated that the temporal window within which I defined synchrony (10 ms) gave me cNEs with STRFs that maximized for the most mutual information (MI) between cNE activity and the stimulus, when compared to other time bin sizes. This corresponded with the upper end of the maximum jitter one might get recording AI neuronal responses to auditory stimuli (Atencio and Schreiner, 2010a; Kaur et al., 2005). I then verified that identified cNEs are equally active in both spontaneous and evoked activity, and that they cannot be trivially explained due to receptive field overlaps. These results were important in

proving the biological significance of cNEs because the cortical column is a fundamental organizing substrate within AI and relates neurons with similar preferred frequencies (Atencio and Schreiner, 2010a, 2010b; Guo et al., 2012); recording from cortical columns therefore increased the likelihood of encountering stimulus-driven synchrony between neurons. Finally, I showed that cNEs were more reliable at encoding the auditory stimulus, and generally contained more information about the presented stimulus than their member neurons or groups of simultaneously recorded neurons of the same size. Taken together, I make a strong case for cNEs as unique and functionally significant units of information processing in AI. This chapter has been peer-reviewed and published (See et al., 2018).

In chapter 3, I further investigate the functional significance of cNEs in AI, and show that cNEs in AI can encode non-auditory information as well. In particular, I identified four major functional categories of cNEs, based on the significance of cNE's and its member neurons' STRFs. These four categories of cNEs have statistically different physical and functional properties, suggesting that there are meaningful biological differences between the different cNE categories. I postulate, given previous studies (Brosch et al., 2015; Francis et al., 2018; Rodgers and DeWeese, 2014), that the cNEs that did not appear to encode spectro-temporal features of the auditory stimulus could encode other aspects of the auditory world, for example, choice in an auditory discrimination task (Francis et al., 2018; Rodgers and DeWeese, 2014) or categorical perception of sounds (Bathellier et al., 2012). I further show that AI neurons produce a heterogeneous set of spikes based on its instantaneous association with different cNEs. This essentially means that AI neurons can multiplex information; their intermixed set of spikes can either encode markedly different spectro-temporal features of the stimulus, or encode both non-spectro-temporal features and spectro-temporal features of the stimulus over the course of the

presentation of the stimulus. Meanwhile, cNEs appear to be relatively stable in encoding a specific type of information, and these results have given the idea that cNEs are the fundamental processing unit of information processing in AI much credence.

Lastly, in chapter 4, I recapitulate the main results from chapter 2 and 3. Importantly, I discuss the immediate future directions and implications of my work to other projects in the lab that are either already in progress or have been planned and also discuss my results in the context of previous work in auditory learning and attention (Fritz et al., 2003, 2005b). Finally, I also discuss how my research has created a useful toolbox for the study of ensembles of neurons and their functional properties in general.

CHAPTER 2

Validation of coordinated neuronal ensembles in primary

auditory cortical columns

2.1 Abstract

The synchronous activity of groups of neurons is increasingly thought to be important in cortical information processing and transmission. However, most studies of processing in the primary auditory cortex (AI) have viewed neurons as independent filters; little is known about how coordinated AI neuronal activity is expressed throughout cortical columns and how it might enhance the processing of auditory information. To address this, we recorded from populations of neurons in AI cortical columns of anesthetized rats and, using dimensionality reduction techniques, identified multiple coordinated neuronal ensembles (cNEs), which are groups of neurons with reliable synchronous activity. We show that cNEs reflect local network configurations with enhanced information encoding properties that cannot be accounted for by stimulus-driven synchronization alone. Furthermore, similar cNEs were identified in both spontaneous and evoked activity, indicating that columnar cNEs are stable functional constructs that may represent principal units of information processing in AI.

2.2 Introduction

How individual neurons work together to encode sensory information and influence behavior remains one of the fundamental questions in systems neuroscience. Single neurons have been traditionally considered to be the basic functional unit of the brain and much of our understanding of how the brain encodes sensory stimuli or motor output comes from decades of single-unit studies. Single-unit activity in isolation, however, is often insufficient to account for observed sensory or motor behaviors (Bizley et al., 2010; Britten et al., 1996; Engineer et al., 2008; Georgopoulos et al., 1986; Herzfeld et al., 2015; Paninski et al., 2004).

Technological advances in large-scale recordings, including calcium imaging and high-density multi-channel electrodes, have allowed the monitoring of the simultaneous activity of large populations of neurons. This has led to the demonstration of coordinated activity within groups of recorded neurons, identified and verified by statistical approaches (Billeh et al., 2014; Gourévitch and Eggermont, 2010; Lopes-dos-Santos et al., 2013; Miller et al., 2014; Peyrache et al., 2010; Pipa et al., 2008).

These concerted neuronal activities are postulated to be local network events that reflect improved associations with decision-making, predictions of perceptual events, memory formation, and behavioral performance over isolated single-unit activity (Bathellier et al., 2012; Bell et al., 2016; Gulati et al., 2014; Ince et al., 2013; Kiani et al., 2014; Laubach et al., 2000; Peyrache et al., 2009; Reimann et al., 2017). Coordinated ensembles have also been postulated to be elementary units of information processing in the brain (for review see Buzsáki, 2010; Harris and Mrsic-Flogel, 2013; Yuste, 2015). However, basic statistical properties of these ensembles,

such as cellular composition, extent, stability, and functional roles, including selectivity of

information extraction and reliability of transmission, are not yet well understood.

This is especially true in the auditory cortex (AC). Most of what we understand about AC

function is based on single-unit and general population analyses. Single-unit studies in the AC

have focused on characterizing receptive fields, treating AC neurons as arrays of (nearly) linear

filters (Aertsen and Johannesma, 1981; Atencio et al., 2012; Calabrese et al., 2011; Thorson et

al., 2015). Population activity in the AC, often based on indiscriminate pooling of single- or

multiple-unit activity, has been shown to correlate with an animal's perception of simple sounds

(Bathellier et al., 2012; Rodgers and DeWeese, 2014) and can be used to decode aspects of

acoustic information (Miller and Recanzone, 2009; Brasselet et al., 2012; Ince et al., 2013; David

and Shamma, 2013; Abrams et al., 2017). However, these studies did not identify groups of

cooperating neurons, treating all simultaneously-recorded neurons as equivalent information-

processing entities. Meanwhile, synchronous activity between local pairs of AC neurons can

reflect shared and unique stimulus aspects, enabling enhanced information encoding (Atencio

and Schreiner, 2013, 2016; Brosch and Schreiner, 1999; Eggermont, 1992, 2006; Eggermont et

al., 2013; Gourévitch and Eggermont, 2010). While neuronal pairs represent the smallest

possible ensemble, larger columnar or distributed networks of synchronously active neurons

could reveal specific and particularly meaningful information-bearing network events (Bathellier

et al., 2012; Bharmauria et al., 2016). One particular issue is whether coordinated ensemble

events can be meaningfully distinguished from coincidental events that result from stimulus

synchrony and overlaps in the receptive fields of neurons. Therefore, to better understand

network encoding, the membership criteria, structure and function of synchronously-firing

neuronal groups must be examined to distinguish between coordinated and coincidental neuronal ensembles.

Here, we identified and characterized coordinated neuronal ensembles (cNEs) from the synchronous activity across populations of neurons within auditory cortical columns. We performed high-density extracellular electrophysiological recordings across several laminae within primary auditory cortical (AI) columns of rats using broadband dynamic noises (Atencio and Schreiner, 2010a, 2010b). The cortical column is a fundamental organizing substrate within the AC, and relates neurons with similar preferred frequencies (Wallace and Palmer, 2007 (gerbil); Atencio and Schreiner, 2010a, 2010b (cat); Guo et al., 2012 (mouse); but see Winkowski and Kanold, 2013). Recording from the column therefore increased the likelihood of encountering stimulus-driven, temporally aligned spiking activity. We chose the ketamine-anesthetized preparation to enhance the stability of the recordings over time and to reduce the influence of state changes on neural networks; the awake (Fritz et al., 2005a; McGinley et al., 2015; Okun et al., 2010; Poulet and Petersen, 2008) or urethane-anesthetized (Clement et al., 2008; Marguet and Harris, 2011; Pagliardini et al., 2013) rat cortex would undergo state changes more frequently, potentially affecting the composition of cNEs. We applied dimensionality reduction techniques to recorded populations and identified cNEs, which consisted of neurons that exhibited sharply synchronous activity (Lopes-dos-Santos et al., 2013). Our results show that cNEs can be accurately and reliably detected, have higher coincident firing than can be explained from receptive-field similarity or second-order (pairwise) correlations, and have active events that occur more frequently than expected from population recordings. cNEs identified from spontaneous and evoked epochs were highly similar, implying that cNEs represent functional networks that are not dependent on stimulus presentation. Our results also show that

cNEs encode auditory information more reliably than single neurons or random groups of neurons with similar receptive fields and support the notion that cNEs represent unique and functionally significant units of processing in AI.

2.3 Results

Heterogeneous functional connectivity between columnar AI neurons

We used high-density recording electrodes (Figure 2.1A) to simultaneously record the activity of populations of neurons across supragranular and granular laminae within AI columns. Presenting dynamic, broadband stimuli with varying statistics (dynamic moving ripple (DMR) or ripple noise (RN); see Materials and methods) allowed us to estimate the spectro-temporal receptive fields (STRFs) of identified neurons (Figure 2.1C). STRFs of neurons from the same columnar penetration showed general similarities in characteristic frequency and frequency tuning, and displayed minor variations in spectral and temporal details (see example in Figure 2.1C), similar to cat cortical columnar recordings (Atencio and Schreiner, 2010a).

We first examined response synchrony between simultaneously recorded single units using Pearson's correlation analysis at 10-ms temporal resolution. Pairs of neurons in the same column with similar STRFs could nonetheless exhibit widely differing pairwise-correlation (PWC) functions (Figure 2.1D). For example, the PWC between pairs of neurons with similar receptive fields, recorded from channels of equal intra-pair distances, could differ in peak correlation magnitude, width and delay (Figure 2.1C and D). Pearson's correlation values (or correlation coefficients), PWC absolute peak delays and sharpness are depicted in Figure 2.2A for one neuron (#15) with all other recorded neurons in the same penetration. Significant PWCs were evident for most pairs in this penetration but low correlation coefficients and broader correlation widths were dominant. Across all penetrations, PWCs were fairly heterogeneous but sharpness and delay covaried with correlation coefficients, accounting for ~50% of the variance (Figure 2.2B–2D).

The degree of synchrony between pairs of neurons often weakens as the distance between neurons within a column increases (Atencio and Schreiner, 2013; Gururangan et al., 2014). We observed similar trends in columnar recordings, although the correlation between these synchrony aspects and pairwise distance only explained ~10–30% of the total variance (R2 values, Figure 2.2E– 2G). Notwithstanding weak spike train correlations, neuronal coordination and synchronization have been widely postulated to play a vital role in information processing and the organization of the nervous system within the cortical column (Atencio and Schreiner, 2010a; Panzeri et al., 2003). We therefore sought to identify larger columnar groups of neurons that reliably share synchronous firing events, resulting in higher correlated rates of firing than could be accounted for by inter-neuronal distances and shared receptive field properties.

Identification of coordinated neuronal ensembles (cNEs)

To identify synchronously firing groups of neurons, or coordinated neuronal ensembles (cNEs), we combined two dimensionality reduction techniques - principal component analysis (PCA) and independent component analysis (ICA), following the approach by Lopes-dos-Santos et al. (2013). This analysis identifies groups of neurons that repeatedly fire synchronous action potentials throughout the duration of the recording. This approach provides a parametric analysis that identifies both the number of cNEs and the neurons that are members of each cNE. Further, simulations show that this approach is applicable to both large- and small-scale recordings, and that it can be applied to spike trains with Poisson characteristics (Lopes-dos-Santos et al., 2013; Peyrache et al., 2010). Application of PCA to the simultaneously recorded spike trains determines the number of cNEs above a significance criterion (see Materials and methods). ICA is then applied to the significant eigenvectors or principal components (PCs) to extract the linear combination of the individual neurons' activity associated with each cNE (Lopes-dos-Santos et

al., 2013). To illustrate this approach, we simulated the spike trains of 8 neurons, where two groups of five neurons had highly synchronous firing, with two neurons being shared between these two groups (Figure 2.3A and B). Each spike train was first binned and z-scored to ensure that cNE identification was not skewed towards neurons with higher firing rates. The correlation coefficients between the spike trains were calculated and represented as a correlation matrix (Figure 2.3B). PCA was applied to the correlation matrix, resulting in a distribution of eigenvalues and corresponding PCs. Each eigenvalue represents the relative importance of each PC. To determine which PCs represented groups of neurons that significantly fired together, we compared the distribution of eigenvalues from the data to the distribution that would be expected for groups of independently firing neurons (Marčenko and Pastur, 1967). In the simulated example, there were two significant eigenvalues reflecting the fact that there were two cNEs (Figure 2.3C). However, because neurons 4 and 5 were shared between both cNE #1 and #2, the PCs did not accurately represent the true membership of the two cNEs due to PCA's variance maximization framework (Figure 2.3D). To resolve this ambiguity, we applied ICA to the two most significant PCs, which resulted in two independent components (ICs) that accurately identified the two cNEs (Figure 2.3F). Both the PCs and ICs can be used to recreate the activity of the cNEs over time by projecting them back onto the spike matrix. However, since PCA did not separate the cNEs properly (Figure 2.3E), only the ICA-derived activity correctly reflected the spiking activity of each cNE (Figure 2.3G).

Applying this cNE detection algorithm to an actual AI columnar recording with 38 isolated neurons (Figure 2.4D), we identified nine separate cNEs (Figure 2.4B) with significantly contributing members ranging from 2 to 6 neurons. Each column in Figure 2.4B shows the color-coded IC weight for all neurons in each cNE. Neuronal membership in a cNE was assigned by

exceeding a statistical criterion of the IC weight for each neuron in the penetration. Green bars in

the weight plot (Figure 2.4B) indicate neurons with significant IC weights. To determine the

threshold criterion, we circularly shuffled spike trains within a spike matrix independently over

100 different iterations, ran PCA and ICA on these shuffled spike matrices, and set the threshold

at which a neuron belongs to a cNE as 1.5 standard deviations above and below the mean of the

distribution of the resulting IC values (Figure 2.4C, see Materials and methods). We used a

similar method to determine the minimum magnitude of cNE activity during the recording period

that counts as a significant cNE event (Figure 2.4E, see Materials and methods). For example,

the cNE analysis revealed that almost every time cNE #2 was deemed active, at least three

members (out of a total of 6) of that cNE had to be active (Figure 2.4D and E).

The cNE detection algorithm typically identified several independent cNEs, each

capturing moments of joint activity between several of its members. In the case of the example

recording, 9 cNEs were extracted. Since cNEs were extracted via ICA, the activities of the

different cNEs are necessarily independent of one another across the whole recording period.

Yet, as can be seen in the bottom panel of Figure 2.4E, cNEs can fire together with other cNEs

(e.g., cNE #2 (red trace) and cNE #3 (green trace) at the 1.1 s and 1.75 s time marks). The cNEs

identified in 15 additional columnar recordings are shown in Figure 2.S1.

We next tested if the identified cNEs were stable and reliable estimates of coordinated

group activity that persisted throughout the entire recording. First, we wanted to determine if the

cNEs identified from one half of a recording were similar to those identified independently from

the other half. We split the recordings into 10 equal parts and applied the cNE algorithm to both

sets of interleaved parts of the recording (Figure 2.S2A). The halves were interleaved to

compensate for potential recording instabilities. The correlation coefficients for each of the two

sets of independent components were compared (Figure 2.S2B). To calculate the significance of all correlations within each penetration, spike trains were circularly shuffled, and processed with the cNE detection algorithm (similar to the method to determine the threshold criterion for neuronal membership in cNEs). The resulting ICs were correlated to give a null distribution, and the 99th percentile for each null distribution was set as the significance threshold (Figure 2.S2C). Approximately 96% of independent components had significant matches (Figure 2.S2D and E), indicating that the two interleaved parts of the recording contained the same independent components and that the cNEs were present throughout the entire recording.

Secondly, we wanted to test if cNEs detected in a subset of the recorded activity would be able to accurately estimate the cNE activity of the rest of the recording. In this case, we split the recording into two datasets for training (first half) and for testing (second half). Applying the PCA/ICA approach to each dataset, we obtained two corresponding sets of independent components, and projected both sets onto the test data to extract training and test cNE activities (Figure 2.S3A). We then correlated the training and test cNE activities (Figure 2.S3B) and took the maximum correlation value for each training cNE as 'matched correlations' (Figure 2.S3D). The remaining correlation values for all penetrations were treated as 'non-matched correlations' and formed the null distribution, and the 95th percentile of that null distribution was the significance threshold (Figure 2.S3C). Across all penetrations, approximately 83% of training cNE activity had significant matches to test cNE activity. Thus, cNEs obtained from the first half of the recordings were sufficient to predict the activity of cNEs in the latter half of the recording. Together, these tests show that derived cNEs are stable and reliable descriptors of coordinated neuronal activity.

18

cNE identity is dependent on time bin size

In previous studies, the time interval for defining synchronous firing varied widely, from a few milliseconds to several hundred milliseconds (see Table 2.1 for a selection of example studies). Reasons for the different choices of time ranges were related to the temporal resolution permitted by the applied methods and the physical distance between considered neuronal elements. We chose a temporal resolution consistent with the time needed for auditory information to traverse the interlaminar microcircuit in AI columns. Using the propagation speed found in the cat and rat, the time required to traverse a rat AI column of approximately 1500 µm in extent is ~10 ms (Atencio and Schreiner, 2010a; Kubota et al., 1997). The upper limit was defined by the need to avoid stimulus-based temporal properties; the dynamic moving ripple stimulus contained temporal envelope fluctuations up to 40 Hz, resulting in a minimum temporal resolution of 25 ms. We therefore selected 10 ms as a reasonable window of within-column synchrony. Inclusion of synchrony with extracolumnar sources and targets may require analysis with a different window size.

To determine the functional significance of the chosen time window, we binned spikes into varying time bin sizes, obtained cNE activity for each bin size, and then estimated the receptive field information of cNEs detected based on the bin sizes (see Materials and methods). Receptive field information peaked at bin sizes of approximately 10 - 15 ms (Figure 2.5A), implying that the temporal resolution used to obtain cNEs provided the most information about the receptive field preferences of groups of synchronously firing neurons in AI, further supporting our choice. This temporal resolution also agrees with other studies on the optimal window for dendritic integration (Branco and Häusser, 2011; Polsky et al., 2004; Williams and

Stuart, 2002) and spike-timing dependent plasticity (Bi and Poo, 1998; Celikel et al., 2004; Markram, 1997).

The size of the time window also affects the number and size of identified cNEs. We would expect that as temporal resolution decreases and the size of the window that defines synchrony increases, more concurrently active neuronal responses are included in each cNE. cNEs that might have been deemed separate in 10-ms time bins might instead be classified as a single cNE when the data is processed using longer time bins, leading to fewer but larger cNEs. Indeed, we found that as the spike train temporal bin size increased, the number of recovered cNEs decreased (Figure 2.5B) and the number of member neurons in each cNE increased (Figure 2.5C).

Synchrony between cNE members cannot be fully explained by receptive field overlap

Columnar neurons show significant heterogeneity in their PWCs (Figures 2.1D and 2.2). First, we sought to confirm that spike train PWCs differ depending on whether pairs of neurons belong to the same cNE. This is to be expected given that cNE membership was based on Pearson's correlation coefficients, which are highly correlated with PWC sharpness (Figure 2.2B). Secondly, we wanted to determine whether the degree of spiking synchrony between neurons from the same cNE can be sufficiently explained by the degree of receptive field similarity between the neurons. Comparing STRF PWCs to spike train PWCs (Figure 2.6A–6F) revealed that the spike train PWCs (Figure 2.6C, histogram) were much narrower than the STRF PWCs (Figure 2.6C, solid line) for neuronal pairs within the same cNE. In contrast, for non-cNE pairs, the spike train PWCs were often as wide as the STRF PWCs (Figure 2.6F).

To compare PWCs across the population, we assessed correlation sharpness by first estimating the peak delay (Figure 2.6G) and then folding the PWCs around that delay (Figure 2.6H). We estimated correlation sharpness as the width that accounted for half the spike count in the PWC histogram (Figure 2.6I). The PWCs were significantly narrower for cNE pairs than for non-cNE pairs ($p < 0.001$, Mann-Whitney U test; Figure 2.6J; see Figure 2.S4 for each of the other penetrations). The difference in sharpness between cNE pairs and non-cNE pairs was consistent with the high degree of correlation between PWC sharpness and spike train correlation (Figure 2.2B), the latter of which was used to define cNEs. The median PWC sharpness for member neurons was ~2 ms, reflecting the central peak of a highly synchronized portion of spikes (Figure 2.6C and G). However, the tail-portion of member PWCs do contain less strictly synchronized spiking events, suggestive of a mix of highly and loosely synchronized events among cNE members.

Together, these observations indicate that spike train PWCs between members of the same cNE show strong evidence of strict synchrony that cannot be explained by receptive field overlap alone. PWCs between non-cNE pairs display loose synchrony that largely reflects coincidental activity expected from independent neurons with significant receptive field overlap.

cNEs are not depth-biased and reflect local circuitry

Across 16 penetrations, we identified a total of 104 cNEs (Figure 2.4, Figure 2.S1), obtained in response to either DMR or RN broadband stimuli. The average number of cNEs per penetration was 6.5 ± 2.0, but the higher the number of neurons isolated in a penetration, the higher the number of identified cNEs (Figure 2.7A). For each penetration, the number of cNEs was ~15% of the recorded number of neurons. The mean number of neurons in cNEs was 7.4 ± 2.5 neurons. This again depended on the number of isolated neurons (Figure 2.7B), with mean

cNE size ~17% of the total number of recorded neurons in the column. However, the covariance between mean cNE size and the number of isolated neurons (Figure 2.7B; $R^2 = 0.33$) was not as strong as the covariance between the number of identified cNEs and the number of isolated neurons (Figure 2.7A; $R^2 = 0.74$). The IC weight of the vast majority of cNE member neurons was positive, that is, the neurons were co-activated with other members during instances of cNE events. However, 9 cNEs (8.3%) also contained at least one neuron with a negative weight, indicating that such neurons had to be inactive at the time of a cNE event. Of the 655 neurons that were isolated, a majority (~68%) belonged to a single cNE, ~9% did not belong to any cNE, and ~23% belonged to multiple cNEs (Figure 2.7C). A majority of pairs of neurons within the same cNE (~82%) tended to be recorded from channels separated by < 200 µm (Figure 2.7D), indicating that cNEs are mostly spatially confined. These assessments were not biased by recording depth, since cNEs were found across the entire explored columnar range and showed no bias to any particular depth (Figure 2.7E). Most isolated cNEs (~70%) spanned a depth range of 150 µm or less of the cortical column, again indicating that cNE activity reflects local circuit properties, even though there were a few cNEs that spanned most of the measured cortical column (Figure 2.7F).

cNEs persist across epochs of spontaneous and evoked activity

Next, we wanted to determine if cNEs are dependent on stimulus-evoked activity or whether they are similar to cNEs obtained for spontaneous activity. Previous work in sensory cortices has shown that patterns of spontaneous activity are often similar to that of evoked activity (Jermakowicz et al., 2008; Luczak et al., 2009). To investigate if this is true of cNEs, we recorded contiguous epochs of spontaneous and evoked activity, separately processed the spontaneous and evoked data with the cNE detection algorithm, and compared the resulting

spontaneous and evoked cNEs from each recording (Figure 2.8A and B). Using previously

described methods (Figure 2.S2), we defined a null distribution for each recording, and set the

significance threshold at the 99th percentile (Figure 2.8C). Across all 8 recordings comprising

contiguous epochs of spontaneous and evoked activity we found significant matches for ~72% of

spontaneous and evoked cNEs (Figure 2.8D and E). This implies that cNEs largely reflect

functional connectivity and are not simply a reflection of stimulus-induced synchrony of neurons

with overlapping receptive fields.

cNEs show enhanced information processing

Stimulus-evoked synchronized spiking events can carry more information about the

stimulus than the individual spikes of each member of the pair (Atencio and Schreiner, 2013;

Brenner et al., 2000). To test whether cNE events also convey more stimulus information than

the spikes of individual neurons that comprise each cNE, we assessed the mutual information

(MI) for both types of events using the responses to fifty 5-s repeats of RN or DMR (Brenner et

al., 2000, see Materials and methods). Intuitively, this method of determining MI is a proxy for

the reliability of responses over stimulus trials. The lower the response reliability of repeated

stimuli, the flatter the resulting PSTH, and the lower the reflected MI (de Ruyter van Steveninck

et al., 1997; Shih et al., 2011; Strong et al., 1998). Spikes of each cNE member neuron were

compared against the events of the cNE that each neuron belonged to (Figure 2.9A and B). cNE

events were extracted by setting a threshold for the activity profile (Figure 2.4E, see Materials

and methods). For each trial-based comparison between a cNE and one of its member neurons,

the entity with the higher spike or event count was sub-sampled to maintain equal number of

spikes or events and ensure unbiased MI comparisons. The cNE event rasters showed less trial-

to-trial variability than individual neuron rasters (Figure 2.9A and B), and the MI conveyed by

cNE events was, accordingly, ~10% higher than the MI of each member neuron (p < 0.001, Wilcoxon signed-rank test; Figure 2.9C). Correspondingly, the ratio of the MIs (cNE/neuron) was significantly larger than 1 (median ± median absolute deviation = 1.11 ± 0.13; p < 0.001, Wilcoxon signed-rank test; Figure 2.9D).

Next, we established that the higher MI attributed to cNE events was not due to the fact that these cNE events integrated spikes across multiple neurons and, by virtue of a larger neuronal count, carried more MI. We therefore compared the MI for spikes of random groups of neurons (Figure 2.S5A) to the MI for spikes of cNE members (Figure 2.S5B). Each random group of neurons consisted of at least one neuron from the cNE (size N) that it was being compared against, along with N − 1 neurons that were pseudo-randomly selected from non-cNE members (i.e., neurons that were not part of the same cNE as the subject neuron; see Materials and methods). We found that the MI of spikes from cNE members was significantly higher than the MI of spikes from random groups of neurons (p < 0.001, Wilcoxon signed-rank test; Figure 2.S5D). For each cNE-random group comparison, we also calculated a uniqueness index (UI), which quantifies the proportion of times the random group MI exceeded the cNE MI value, $UI = (1 - \frac{x}{100})$, where x is the percentile at which the cNE value was found in comparison to the random group MI distribution (Figure 2.S5C; n = 501). The distribution of UI values was significantly skewed towards 0 and was significantly different from an expected uniform distribution that would be obtained if cNEs were composed of randomly selected neurons from the population recordings (p < 0.001, Kolmogorov–Smirnov test). Together, these results show that cNEs respond more reliably to the stimulus than its member neurons, suggesting that they might have important functional roles in information processing and transmission in AI.

We were also interested to see how noise correlations (for review, see Cohen and Kohn, 2011; Kohn et al., 2016) corresponded to neuronal pairs within detected cNEs. Over the 16 columnar recordings, pairs of neurons that were members of the same cNE had significantly higher noise correlations than pairs of neurons not within the same cNE (Figure 2.S6). This is consistent with common synaptic input being one potential source of noise correlations (Kanitscheider et al., 2015; Kohn and Smith, 2005; Shadlen and Newsome, 1998). Since neurons in the same cNE have reliable synchronous activity, it is likely that they receive more similar common synaptic inputs than neurons not within the same cNE.

To determine the kind of information being enhanced by cNEs, we compared the STRFs (calculated via spike-triggered or event-triggered averages) of cNEs and their member neurons. Spike or event trains were subsampled so that they had equal numbers of spikes or events. cNE STRFs had stronger excitatory and inhibitory subfields than their member neurons (Figure 2.10A). Compared to neuronal STRFs, cNE STRFs had higher peak-trough differences (PTD; $p < 0.001$, Wilcoxon signed-rank test; Figure 2.10B) and MI ($p < 0.001$, Wilcoxon signed-rank test; Figure 2.10C). Thus, cNE STRFs conveyed more information about the stimulus than neuronal STRFs by increasing the signal-to-noise ratio represented in their STRFs.

To show that the increase in information conveyed by cNE STRFs was not trivially due to the fact that cNEs integrate over multiple neurons with similar STRFs and, thus, must convey more information, we also compared multi-unit STRFs of cNE member neurons against multi-unit STRFs of random group of neurons (Figure 2.S7A). Across the entire population, the STRF PTD for cNE members was significantly higher than that of the STRF PTD for random groups of neurons ($p < 0.001$, Wilcoxon signed-rank test; Figure 2.S7C). The STRF MI of cNE member neurons was also significantly higher than that of random groups of neurons ($p < 0.001$,

Wilcoxon signed-rank test; Figure 2.S7E). For both STRF PTD and MI, we found that the distribution of UI values was significantly skewed towards 0, and was significantly different from a uniform distribution, which would be obtained if cNEs were composed of randomly selected neurons from the population recordings ($p < 0.001$, Kolmogorov–Smirnov test).

Taken together, the MI results (Figures 2.9, 2.10, 2.S5 and 2.S7) suggest that cNEs enhance information processing by reducing the number of stimulus-independent events or increasing the ratio of stimulus-dependent events. This is reflected in the increase in response reliability and magnitude of the excitatory and inhibitory subfields in the cNE STRF compared against its member neurons.

cNE events are not fully accounted for by stimulus synchronization or receptive field overlap

Finally, we investigated whether the coordinated responses observed in detected cNEs can be accounted for by stimulus-driven response synchronization resulting from either receptive field overlap of cNE member neurons or by second-order (pairwise) correlations. First, we repeated a 10-minute ripple noise stimulus 15 times and examined units that responded to each stimulus repetition with firing rates that varied by less than 30% across adjacent ripple presentations. We isolated 24 units that satisfied these criteria (Figure 2.S8). For the simultaneously collected ('real') spike trains, we randomly selected 15 of the 24 neurons from one repetition at a time and calculated cNEs (for a total of 100 unique combinations). For the non-simultaneous ('surrogate') spike train matrix, we selected the same 100 unique combinations of 15 neurons, but the spike train of each neuron was extracted from a different stimulus repetition (Figure 2.S8). If the cNEs resulted purely from stimulus-driven synchronization, the simultaneous/'real' spike train matrices should have cNE statistics that were similar to that of the

non-simultaneous/'surrogate' spike train matrices. We found, however, that the 'surrogate' spike

train matrices had fewer cNEs (Figure 2.11A) and that the detected cNEs had fewer member

neurons (Figure 2.11B). This suggests that even though some stimulus-driven synchronization,

including synchronization by features of the stimulus not represented in STRFs, is present in the

non-simultaneous spike trains, it is not the dominant contributor to synchrony observed in the

simultaneous recordings.

We next modeled spike trains to confirm that receptive field similarities among cNE

members and second-order correlations did not account for the high degree of observed cNE

synchrony. In a first model ('shuffled'), we circularly shuffled each of the simultaneously

recorded spike trains, breaking the temporal relationships between neurons as well as the match

between stimulus and receptive field at each spike's occurrence while preserving inter-spike

intervals (ISIs) for each neuron (see Materials and methods).

In a second model ('preserved receptive field' or 'PR'), we again shuffled the spike times

of neurons but in a way that preserved the match between stimulus and receptive fields at the

new spike times (Figure 2.S9A). To achieve that, we calculated the similarity between stimulus

segments and neuronal STRFs (projection value) and shuffled spikes between time bins with

similar projection values (Figure 2.S9A, see Materials and methods). This procedure created

surrogate spike trains that had shuffled spike times but similar receptive fields to the original

spike trains.

In the third model, we used a parametric analysis to investigate if second-order

correlations were sufficient to describe the degree of synchrony we see in cNE members. We

used a 'dichotomized Gaussian' ('DG') model (Macke et al., 2011, 2009, see Materials and

methods), which creates spike trains that match both the firing rate of each neuron and the PWC

between each pair of neurons. Thus, this model accounts for the overall activity of neurons up to second-order correlations but does not contain information about higher-order correlations.

To validate that the models behaved as expected, we calculated the STRFs for each neuron in each dataset via spike-triggered averaging. The 'PR' model preserved STRFs relative to STRFs derived from the real spike trains (Figure 2.S10A), indicating that the long-term receptive field information was accurately preserved. As expected, the 'shuffled' and 'DG' models did not provide well-defined STRFs (Figure 2.S10A). When STRF similarity between the three different models and the real spike train was calculated, the 'PR' model had a high median correlation value (~0.74) while the 'shuffled' and 'DG' models had values close to zero (Figure 2.S10B). We also calculated PWCs for pairs of neurons in the real spike train and the models and found that the 'shuffled' and 'PR' models captured very little of the correlation between pairs of neurons in the real spike train, while the 'DG' model only captured the correlation between pairs of neurons in the real spike train at zero delay (Figure 2.S10C). This is expected because the 'shuffled' model breaks up all inter-neuronal correlations; the 'PR' model only preserves receptive fields, and PWCs between the temporal profiles of STRFs are broad (Figure 2.6); and the 'DG' model is based on correlations between pairs of neurons. Since the correlation value is a single value estimated across the entire spike train, it does not take into account intra-spike train synchrony and does not model delay-dependent temporal correlations. However, since PWCs between pairs of neurons in the same cNE are sharp, have mean sharpness values of 3.39 ± 1.21 ms (Figure 2.6G – 6J, Figure 2.S4), and mean absolute peak delays of 0.76 $\pm$ 1.74 ms, modeling spike trains using the 'DG' model at 5-ms temporal resolution effectively encompasses the temporal correlations between neurons from the same cNE (Figure 2.S10C).

Surrogate spike trains from the 'shuffled' and 'PR' models had significantly fewer cNEs (Figure 2.11C) than the real data. On the other hand, the 'DG' model overestimates the number of detected cNEs (Figure 2.11C). Despite that, all three models, including the 'DG' model, resulted in cNEs that were smaller in size (i.e., fewer member neurons, Figure 2.11D) than the real data (n = 16 penetrations; mean values for each penetration were calculated over 200 iterations for each model). These results imply that the observed synchrony represents unique network events that cannot be explained simply by second-order correlations or by synchrony reflective of the underlying similarity in the receptive fields of neurons.

Finally, to estimate the degree of coincident firing between cNE member neurons, we calculated coincidence ratios (CRs) between groups of neurons. The CR was defined as the ratio of the number of time bins in which all neurons in the pre-defined group spike together, to the number of time bins in which the neuron with the lowest firing rate is active. When the whole group is active each time the neuron with the lowest firing rate spikes, the CR value will be one (Figure 2.S11A). This measure allows us to compare the probability that several neurons in a cNE will fire together and can be thought of as a generalization of the dot product between two spike trains to multiple spike trains. Using this measure, we directly compared our responses across different presentations of the same stimulus and the simulated spike trains (1000 iterations each) to the real data (Figure 2.11E and F), by estimating CRs for 400 unique groups of 2, 3, 4, and 5 neurons within the same cNE (Figure 2.S11B). Since neuronal identity was preserved in all simulations and the repeated stimulus data, we normalized each group of neurons by the CR of the real data, to determine how well the synchrony from simulated data approximated the synchrony from the real data. The normalized CR for the repeated stimulus data ('repeat'), 'shuffled', 'PR' and 'DG' models were significantly less than one for all groups (Figure 2.11E

and F). For groups containing more than three members, the CR values for most control models were always near or at 0 (not shown). These results suggest that member neurons of cNEs have more coincident activity than expected, further supporting the hypothesis that coordinated cNE events are driven by higher-than-second-order correlations and cannot be explained simply via stimulus-induced synchronization or shared receptive field properties, including those not reflected in STRFs.

2.4 Discussion

A key goal of this study was to test the hypothesis that auditory processing of an ongoing stimulus is better interpreted as sequences of coordinated activity among small subsets of neurons in the cortical column, rather than independent spikes distributed across the same subsets of neurons. We were able to identify, for each columnar recording in rat AI, several groups of neurons with coordinated activity (cNEs) in response to broadband stimuli with dynamic spectro-temporal envelopes. Each identified cNE comprised a subset of ~15-20% of the simultaneously recorded neurons. The spatial distribution of neurons within most cNEs was contained within < 250 □m of columnar depth, approximately corresponding to the width of granular and supragranular layers. However, several cNEs encompassed neurons distributed across multiple layers.

We isolated and identified groups of neurons in AI columns that showed coordinated spiking activity by using dimensionality-reduction techniques (Lopes-dos-Santos et al., 2013). The group spiking events associated with each cNE were more reliable and conveyed more information about the stimulus than spikes of individual neurons or random groups of neurons (matched to cNE size). We verified that identified cNEs were not chance constructs as a result of stimulus-driven synchronization alone by showing that the frequency of occurrence of cNE events was higher than predicted from basic receptive field similarities within the column. Since a high percentage of cNEs had matching neuronal composition during evoked and spontaneous activations, these findings support the concept that cNEs are stable local units of information processing in the cortical column.

Challenges in detecting cNEs in AI

Detecting or identifying cNEs in sensory cortex is difficult due to the multidimensional nature of population recordings. The first difficulty is implementing a method that may estimate the number and members of ensembles from high-dimensional data. Another difficulty is that any method will rely on responses occurring over chosen time windows and may be confounded with the temporal information processing performed by the region of interest. Non-auditory responses are more easily examined for ensemble activity because stimuli that drive the neurons in these regions are not functions of time. While the visual scene may vary over time, it is not dominated by rapid temporal information. Since visual images are predominantly static, visual cortical studies have used two-photon calcium imaging, which measures a response signal over an integration window of 125-250 ms (Table 2.1; Miller et al., 2014; Carrillo-Reid et al., 2015). Such imaging approaches and time windows cannot be utilized in the auditory cortex, since auditory cortical neurons are responsive to envelope modulations occurring over tens of milliseconds. Only static stimuli, such as pure tones or spectral ripple stimuli, could be utilized for auditory imaging studies. Thus, it is uniquely challenging to identify ensembles in auditory cortical regions due to the confound between analysis windows and temporal stimulus preferences.

Once estimated, interpreting auditory cortical ensembles poses additional challenges. Visual cortical receptive fields are often estimated over long time delays, while auditory cortical receptive fields based on dynamic stimuli necessarily contain more precise timing information. Thus, ensemble synchrony might be due to either stimulus-driven receptive field similarity, or it may be due to specific network configurations that favor joint activity, such as high anatomical or functional connectivity. To determine if ensembles represent more than the consequences of

receptive field similarity, we utilized a variety of controls and showed that while ensemble

activity can be related to stimulus preferences, it cannot be wholly explained through similarities

in basic receptive field processing (Figures 2.6 and 2.11). Furthermore, the observation that

spontaneous activity engages local networks that are very similar to those engaged by externally

driven activity (Figure 2.8) supports the idea of the utilization of basic network building blocks

(Hebb, 1949) in the transition from representation to interpretation to behavior. cNEs appear to

be stable, functionally connected subsets of neurons that are embedded in broader columnar and

horizontal networks whose activity reliably signifies specific auditory information.

Detection and identification of cNEs

Even though technologies for recording large populations of neurons have been growing

exponentially (Blanche et al., 2005; Du et al., 2011; Rios et al., 2016; Stevenson and Kording,

2011), the ability to make meaningful observations about network activity or ensembles of

neurons has been hampered by the lack of analytic tools to detect and identify such constructs.

New methods have more recently been proposed to identify functionally meaningful groups or

ensembles of neurons (Billeh et al., 2014; Gourévitch and Eggermont, 2010; Lopes-dos-Santos et

al., 2013; Miller et al., 2014; Peyrache et al., 2010; Pipa et al., 2008), either as neurons that fire

synchronously within a predefined time window (Harris et al., 2003), or as neurons with fixed

sequential firing patterns over time, akin to synfire chains (Ikegaya et al., 2004). Columnar AI

responses to stimuli (Atencio and Schreiner, 2010a; Kaur et al., 2005) or thalamocortical

excitation (Krause et al., 2014) have been previously shown to occur within a narrow time

window (<10 ms), and we have defined groups of cooperating neurons as neurons that spike

synchronously within a 10-ms time window. Previous studies have demonstrated that the

analysis of (near) zero-lag synchrony can provide rich insights into network function (Atencio

and Schreiner, 2013; Harris, 2005; Nicolelis et al., 1995), and is more easily identified by statistical methods than when considering sequential firing patterns over time (Peyrache et al., 2009). Finally, if we consider that cNEs are dynamically assembled, and that connectivity between cNE members can change over time (Buzsáki, 2010; Reimann et al., 2017), it is essential to follow them at rapid time scales, which is intractable if large time lags are considered.

We made use of dimensionality reduction techniques, namely PCA and ICA (Lopes-dos-Santos et al., 2013), to detect cNEs. This approach extracts the number of significant cNEs, assigns weights to each neuron's contribution to each cNE, and gives an intuitive measure of cNE activity. The identification of significant cNEs was also calculated using established results from random matrix theory (Marčenko and Pastur, 1967), and therefore reduces the reliance on surrogate randomization methods. The method is mostly linear; ICA uses nonlinear equations to quantify negentropy (a measure of distance between a distribution and the Gaussian distribution), but this is highly optimized by the fast ICA algorithm (Hyvärinen and Oja, 1997), making it significantly faster than most nonlinear algorithms. Since we were focusing on the cortical column, we limited our examination to synchronous spiking within 10-ms time windows. The short time window that we used de-emphasizes the need to assess neuronal ensembles as a function of the lag between time bins (Russo and Durstewitz, 2017). Furthermore, since our neuronal firing rates are approximately stationary, we did not correct for nonstationarity (Russo and Durstewitz, 2017).

cNEs are distinct from cell assemblies or simple populations of neurons

Because we defined cNEs via synchronous firing within temporal bins, we termed these constructs 'coordinated neuronal ensembles' instead of the more commonly used term 'cell

assemblies'. Cell assemblies, as first defined in Donald Hebb's seminal work (Hebb, 1949), were hypothesized to be fixed units of neurons that are highly and strongly interconnected, and that the activation of a few of its members would be sufficient to reliably cause the activation of the whole assembly. Hebb's cell assemblies also relate to large scale behavioral phenomenon (Hebb, 1949) and have been proposed to require a read-out mechanism (Buzsáki, 2010). Since our groups of neurons have not been linked to global aspects of behavior, and do not include the activity of potential target neurons outside the column, we chose to use the term 'coordinated neuronal ensembles' to distinguish it from the more functionally related term 'cell assemblies'.

Another distinction is that the identity of neurons contributing to each cNE is important and their roles should be distinguished from neurons accumulated for the purpose of deriving simple population codes (Bizley et al., 2010; Chakraborty et al., 2007; Furukawa et al., 2000; Rodgers and DeWeese, 2014). We showed that neurons classified as belonging to the same cNE were significantly more synchronous with one another than those from simultaneously recorded non-members (Figures 2.6 and 2.11). We also showed that identified cNEs tended to convey more information about the stimulus than simultaneously recorded, arbitrary groups of neurons of the same size (Figure 2.S5 and Figure 2.S7). Altogether, we demonstrated that cNEs are a subset of simultaneously-recorded neurons that have sharply synchronized activity with enhanced information processing abilities.

cNE activity does not require auditory stimuli

We demonstrated that the majority of cNEs (~72%) for spontaneous and evoked activity were significantly matched in member neuron composition (Figure 2.8). This implies that cNEs are stable functional entities that reflect stimulus selectivity but do not require external stimuli to be activated. This also supports previous findings in sensory cortices that patterns of spontaneous

activity can be similar to those of evoked activity (Jermakowicz et al., 2008; Luczak et al., 2009), albeit at much finer timescales. These underlying functional network units have also been hypothesized to reflect constraints that cortical organization has imposed on the local propagation of neuronal activity (for in-depth reviews see Luczak et al., 2015; Luczak and MacLean, 2012).

The identity and stability of cNEs across spontaneous and evoked activity has only been established over relatively short periods of time (usually less than an hour under anesthesia). Previous studies have shown that exposure to specific stimuli tended to bias spontaneous activity towards the firing patterns seen in activity evoked by those specific stimuli (Bermudez Contreras et al., 2013; Eagleman and Dragoi, 2012; Han et al., 2008). This suggests that cNEs might similarly be mutable by dominant aspects of the auditory environment. By extension, we speculate that cNE composition and functional identity can be influenced by learning or any plasticity-inducing experiences. These potential effects need to be addressed in future work.

cNEs cannot be accounted for by stimulus-driven synchronizations

A fundamental question regarding the identified cNEs is whether they were constructs that arose from the trivial fact that their member neurons all had similar receptive fields and, therefore, were simultaneously driven by the same stimulus. AI neurons in a column have been previously shown to have fairly closely related STRFs (Atencio and Schreiner, 2010a, 2010b; Guo et al., 2012; Wallace and Palmer, 2007), making stimulus-driven synchronization a likely contributor to the observed synchrony of cNEs.

We demonstrated, however, that cNE membership was not solely accounted for by receptive field overlap using several independent methods. Spike train PWCs between pairs of

neurons in the same cNE are significantly narrower than that of STRF PWCs (Figure 2.6). Assuming that pairs of neurons are simply synchronized by the stimulus, it follows that the spike PWCs would have similar widths to that of their receptive field PWCs. This is often true for pairs of neurons not within the same cNE (Figure 2.6D – F), putative examples of neurons that are only synchronized due to their receptive field overlap.

Moreover, we showed that cNEs detected during evoked activity were often also detected during spontaneous activity (Figure 2.8). Consequently, the evoked cNEs that had significant matches to spontaneous cNEs did not require input synchronization by a stimulus.

Another argument against stimulus-driven synchrony as the main reason for the observed synchrony is the lower degree of synchrony between independently firing neurons than that of the coordinated firing of cNE neurons (Figure 2.S8). The reduced synchrony, expressed in terms of lower cNE numbers and smaller cNE sizes for non-simultaneous spike matrices ('surrogate' spike matrices; Figure 2.11A and B), indicates that basic receptive field similarity and overlap are insufficient to account for the observed coordinated firing patterns.

Finally, we modelled stimulus-driven synchrony with the 'preserved receptive field' ('PR') model (Figure 2.S9A). The idea behind this model was to shuffle the spike times of each neuron only among time bins associated with the same degree of similarity between the stimulus and receptive field (projection values). Given the probabilistic nature of neuronal spiking even for the highest projection values, the 'PR' model had, unsurprisingly, a lower number of cNEs per spike matrix and fewer neurons per cNE (Figure 2.11C and D). The CR for the 'PR' model was also significantly less than that of the real data for all sizes of sets of neurons (Figure 2.11E and F). These observations further support the notion that multi-neuronal coordinated activity requires conditions beyond receptive field overlap. The linear STRFs used here are, of course,

not a complete representation of the stimulus preference of a neuron. Additional stimulus features may activate neurons, often in nonlinear fashion, as has been demonstrated using multi-filter approaches (Atencio et al., 2008; Harper et al., 2016). Even though the main (linear) filters appear to carry the most information, we cannot completely exclude the possibility that shared but 'hidden' receptive field properties, not reflected in either basic or advanced multi-filter STRFs, contribute to the formation and activation of cNEs. However, these contributions cannot be very substantial, as indicated by our model-free analysis of cNE synchrony (Figure 2.11A, B and Figure 2.S8). Further determinants of cNE events are likely governed by more globally defined network states, including top-down influences (Harris and Thiele, 2011; McGinley et al., 2015; Okun et al., 2010).

cNEs reflect higher-order correlations

To test if second-order correlations were sufficient to account for the synchrony observed between cNE neurons, we used the 'DG' model (Figure 2.11C–F). The 'DG' model is fully specified by neuronal firing rates and the covariance of actual spike trains. It produces a set of spike trains that reproduces these statistics, but does not contain information about higher-order correlations (Macke et al., 2011, 2009, see Materials and methods). The 'DG' model was able to capture correlations between all possible pairs of neurons but was unable to directly replicate the correlations between three or more neurons (Figure 2.S10C). The lack of higher-order correlations resulted in smaller cNE sizes (Figure 2.11D) and lower CRs for three or more neurons that were members of the same cNE (Figure 2.11F). The lower CR for pairs of neurons from the same cNE (Figure 2.11E) shows that the 'DG' model incompletely captures the columnar synchrony of AI neurons (Figure 2.S10C).

That second-order (pairwise) correlations were unable to recapitulate the synchrony seen between cNE neurons should not be surprising. The premise of the cNE algorithm (Lopes-dos-Santos et al., 2013) is to detect groups of neurons with reliable synchronous activity and identify these synchronous events, which are likely driven by higher-order correlations (Montani et al., 2009; Ohiorhenuan et al., 2010; Yu et al., 2011). These higher-order correlations could be a result of common synaptic input (Shadlen and Newsome, 1998), often shared by neighboring cortical neurons (Song et al., 2005; Yoshimura et al., 2005), or a result of global fluctuations driven by top-down inputs (Goris et al., 2014; Schölvinck et al., 2015). cNEs, and the statistics that we have used to describe them, are hence a useful tool to study higher-order correlations and the biological phenomena that drive them.

cNEs are functional units that reliably represent and transmit information

Our data suggest that cNEs may be conveyors of specific environmental information that is either externally triggered or internally engaged. Acoustically driven cNEs conveyed more information in their response events than individual neurons or randomly selected groups of neurons. Since information necessarily includes timing information, this implies that cNEs are tightly locked to stimulus features that trigger ensemble responses. Additionally, since information increases as response precision increases (Figure 2.9 and Figure 2.S5), our results show that cNEs are more temporally reliable.

One way to examine a subset of stimulus features that trigger ensemble responses was to compare cNE STRFs against neuronal STRFs. To that end, we showed that cNE STRFs had enhanced excitatory and inhibitory subfields when compared to those of member neurons or randomly selected groups of neurons (Figure 2.10 and Figure 2.S7). This phenomenon could be explained by a reduction in the number of events during periods where the stimulus and the

STRFs are relatively uncorrelated, or by an increase in the number of events that respond to stimulus segments with high projection values. We propose that both ideas are likely and contribute to the observed phenomena. cNE events only require a subset of member neurons to fire together (Figure 2.4D and E). Furthermore, from one cNE event to another, the neurons that contribute to any event can vary among the cNE members. Given the probabilistic nature of neuronal spiking even for the highest projection values, it suggests that cNEs integrate complementary spiking information from member neurons during stimulus presentations to increase the number of events that respond to stimulus segments with high projection values and to better encode the stimulus. On the other hand, spikes that occur during stimulus segments with low projection values do not tend to be synchronous with spikes from other neurons and are less likely to be associated with cNE events (see Figures 2.4D, E, 2.9A and B).

Altogether, cNEs appear to more reliably encode features of auditory stimuli. If one purpose of synchronous firing by groups of neurons is to produce a more efficacious downstream effect than can be accomplished by a single neuron (Buzsáki, 2010), then cNEs may be the functional units that serve to maintain and/or enhance the fidelity of the encoding of relevant sound features. They are likely part of larger networks that can span more than a column and extend over several hierarchical levels. The relative constancy of cNEs can provide stable spatio-temporal firing patterns with transient multi-neuronal synchrony that may, among other functions, signal salient spectro-temporal events (Hopfield and Brody, 2001) or contribute to the learning and extraction of stimulus categorization (Higgins et al., 2017).

Conclusion

In recent years, studies have identified neuronal ensembles or cell assemblies based on sophisticated statistical analyses and have ascribed functions to them based on observations of

the behaviors of the identified constructs (Bathellier et al., 2012; Ikegaya et al., 2004). We extend these results to columnar AI responses by identifying cNEs and rigorously examining alternative hypotheses for their generation. Synchrony driven by the overlap of STRFs appears to be insufficient to account for the synchrony observed between neurons in the same cNE. Since we have established that AI cNEs are also found in spontaneous activity and cannot be reduced to stimulus synchronization, further elucidating the formation of cNEs and their roles in auditory information processing and transmission are necessary next steps.

2.5 Materials and methods

Electrophysiology

All experiments were approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco. Female Sprague-Dawley rats (200 – 300 g, 2 – 3 months old; RRID:MGI:5651135) were anesthetized with a mixture of ketamine and xylazine, supplemented with dexamethasone, atropine and meloxicam. A tracheotomy was performed to stabilize the rat's breathing. Lidocaine was injected at the sites of the tracheotomy and craniotomy before each incision. The skin, bone and dura over the right auditory cortex were removed and the brain was kept moist with silicone oil. A cisternal drain at the cisterna magna was performed to keep the brain from swelling. The primary auditory cortex (AI) was located by using multiunit responses to pure tones of different frequencies (1 – 40 kHz) and intensities (10 – 80 dB). Regions with short latencies (10 – 30 ms) and a tonotopic gradient in the rostrocaudal axis were identified to be AI (Polley et al., 2007). Recordings for all data except Figure 2.8 were made using a 32-channel probe (a1x32-poly3, NeuroNexus, Figure 2.1A), and recording data for Figure 2.8 was made using a 64-channel probe (H3, Cambridge NeuroTech), inserted perpendicular to the cortical surface to a depth of approximately 800 or 1400 µm respectively using a microdrive (David Kopf Instruments). Neural traces were band-pass filtered between 500 and 6000 Hz and were recorded to disk at 20 kHz sampling rate with an Intan RHD2132 Amplifier system. The 32-channel traces were spike-sorted offline with a Bayesian spike-sorting algorithm that included a spike-waveform decomposition procedure to extract single-unit spike timing (Lewicki, 1994). Neurons recorded on neighboring electrode contacts were identified via cross-correlation analysis. The 64-channel traces were spike-sorted offline using MountainSort, a fully automated spike sorter (Chung et al., 2017).

Stimulus

The stimulus was either a dynamic moving ripple or a ripple noise (Atencio et al., 2008; Escabi and Schreiner, 2002). The dynamic moving ripple (DMR) was a temporally varying broadband sound (500 Hz – 40 kHz) made up of approximately 50 sinusoidal carriers per octave, each with randomized phase. The maximum spectral modulation frequency of the DMR was 4 cycles/oct, and the maximum temporal modulation frequency was 40 cycles/s. The maximum modulation depth of the spectro-temporal envelope was 40 dB. Mean intensity was set at 30–50 dB above the average pure tone threshold within a penetration. The ripple noise (RN) stimulus was the sum of 16 independently created DMRs. Both DMR and RN were presented as long and continuous 10-minute stimuli. Spike-triggered averaging of the DMR and RN was used to recover the spectro-temporal receptive field (STRF) of each neuron (Figure 2.1C). 5-second identical segments of either stimulus were also repeatedly presented 50 times, with 0.5 seconds of silence between each repetition (Escabi and Schreiner, 2002).

Detection of coordinated neuronal ensembles (cNEs)

All data analysis was performed in MATLAB (Mathworks). The cNE detection algorithm used in this study was based on dimensionality reduction techniques (Lopes-dos-Santos et al., 2013). Single-unit spike times were binned into 10-ms time bins and normalized via z-scoring. To determine the number of cNEs in each dataset, principal component analysis (PCA) was applied to the z-scored spike matrix to obtain the eigenvalue spectrum. Eigenvalues that exceeded the upper bounds of the Marčenko-Pastur distribution (Marčenko and Pastur, 1967) were deemed significant and represented the number of detected cNEs. The eigenvectors corresponding to significant eigenvalues were then processed using the fast independent component analysis (fastICA) algorithm, which approximates the distance to Gaussianity based

on negentropy (Hyvärinen and Oja, 1997). The resulting independent components (ICs) represent the contribution of each neuron to each cNE. The activity of each cNE was calculated by then projecting the ICs back onto the z-scored spike matrix (Lopes-dos-Santos et al., 2013). To validate cNE membership and threshold cNE activity, the rows of the binned spike matrix were circularly shuffled independently. PCA/ICA was then applied to the resulting shuffled spike matrices. Because correlations between neurons have been broken up by shuffling, no eigenvalues crossed the upper bounds of the Marčenko-Pastur distribution. However, the N-largest eigenvalues, where N was the number of significant eigenvalues derived from the original spike matrix, were processed using ICA. This process was iterated 100 times to get a normal distribution of IC weights for the shuffled condition. The upper and lower thresholds for neuronal membership to cNEs in a particular recording was set as ±1.5 standard deviations from the mean of this distribution. To threshold cNE activity, the ICs that were calculated from the original spike matrix were projected onto the shuffled spike matrices to obtain a normal distribution of cNE activity values. The threshold for considering a cNE to be active was defined by the 99.9th percentile of that distribution.

STRF mutual information (MI)

The estimation of the MI between the STRF and single spikes or events was based on previous studies (Atencio and Schreiner, 2013; Atencio et al., 2008). Each stimulus segment, *s*, that preceded a cNE event or neuronal spike was projected onto the STRF via the dot product $z = s \cdot \text{STRF}$. These projections were binned to form the probability distribution P(z|event). We then formed the prior probability distribution, P(z), by projecting all stimulus segments onto the STRF, regardless of event/spike occurrence. Both probability distributions were then normalized

to the mean, μ, and standard deviation, σ, by x = (z - μ) / σ, to obtain P(x) and P(x|event). The

MI between projections onto cNE STRFs and single events/spikes was then computed by:

$$I = \int dx P(x|event) log_2 \frac{P(x|event)}{P(x)}$$

Each information value was calculated using different fractions of the dataset for each

spike train. To accomplish this, the information values were calculated over 90, 95, 97.5, 99 and

100% of the number of cNE events. This random sampling was iterated 20 times for each

percentage. The means of the information calculated from the iterations of these percentages

were plotted against the inverse of the percentages (1/90, 1/95, etc.). The information values

were then extrapolated to infinite dataset size by fitting a line to the data and taking the ordinate

intersect as the information value for unlimited dataset size. This method was used to determine

the optimal time bin size for cNEs (Figure 2.5A) and to compare the STRF MI between cNEs,

neurons and randomly-selected subsets of neurons (termed 'random groups'; Figure 2.10 and

Figure 2.S7).

To determine the optimal time bin size for cNEs, we binned spike trains at multiple bin

sizes (2, 5, 8, 10, 15, 20, 35, 50 ms; Figure 2.5), and then applied the cNE detection algorithm.

To allow comparison between cNEs, we interpolated the estimated activity vectors to obtain 1-

ms resolution cNE activity vectors. For each of these vectors, we set a threshold such that each

cNE had 5000 events. We then calculated the STRF for each cNE via event-triggered averaging,

and the STRF MI for each bin size (Figure 2.5A).

MI analysis on 5-s repeats

The estimation of MI between the patterns of firing over repeated trials of the same stimulus and the average firing rate of a neuron or group of neurons was based on an earlier study (Brenner et al., 2000). The spike or event trains obtained from single neurons, cNEs or random groups were represented in a matrix where each row represented one of the 50 trials, and each column represented a 10-ms time bin at the same point in the stimulus in each of the 50 trials. The post-stimulus time histogram (PSTH) vector, or r(t), was computed by taking the average number of spikes across each column of the matrix. The vector, r(t), was averaged to get a scalar, $\bar{r}$, or the average firing rate of each neuron, cNE or random group. The MI between the patterns of firing and the average firing rate was then computed by:

$$I = \frac{1}{T}\int_0^T dt \left(\frac{r(t)}{\bar{r}}\right) log_2 \left(\frac{r(t)}{\bar{r}}\right)$$

Each information value was calculated using different fractions of the dataset for each spike raster. To accomplish this, the information values were calculated over 90, 95, 97.5, 99 and 100% of the number of time bins of r(t). This random sampling was iterated 20 times for each percentage. The means of the information calculated from the iterations of these percentages were plotted against the inverse of the percentages (1/90, 1/95, etc.). The information values were then extrapolated to infinite dataset size by fitting a line to the data and taking the ordinate intersect as the information for unlimited dataset size. This method was used to compare the MI provided by the rasters of responses to 5-s repeats of DMR or RN stimuli between cNEs, neurons and randomly-selected subsets of neurons (termed 'random groups'; Figure 2.9 and Figure 2.S5).

MI comparisons between cNE and neurons and cNE and random groups

For each MI comparison between a cNE and a member neuron (Figures 2.9 and 2.10), we sub-sampled either the neuronal spikes or the cNE events such that the number of cNE events was equal to the number of neuronal spikes.

For each MI comparison between a cNE and random groups (Figure 2.S5 and Figure 2.S7), we had to generate random groups and the spike trains of the multi-unit activity of cNEs and random groups. To generate random groups of neurons for cNE/random-group comparisons, we selected one neuron from the cNE it was being compared to and randomly sampled from the rest of the neurons within a penetration, while excluding other member neurons of the cNE. Random groups of neurons that had more than one neuron from one of the other cNEs in the same recording were excluded as well. If 500 unique random groups of neurons could not be obtained (depending on the number of recorded neurons and the number of cNEs), this procedure was repeated to include two neurons from the cNE it was being compared to and the threshold for neurons from other cNEs in the same recording was similarly increased to two. This procedure was repeated until we obtained 500 unique random groups of neurons for each random group-cNE comparison that had as few neurons from within the same cNE as possible.

To generate rasters or spike trains, we summed the spikes of all member neurons of cNEs and neurons of random groups. For each cNE comparison to 500 iterations of random groups, we sub-sampled the number of spikes in the cNE and each of the 500 random groups to the number of spikes of the cNE or random group with the fewest spikes. Any set of cNE and random groups with fewer than 200 spikes after sub-sampling were excluded from the MI analysis to limit spurious comparisons.

Constructing repeat ('surrogate') spike matrices and 'real' spike matrices

To verify that the synchrony observed between neurons in the same cNE cannot be explained via overlapping receptive fields or stimulus synchronization, we compared synchrony between neuronal spike trains from the same presentation of the stimulus ('real') against neuronal spike trains across different presentations of the same stimulus ('surrogate'). A brief overview of the process of constructing 'real' and 'surrogate' spike matrices is depicted in Figure 2.S8. From the 24 units that had reliable firing rates across all 15 repetitions of the 10-min RN stimulus, we randomly selected 100 unique 15-neuron combinations. For each combination, we constructed 15 'real' and 'surrogate' spike matrices. Each 'real' spike matrix was constructed by taking the spike trains of the neurons in each combination from one presentation of the RN stimulus. Since there were 15 presentations of the stimulus, there were 15 'real' spike matrices for each unique 15-neuron combination. Each 'surrogate' spike matrix was constructed by first randomly sorting each 15-neuron combination. The spike train for the first neuron in the combination was then taken from the first presentation of the stimulus, the second neuron from the second presentation, etc., to get one 'surrogate' spike matrix. This was repeated 14 additional times, each time by circularly shifting the randomly sorted 15-neuron combination vector by 1, to get 15 different 'surrogate' spike matrices for each 15-neuron combination. We then calculated the number of cNEs detected and mean cNE size for each of the 'real' and 'surrogate' spike matrices. Finally, by averaging across the 15 spike matrices for each measure (number of cNEs or mean cNE size), label ('real' or 'surrogate') and unique combination, we obtained 100 data points for each measure and label. The results are plotted in Figure 2.11A and B.

Simulated spike trains

We used three simulated spike trains in this study to validate our observations, the 'shuffled', the 'preserved receptive field' ('PR'), and the 'dichotomized Gaussian' ('DG') models. The 'shuffled' model was computed by circularly shifting the rows of the spike matrix (where each row represents the binned spike train of one neuron) by a random number of bins, independently of other rows.

The 'PR' model was simulated based on each neuron's STRF. First, each stimulus segment was correlated with the STRF to obtain P(z), the distribution of projection values without regard to a spike. Next, the full range of P(z) was divided into 15 equal-sized projection value bins. Projection values in each bin represent stimulus segments that were similarly correlated with the STRF. Projection value bins contained multiple projection values, with each of them corresponding to a different time within the stimulus. To obtain the distribution of projection values that corresponded to a spike for each projection value bin, the number of projection values corresponding to spikes was divided by the total number of projections values for that bin regardless of whether a spike occurred. This normalized distribution was P(spk|z). To shuffle spike times with respect to P(spk|z), spikes corresponding to a specific projection value were randomly assigned to different time bins that corresponded to projection values in the same projection value bin. This shuffling procedure preserves the distribution of projection values while creating a random spike train (Figure 2.S9).The 'DG' model simulates spike trains that preserve the experimentally derived mean firing rate for each neuron as well as the pairwise correlation between each pair of neurons (Macke et al., 2009, 2011). Implementation of the 'DG' model with MATLAB can be found in Macke et al., 2009 and is available at https://github.com/mackelab/CorBinian. The 'DG' model requires and produces binary spike

trains and works best if the temporal resolution is high enough so that a maximum of one spike occurs in each time bin. However, since we also wanted to capture the temporal correlations between neurons, we simulated 'DG' spike matrices at 5-ms time resolution. We then binned the spike trains at a temporal resolution of 10 ms and completed the same analyses that were used to process the experimental data.

We iterated each of the 16 recordings 200 times per model and calculated the mean of cNE number and size for each model and recording. These statistics were compared with those of the real spike trains in Figure 2.11C and D.

Calculation of coincidence ratios (CRs)

The CR was defined as the ratio of the number of time bins in which all neurons in the pre-defined group spike together, to the number of time bins in which the neuron with the lowest firing rate is active. An illustration of this calculation is shown in Figure 2.S11A. 400 unique combinations of 2, 3, 4 or 5 neurons within the same cNE were randomly selected. For each combination of neurons, we obtained one 'real' and one 'repeat' CR values. 'Real' CRs were calculated by computing the CR between neuronal spike trains from one presentation of the stimulus. 'Repeat' CRs were calculated by computing the CR between neuronal spike trains from different presentations of the stimulus. Each combination of neurons also had 1000 'shuffle', 'PR' and 'DG' CR values. These were computed by iterating the shuffle and PR models 1000 times and calculating the CR for each iteration of each model. The median CR value for each model was calculated and compared to the single 'real' and 'repeat' CR values. An example of this comparison for a 2-neuron combination is shown in Figure 2.S11B. CR values for 2- and 3-neuron combinations are plotted in Figure 2.11E and F. Each set of CR values is normalized to

the 'real' CR value to determine how much of the real coincidence can be approximated by each

data manipulation (Figure 2.11E and F).

2.6 Acknowledgements and author contributions

Author contributions

Jermyn Z See, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Craig A Atencio, Conceptualization, Formal analysis, Investigation, Methodology, Writing—review and editing; Vikaas S Sohal, Supervision, Writing—review and editing; Christoph E Schreiner, Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Methodology, Project administration, Writing—review and editing

2.7 Figures
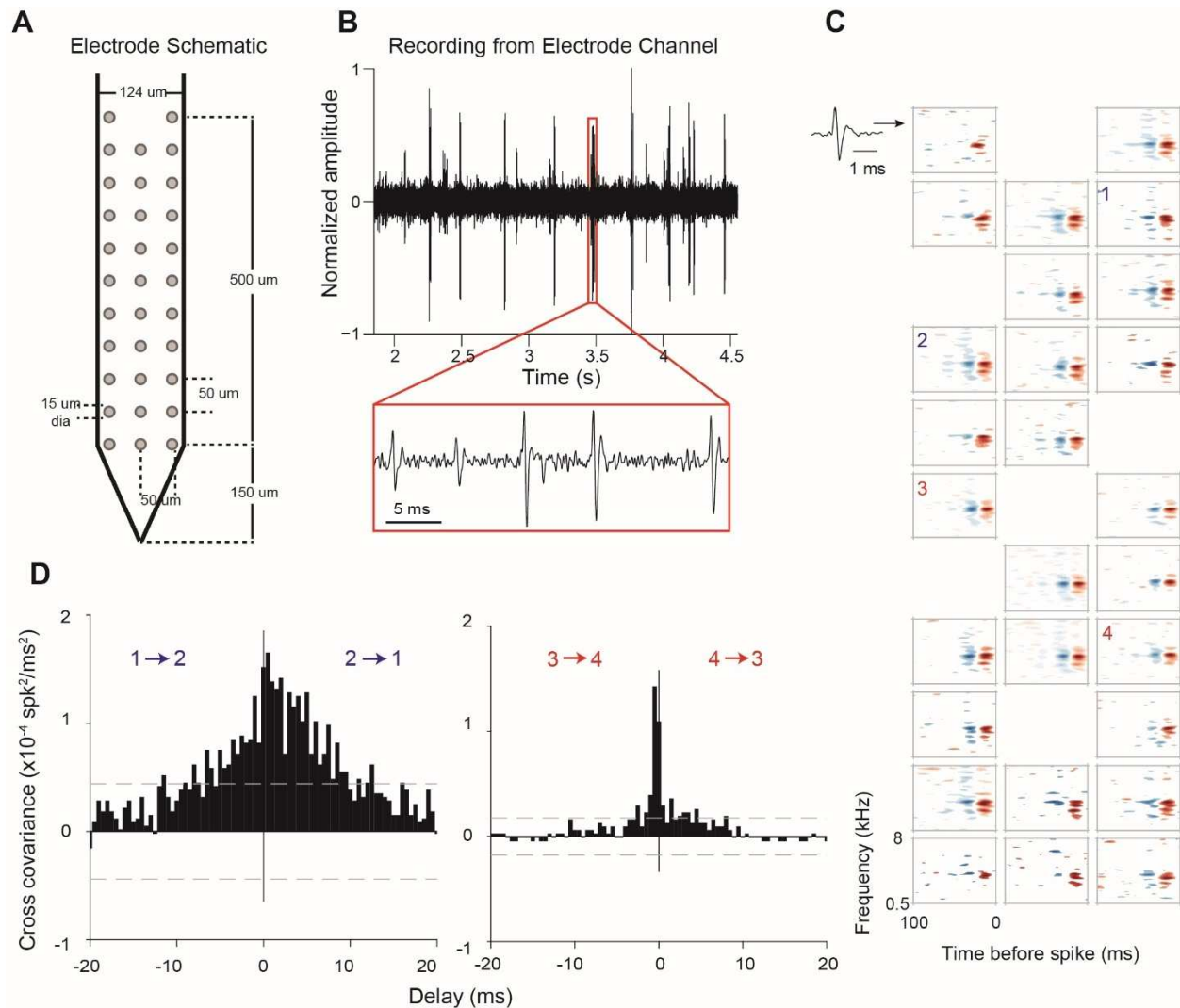
Figure 2.1. Electrode used, sample recording, STRFs, and pairwise-correlations.
(A) Electrode schematic. **(B)** Sample recording from one electrode channel. **(C)** Sample STRFs obtained after spike sorting. Numbers 1 - 4 indicate the positions and STRFs of pairs of neurons whose pairwise correlations are plotted in (D). **(D)** Example pairwise correlations from two pairs of neurons (1-2 and 3-4). Neurons 3 and 4 exhibit a sharper pairwise correlation than neurons 1 and 2 despite having approximately similar STRFs and pairwise distances. Dashed lines represent 99% confidence intervals.
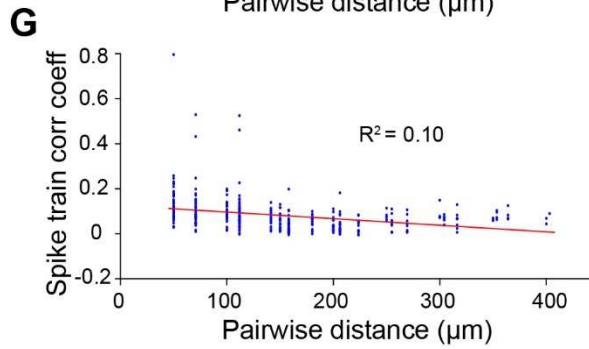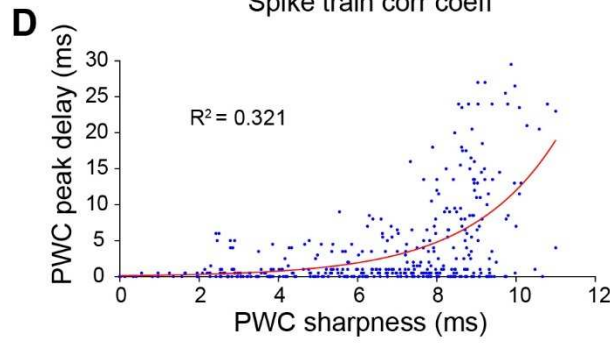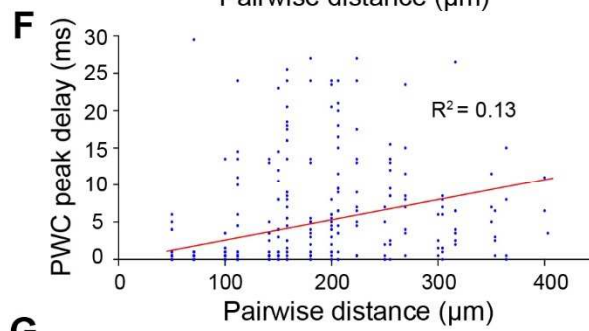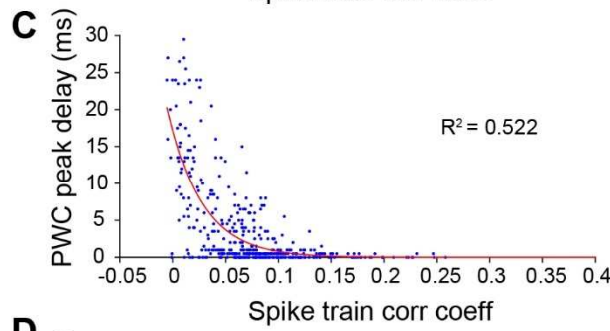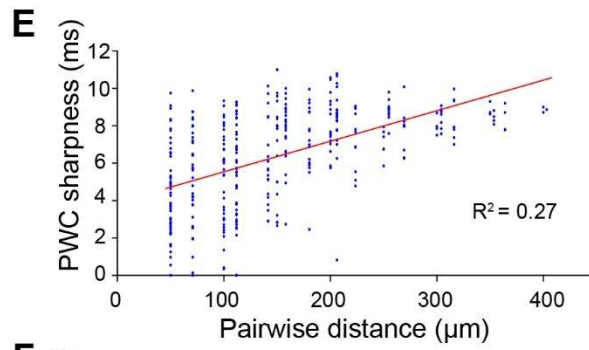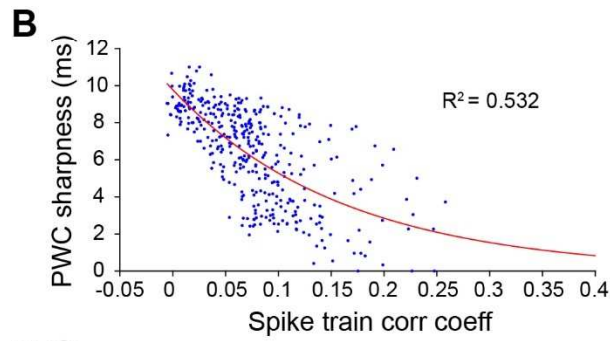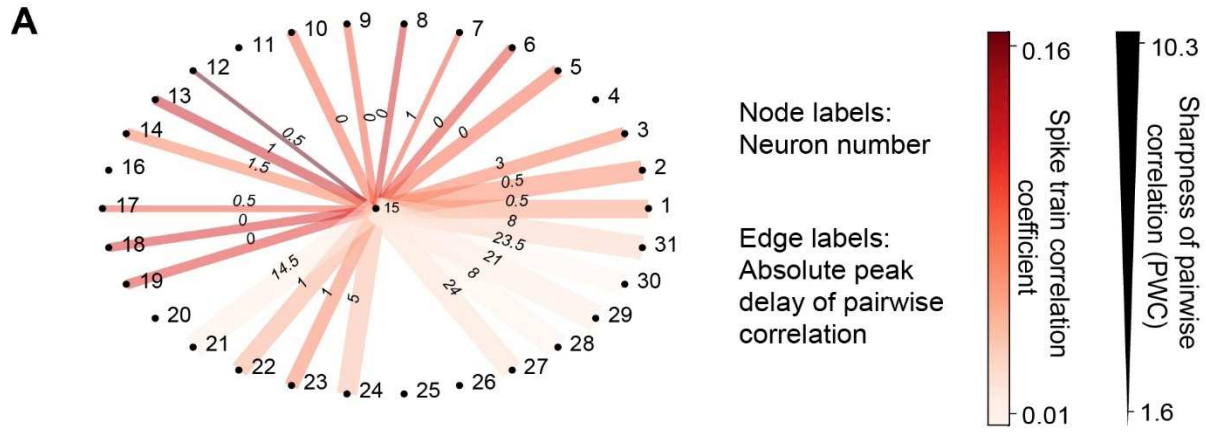
Figure 2.2. Methods for calculating neuronal correlation.
(A) An example neuron's (#15) correlation with other neurons in a dataset of 31 neurons measured in 3 different ways. The node labels (outer circle numbers) represent the neuron numbers. The edge labels (inner circle numbers) represent the absolute peak delay of the pairwise correlations (PWCs). The color of the edges represents the spike train correlation coefficient (Pearson's correlation) between pairs of neurons. The thickness of the edges represents the sharpness of the PWCs (see Figures 2.6G - I for an illustration of how this is computed). **(B)** Sharpness of PWCs against spike train correlation coefficient. **(C)** Peak delay of PWCs against spike train correlation coefficient. **(D)** Peak delay against sharpness of PWCs. All measures of correlation are highly correlated with one another. **(E-G)** Sharpness of PWCs (E), peak delay of PWCs (F) and spike train correlation between pairs of neurons (G) against pairwise distance. Pairwise distance explains only a small fraction of the variance seen in the different measures of correlation between pairs of neurons.
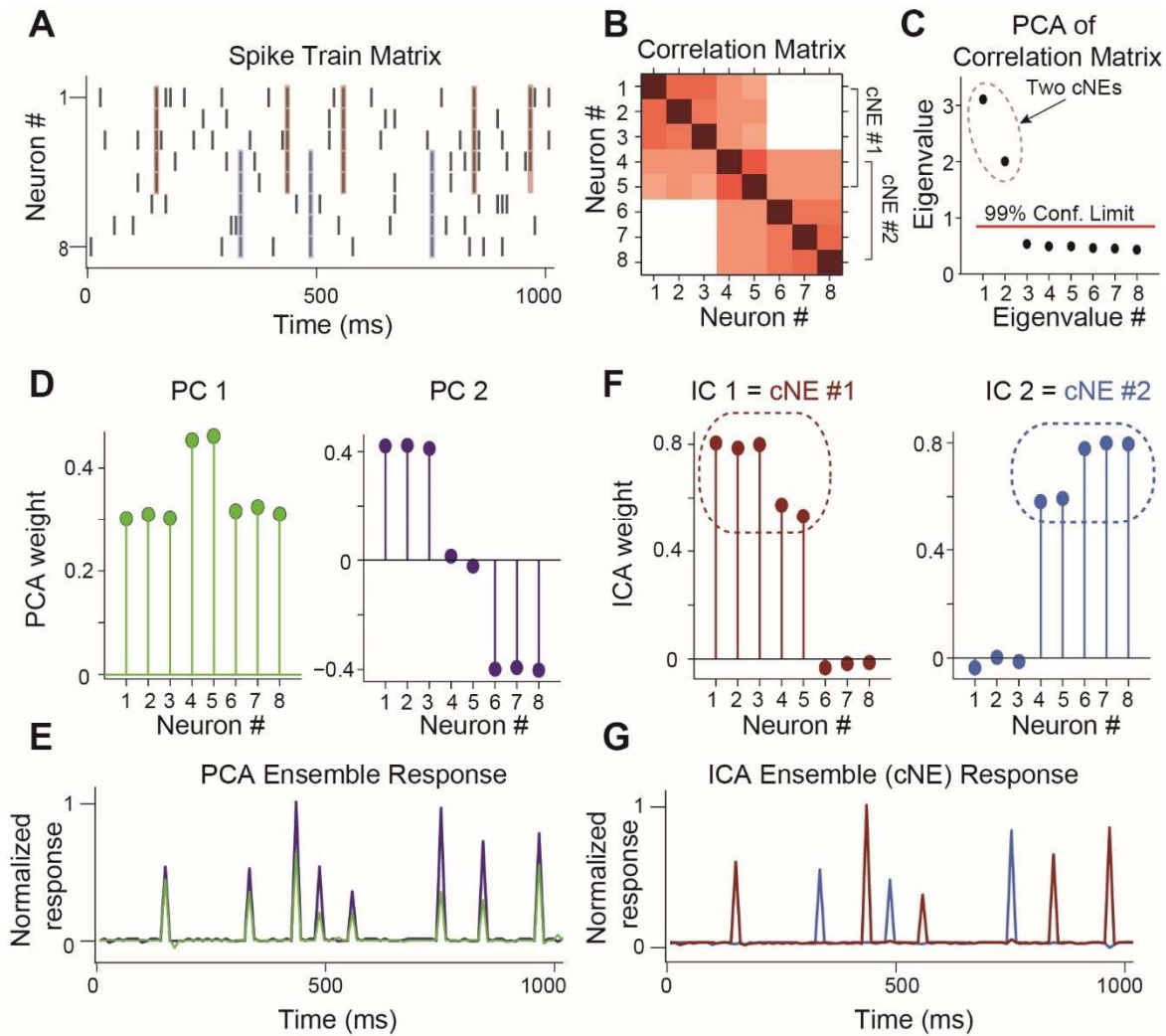
Figure 2.3. Simulated data illustrating the coordinated neuronal ensemble (cNE) detection algorithm.
(A) Raster of the simulated data. cNE #1 consists of neurons 1 to 5 while cNE #2 consists of neurons 5 to 8. Neurons 4 and 5 belong to both cNEs. Coincident spiking events between neurons in cNE #1 and cNE #2 are highlighted in red and blue respectively. **(B)** Autocorrelation matrix of spike train matrix from (A). **(C)** Eigenvalues from applying PCA to the autocorrelation matrix. The top two eigenvalues are significant and represent the number of cNEs. **(D)** The eigenvectors (or principal components, PCs) corresponding to the two significant eigenvalues from (C). The PCs do not represent the two cNEs denoted in (A) and (B). **(E)** Ensemble responses calculated from the projection of the PCs onto the spike train matrix. The PCA ensembles do not separate the activities of cNE #1 and #2 shown in (A). **(F)** Independent components (ICs) obtained after applying ICA to the two significant eigenvectors from (C). The ICs accurately represent the two cNEs denoted in (A) and (B). **(G)** Ensemble responses calculated from the projection of the ICs onto the spike train matrix. The ICA ensembles successfully resolve the activities of cNE #1 and #2 shown in (A).

56
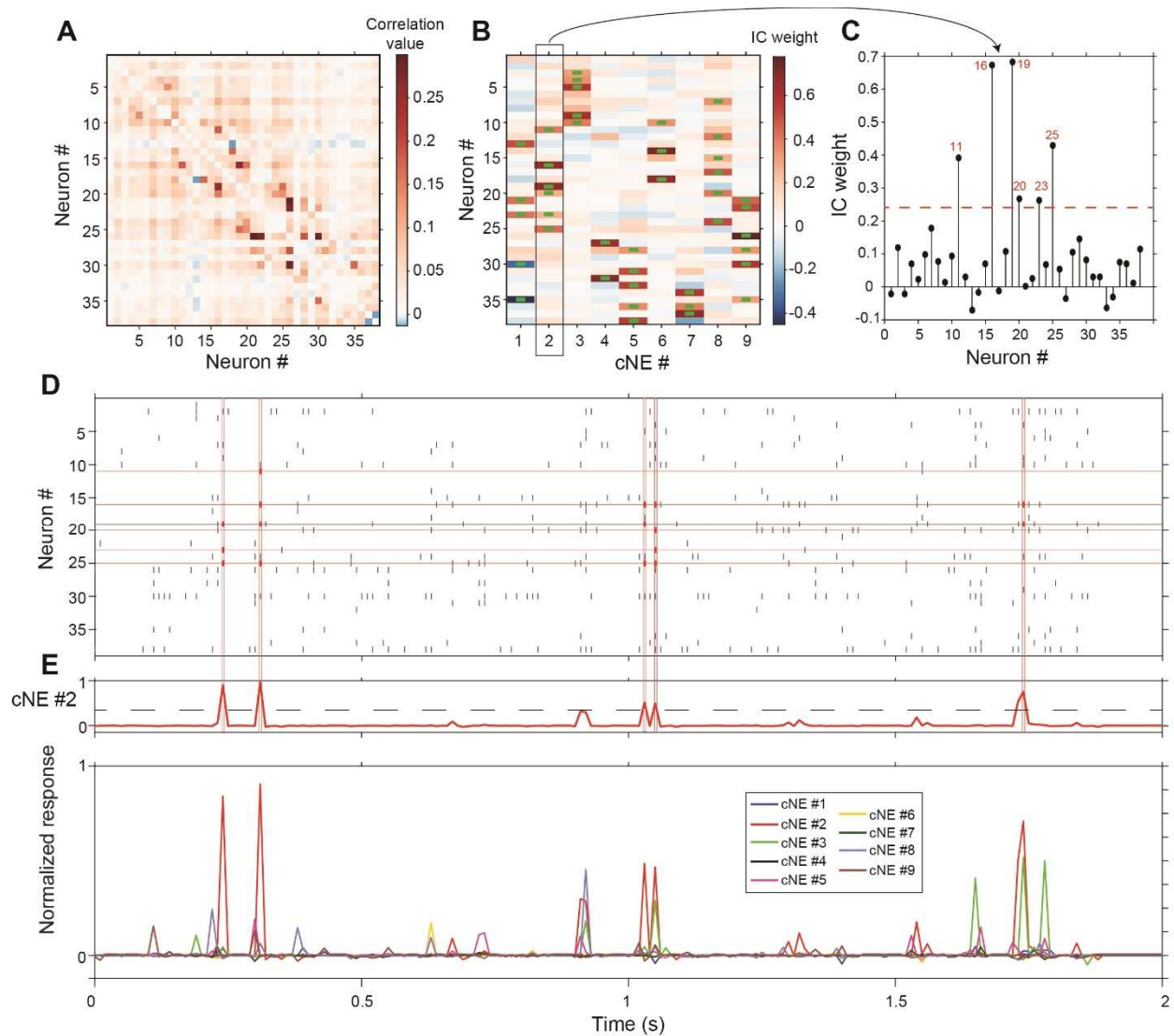
Figure 2.4. cNE detection algorithm for AI neurons.
(**A**) Autocorrelation matrix of spike train matrix from (D). The diagonal has been set to 0. (**B**) IC weight for each neuron in each cNE. Each column represents one cNE; each row represents one neuron. The IC weight represents the contribution of each neuron to each cNE. The green bars represent neurons that are part of a cNE. (**C**) IC 2 (cNE #2). The red dotted line is the threshold for cNE membership determined by Monte-Carlo methods. The red numbers are the numbers of neurons that are members of the cNE. (**D**) Spike raster of 2-seconds of real data. Spikes that contribute to instances of cNE #2 activity in (E) are in red. (**E**) (top) Trace of cNE #2 activity. The black dotted line represents the threshold for cNE #2 activity estimated via Monte Carlo methods. Peaks that cross the threshold correspond to the coincident neuronal spikes highlighted in red in (D). (bottom) Activity traces of the 9 cNEs recovered from this dataset.
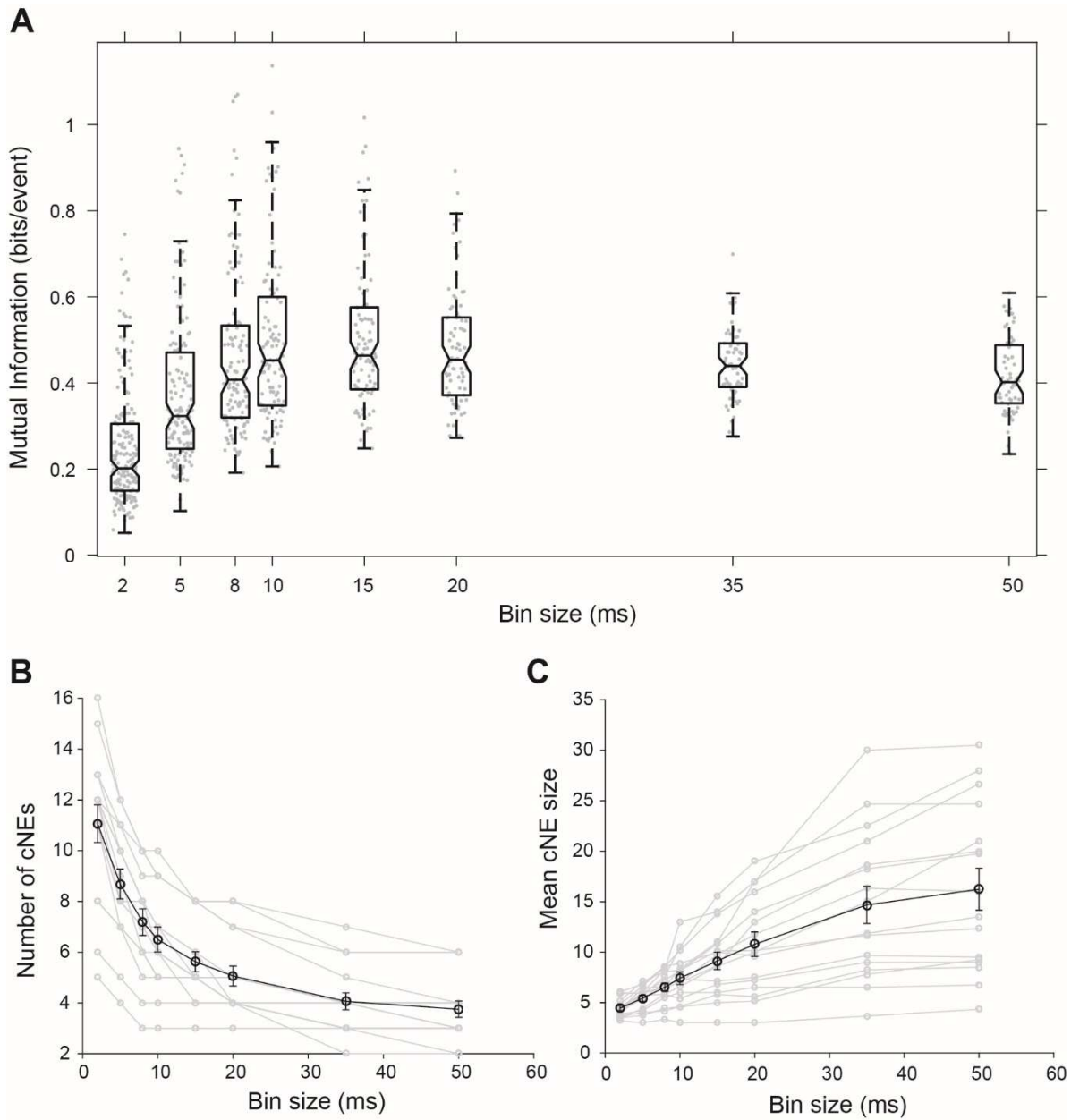
Figure 2.5. Effect of bin size on cNE properties
Effect of bin size on cNE properties. (**A**) Mutual information between the STRF and the stimulus for different bin sizes. The highest mutual information occurred for bin sizes of approximately 10 to 15 ms. (**B**) Number of detected cNEs decreases with increasing bin size. (**C**) Mean cNE size increases with increasing bin size.

A Neurons within same cNE
Site 6 Neuron 32
Site 6 Neuron 40

B STRF1 mean
STRF2 mean

C STRF correlation vs spike correlation

D Neurons NOT within same cNE
Site 6 Neuron 3
Site 6 Neuron 37

E STRF1 mean
STRF2 mean

F STRF correlation vs spike correlation

G Neurons within same cNE
Neurons NOT within same cNE

H

I

J

59
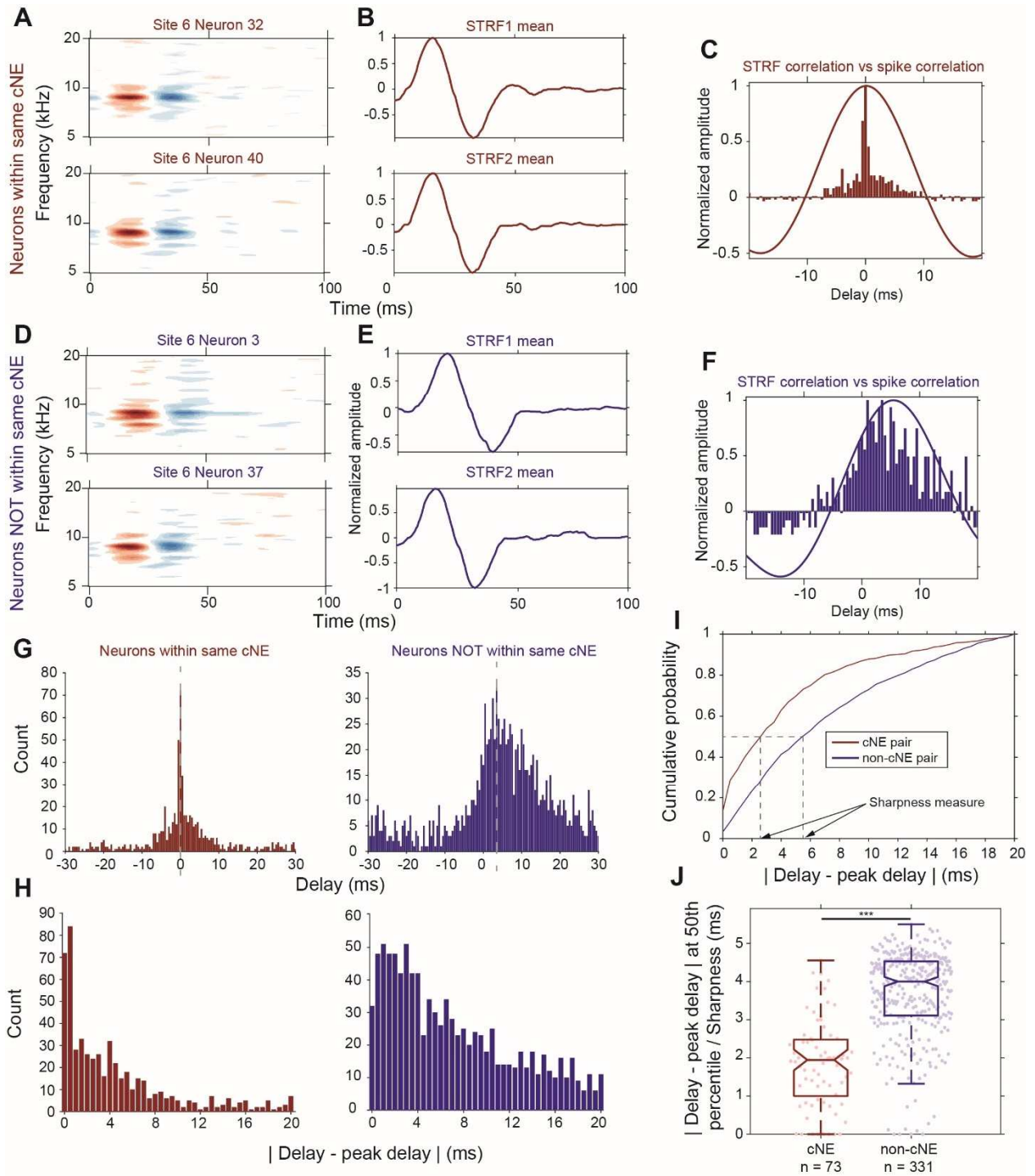
Figure 2.6. Synchrony between cNE members cannot be fully explained by receptive field overlap.
(A-C) Sample data for two neurons within the same cNE. **(D-F)** Sample data for two neurons not within the same cNE (but from the same recording). **(A, D)** STRFs for the two pairs of neurons. **(B, E)** Mean temporal profiles of each STRF in (A, D), obtained by averaging across the frequency axis of STRFs at every point on the time axis. **(C, F)** Comparisons between STRF PWCs and spike train PWCs. Neurons within the same cNE exhibit sharper spike train PWCs than their STRF PWCs. Neurons not within the same cNE exhibit spike train PWCs that are as wide as their STRF PWCs. **(G-I)** Method for calculating sharpness of a spike train PWC. **(G)** Non-normalized spike train PWCs from (C) and (F). Grey dotted lines mark the peak delay (peaks of PWCs). **(H)** PWCs folded around their peak delays (grey dotted lines in (G)). **(I)** CDFs of the distributions in (H). Black dotted lines represent the time from the peak delay at the median of the distribution, which are the sharpness values for the two example PWCs in (G). **(J)** Sharpness comparisons for one penetration. Spike train PWCs of neurons within the same cNE are sharper than those of neurons not within the same cNE. ***p < 0.001, Mann-Whitney U test.
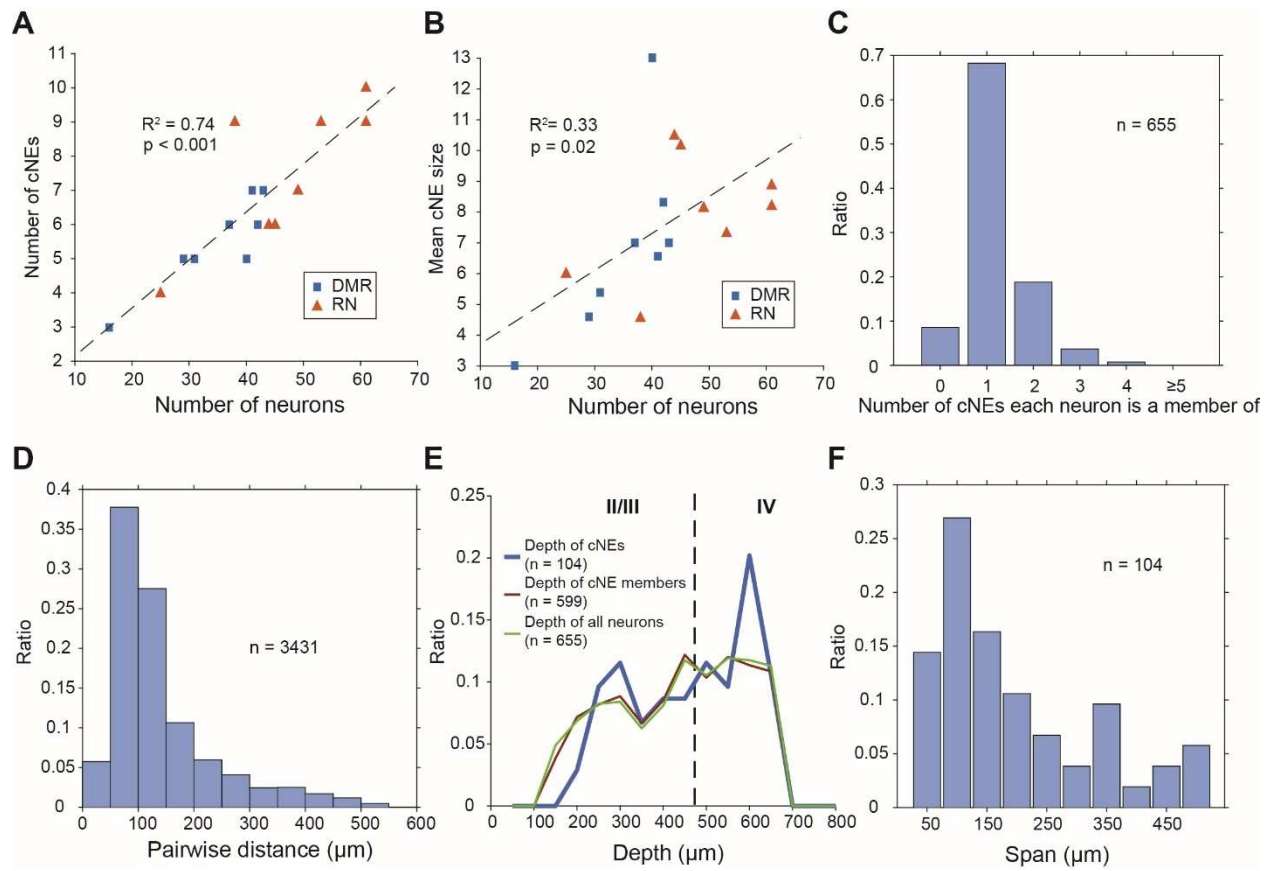
Figure 2.7. Properties of cNEs computed using 10-ms time bins.
(**A**) Number of cNEs detected increases with the number of neurons recorded. (**B**) Mean cNE size increases with the number of neurons recorded. (**C**) Most neurons recorded (~68%) belong to only 1 cNE. A small subset of neurons (~9%) did not belong to any cNE. ~23% of neurons belonged to multiple cNEs. (**D**) Pairwise distances between neurons in the same cNE. Most pairs of neurons in the same cNE (~82%) were recorded within < 200 μm of each other. (**E**) Depth of cNEs, of neurons within cNEs and of all isolated neurons. The depth of each cNE was calculated by taking the median depth of each cNE's member neurons. The depth of neurons within cNEs is the depth of all neurons found in at least one cNE, i.e. all recorded neurons except those found in the 0 bin in (C). The dotted line indicates the putative boundary between layers II/III and layer IV based on Szymanski et al., 2009. There was no depth bias for cNE neurons, i.e. the distributions of the depth of cNE members and that of all recorded neurons were similar. (**F**) cNE span, determined by the difference between the maximum and minimum depth of member neurons in each cNE. Most cNEs (~70%) had a span of 150 μm or less.
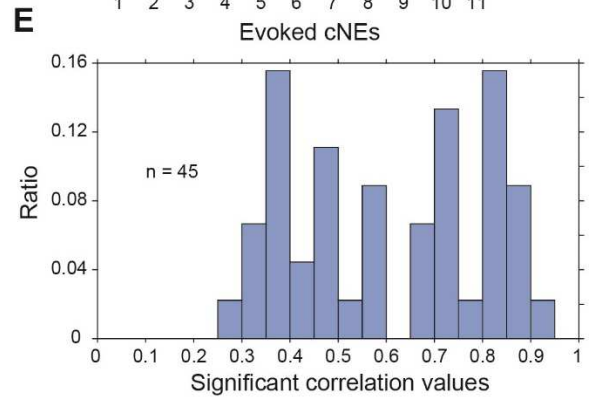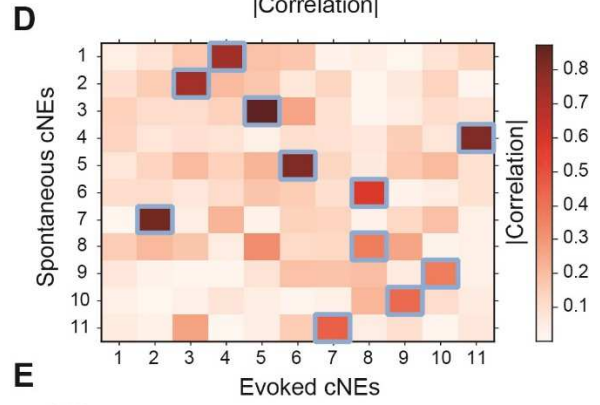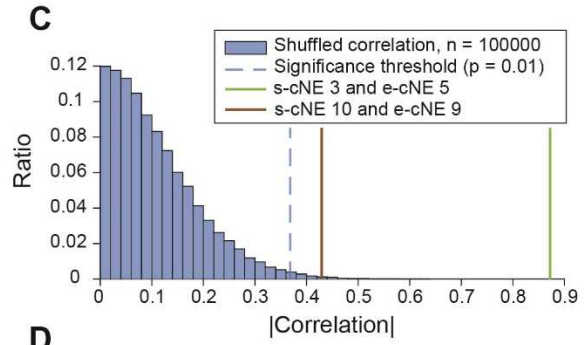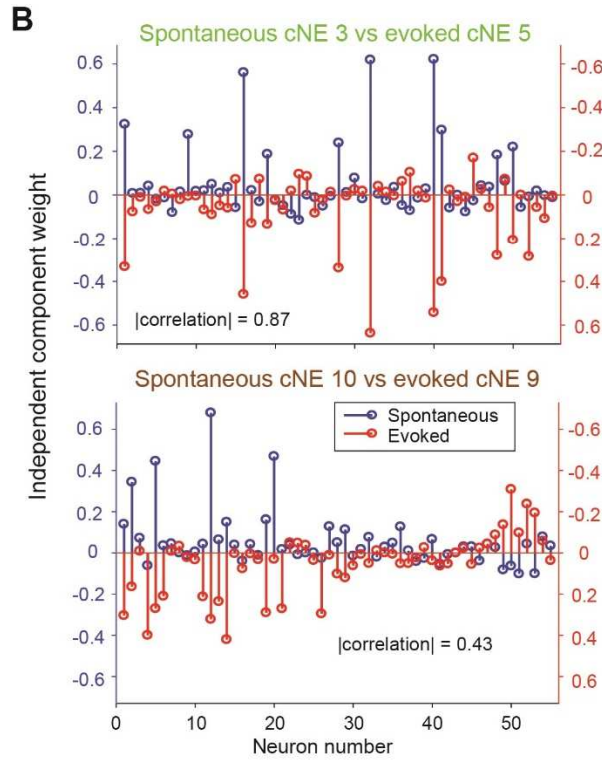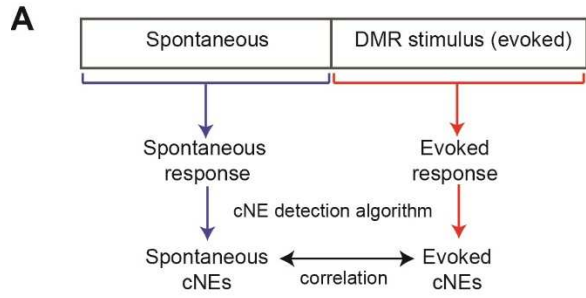
Figure 2.8. cNEs are highly preserved across spontaneous and evoked activity.
(A) Illustration of the procedure for comparing spontaneous cNEs with evoked cNEs. **(B)**
Examples of pairs of spontaneous-evoked cNEs from one penetration. cNEs for each pair of
examples have values with opposite signs for illustration purposes. (top) A pair of cNEs with an
absolute correlation value of 0.87, implying that this particular cNE was highly preserved
regardless of the presence of a stimulus. (bottom) A pair of cNEs with a lower absolute
correlation value of 0.43. Almost all of the neurons with high weights were preserved and the
low correlation value can be attributed to the noise in the neurons with low weights. **(C)** The two
pairs of cNE examples in (B) were significantly matched. See text for how the null distributions
(blue bars) were calculated. Setting a significant threshold of $p = 0.01$ for the null distribution
(blue dashed line) revealed that both comparisons in (B) were significantly matched (green and
brown solid lines). **(D)** Absolute correlation values between the weights of spontaneous and
evoked cNEs in one penetration (that includes the two examples in (B)). Correlation values
enclosed by blue squares indicate the correlation values that are higher than the significance
threshold in (C). **(E)** Significant correlation values across all penetrations with contiguous
spontaneous and DMR-evoked recordings. Approximately 72% of cNEs (both spontaneous and
evoked) had significant matches.
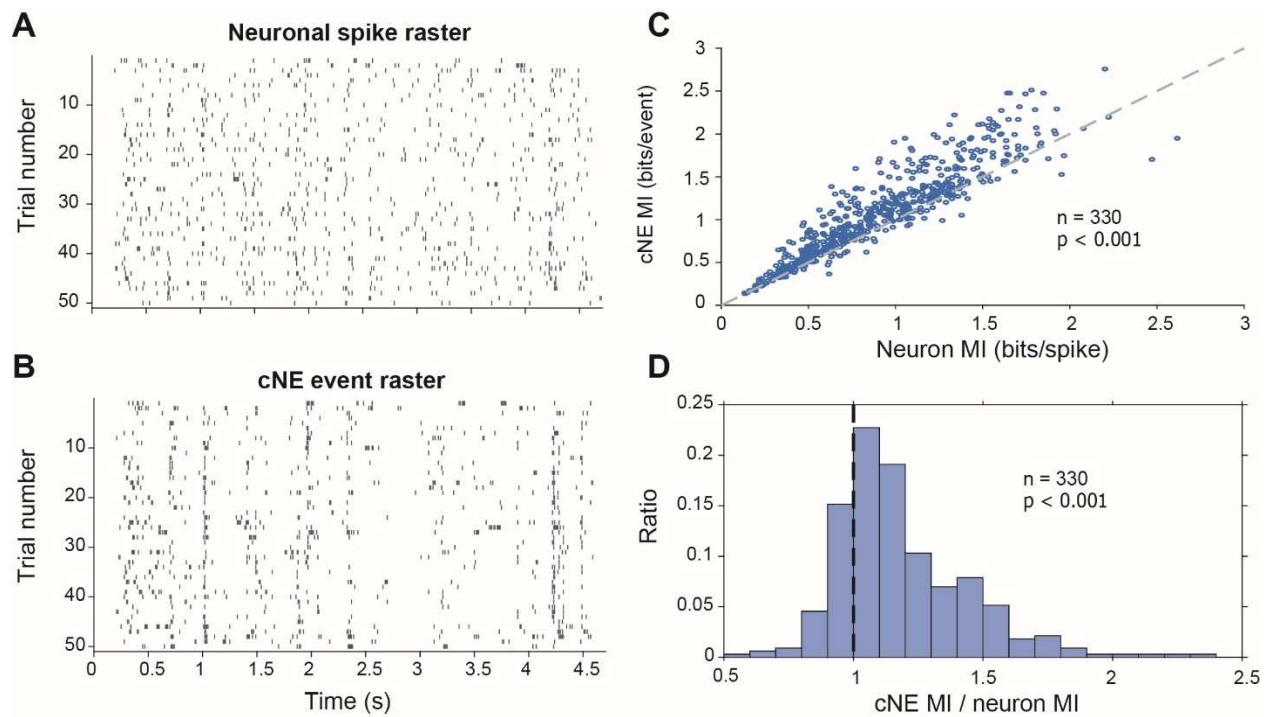
Figure 2.9. **Mutual information (MI) carried by patterns of cNE events is higher than single neuron MI.**
(A) Raster for neuronal spikes over 50 presentations of a ~5-second long DMR or RN stimulus.
(B) Raster for cNE events over 50 presentations of the stimulus. The neuron in (A) is a member of the cNE in (B). Random sampling was used to ensure equivalent spike or event counts for each comparison. **(C)** Population data for cNE MI against single neuron MI. cNE MI was significantly higher than that of the cNE's constituent neurons. **(D)** cNE MI / neuron MI. Wilcoxon signed-rank test was used in (C) and (D).
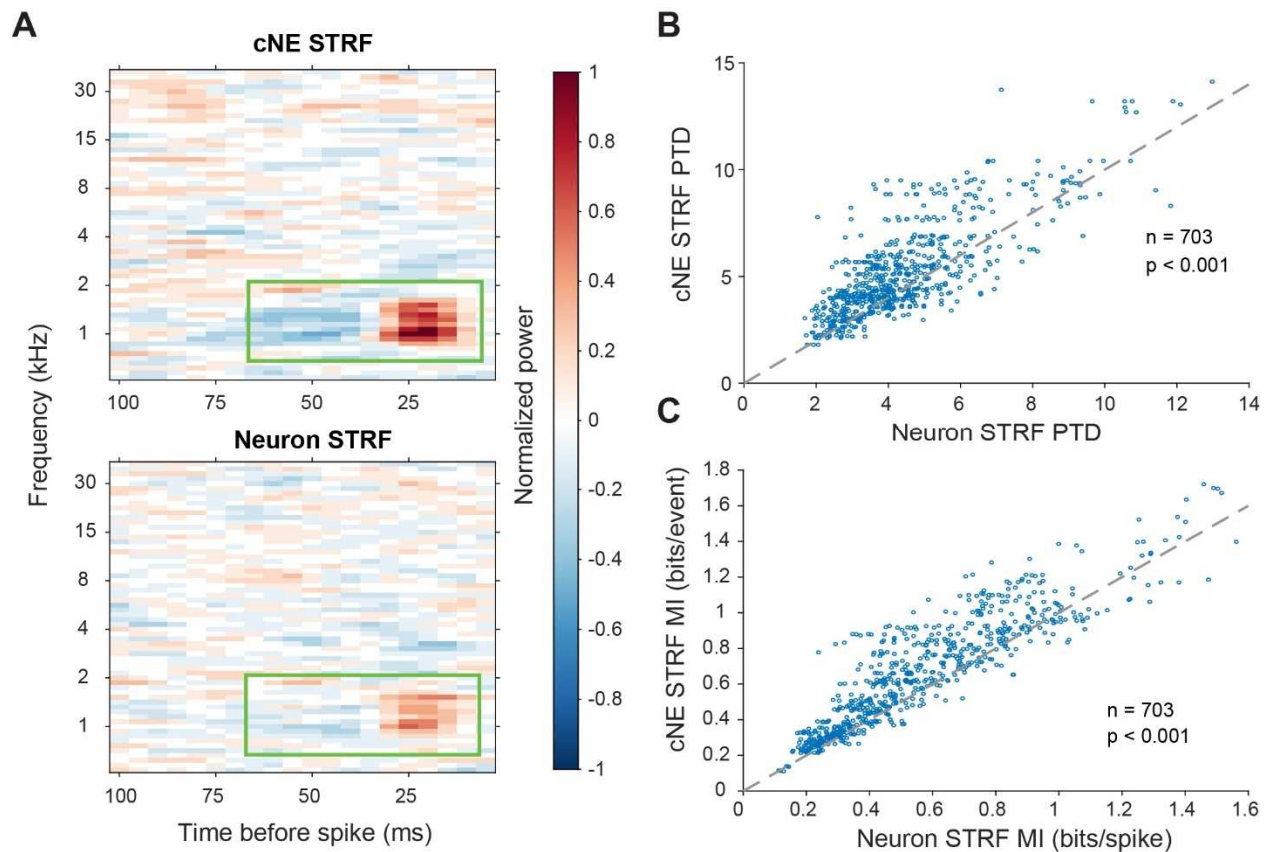
Figure 2.10. cNE STRFs have enhanced features compared to STRFs of member neurons.
(A) Example comparison between a cNE's STRF (top) and a member neuron's STRF (bottom). Random sampling was used to estimate STRFs with equivalent spike counts. The color scales for the sample STRFs have been normalized. The cNE STRF has stronger excitatory and inhibitory subfields than the neuronal STRF (highlighted by green boxes). **(B)** Over the entire population, cNE STRFs had a higher peak-trough difference (PTD) than neuronal STRFs. PTD was the difference between the highest and lowest values, divided by the number of spikes/events. **(C)** Similarly, cNEs had higher mutual information (MI) between their STRFs and single events than neurons had between their STRFs and single spikes. Wilcoxon signed-rank test was used in (B) and (C).

A Mean number of cNEs for repeated stimulus

B Mean cNE size for repeated stimulus

C Mean number of cNEs (real vs models)

D Mean cNEs size (real vs models)

E Coincidence ratio for sets of 2 neurons

F Coincidence ratio for sets of 3 neurons
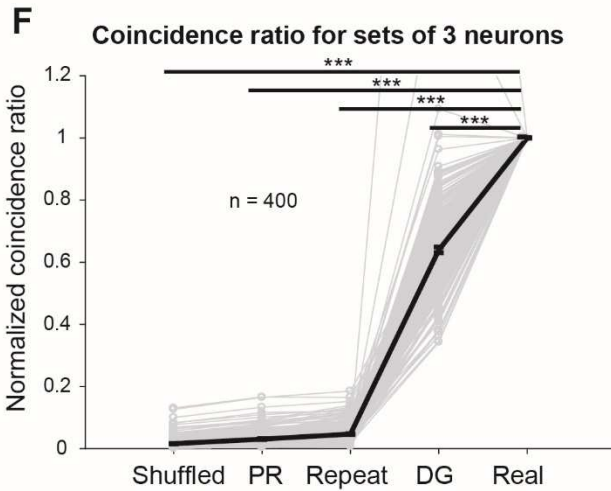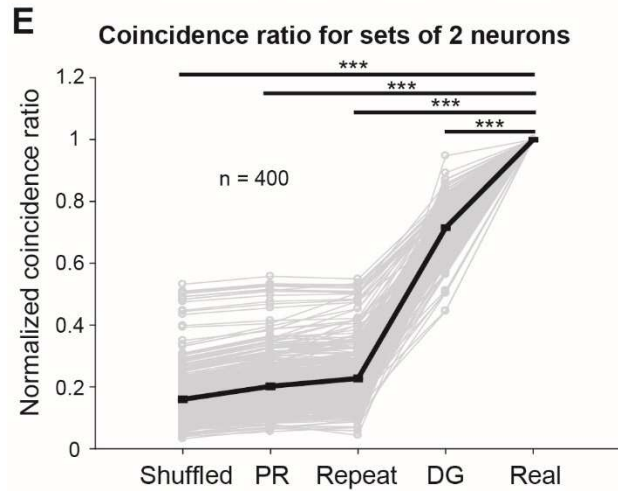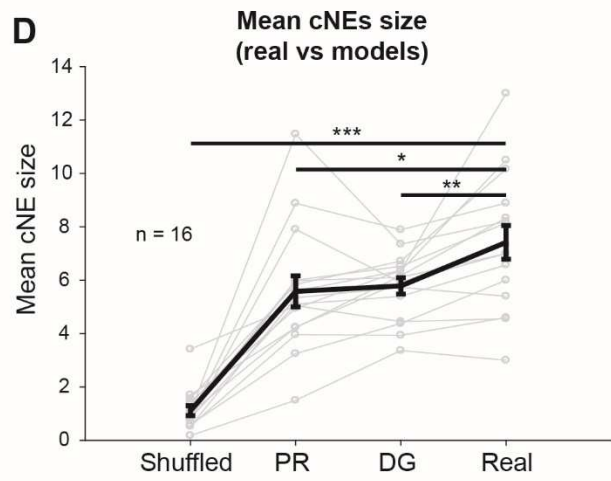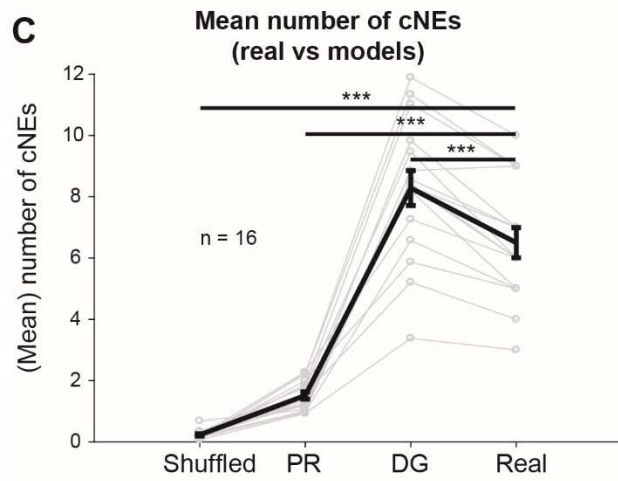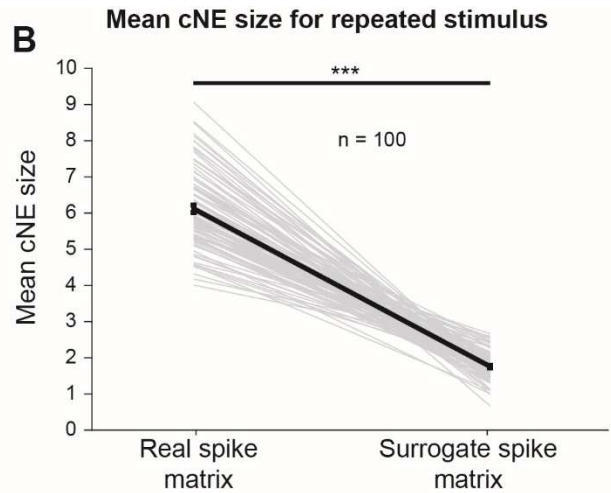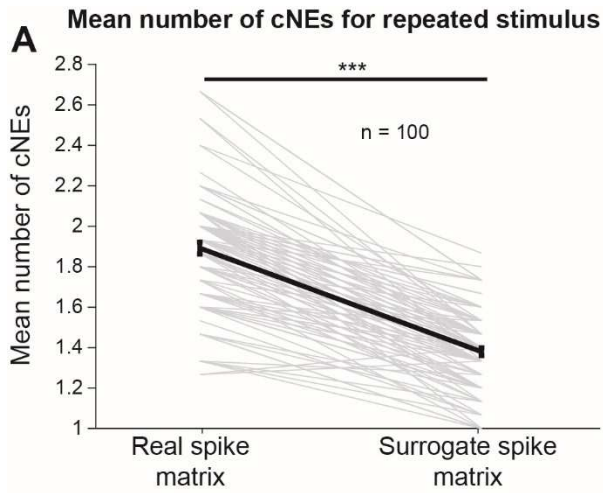
Figure 2.11. Receptive field similarity is insufficient to explain the coincident spiking between neurons in the same cNE.

(A) The number of cNEs identified in 'real' spike train matrices is significantly higher than that of 'surrogate' spike train matrices. **(B)** Mean size of cNEs identified in 'real' spike train matrices is significantly larger than that of 'surrogate' spike train matrices. See Figure 10 - figure supplement 1 for details on how the spike train matrices were constructed. Each paired data set represents one unique combination of 15 neurons. **(C)** The number of cNEs identified in the 'real' data is significantly higher than that of 'shuffled' and 'PR' models, but is significantly lower than that of the 'DG' model. **(D)** The mean size of cNEs identified in the 'real' data is significantly larger than that of 'shuffled', 'PR' and 'DG' models. See Figure 10 - figure supplement 2A for how the 'PR' model was computed. Each data set represents the mean number of cNEs and the mean cNE size for one recording over 200 iterations. **(E, F)** The coincidence ratio between neurons from different repeats of the same stimulus ('repeat'), or from the 'shuffled', 'PR' and 'DG' models is significantly less than the coincidence seen between neurons from the same cNE in the 'real' data. See Figure 10 - figure supplement 4 for how the coincidence ratio was calculated. Each data set represents a unique combination of 2 or 3 neurons and was normalized by the coincidence ratio between neurons from the 'real' spike matrix. Sets of more than 3 neurons had coincidence ratios that were close to zero for most models and are not shown. * $p < 0.05$, ** $p < 0.01$ *** $p < 0.001$, paired t-test with Bonferroni correction.

Table 2.1. Example studies that define neuronal synchrony using different methods in different brain areas.
The temporal resolution at which synchrony is defined ranges from 5 ms to 1.4 s. Legend for brain areas – A1: primary auditory cortex, AC: auditory cortex, ACC: anterior cingulate cortex, BF: basal forebrain, DG: dentate gyrus, lAG: lateral agranular cortical area, LPFC: lateral prefrontal cortex, M1: primary motor cortex, mAG: medial agranular cortical area, mPFC: medial prefrontal cortex, PAF: posterior auditory field, PP: posterior parietal cortex, PrV: principal trigeminal brainstem nucleus, S1: primary somatosensory cortex, V1: primary visual cortex, Vg: trigeminal ganglion, VPM: ventral posterior medial thalamus.

Bin size (ms)	Brain area	Method/Analyses	Reference
5	CA1, DG, S1, V1	PCA/ICA, graph theory	(Almeida-Filho et al., 2014)
5 - 20	mAG, lAG	PCA/ICA	(Laubach et al., 2000)
10	A1, PAF	Pairwise cross-correlations	(Eggermont, 2006)
10	Retina	Maximum entropy models	(Granot-Atedgi et al., 2013)
10	CA1	PCA/ICA	(Lopes-dos-Santos et al., 2013)
10	Vg, PrV, VPM, S1, M1	PCA	(Nicolelis et al., 1995)
10 - 15	V1	Maximum entropy models	(Ohiorhenuan et al., 2010)
25	CA1	Population vectors, regression	(Harris et al., 2003)
50	V1	Logistic regression, population response	(Berens et al., 2012)
55	BF	Generalized linear models	(Tingley et al., 2015)
100	S1	Dynamic bayesian networks	(Eldawlatly and Oweiss, 2011)
100	CA1	Correlation, population vectors	(Kudrimoti et al., 1999)
100	mPFC	PCA	(Peyrache et al., 2010)
100	M1, PP, PMd	Artificial neural networks	(Wessberg et al., 2000)
125 - 250	V1	Singular value decomposition	(Carrillo-Reid et al., 2015)
125 - 250	V1	Monte Carlo	(Miller et al., 2014)
200	ACC	Populations state space, trajectories	(Balaguer-Ballester et al., 2011)
200	CA1	k-means clustering	(Malvache et al., 2016)
200 - 500	Corticostriatal slices	Population vectors	(Carrillo-Reid et al., 2011)
250	AC	Support vector machines, population response	(Bathellier et al., 2012)
400	LPFC	SVM decoding	(Tremblay et al., 2015)
650 - 1400	V1	Monte Carlo	(Ikegaya et al., 2004)
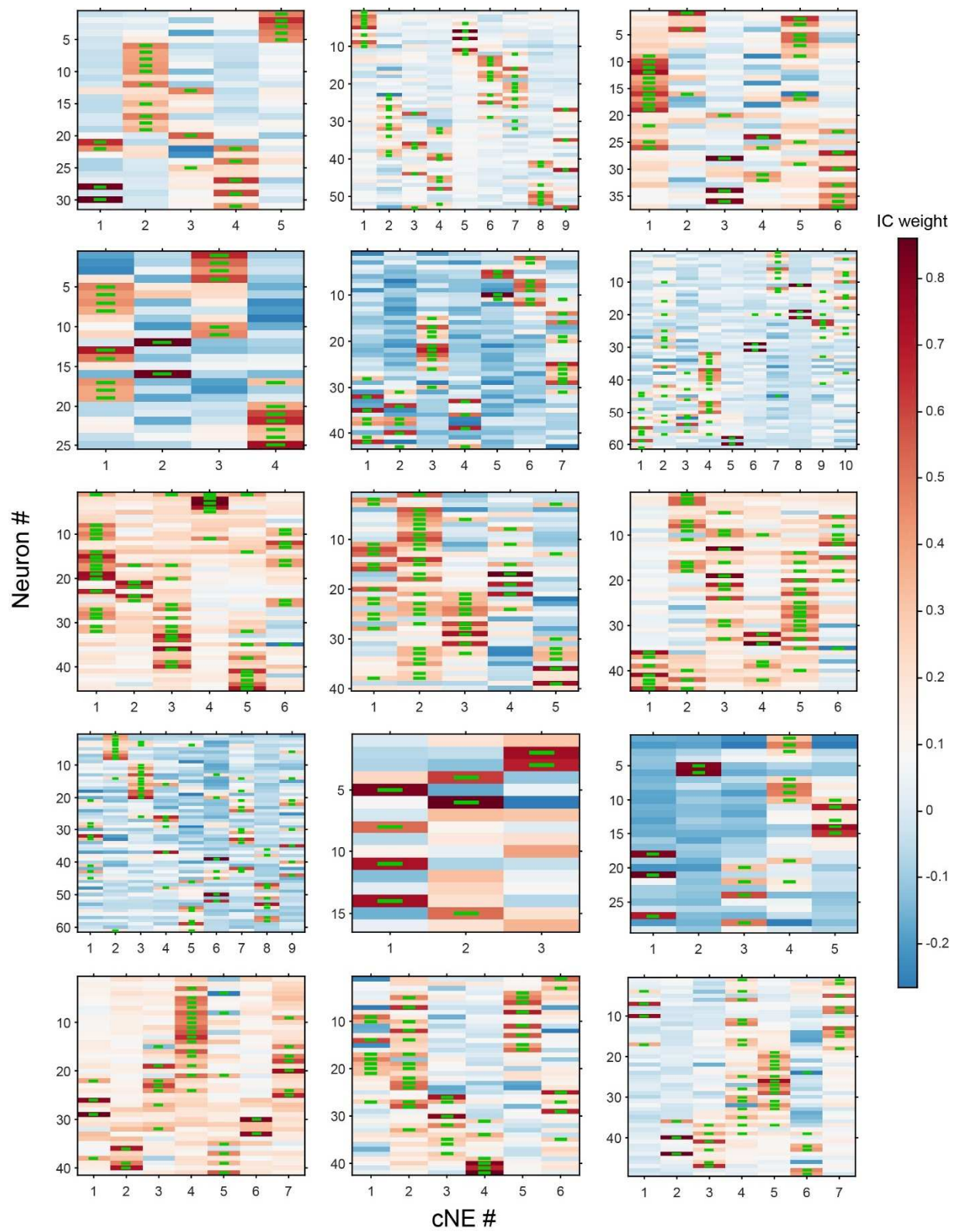
Figure 2.S1. IC weight for each neuron in each cNE for all datasets used.
Green bars represent neurons that belong to a particular cNE.
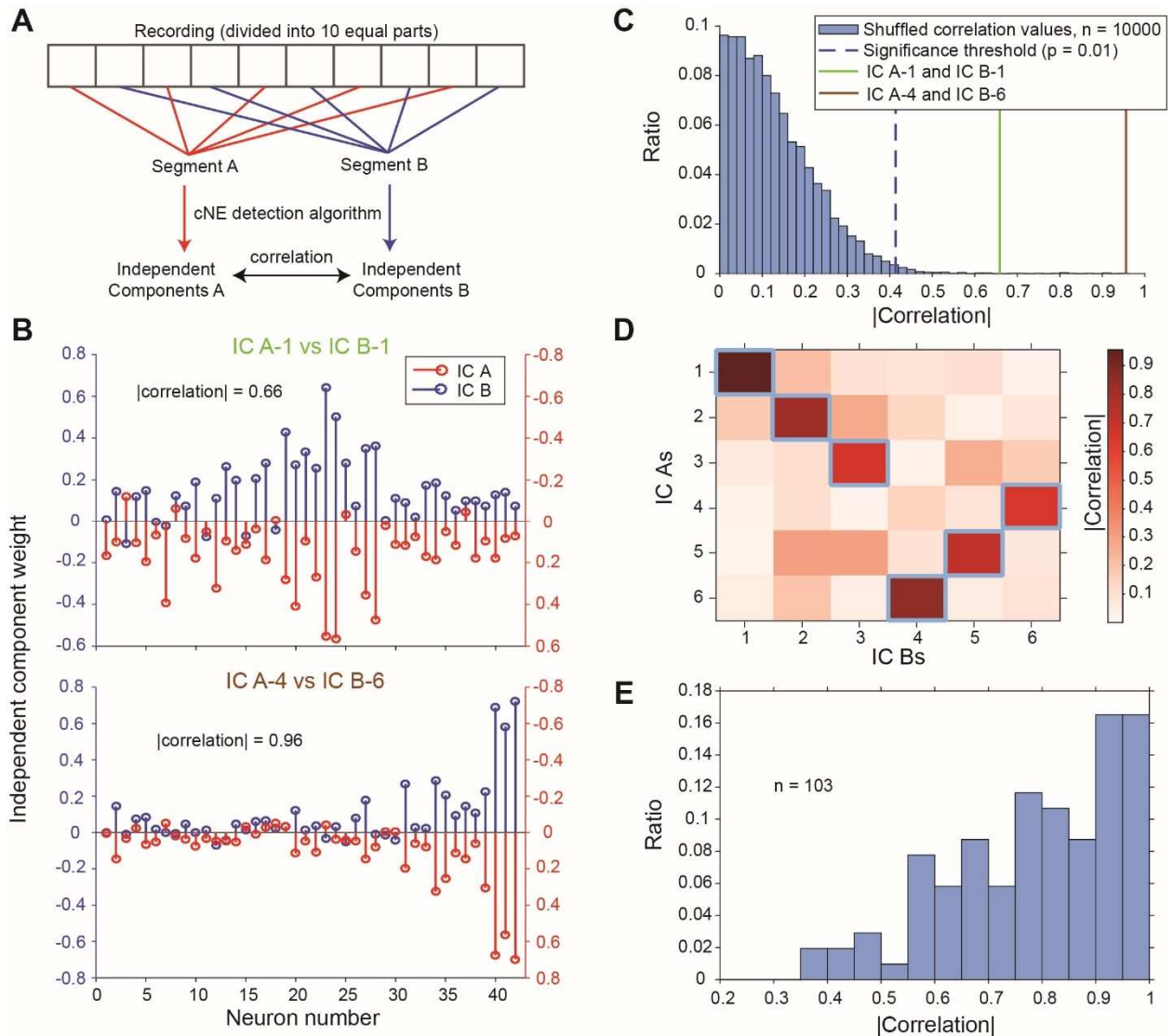
Figure 2.S2. Detected cNEs are stable throughout the recording.
(**A**) Diagram illustrating how independent components (ICs) of two segments were calculated and compared. (**B**) Example pairs of ICs from one penetration. Each pair of ICs had opposite signs for illustration purposes. (top) Example of a pair of ICs with an absolute correlation value of 0.66. Almost all of the neurons with high weights were preserved and the low correlation value can be attributed to the noise in the neurons with low weights. (**C**) The two pairs of IC examples in (B) were significantly matched. See text for how the null distribution (blue bars) were calculated. Setting a significant threshold of p = 0.01 for the null distribution (blue dashed line), both comparisons in (B) were significantly matched (green and brown solid lines). (**D**) Absolute correlation values between the weights of IC A and IC B in one penetration (that includes the two examples in (B)). Correlation values enclosed by blue squares indicate the correlation values that are higher than the significance threshold in (C). (**E**) Significant correlation values obtained across all penetrations. Approximately 96% of IC As and IC Bs had significant matches.

A Recording (divided into 2 halves)

Training Data → cNE detection algorithm → Test Data

Training Independent Components / Test Independent Components

Project onto 2nd half of recording

Training cNE Activity ← correlation → Test cNE Activity

B Training cNE 6 vs Test cNE 4

training
test

correlation = 0.99

Training cNE 4 vs Test cNE 5

correlation = 0.85

cNE activity

Time (s)

C

Non-matched correlations, n = 620
Significance threshold, p = 0.05
Training cNE 6 and Test cNE 4
Training cNE 4 and Test cNE 5

Ratio

Correlation

D

Training cNE activity

Test cNE activity

Correlation

E

n = 85

Ratio

Correlation

71

Figure 2.S3. Detected cNEs are stable throughout the recording.
(**A**) Diagram illustrating how training cNE and test cNE activities were calculated and compared. (**B**) 10-second snippets of example pairs of training/test cNE activities from one penetration. Test cNE activity was offset by 50 ms for illustration purposes. (**C**) Correlation values between non-matched pairs, which are non-maximum values for each column and row (values not surrounded by blue squares in the example in (D)), of training and test cNE activities. The significance threshold for matched pairs (maximum correlation value in each column and row; blue squares in the example in (D)) was set at $p = 0.05$ (blue dashed line). The green and brown solid lines represent the correlation values of the matched examples in (B) and are significant. (**D**) Correlation values between the training and test cNE activities in one penetration (including the two examples in (B)). Correlation values enclosed by blue squares indicate the correlation values that are higher than the significance threshold in (C). (**E**) Significant correlation values across all penetrations. There were 102 training cNEs and approximately 83% of them had significant matches to the test cNEs.

Y-axis label: | Delay - peak delay | at 50th percentile / Sharpness (ms)
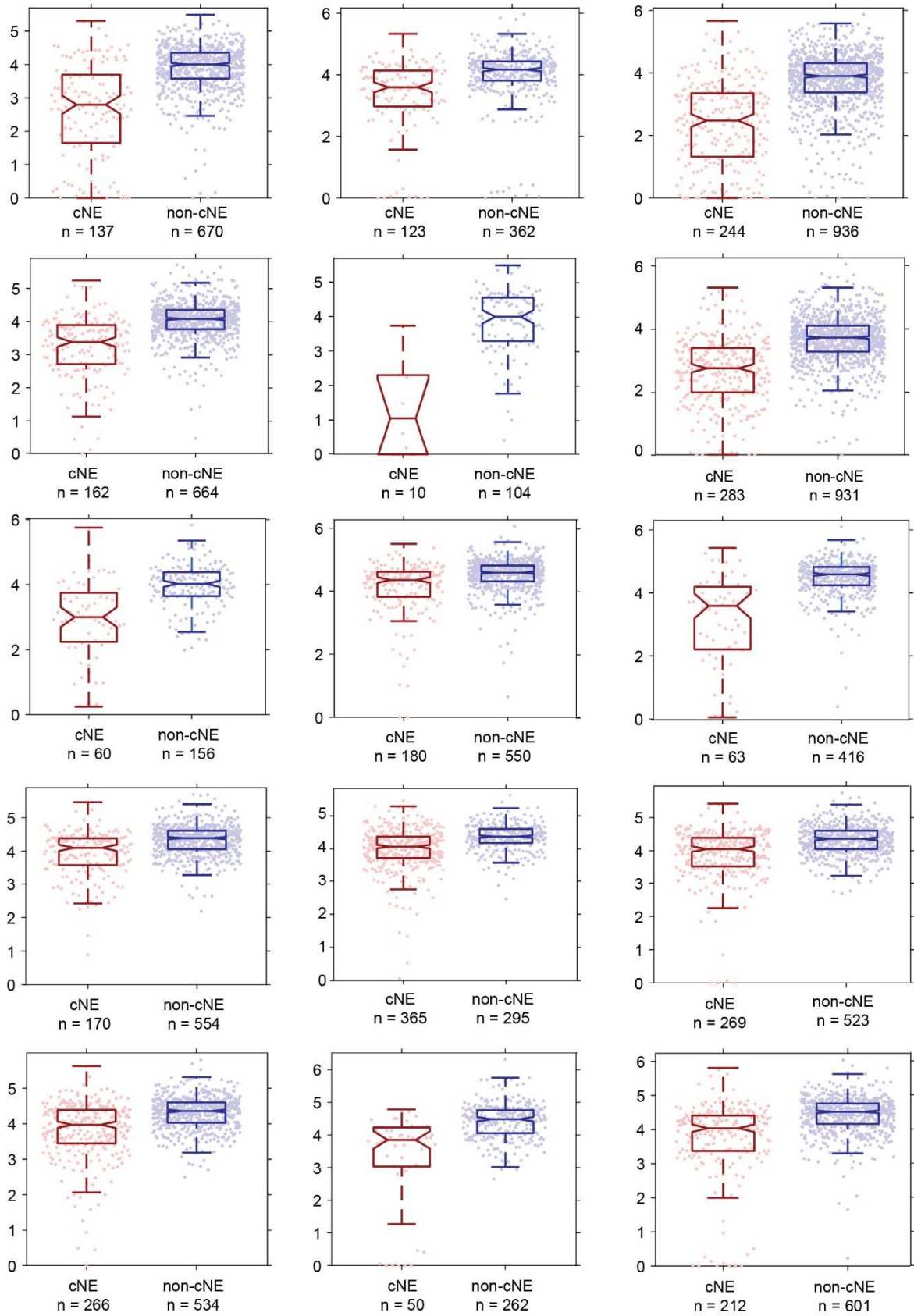
Figure 2.S4. Comparison of sharpness of PWC functions of pairs within the same cNE against that of pairs not within the same cNE for all datasets.
For all comparisons, the sharpness values for cNE pairs were significantly less than that for non-CNE pairs (p < 0.001, Mann-Whitney U test).
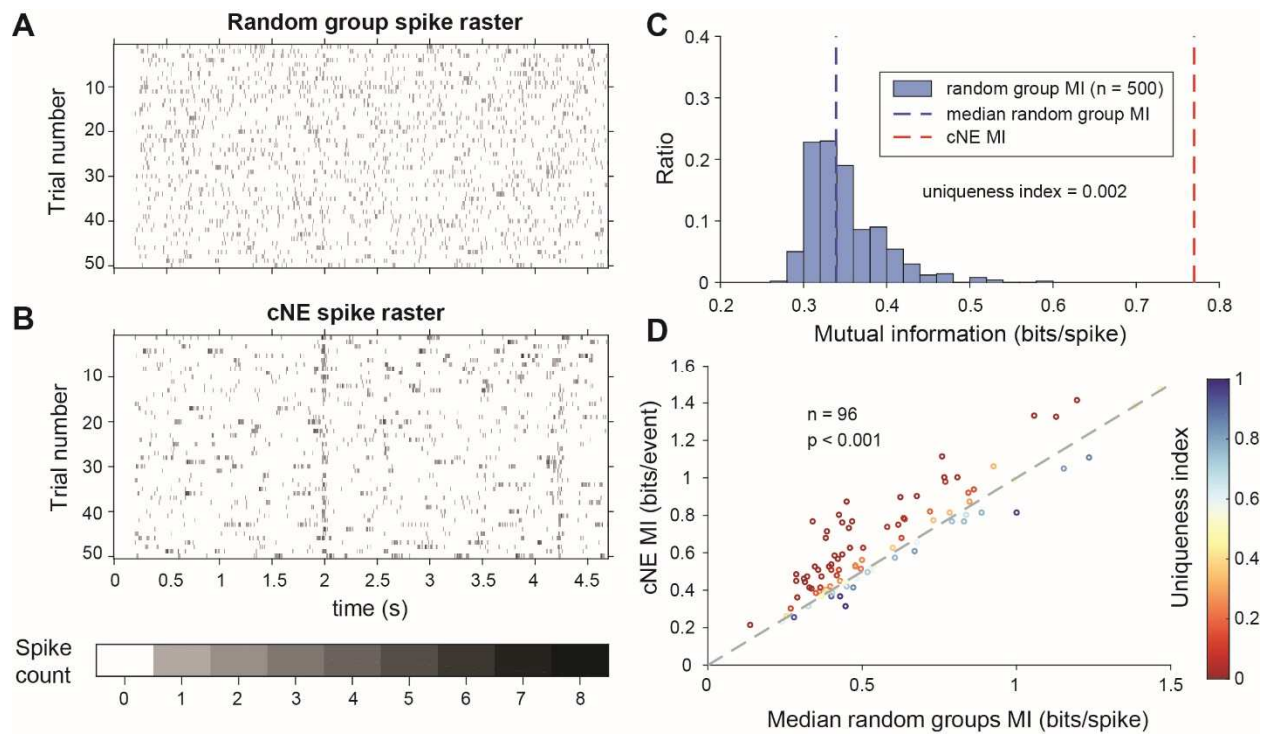
Figure 2.S5. Mutual information (MI) carried by patterns of cNE members' spikes is higher than that of random groups of neurons.
(A) Raster for the summed spikes of a random group of neurons over 50 presentations of a ~5-second long DMR or RN stimulus. Each random group of neurons has one neuron in common with the cNE it is being compared to in (B). **(B)** Raster for the summed spikes of all members of a cNE over 50 presentations of the stimulus. Random sampling was used to ensure equivalent spike counts for each comparison. **(C)** MI values for 500 randomly chosen groups (blue bars) and cNE (red dashed line), for the example in (B). The uniqueness index is the proportion of the entire distribution (blue bars and red dashed line; n = 501) that is greater than or equal to the cNE MI value (red dashed line). The minimum possible uniqueness index is therefore 0.002. **(D)** Population data for cNE MI against the median of the random group MI (blue dashed line in (C)). cNE MI was significantly higher than the MI for random groups of neurons. The uniqueness index for each cNE-random group comparison is colored according to the displayed color bar. Wilcoxon signed-rank test was used in (D).
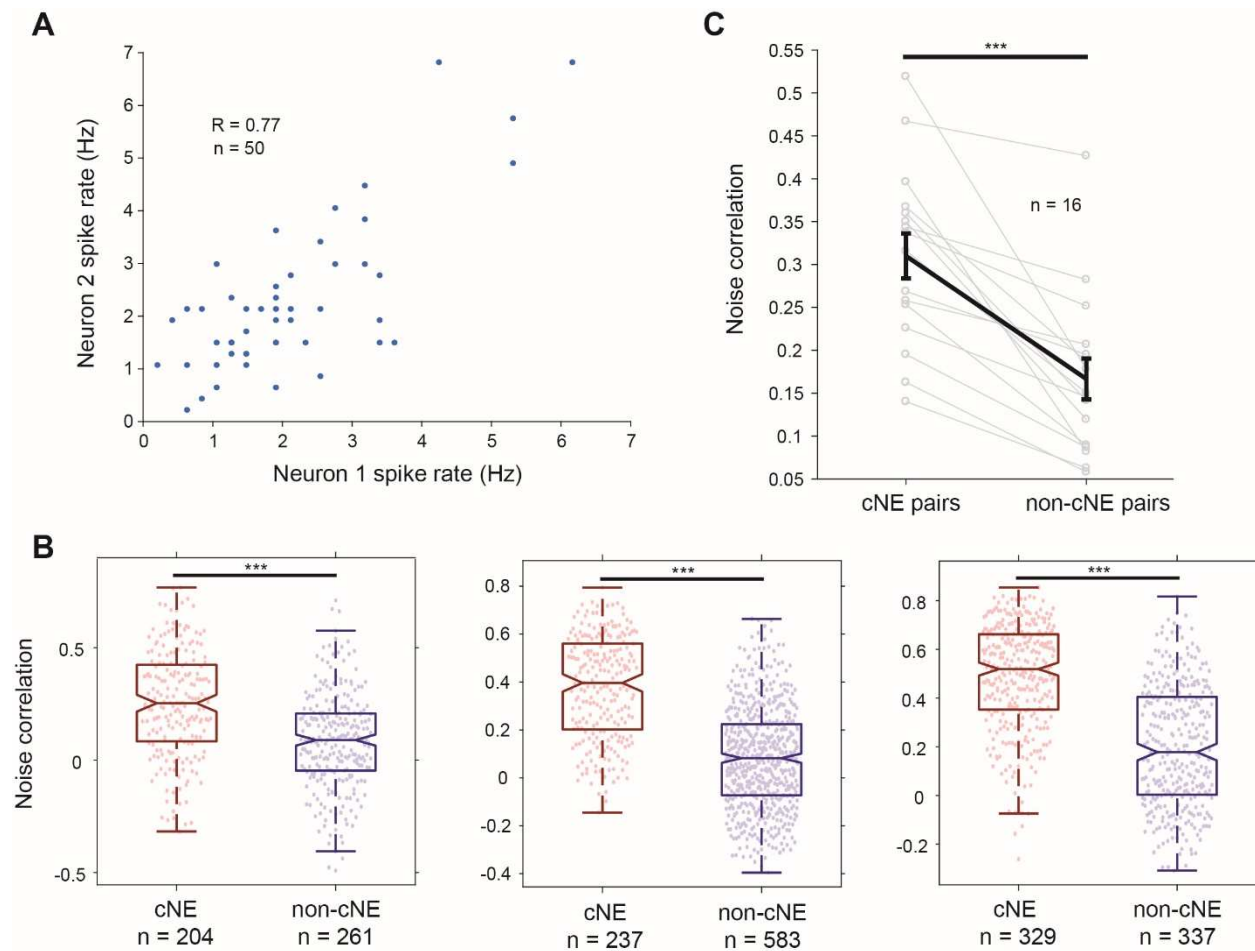
Figure 2.S6. Noise correlations between cNE pairs were higher than those between non-cNE pairs.
(A) Example spike rates for a pair of neurons over 50 trials of the short, repeated stimulus. The noise correlation was calculated as the Pearson's correlation coefficient (R) between the two neurons' spike rates over different trials. **(B)** Examples of raw noise correlation data from 3 of the 16 recordings. All recordings had noise correlations between cNE pairs that were significantly higher than noise correlations between non-cNE pairs. **(C)** Summary figure for all 16 recordings. The median noise correlations for each group and recording are plotted. *** p < 0.001, Mann-Whitney U-test was used for (B), paired t-test was used for (C).
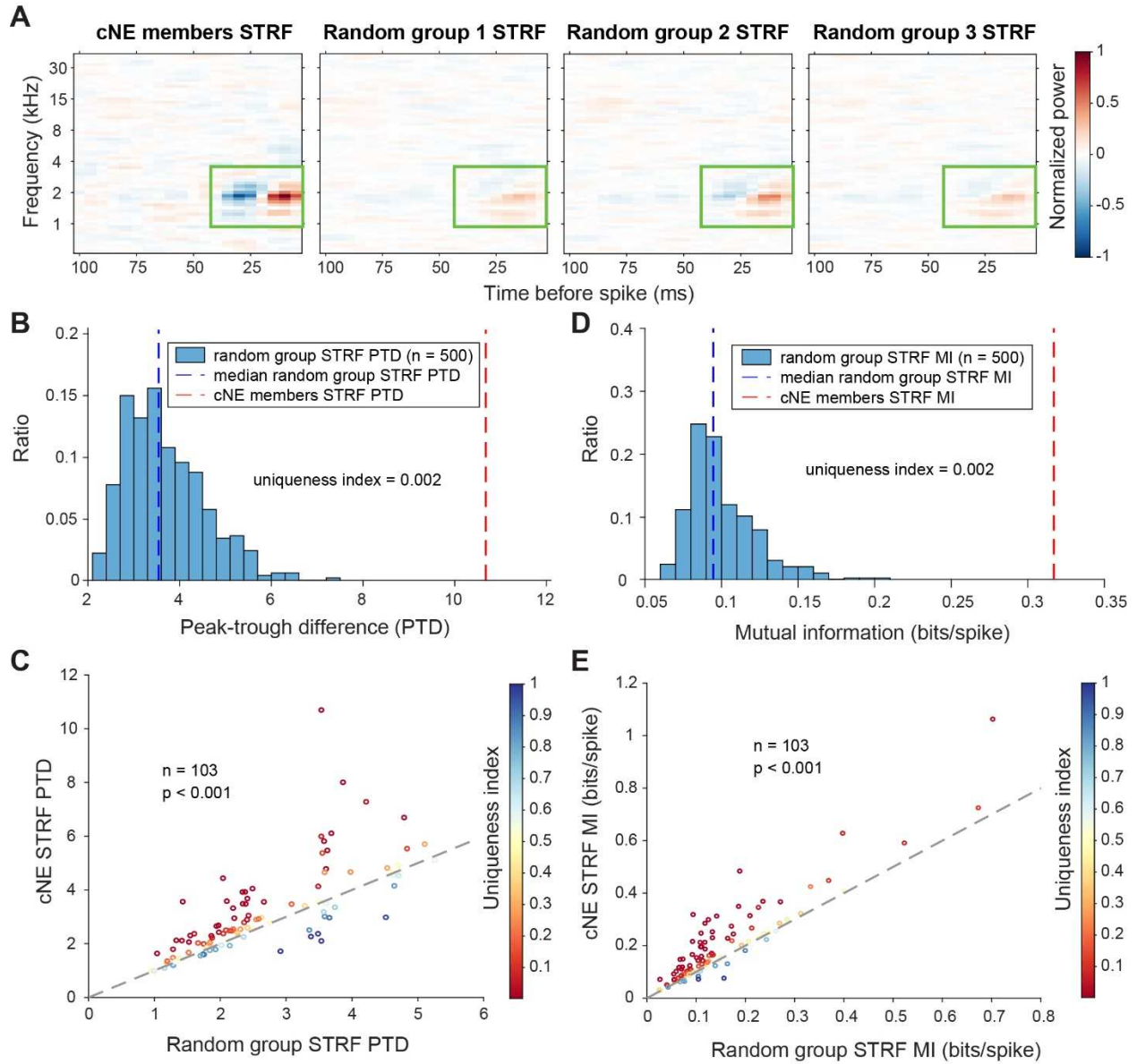
Figure 2.S7. cNE members have multi-unit STRFs with enhanced features compared to the multi-unit STRFs of random groups of neurons.
(A) Sample of a comparison between the multi-unit STRF of members of a cNE (left-most) and the multi-unit STRFs of random groups of neurons the same size as the cNE. The number of spikes used to calculate the STRFs was equalized via random sampling. The color scales for the sample STRFs have also been normalized. The STRF for cNE members shows stronger excitatory and inhibitory subfields than the STRFs for random groups (highlighted by green boxes). **(B)** Peak-trough difference (PTD) values for 500 iterations of random groups (blue bars) and cNE (red dashed line), for the example in (A). The uniqueness index is the proportion of the entire distribution (blue bars and red dashed line; n = 501) that is greater than or equal to the cNE MI (red dashed line). The minimum possible uniqueness index is therefore 0.002. **(C)** Population data for cNE PTD against the median of random group PTD (blue dashed line in (B)). cNE PTD was significantly higher than that of random groups of neurons. The uniqueness index for each cNE-random group comparison is colored according to the displayed color bar. **(D)** Similar to (B), but for mutual information (MI) between the multi-unit STRFs and single spikes. **(E)** Similar to (C), but for MI. cNE MI was significantly higher than that of random groups of neurons. Wilcoxon signed-rank test was used in (C) and (E).
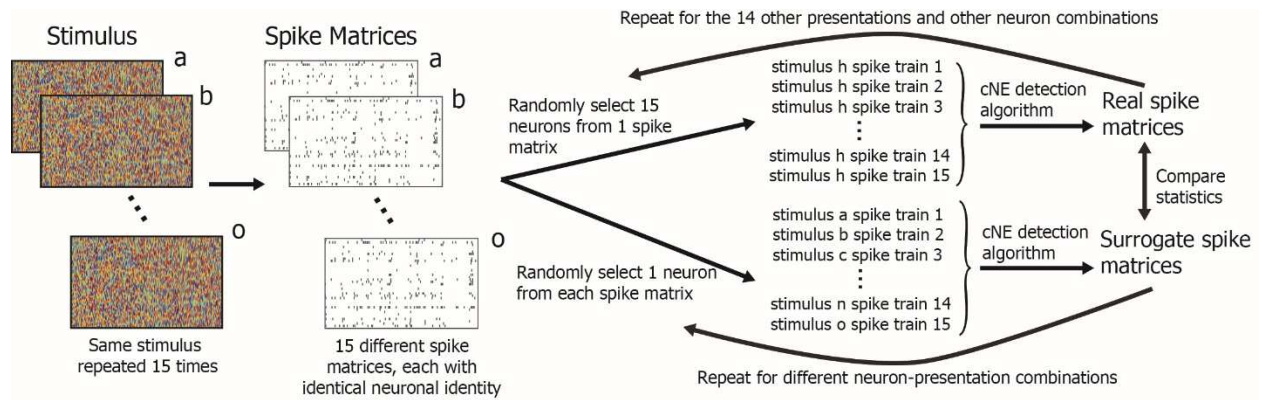
Figure 2.S8. Illustration of the process of constructing spike trains from different presentations of the same stimuli to obtain 'real' and 'surrogate' spike matrices. The statistics of these spike matrices are presented in Figure 2.10A and B.
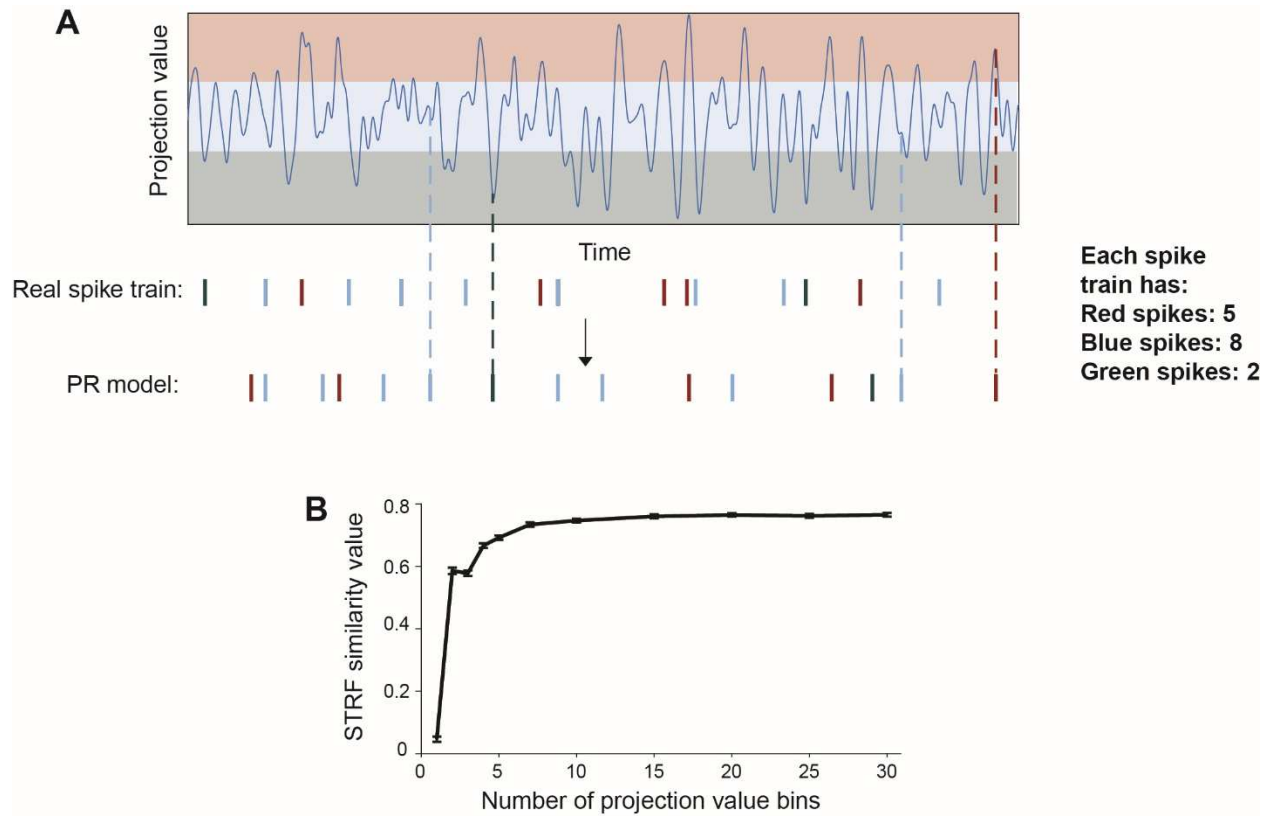
Figure 2.S9. Method for the 'preserved receptive field' ('PR') model.
(A) Illustration of how the 'PR' model was computed. The similarity between stimulus segments and each neuron's receptive field (projection value) was calculated and binned. The red, blue and green shades represent the 3 bins in this simple illustration, even though in the actual model 15 bins were used. The spikes, each belonging to a projection value bin (and colored correspondingly) were then randomly reassigned to another time bin within the same projection value. **(B)** STRF similarity values between STRFs obtained from the 'PR' model and the STRFs of the 'real' spike trains against the number of projection value bins used in computing the 'PR' model. The similarity values plateaued after 15 projection value bins, justifying the use of 15 projection value bins.
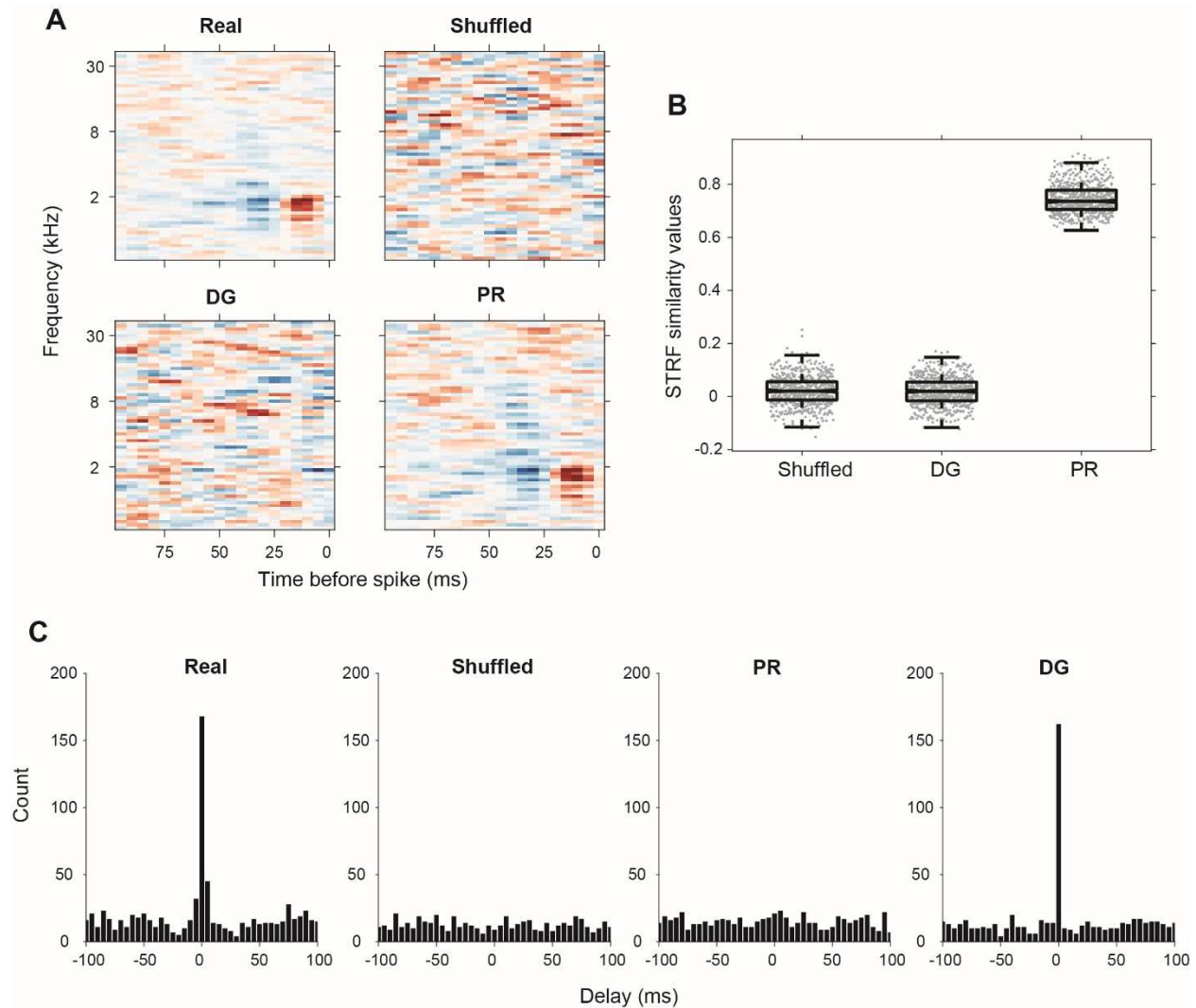
Figure 2.S10. Verification of the 'shuffled', 'PR', and 'DG' models.
(**A**) Sample STRFs for one neuron from the 'real' spike train, 'shuffled', 'DG' and 'PR' models. The 'PR' model has similar STRFs to that of the 'real' spike trains while the 'shuffled' and 'DG' models do not. (**B**) STRF similarity values between neuronal spike trains from the 'shuffled', 'DG' and 'PR' models and the 'real' spike trains. The STRFs obtained from the 'PR' model spike trains were highly similar to the STRFs obtained from 'real' spike trains, while the 'shuffled' and 'DG' models have STRFs that are uncorrelated with the 'real' STRF. (**C**) PWC functions at 5-ms temporal resolution for one pair of neurons. The correlations in the 'shuffled' and 'PR' models have been completely broken. The 'DG' model preserves the correlation at zero delay but not at other time delays. Using 5-ms time bins captures almost all the temporal correlations between two neurons in a real spike train.

A

Neuron A:	1	0	1	0	1	0	0	1	0	1	Sum A: 5
Neuron B:	1	0	0	1	0	1	0	1	0	0	Sum B: 4
Neuron C:	1	1	0	1	0	1	0	1	0	1	Sum C: 6

$$\text{Coincidence ratio} = \frac{2}{4} = 0.5$$

B
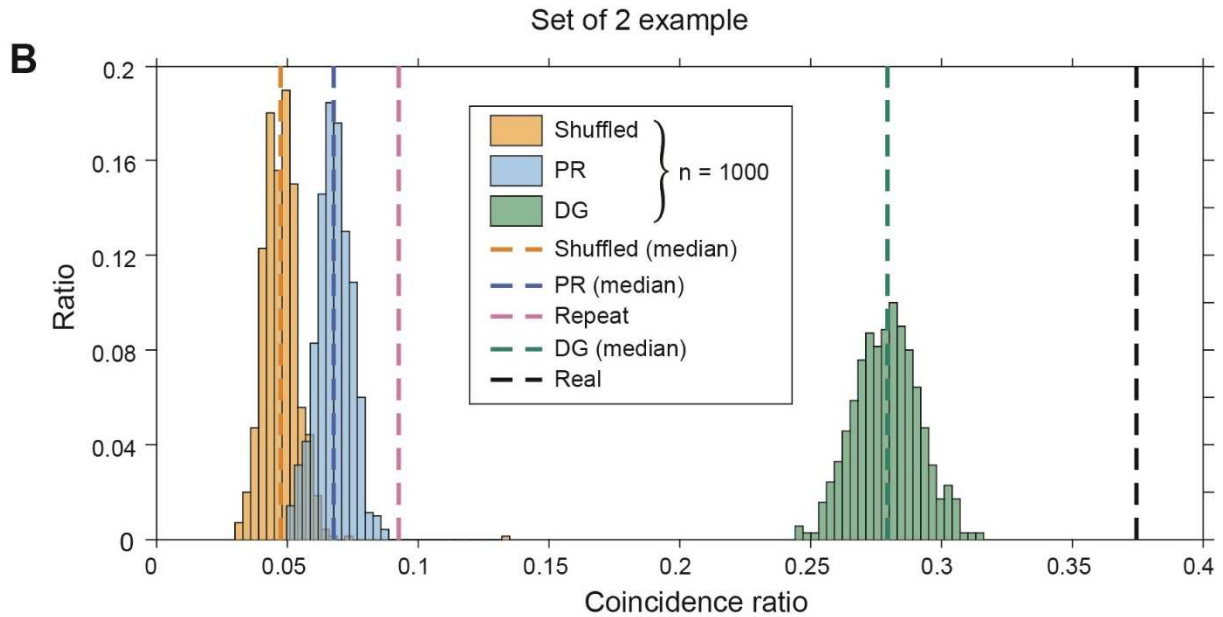
Figure 2.S11. Calculation of coincidence ratio between groups of neurons in the same cNE. **(A)** Illustration of how coincidence ratio (CR) was calculated for a set of 3 neurons. The maximum CR is 1 (when all neurons spike whenever the neuron with the lowest firing rate spikes) and the minimum CR is 0 (when there are no coincident bins). **(B)** Sample CR values for an example set of 2 neurons. The 'repeat' model compares two pairs of spike trains from different presentations of the same stimulus. Distributions represent 1000 iterations of the 'shuffled', 'PR' and 'DG' models and associated coincidence ratios. Together with 'repeat' and 'real' values, these values comprise one data set in Figure 2.11E.

CHAPTER 3

Differential encoding of auditory information by coordinated neuronal ensembles in the primary auditory cortex

3.1 Abstract

Neuronal activity in the auditory cortex is often highly synchronous between neighboring neurons at millisecond precision. Such coordinated activity is thought to be crucial for information processing, likely reflecting the convergence of top-down and bottom-up contributions. We identified coordinated neuronal ensembles (cNEs) within primary auditory cortical (AI) columns and assessed their functional properties relative to the contributing neurons. Nearly half of AI cNEs showed robust spectro-temporal receptive fields, representing bottom-up information. The remaining cNEs showed little or no acoustic feature selectivity, potentially reflecting internally generated influences. Individual neurons were found to participate in the activity of several cNEs with either high or absent feature selectivity, producing a heterogeneous set of spikes identifiable by instantaneous association with different cNE events. This suggests that auditory cNEs can represent either stimulus-specific, feed-forward aspects or stimulus-unspecific, feedback information. By contrast, single neuron spike trains are a multiplexed representation of multiple sensory information aspects.

3.2 Introduction

Neurons in the cortex have been shown to be highly interconnected (Braitenberg and Schüz, 1991). As such, the activity of any individual neuron is strongly correlated with some of its neighbors, and this coordinated activity is thought to underlie information processing and transmission throughout the cortex (Bathellier et al., 2012; DeNardo et al., 2019; Gulati et al., 2014; Jacobs et al., 2015; Kiani et al., 2014; Reimann et al., 2017; Roland et al., 2017; Tremblay et al., 2015). This is due in part to the robustness of information representation or encoding in neural ensembles over single neurons (See et al., 2018), and in part to the increased likelihood of information transmission downstream when groups of neurons fire synchronously (Matsumura et al., 1996; Stevens and Zador, 1998).

In the primary auditory cortex (AI), coordinated activity at different time resolutions has been observed (Atencio and Schreiner, 2013; Brosch and Schreiner, 1999; Eggermont, 2006; Eggermont et al., 2013; Gourévitch and Eggermont, 2010; See et al., 2018). We have previously shown that coordinated neuronal ensembles (cNEs), defined as groups of neurons with reliable and precise synchronous activity, are stable functional constructs. They are present in both spontaneous and evoked activity, and can encode more information about the auditory stimulus than single neurons or random groups of neurons, suggesting that they may represent the principal unit of information processing in AI (See et al., 2018).

Meanwhile, other studies in AI have shown the existence of populations of neurons that do not appear to encode spectral or temporal features of the auditory stimulus (Hromádka et al., 2008), but instead encode other aspects of the auditory world, such as the choice of an animal in an auditory discrimination task (Francis et al., 2018; Rodgers and DeWeese, 2014), the

categorization of a sound based on an animal's percept (Bathellier et al., 2012), or associated

multisensory or cognitive cues (Brosch et al., 2015; Morrill and Hasenstaub, 2018). In a few of

these studies, these neurons have also been shown to have coincident activity, akin to cNEs, and

these constructs encode the animal's behavior much better than possible by single neurons

(Bathellier et al., 2012; Francis et al., 2018).

Here, we characterize the functional properties of cNEs in rat AI, anesthetized with

ketamine/xylazine, while presenting dynamic broadband stimuli. We identified cNEs using

dimensionality reduction techniques and isolated cNE events, using them to assess spectro-

temporal information processing. We find that there are different functional categories of cNEs –

some with improved spectro-temporal information processing over that of member neurons,

while others seem agnostic to the presented stimuli. By tagging individual spikes of single

neurons to cNE events we find that AI neurons produce a functionally heterogeneous set of

spikes that can be grouped according to instantaneous association with different cNEs. These

subsets of spikes can differentially encode explicit auditory information, or other, potentially top-

down, information according to the cNEs to which they are associated. By linking individual

neural events to the activity of local networks, we are able to perform a more refined analysis of

cortical processing principles.

3.3 Results

cNEs can be stimulus-dependent or stimulus-independent

A main goal of this study was to compare the stimulus-related properties of cNEs and their contributing single neurons. We densely sampled the activity of populations of neurons in cortical columns of rat primary auditory cortex (AI) with high-density recording arrays that spanned the entire cortical depth (see Materials and methods). Sampled sites were previously mapped and shown to have multi-unit activity that had well-defined tuning curves when presented with pure tones. Rats were anesthetized with a cocktail of ketamine and xylazine and dynamic broadband stimuli (dynamic moving ripple - DMR, or ripple noise - RN; see Materials and methods; Escabi and Schreiner, 2002) were presented while neuronal activity was recorded. Spectro-temporal receptive fields (STRFs) of well-isolated neurons were then estimated via spike-triggered averaging (Figure 3.S1A). For each STRF, we calculated a reliability index metric (RI; Figure 3.S1B; see Materials and methods; Escabí et al., 2014), a measure of the reproducibility of the STRF from various segments of a neuron's evoked activity. RI is a proxy for the homogeneity of neuronal spikes with respect to the presented stimulus; the higher the RI, the more similar the stimulus segments that are encoded by the spikes. Significance of any STRF was determined by circularly shuffling spike trains and calculating the RI of the shuffled spike trains, generating a null distribution ($p < 0.05$; Figure 3.S1C).

cNEs, defined as neurons that fire synchronously and reliably, were determined from simultaneously recorded neurons via principal component and independent component analyses (Figure 3.S1D; see Materials and methods; Lopes-dos-Santos et al., 2013; See et al., 2018). By projecting the independent components (ICs) back onto the spike train matrix (Figure 3.S1D) a

binary cNE event train was reconstructed by calculating a response threshold via surrogate methods (see Materials and methods). From those events we calculate each cNE's STRF and determined its significance (Figure 3.S1A – C). An example columnar penetration with STRFs of neurons and two cNEs is illustrated in Figure 3.1. The neuronal STRFs were based on all well-isolated spikes obtained during the stimulus presentation.

We found that neurons that contribute to a cNE had similar, but not necessarily identical, STRFs as the cNE as previously reported (see cNE 1 in Figure 3.1; See et al., 2018). However, we also found that many cNEs had non-significant STRFs, with contributions of member neurons that either had also non-significant or significant STRFs based on all their spikes (see cNE 8, Figure 3.1). Because cNEs are made up of neurons that fire together reliably and with high temporal precision, these non-significant cNEs are not trivially a result of "noise spikes", but that cNE events likely encode some information that is either not a tightly phase-locked portion of the stimulus spectrogram or represent other stimulus-associated information potentially including top-down information about animal state, task aspects, or other relevant modulatory influences .

Classification of cNEs based on STRF significance

Based on the significance of neuronal and cNE STRFs, we observed 4 broad categories of cNEs. *Facilitative cNEs* are cooperative ensembles with significant STRFs, while a majority, if not all, of their member neurons also have significant STRFs (Figure 3.2A – C). We have previously identified these cNEs and interpreted them as canonical processing units that enhance the representation of auditory stimuli over single neurons and random groups of neurons (See et al., 2018). *Stimulus-independent cNEs* are cNEs without significant STRFs, with member neurons that generally do not have significant STRFs but nevertheless contain subsets of spikes

with high synchrony to other cNE members (Figure 3.2D – F). We also identified two other, less common types of cNEs: *constructive cNEs* have significant STRFs while most of its member neurons do not (Figure 3.2G – I) and *non-constructive cNEs* do not have significant STRFs while most of its member neurons do (Figure 3.2J – L).

We grouped the entire population of identified cNEs into these four categories via their STRF significance and the STRF significance of their neuronal members (all spikes). cNEs with significant STRFs were classified as a facilitative cNE if ≥50% of its member neurons had significant STRFs; if <50% of its member neurons had significant STRFs, it was classified as a constructive cNE. cNEs without a significant STRF were classified as stimulus-independent if ≤50% of its member neurons had significant STRFs; if >50% of its member neurons had significant STRFs, it was classified as a non-constructive cNE (Figure 3.3A). Across all identified cNEs (N=487), stimulus-independent cNEs were most common (52%), followed by facilitative cNEs (33%), constructive cNEs (11%) and non-constructive cNEs (4%; Figure 3.3B).

When we compared the mean difference between the reliability index (RI) of cNEs and their member neurons, with the mean RI of member neurons, the groups were well-clustered; facilitative cNE generally had member neurons with high average RI (>0.1); most constructive cNEs had mean differences above 0 while most non-constructive cNEs had mean difference below zero; and most stimulus-independent cNEs had member neurons with average RI close to zero (<0.02; Figure 3.3C). The four categories of cNEs also had distinct physical and functional properties. Stimulus-independent and non-constructive cNEs (i.e., non-significant cNEs) had member neurons that were located deeper in the cortical column, followed by facilitative and constructive cNEs (significant cNEs; Figures 3.3D, 3.S2C). Facilitative cNEs were the largest in size, and were significantly larger than constructive or stimulus-independent cNEs (Figure 3.3E).

cNEs with significant STRFs tended to be larger in comparison to cNEs with non-significant STRFs (Figure 3.S2D). Constructive cNEs had neuronal member pairs with the sharpest cross-correlation functions, which means that they had the tightest spike correlation in time, followed by facilitative cNEs, while the stimulus-independent and non-constructive cNEs had broader cross-correlation functions (Figures 3.3F and 3.S2E). When comparing stimulus driven cNEs with their corresponding spontaneous cNEs (See et al., 2018), non-significant cNEs (correlation value = 0.63 ± 0.20) were not statistically more similar to their spontaneous counterparts than significant cNEs (correlation value = 0.63 ± 0.26). In other words, both significant and non-significant evoked cNEs are stable entities equally well-represented in driven and spontaneous activity. These comparisons of some physical and functional properties of the different cNE categories suggest that the proposed segregation of cNE types is a meaningful one, and supports the idea that multiple modes of information processing exist in AI expressed in the activity of local, synchronized networks.

STRF properties of cNE categories

We previously reported that cNEs in AI enhanced the encoding of auditory information over single neurons or random groups of neurons (See et al., 2018). In that study, we did not select neurons and cNEs based on the significance of their STRFs. Here, we compared the RI of facilitative cNEs, by definition with significant STRFs, against the average RI of member neurons also with significant STRFs, to perform an unbiased analysis of the stimulus encoding properties. Facilitative cNEs had a significantly higher RI than their member neurons, with ~67% of the data having a higher RI in cNEs than in neurons (Figure 3.4A), reinforcing the idea that stimulus-dependent cNEs encode acoustic features more reliably than their member neurons. Since we compared the mean RI of significant member neurons against cNE RI, we wanted to

assess if the increased stimulus-encoding capacity was driven primarily by the neuronal members with the strongest STRFs (i.e., neurons with the highest RI). We calculated the difference between cNE RI and the RI of individual member neuron with either the highest or lowest IC weight (above threshold for membership; Figure 3.S3A). If the neurons with the highest IC weight contributed the most to each cNE, we would expect the RI difference to be small and potentially negative. We did not observe any significant difference between the RI difference of the maximum IC weight neuron and the minimum IC weight neuron (Figure 3.S3A). When comparing the reliability index difference against the IC weight of all significant neuronal members of all facilitative cNEs, we saw a slight trend towards lower RI difference values as IC weight increases, but this dependency only accounted for about 0.02 of the variance in the dataset (Figure 3.S3B; $R^2 = 0.024$). Thus, the encoding of acoustic information by cNEs is not dominated by single member neurons with the most significant STRFs, but appears to reflect coordination across all member neurons.

We used this comparison in other cNE categories to verify that the cNE classification was accurate and meaningful. Similar to the above comparison, the RI of cNEs were compared against significant or non-significant member neurons, depending on which group formed the majority of member neurons. Stimulus-independent cNEs were compared against their non-significant member neurons, and both cNE RI and average neuronal RI were clustered around 0 ($|RI| < 0.06$; Figure 3.4B), with ~56% of the data having higher cNE RI than neuronal RI. Constructive cNEs were compared against their non-significant member neurons, and cNE RI was significantly higher than that of their member neurons (Figure 3.4C; 100% of data had cNE RI greater than neuronal RI). Non-constructive cNEs were compared against their significant member neurons, and cNE RI was significantly lower than that of their member neurons (Figure

3.4D; only 21% had cNE RIs exceeding that of member neurons). These results verified that our classification method was robust and functionally consistent.

The next question was whether columnar neurons with similar STRFs were preferentially selected for by a given cNE. Because only facilitative and non-constructive cNEs had a majority of member neurons with significant STRFs, we focused on these two cNE categories. We found indeed that STRF similarity between pairs of neuronal members in facilitative cNEs were higher than that of pairs of simultaneously-recorded columnar neurons that were not contributing to the same cNE (Figure 3.4E). Surprisingly, the opposite was seen for non-constructive cNEs; pairs of neuronal members in non-constructive cNEs had lower STRF similarity than pairs of neurons not in the same cNE (Figure 3.4E). Neuronal members in facilitative cNEs having more similar receptive fields than the general population of neurons supports the idea of more selective and less noisy encoding by coordinated, and potentially more strongly connected, ensembles. Meanwhile, non-constructive cNE members having less similar STRFs than the general population of neurons supports the idea that these cNEs - and potentially stimulus-independent cNE - perform a different, potentially higher-level task than enhancing stimulus-specific information and may reflect other processing aspects, e.g. related to behavioral task structure, predictive coding, multi-modal inputs, or other non-sensory, state-dependent modulations (see Discussion).

cNE-associated spiking-events have superior stimulus-encoding properties

We next wanted to parse the mechanism by which cNEs enhance auditory information. We know from previous studies that neurons with significant STRFs respond to the stimulus in a non-linear manner – the more similar the match between the neuron's STRF and a stimulus segment (projection value), the higher the probability of firing, once it passes a threshold value

for the projection value (Atencio and Schreiner, 2012). In other words, some neuronal spikes correspond to poorer matches between the stimulus and the STRF than others as reflected in the shape of the input/output nonlinearity. Hence, we tested the hypothesis that cNEs select for neuronal member spikes associated with strong matches between stimulus segments and each neuron's STRF. We split each neuron's spike train into two different subsets: the cNE-associated spike train (cNE-a) includes only spikes that correspond to cNE events and the cNE-independent associated spike train (cNE-i) that includes the remaining spikes (Figure 3.5A and B). After sub-sampling the spike trains such that each contained equal number of spikes, we found that the STRFs corresponding to the cNE-a spikes had the strongest STRFs, followed by the STRFs from all spikes, and then the STRFs from cNE-i spikes (Figure 3.5C).

To ensure a fair comparison across the entire population of recorded neurons and detected cNEs, we compared spikes of neurons that had significant STRFs, and were part of at least one significant cNE only. We also took the additional step of ensuring that the spike waveforms of cNE-i and cNE-a subsets were similar by comparing the distribution of their spike waveforms (Figure 3.S4A and B), to confirm that the segregated subsets of spikes originated from the same neuron. We found that RIs for cNE-a spikes were higher than RIs for cNE-i spikes (Figure 3.5D) and that they were also significantly higher than that of all spikes in a train (Figure 3.5E), suggesting that facilitative and constructive cNEs select for spikes that encode stimulus segments with high projection values. Using RI as a metric for evaluating STRF strength shows that cNE-a spikes are much more reliable in encoding the same stimulus parameters over time than cNE-i spikes.

93

Spikes of a neuron associated with several cNEs can differ in their stimulus-encoding properties

A single neuron can be a member of more than one cNE (see Figure 3.1). We found that ~20% of neurons in our sample contributed spikes to more than one cNE (Figure 3.7E). Do spikes that are associated with the activity of different cNEs encode the same stimulus aspect? To answer this question, we looked at neurons with significant STRFs that were part of two or more significant cNEs ("shared neurons"; Figure 3.6A). For each "shared neuron", we split its spike train into two subsets, each representing spikes that were uniquely associated with one of the cNEs (Figure 3.6B). We sub-sampled the trains to equalize the number of events and derived their STRFs (Figure 3.6C). In general, STRFs from spikes of the same neuron but associated with different cNEs had similar frequency preferences but differed in their distribution of excitatory and inhibitory subfields (Figure 3.6C). This difference is reflected in each subset's modulation transfer function (MTF; calculated by taking the two-dimensional Fourier transform of each STRF; see Atencio and Schreiner, 2010 for more details) that captures the encoded spectral and temporal modulation frequencies. In the sample neuron, the spectral and temporal modulation profiles of the subsets revealed lower preferred spectral and temporal modulation frequencies for spikes associated with cNE 4 compared to spikes of the same neuron associated with cNE 3 (Figure 3.6D).

To quantify STRF differences between two cNE-associated spike trains of a "shared neuron", we compared the similarity of the real STRFs against a null distribution derived from combining spikes from each pair of STRFs, randomly sorting them into two spike trains, and calculating the STRF similarity over multiple iterations (Figure 3.S5A). A "uniqueness index" was then assigned by the percentile of the real STRF similarity value relative to the null

distribution (Figure 3.S5B). If the segregation of spikes by unique association with a cNE is functionally significant, then the resulting STRFs would be more dissimilar than would be expected by chance. Indeed, across the entire population of "shared neurons", nearly 50% had "uniqueness index" values of <0.05, and 26% had "uniqueness index" values of <0.001. For the population, the real STRF similarity values were significantly lower than the median of the null distributions (paired t-test, $p < 0.0001$; Figure 3.6E), implying that "shared neurons" can encode quite different stimulus parameters, mostly distinguished by their spectro-temporal modulation profiles, based on association with different cNEs.

To ensure that the significant differences we have observed in the STRFs of subsets of neuronal spikes cannot be trivially attributed to the differences in their cNEs' STRFs, we compared the STRF similarity between cNEs and neuronal subsets, sub-sampled to give equal numbers of events or spikes (Figure 3.S6A), and found that the neuronal subset STRF similarity was significantly lower than that of the corresponding cNE STRF similarity (paired t-test, $p<0.0001$; Figure 3.S6B).

Taken together, spikes from a single neuron that are associated with different cNEs can significantly differ in their STRFs and corresponding modulation preferences. This implies that STRFs derived from the full spike train of a neuron only represent an upper bound for estimating stimulus preference. A neuron's spike-to-spike variations in stimulus preference can thus be traced to rapid changes in the association with different local networks.

Leveraging on our cNE categories and our improved understanding of the potentially heterogeneous functional nature of events in a spike train, we investigated if neurons that were part of a cNE with a significant STRF (facilitative or constructive) and also a non-significant cNE (stimulus-independent or non-constructive) had subset of spikes that, correspondingly, were

95

stimulus-dependent and stimulus-independent, based on unique association with either a significant and non-significant cNE respectively (Figure 3.7A and B). As expected, spike subsets associated with non-significant cNEs also revealed non-significant STRFs while subsets associated with a significant cNE had significant STRFs, despite the fact that the subsets of spikes came from the same neuron (Figure 3.7C). This difference is also reflected in the RI, with the subset of spikes associated with significant cNEs having significantly greater RI than that of the subset of spikes associated with non-significant cNEs (Figure 3.7D). This confirms that individual spikes of a neuron that are associated with different cNEs can not only differ in what acoustic feature to encode (see Figure 3.6), but also can differ in whether it encodes a specific stimulus feature at all.

Across the entire population of recorded neurons (n = 2716), only ~20% of neurons were part of more than 2 cNEs (Figure 3.7E). Out of those neurons (n = 534), ~69% were part of 2 or more of the same category of cNE, while ~31% were part of two or three unique categories of cNEs (Figure 3.7F – H). For neurons participating in only one type of cNE, the cNE type was about equally distributed between stimulus-independent and facilitative cNEs (Figure 3.7G). For neurons that participated in multiple unique types of cNEs (Figure 3.7H), almost all combinations of cNE types were observed, with the most common combination being stimulus-independent cNEs and either constructive or facilitative cNEs (~62%).

Taken together, tagging individual spikes within a spike train of an individual neuron by the membership to different cNEs in the cortical column reveals that spike-by-spike information can vary greatly with regard to the nature of what is conveyed about the stimulus from highly specific frequency and envelope information to other, currently unknown aspects not obviously linked to the momentary stimulus waveform.

3.4 Discussion

The aim of this study was to characterize functional properties of cNEs in the cortical column and to enhance our understanding of potential biological interpretations of these functional, multi-neuronal constructs that exhibit enhanced information processing capabilities in AI (See et al., 2018). To achieve this, we classified cNEs and their associated neurons in this predominantly sensory cortical area based on their ability to encode specific acoustic features as reflected in their STRFs. We found that about half of cNEs had STRFs with high specificity to acoustic features with the remaining cNEs showing seemingly stimulus-unspecific activity. The observation that a given neuron can be associated with several cNEs enabled us to identify distinct subsets of spikes based on their association with different, stimulus-specific or unspecific cNE events. This revealed that a given AI neuron can generate functionally heterogenous, intermixed sets of spikes that encode markedly different pieces of information in the presence of an ongoing stimulus. The findings support the notion that cNEs can be considered a fundamental processing unit with coordinated ensemble events encoding a relatively specific type of information. By contrast, spike trains of individual neurons process, through multiplexing, several pieces of stimulus-associated information.

Classification of cNEs

In this study, we classified cNEs into four distinct groups (facilitative, constructive, stimulus-independent and non-constructive) based on the significance of cNE STRFs and the proportion of member neurons with significant STRFs (Figure 3.3). The cutoff between facilitative and constructive cNEs and between stimulus-independent and non-constructive cNEs was set at the 50% level (Figure 3.3A). While there were functional and physical differences

between the four cNE types classified in this manner (Figure 3.3C – G), this cutoff was ultimately arbitrary and done for the convenience and power of data analysis. In the broader schema, there are two general types of cNEs – stimulus information-enhancing cNEs (comprising facilitative and constructive cNEs) and stimulus-agnostic cNEs (comprising stimulus-independent and non-constructive cNEs) as differentiated by the significance of cNE STRFs. In each of these two broad cNE types, however, we postulate that there exist different degrees of enhancing or diminishing the encoding of specific stimulus aspects, when compared to their member neurons. For instance, the classification of facilitative and constructive cNEs constitutes a spectrum from slightly enhancing auditory information (i.e., when the cNE and all member neurons have significant STRFs), to strongly enhancing auditory information (i.e., when the cNE has a significant STRF but not any of its member neurons). The flip side is equally diverse, with highly non-constructive cNEs (i.e., when the cNE has no significant STRF while all its members do) that are strongly stimulus-unspecific, while the most stimulus-independent of cNEs i.e., neither the cNE nor its member neurons have significant STRFs) reflect an apparent stimulus-independence of its member neurons.

The observation that ~56% of cNEs showed non-significant STRFs suggests that their prevalence is comparable to that of non-significant individual neurons with an occurrence rate of ~62%. It should be noted that the lack of an STRF does not mean that the neuron/cNE is not responsive to some kind of information related to or conveyed by the stimulus. This is especially true given the fact that multi-unit, pure-tone mapping at recorded sites showed that multi-unit activity at these sites produced well-defined tuning curves. The high percentage of cNEs without a clear relation to acoustic stimulus features indicates a strong contribution of cortical activity that may relate to other aspects of cortical processing, including stimulus or behavioral context

and other top-down influences, such as decision making, memory, and prediction (Brosch et al., 2015; Rodgers and DeWeese, 2014; Teufel and Nanay, 2017). Single neuron recordings in previous studies had pointed to the frequent encounter of unresponsive neurons, at least with regard to modulations in firing rate by relatively simple sounds (Bendor and Wang, 2005; Hromádka et al., 2008). This was interpreted as evidence of sparse coding at least at the level of firing strength. It remains unresolved, however, whether the firing rate of cNEs is an appropriate measure for sparseness given that it is the synchrony among neurons that defines cNE operation with firing rate a measure of salience rather than activation per se. While it is possible, it is rather unlikely that stimulus-independent cNEs are based on random events without biological meaning. The statistical probability that all member neurons of a cNE fire together by chance is exceedingly small (See et al., 2018).

We have previously discussed how information-enhancing cNEs are useful for auditory processing in AI, including roles in enhancing neural signal-to-noise ratio, stimulus selectivity, and signal transmission (See et al., 2018). But what is the functional significance of ostensibly stimulus-agnostic cNEs? Many studies have shown that populations of AI neurons can encode the recognition of a stimulus, or the decision of the animal in a behavioral task (Bathellier et al., 2012; Francis et al., 2018; Rodgers and DeWeese, 2014). As such, we speculate that stimulus-independent and non-constructive cNEs are the result of the convergence of top-down inputs from higher-order cortical areas, and may represent coordinated activity in neurons that encode for higher-order, conceptually-driven auditory functions including stimulus contexts and probabilities, predictive signaling, expectations, motivation, memory, and task-related information (for review see Buzsáki, 2010; Chambers and Rumpel, 2017; Harris and Mrsic-

Flogel, 2013; Yuste, 2015). Identifying these potential information-bearing aspects of the stimulus-independent cNEs requires future studies in awake and behaving animals.

Selection of neurons by facilitative and non-constructive cNEs

Non-constructive and facilitative cNEs share one major similarity – a majority of their member neurons have significant STRFs. Based on this, we analyzed the pairwise similarity of neurons in facilitative or non-constructive cNEs and found that facilitative cNEs contain neurons with more similar STRFs than the rest of the neurons from the same recording while non-constructive cNEs contain neurons with less similar STRFs than the rest of the neurons from the same recording (Figure 3.4E – G). In light of the idea that stimulus information-enhancing and stimulus agnostic cNEs might be performing different functions, we speculate that the facilitative cNEs and non-constructive cNEs are similarly performing very different roles; facilitative cNEs predominantly receive input from bottom-up sources, while non-constructive cNEs appear to be more strongly influenced by top-down projections. This concept that sensory cortical processing comprises different kinds of information is not surprising, because non-constructive cNEs appear to integrate the activity of neurons with different STRFs, which can be generalized to neurons that transmit different types of information, since neuronal heterogeneity has been shown to enhance the efficiency of encoding (Chelaru and Dragoi, 2008; Marsat and Maler, 2010; Osborne et al., 2008; Winkowski and Kanold, 2013). This might account for the presence of a high percentage of cNEs in AI that might code information beyond acoustic features (Bathellier et al., 2012; Francis et al., 2018; Rodgers and DeWeese, 2014).

Heterogeneity of neuronal spikes and the efficacy of neuronal coding

Based on temporal association of individual spikes in neurons to cNE events, we were able to assign spikes in a given neuron to subgroups that encode different types and degrees of information (Figures 3.5 – 3.7). The ability to functionally tag individual spikes in a train from a single neuron reveals that that neuron can virtually simultaneously be associated with different spectro-temporal stimulus features as well as with spectro-temporally unspecified processing aspects. The idea of multiple neuronal codes is not new, and several studies have shown that this can take place in terms of the coding of different stimulus features (Lankarany et al., 2019; Walker et al., 2011), different temporal scales of a stimulus (Atencio et al., 2008; Kayser et al., 2009), or different types of coding, e.g., rate coding vs. temporal coding (Ainsworth et al., 2012; Harvey et al., 2013), or spikes vs. local field potentials (Belitski et al., 2008; Perel et al., 2015). What we were able to show here appears to be a different type of multiplexing, one at a much finer scale, where individual spikes could encode different types of information regarding a stimulus and representation of the stimulus (STRFs) using the same currency (neural spikes) but in the context of coordinated network activity (cNE events). The ability to isolate subgroups of neuronal spikes with differing functional association is due to the fact that we could isolate cNEs, and it provides the link between multiplexing in neuronal coding and network activity. Spike-to-spike multiplexing can increase the encoding capacity of neuronal responses and enables stable representation of distinct stimuli aspects.

Functional reliability of cNEs and neurons

In our previous report, we made a case for cNEs over single neurons as the principal unit of information processing in AI based on its functional reliability (See et al., 2018). We define functional reliability in AI as the ability of the neuron/cNE to maintain and/or enhance the

fidelity of the encoding of relevant sound features, with the reliability index of an STRF as a proxy. We added that even after selecting only for neurons and cNEs with significant STRFs (facilitative cNEs), cNE STRFs had higher reliability indices than their member neurons (Figure 3.4A). We further showed that the functional association of a neuron's spikes can be heterogenous, i.e., they can be related to different cNE types and encode differing STRFs (Figures 3.5 – 3.7). This renders the average spike-associated information inherently unreliable for identifying what stimulus features are encoded over the whole duration of the stimulus. Neurons that take part in at least one significant and one non-significant cNE are examples that some spikes encode spectro-temporal features and other spikes may encode other, not spectro-temporally specified types of information (Figure 3.7E and F). By contrast, cNE events appear to encode only one type of these conceptually different stimulus information aspects by reading out the jointly associated events from several neurons. This observation is consistent with ideas that reconcile the stability of critical brain functions with the volatility of synaptic and single-neuron properties: despite the underlying dynamism of neurons and synapses that are essential for learning and adapting, ensembles of neurons have attractor-like dynamics that provide stable outputs (Chambers and Rumpel, 2017; Laje and Buonomano, 2013). Taken together, these observations reinforce the idea that cNEs, due to their higher information specificity, could be considered the principal unit of information processing in AI.

Relation of spike heterogeneity to arousal and behavioral training

The functional heterogeneity of neuronal spikes in AI has likely behavioral implications. We would predict that behavioral requirements would modulate which specific aspect among the several intermingled information aspects of a neuron is most useful. This could be achieved by increasing the firing rate associated with one aspects over that of other aspects with

consequences for the average spike-triggered activity. Previous work has shown that the attentional state of an animal was sufficient to cause clear and rapid changes in a neuron's receptive field (Fritz et al., 2003, 2005b; Yin et al., 2014). This effect has been shown to be reversible when an animal stops attending to the task, begins passively listening, or is forced to change its decision-making goal. This has been attributed to "rapid plasticity" in AI (Fritz et al., 2003, 2005b; Yin et al., 2014). We propose that this could instead be due to changes in the heterogeneity of spikes that we have described above. Since neurons can adopt somewhat different spectro-temporal receptive field attributes when associated with different cNEs (Figure 3.6), rapid changes in the dominant (average) receptive fields can occur when an animal engages in a detection or discrimination task that biases a neuron's activity toward the most useful cNE, potentially selected for via top-down modulations. This idea is reminiscent of how novel environments 'select' pre-existing temporal sequences in the hippocampus to represent the novel environments instead of creating these temporal sequences *de novo* (Dragoi and Tonegawa, 2013). Hence, rapid plasticity (Fritz et al., 2003, 2005b; Yin et al., 2014) could be interpreted as a reinforcement of more useful network configurations that are already in place. Synaptic plastic changes would not be required at first, although prolonged behavioral training likely will solidify those selections via regular synaptic plasticity. This hypothesis is also compatible with the fact that trial-to-trial variability to a repeated stimulus decreases with increased arousal (Cohen and Maunsell, 2009; Mitchell et al., 2009), and may represent the bias of neurons to fire in a limited or fixed network configuration, dictated by neuromodulation or higher-order cortical influences.

Taken together, we postulate that auditory cortical processing is based (i) on neurons with functionally more heterogeneous spikes and (ii) local networks or cNEs with functionally more homogeneous events focusing on one of several different types of stimulus-associated

information. Testing is tractable for larger populations of neurons. If we could record from all

neurons in AI, we predict that we would see that all neurons would participate in multiple cNEs

and that each neuron would participate in a mixture of what we currently call spectrotemporally-

specific and -unspecific stimulus aspects. The nature of the "unspecific" parts still need to be

determined in detail, while animals are engaged in various tasks or are in different states.

Identifying the different types of information encoded by cNEs would allow identifying the

manifestation and role of different sensory processing aspects for tracing efficient information

transmission and integration between top-down cognitive areas and bottom-up auditory areas.

Each neuronal spike would be associated with a firing mode or be part of a functionally

identifiable "network configuration". The identity of cNEs would mostly be stable, especially

with the "core members" of the cNE, but more marginally associated members might drop out as

the state and/or arousal level in an animal changes, in concordance with the decreased noise

correlation seen in awake behaving animals (Ni et al., 2018; Stringer et al., 2016). Increased

arousal would lead to smaller but more reliable cNEs, as top-down influence would cause

networks to be biased towards those that are optimized for carrying out the task at hand.

3.5 Materials and methods

Electrophysiology

Detailed experimental procedures are described in a previous study (See et al., 2018). All experiments were approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco. 16 female Sprague-Dawley rats (200–300 g, 2–3 months old; RRID: MGI:5651135) were anesthetized with a mixture of ketamine and xylazine. The primary auditory cortex (AI) was mapped by using multi-unit responses to pure tones of different frequencies (1–40 kHz) and intensities (10–80 dB). Regions with short latencies (10–30 ms) and a tonotopic gradient in the rostrocaudal axis were identified to be AI (Polley et al., 2007). Recordings were made using either one of three 32-channel probes (a1x32-poly1, a1x32-poly2, a1x32-poly3; NeuroNexus), or a 64-channel probe (H3; Cambridge NeuroTech), inserted perpendicular to the cortical surface to a depth of approximately 800 - 1400 µm using a microdrive (David Kopf Instruments). Neural traces were band-pass filtered between 500 and 6000 Hz and were recorded to disk at 20 kHz sampling rate with an Intan RHD2132 Amplifier system. Traces were spike-sorted offline using MountainSort, a fully automated spike sorter (Chung et al., 2017).

Stimulus

The stimulus was either a dynamic moving ripple or a ripple noise (Escabi and Schreiner, 2002; See et al., 2018). The dynamic moving ripple (DMR) was a temporally varying broadband sound (500 Hz – 40 kHz) made up of approximately 50 sinusoidal carriers per octave, each with randomized phase. The maximum spectral modulation frequency of the DMR was 4 cycles/oct, and the maximum temporal modulation frequency was 40 cycles/s. The maximum modulation

depth of the spectro-temporal envelope was 40 dB. Mean intensity was set at 30–50 dB above the average pure tone threshold within a penetration. The ripple noise (RN) stimulus was the sum of 16 independently created DMRs. Both DMR and RN were presented as long and continuous 10-, 20- or 30-min stimuli. Spectro-temporal receptive fields (STRFs) were calculated by spike triggered averaging of the DMR and RN stimuli (Figure 3.S1A).

Coordinated neuronal ensemble (cNE) detection algorithm

All data analysis was performed in MATLAB (Mathworks). Complete details of the cNE detection algorithm used in this study is described in a previous study (See et al., 2018). The cNE detection algorithm used in this study was based on dimensionality reduction techniques (Lopes-dos-Santos et al., 2013). Principal component analysis (PCA) was applied to the z-scored spike matrix (10-ms bins) to obtain an eigenvalue spectrum. Eigenvalues that exceeded the upper bounds of the Marčenko-Pastur distribution (Marčenko and Pastur, 1967) were deemed significant and represented the number of detected cNEs. The eigenvectors corresponding to significant eigenvalues were then processed using independent component analysis (ICA). The resulting independent components (ICs) represent the contribution of each neuron to each cNE. The activity of each cNE was calculated by then projecting the ICs back onto the z-scored spike matrix (Lopes-dos-Santos et al., 2013). cNE membership and binarized cNE activity were determined using surrogate methods, which are detailed in a previous study (See et al., 2018).

Reliability index (RI) and STRF significance

A reliability index (RI) metric, adapted from Escabí et al., 2014, was used to quantify the strength of STRFs from both neurons and cNEs. To calculate RI, all spike or cNE event trains were divided into 1-minute segments, and the STRF of each 1-minute segment was calculated.

The set of 1-minute segments was split into two equal groups (segment A and B) and two STRFs (STRF A and B) were obtained by addition of the STRFs from each of the corresponding 1-minute segments. The similarity of STRF A and B was then calculated using Pearson's correlation. This STRF similarity was calculated over 100 iterations, with different combinations of 1-minute segments for each iteration. RI was then calculated as the mean STRF similarity over the 100 iterations (Figure 3.S1B).

To determine the significance of each STRF, each corresponding spike or cNE event train was circularly shuffled by increasing multiples of N/100, where N is the number of time bins for each spike or cNE event train, 100 times. For each shuffle, 100 STRF similarity values were calculated as described above, giving rise to 100,000 STRF similarity values, which formed the null distribution of RI. A p-value was calculated by quantifying the proportion of the total distribution (null distribution and real RI value) that was larger than or equal to the real RI value. STRFs with $p < 0.05$ were considered significant, while STRFs with $p \geq 0.05$ were not significant (Figure 3.S1C).

For all direct comparisons of RI between two different groups (Figures 3.3C, 3.4A – D, 3.5D, 3.5E and 3.7D), the group with the larger number of spikes or events was subsampled to match the number of spikes or events of the smaller group. Each random subsampling was done 20 times, and the mean RI of the 20 iterations was used as the RI value for comparison.

Pairwise cross-correlation sharpness

The pairwise cross-correlation sharpness measure (Figures 3.3F and 3.S2E) was calculated based on a previous study (See et al., 2018). For each sharpness value, a pairwise cross-correlation function between two spike or cNE event train was calculated with a time

resolution of 0.5 ms and a maximum delay of 50 ms. The peak delay for each pairwise cross-correlation function was estimated and each function was folded around that delay, to a maximum of 20 ms. Sharpness was then estimated by the delay that accounts for half the spike count in each folded cross-correlation function. For an illustration of this calculation, see See et al., 2018.

3.6 Acknowledgements and author contributions

Acknowledgements

Author contributions

3.7 Figures

Figure 3.1. Example penetration with two sample cNEs.
Depth of some neurons and their respective STRFs are shown. Two example cNEs are represented. cNE 1 (orange) is mostly made up of neuronal members with significant STRFs, and also has a significant STRF. cNE 8 (blue) is mostly made up of neuronal members with no significant STRFs, and also does not have a significant STRF. One neuron, highlighted with a red square at a depth of approximately 600 µm, is part of both cNE 1 and cNE 8.
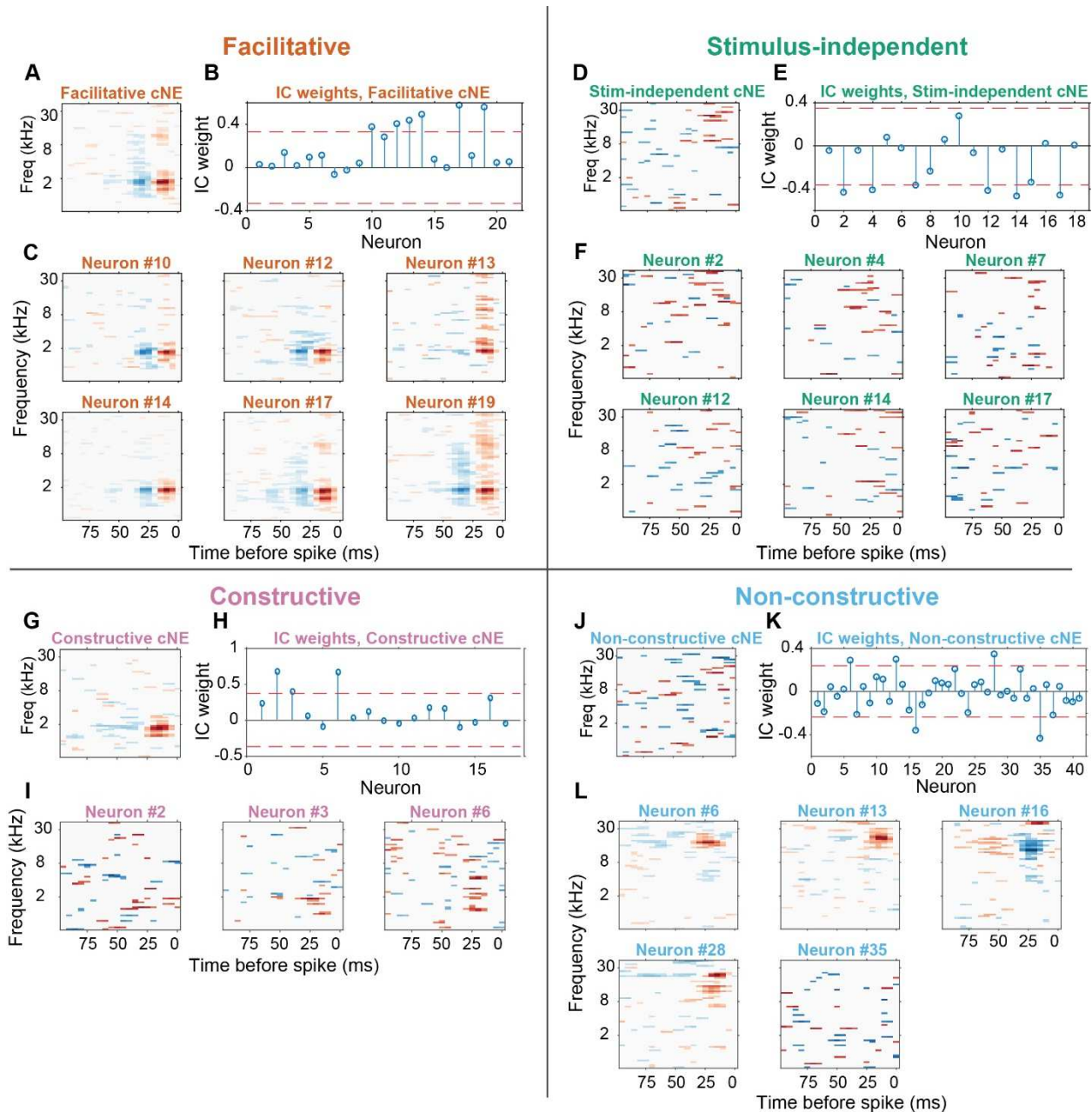
Figure 3.2. STRFs of cNEs and member neurons.
(**A-C**) Facilitative cNE, where the cNE and most of its member neurons have significant STRFs.
(**D-F**) Stimulus-independent cNE, where the cNE and most of its member neurons have no significant STRFs. (**G-I**) Constructive cNE, where the cNE has a significant STRF but most of its member neurons have non-significant STRFs. (**J-L**) Non-constructive cNE, where the cNE does not have a significant STRF but most of its member neurons do have significant STRFs. (**A, D, G, J**) STRFs of the cNEs. (**B, E, H, K**) IC weights for each cNE. The magnitude of the IC weights represents the contribution of each neuron to the cNE. Red dashed lines: thresholds that determine if a neuron is a member of the cNE. (**C, F, I, L**) STRFs of member neurons.
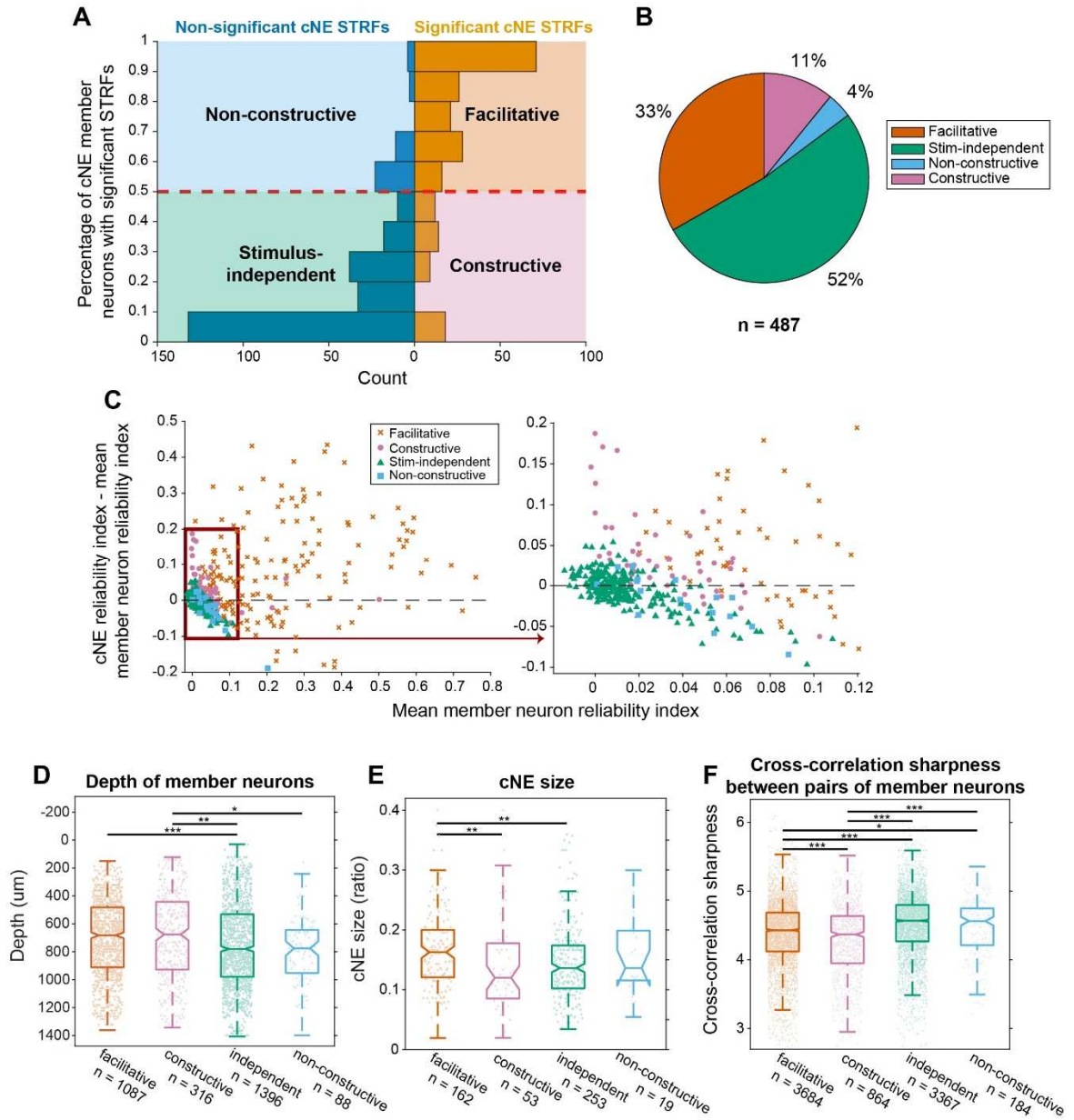
Figure 3.3. Statistics and properties of the different cNE categories.
(A) Illustration of how cNEs were categorized. cNEs with significant STRFs and ≥50% of member neurons with significant STRFs were facilitative cNEs; cNEs with significant STRFs and <50% of member neurons with significant STRFs were constructive; cNEs with non-significant STRFs and ≤50% of member neurons with significant STRFs were stimulus-independent; and cNEs with non-significant STRFs and >50% of member neurons with significant STRFs were non-constructive. **(B)** Pie chart illustrating the breakdown in percentages of the four different types of cNEs. **(C)** (Left) Mean of the difference between cNE RI and member neuron RI vs mean RI of all member neurons. Only facilitative cNEs exist outside the red square, apart from a few exceptions. (Right) Enlarged version of left graph, indicated by the red square. Constructive cNEs mostly fall above the black dashed line (i.e. cNE RI > member neuron RI), stimulus-independent cNEs fall on either side of the black dashed line and are mostly clustered at low values around 0, while non-constructive cNEs cluster below the black dashed line. The 4 different groups exist mostly in different clusters. **(D)** Depth of member neurons between different types of cNEs. Constructive and facilitative cNEs were found at a shallower depth than stimulus-independent and non-constructive cNEs. **(E)** cNE size, calculated by the number of neuronal members as a fraction of the total number of neurons recorded. Facilitative cNEs were larger than constructive or stimulus-independent cNEs. **(F)** Cross-correlation sharpness (STAR Methods) between pairs of neurons from the same cNE. Stimulus-independent and non-constructive cNEs had broader cross-correlation functions than facilitative or constructive cNEs. *p < 0.05, **p < 0.01, ***p < 0.001, Mann-Whitney U test.
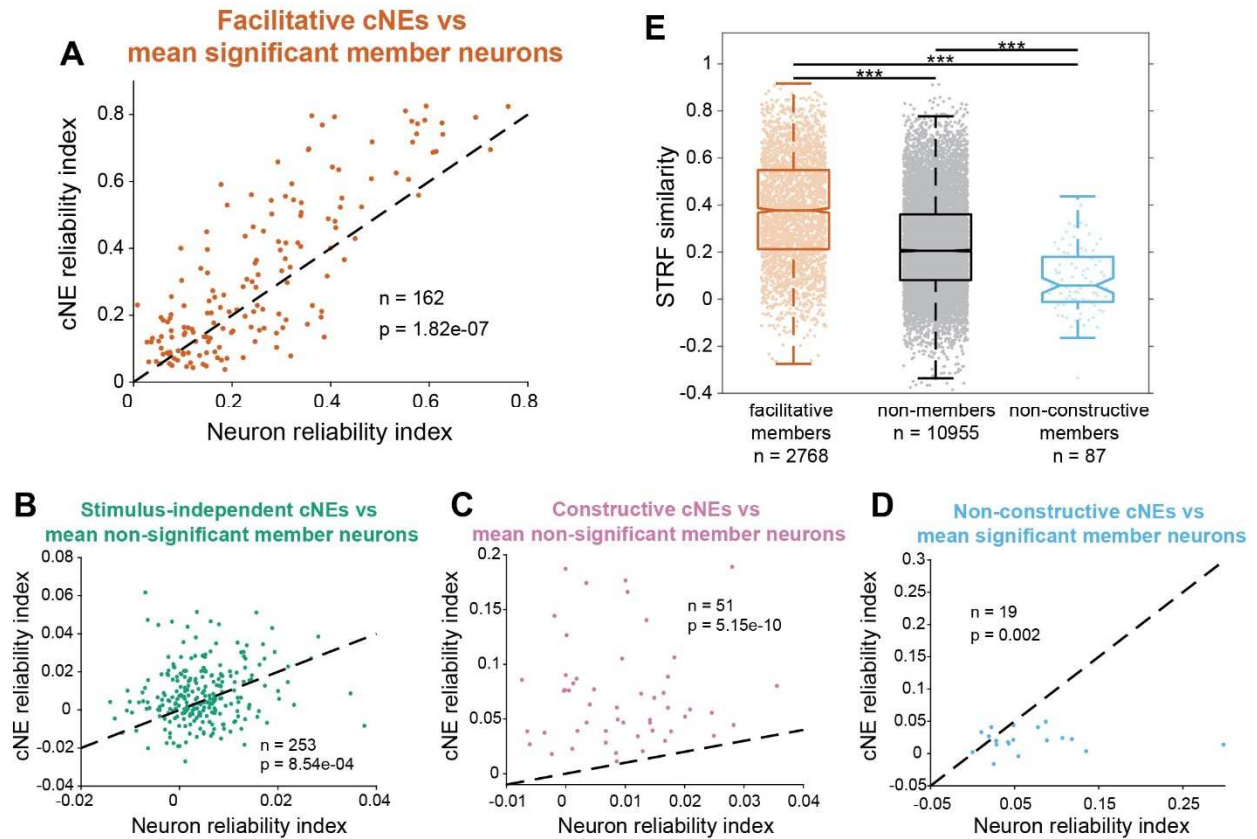
Figure 3.4. Functional relationship between cNE and member neurons.
(A) RI between cNE STRFs and the mean of RI of member neurons with significant STRFs for facilitative cNEs. ≈67% of cNEs had higher reliability indices than that of their member neurons. **(B)** Same as (A), but for neurons without significant STRFs and for stimulus-independent cNEs. ≈56% of cNEs had higher reliability indices than that of their member neurons. **(C)** Same as (A), but for neurons without significant STRFs and for constructive cNEs. 100% of cNEs had higher reliability indices than that of their member neurons. **(D)** Same as (A), but for neurons with significant STRFs and non-constructive cNEs. ≈21% of cNEs had higher reliability indices than that of their member neurons. **(E)** STRF similarity between pairs of neurons within the same facilitative cNE, non-constructive cNEs, and neurons that were not part of the same cNE but recorded simultaneously. Facilitative cNE members had more similar STRFs than that of non-members while non-constructive cNE members had less similar STRFs that that of non-members. **(A-D)** Wilcoxon signed-rank test; **(E-G)** ***p < 0.001; Mann-Whitney U test.
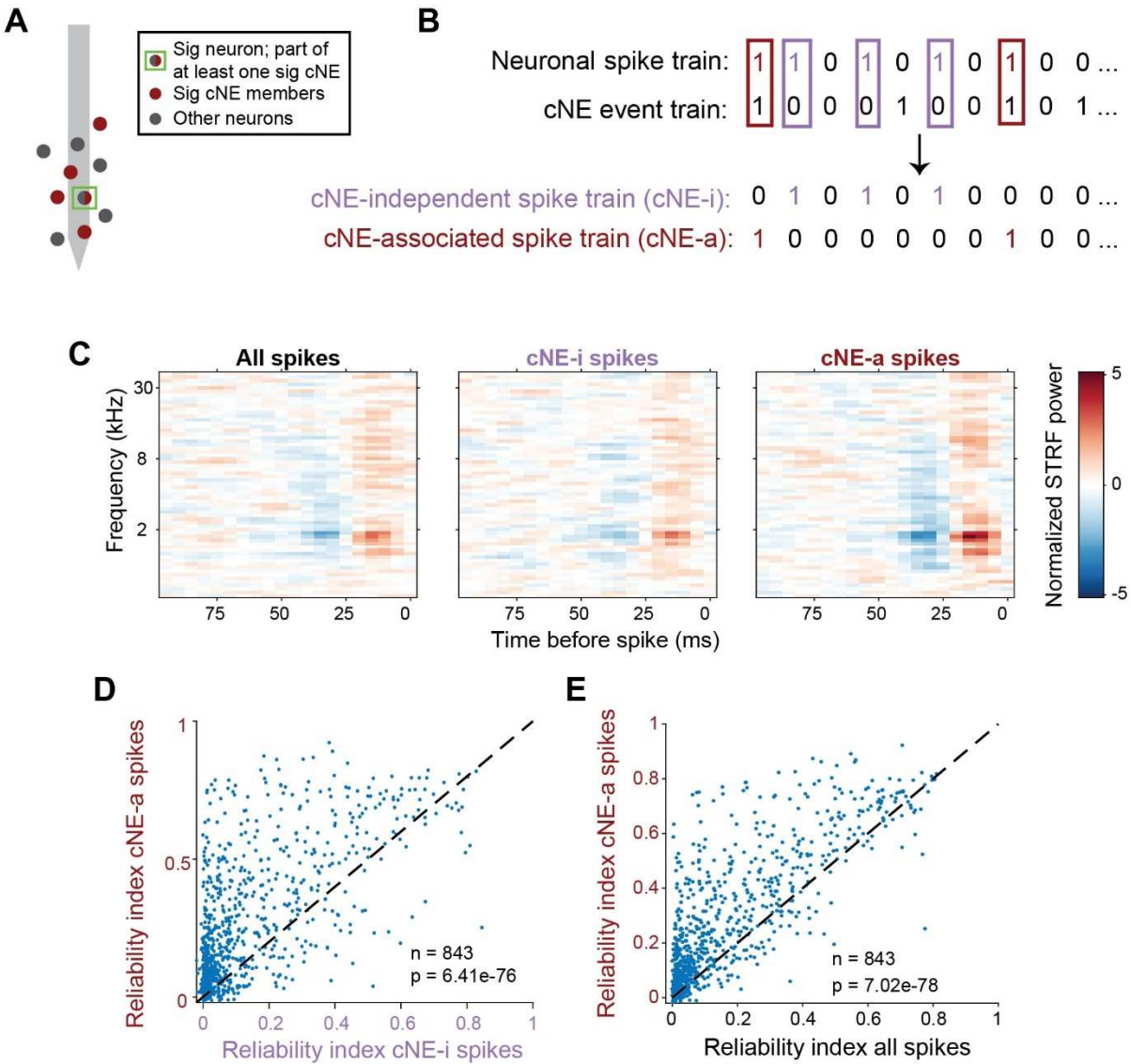
Figure 3.5. cNE-associated (cNE-a) spikes are more stimulus-dependent that cNE-independent (cNE-i) spikes.
(**A**) Schematic of a sample neuron (green square) with a significant STRF that is part of at least one significant cNE. This neuron is part red and part grey to illustrate the splitting of its spikes into smaller subsets. (**B**) Illustration of a neuronal spike train split into cNE-i and cNE-a subset spike trains. (**C**) Sample STRFs for a neuron's entire spike train, cNE-i and cNE-a subset spike trains. Spike trains were sub-sampled to get equal number of spikes across all three spike trains. STRF power was also normalized across all three STRFs. The cNE-a STRF had the highest power, followed by the STRF for all spikes and the cNE-i STRF. (**D**) Across the entire population, cNE-a spikes had higher RI than cNE-i spikes (≈80% above black dashed unity line). (**E**) cNE-a spikes also had higher RI than all spikes (≈79% above black dashed unity line). Wilcoxon signed-rank test was used in (**D**) and (**E**).
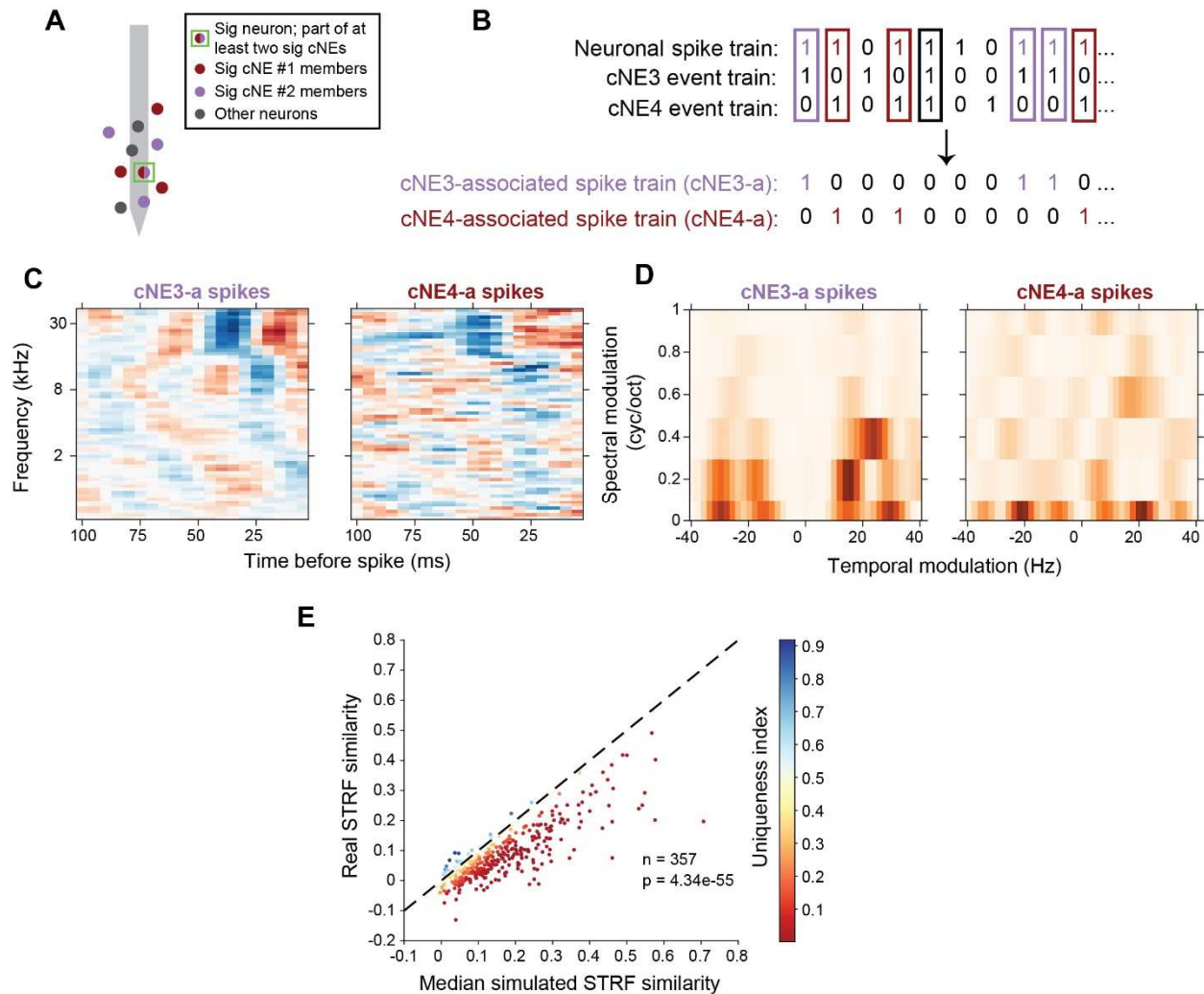
Figure 3.6. STRFs generated by cNE-a subsets of spikes from the same neuron are dissimilar.
(**A**) Schematic of a sample neuron (green square) with a significant STRF that is part of at least two significant cNEs. This neuron is part red and part purple to illustrate the splitting of its spikes into smaller subsets. (**B**) Illustration of a neuronal spike train split into two cNE-a subsets. Purple: spikes associated with cNE3; red: spikes associated with cNE4; spikes associated with both cNEs (black rectangle) or neither of the cNEs are omitted. (**C**) Sample STRFs from the two subset spike trains. The STRFs differ in their temporal profile. (**D**) Sample MTFs from the two subset spike trains. The preferred spectral modulation and temporal modulation frequencies between the two subsets are different. (**E**) Real STRF similarity vs median simulated STRF similarity (null distribution; Figure 3.S5A and B) across the entire population. Real STRF similarity is significantly lower than the null distribution (≈90% below black dashed unity line), implying that spikes associated with different cNEs generated STRFs that were significantly different than that by chance. Wilcoxon signed-rank test was used in (**E**).

A Sig neuron; part of at least one sig cNE and one non-sig cNE
● Sig cNE members
● Non-sig cNE members
● Other neurons

B
Neuronal spike train: 1 1 1 1 0 1 1 0 1 1 ...
Non-significant-cNE-a event train: 0 1 1 0 1 0 1 1 1 0 ...
Significant-cNE-a event train: 1 0 0 1 1 0 1 1 0 1 ...

Non-significant-cNE-a spike train: 0 1 1 0 0 0 0 0 1 0 ...
Significant-cNE-a spike train: 1 0 0 1 0 0 0 0 0 1 ...

C
cNE3-a spikes (stim-independent)
cNE5-a spikes (facilitative)

Frequency (kHz)
Time before spike (ms)
Normalized STRF power

D
n = 61
p = 3.09e-07
Sig-cNE-a reliability index
Non-sig-cNE-a reliability index

E
n = 2716
Ratio
cNE membership

F
n = 534
Ratio
cNE membership, number of unique cNE types
One Two Three

G
n = 370
Ratio
SI F C NC
Unique cNE types

H
n = 164
Ratio
C/SI F/SI C/F NC/SI F/NC C/F/SI C/NC/SI F/NC/SI
Unique cNE type combinations

117

Figure 3.7. STRFs generated by neuronal spikes from the same neuron associated with non-significant cNEs are less stimulus-dependent than those associated with significant cNEs.

(A) Schematic of a sample neuron (green square) that is part of at least one significant cNE and one non-significant cNE. This neuron is part blue and part orange to illustrate the splitting of its spikes into smaller subsets. **(B)** Illustration of a neuronal spike train split into two cNE-a subsets. Blue: spikes associated with the a non-significant cNE; orange: spikes associated with a significant cNE; spikes associated with both cNEs (black rectangle) or neither of the cNEs are omitted. **(C)** Sample STRFs generated from neuronal spikes associated with a stimulus-independent cNE or a facilitative cNE. Spike trains were sub-sampled to get equal numbers of spikes between the two subsets, and the STRF power was normalized. The stimulus-independent cNE-associated STRF was not significant while the facilitative cNE-associated STRF was, despite the fact that the spikes originated from the same neuron. **(D)** RI for significant cNE-a spikes was significantly higher than RI for non-significant cNE-a spikes across the entire population (≈79% above black dashed unity line). **(E)** Histogram of cNE membership for all recorded neurons. ≈20% of all recorded neurons were part of 2 or more cNEs. The red box indicates the total number of neurons represented in **(F)**. **(F)** Histogram of unique categories of cNE membership for all neurons that are part of 2 or more cNEs. ≈31% of all neurons that were part of two or more cNEs (≈6% of all recorded neurons) were also part of two or more different categories of cNE, while the other ≈69% were part of only one type of cNE (≈14% of all recorded neurons). The green box represents the total number of neurons represented in **(G)** and the red box represents the total number of neurons represented in **(H)**. **(G)** Histogram of cNE categories for neurons that were part of two or more cNEs of the same category. **(H)** Histogram of cNE category combinations for neurons that were part of two or more unique categories of cNEs. **(G – H)** SI: stimulus-independent; F: facilitative; C: constructive; NC: non-constructive.

A

Stimulus

Freq (kHz)

Time (s)

Spike-triggered averaging

Estimated STRF

Freq (kHz)

Time before spike (ms)

B

Recording (divided into equal parts)

Randomly sampled equally

Segment A   Segment B

STA calculation

STRF A ⟷ correlation ⟷ STRF B

Reliability index
(averaged over 100 iterations
of random sampling)

C

Circularly shuffle
spike trains and
calculate STRFs,
100 iterations

Freq (kHz)

Time before spike (ms)

Calculate STRF reliability
index, 100 different
combinations for each
shuffled STA (n = 10000)

Calculate STRF
reliability index for real
data

Ratio

n = 10000

Reliability index

p < 0.05   **Significant STRF**

p ≥ 0.05   **Non-significant STRF**

D

Spike Train Matrix

Neuron #

Time (ms)

PCA/ICA

IC 1 = cNE 1   **IC 2 = cNE 2**

ICA weight

Neuron #

Projection

ICA Ensemble (cNE) Response

Normalized
response

Time (ms)

119

Figure 3.S1. Main experimental and analytical methods.
(A) Stimulus sample and derived receptive field. (left) Spectrogram of the dynamic moving ripple stimulus, a temporally varying broadband sound made up of approximately 50 sinusoidal carriers per octave. (right) Spectral-temporal receptive field (STRF), calculated by spike triggered averaging. **(B)** Calculation of the RI to determine if a STRF was significant. Each spike or event train was split up into two random but equal groups and had their STRFs calculated. The RI is the average of the similarity between each pair of STRFs over 100 iterations. **(C)** Method of determining the significance of an STRF. Each spike train was circularly shuffled randomly over 100 iterations to generate 100 pairs of null STRFs. Every possible pairwise combination for STRF similarity was calculated to generated a null distribution (dark green bars) and the null distribution was compared against the real STRF similarity calculated in **(B)** to get a p-value. A STRF is significant if $p < 0.05$. **(D)** Toy model illustrating the detection method for cNEs (see Materials and methods). (top) A spike train matrix is binned, z-scored. Neurons $1 - 5$ and neurons $4 - 8$ were modeled to have a high probability of coordinated activity. (middle) The spike trains were processed with principal component analysis and then independent component analysis to determine cNE membership. IC 1 and 2 identify neurons $1 - 5$ and neurons $4 - 8$ as cNEs respectively. (bottom) The ICs were projected back onto the z-scored spike train matrix to determine cNE activity, and these peaks match the periods of coordinated activity seen in the spike matrix.
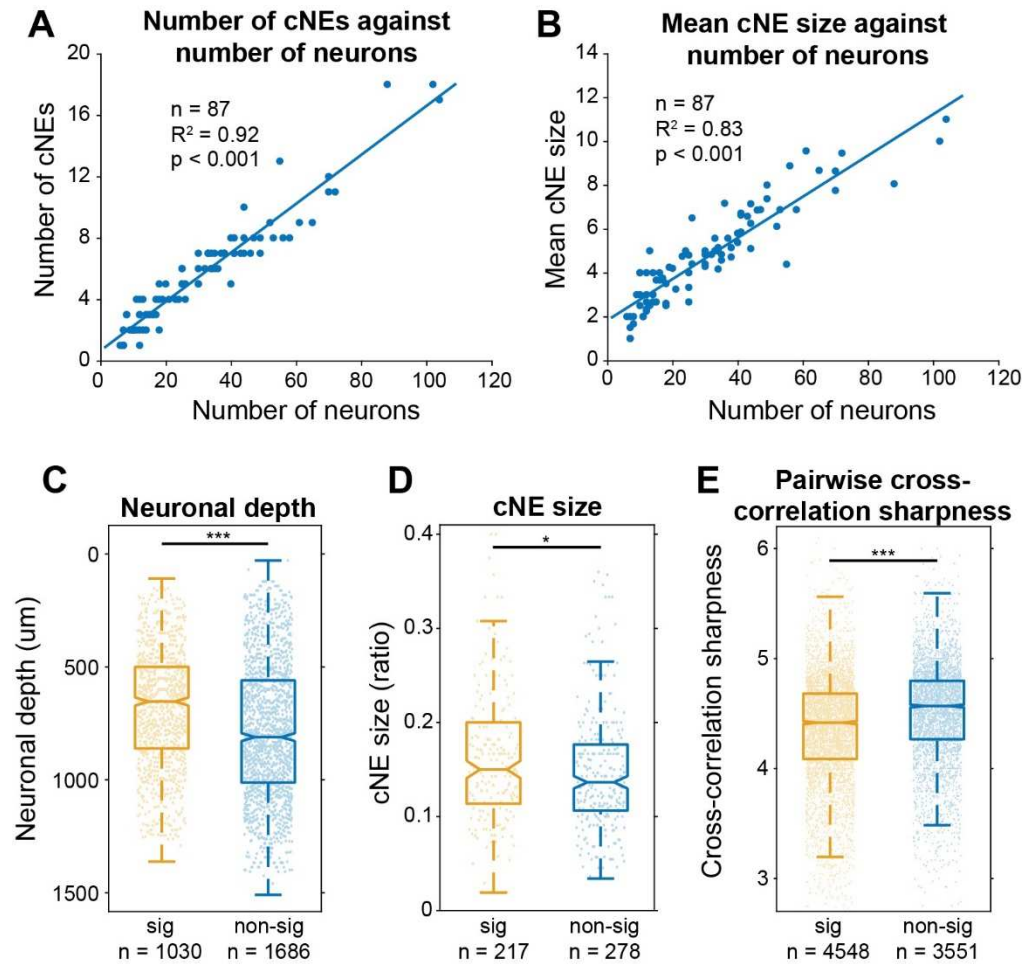
Figure 3.S2. Statistics of significant and non-significant member neurons and cNEs.
(A) The number of cNEs identified is positively correlated with the number of neurons recorded.
(B) The mean cNE size is also positively correlated with the number of neurons recorded. **(C)**
Neurons with significant STRFs were found at a shallower depth than that of neurons with non-significant STRFs. **(D)** cNEs with significant STRFs made up a larger proportion of recorded neurons than cNEs with non-significant STRFs. **(E)** Cross-correlation functions between neuronal members of significant cNEs were sharper than that of non-significant cNEs. *p < 0.05,
***p < 0.001, Mann-Whitney U test for **(C – E)**.
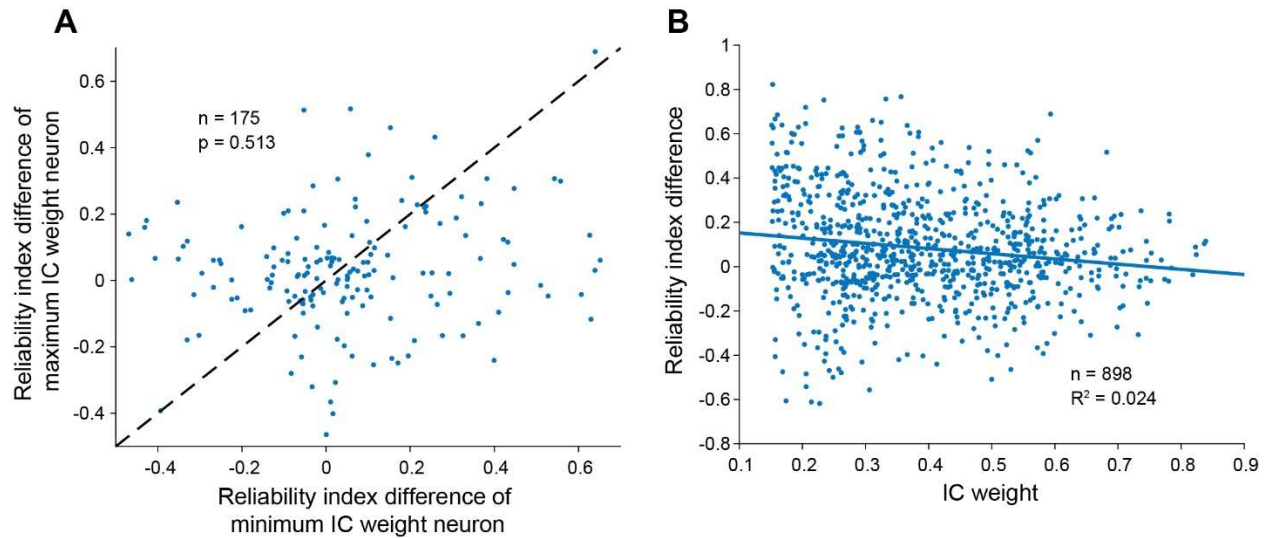
Figure 3.S3. STRFs of facilitative cNEs are not dominated by member neurons with the highest RI.
(A) Difference in RI between a cNE and the member neuron with the highest absolute IC weight vs difference in RI between a cNE and the member neuron with the lowest (but still significant) absolute IC weight. There was no significant disparity in the RI difference between significant neurons with the highest and lowest IC weights. **(B)** Scatter plot of RI difference against absolute IC weight of member neurons. This covariance only accounted for about 0.02 of the variance seen in the data. Wilcoxon signed-rank test was used for **(A).**

A Raw spike waveforms

B Spike waveforms in PC projection space

C Gap statistic, one vs two clusters

n = 2804
p < 100^-10
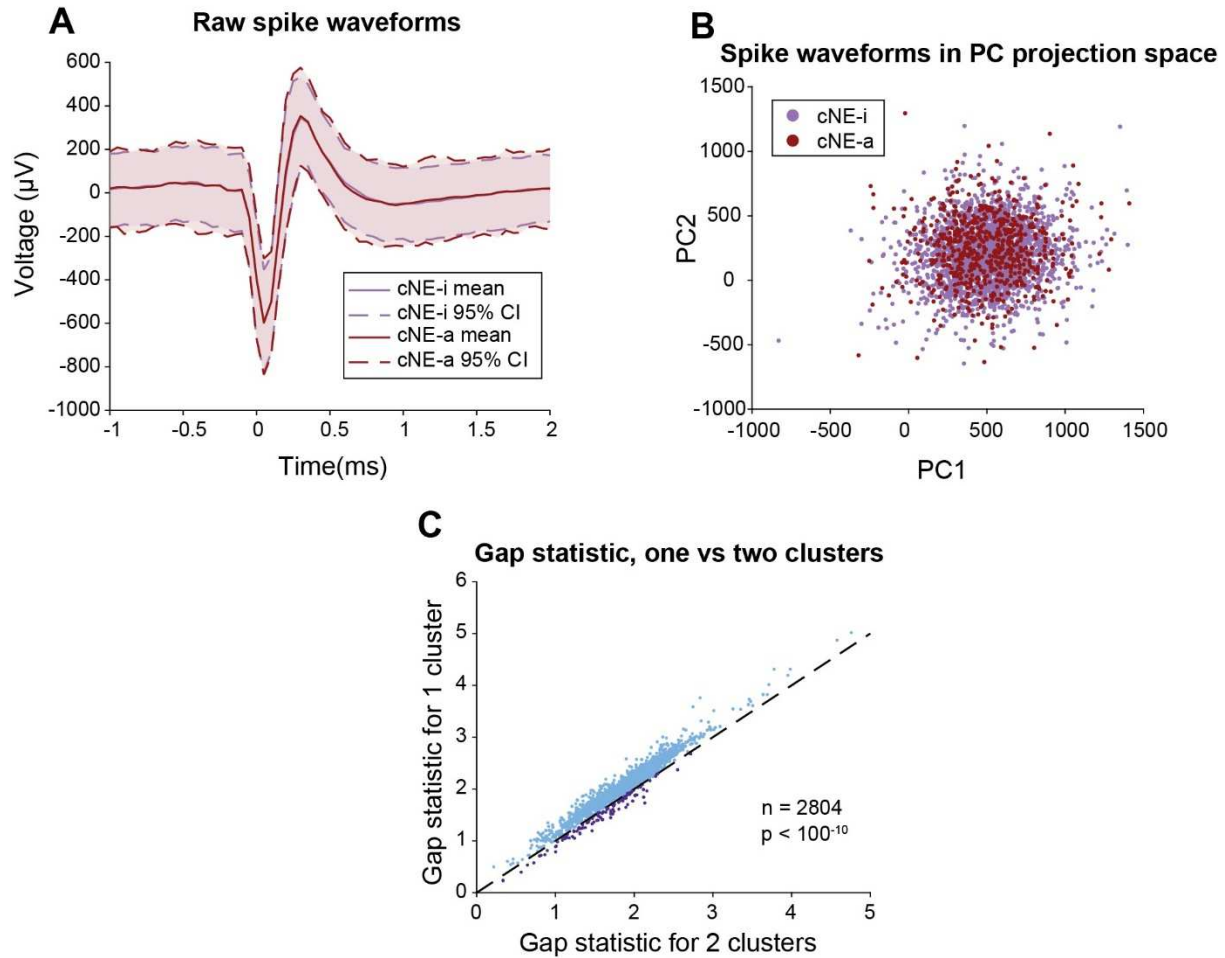
Figure 3.S4. cNE-i and cNE-a spike waveforms cannot be distinguished from each other.
(A) Raw spike waveforms of cNE-i and cNE-a spikes from a sample neuron. **(B)** Spike
waveforms of neuron in **(A)** in two-dimensional PC projection space. cNE-i spikes are labeled in
purple and cNE-a spikes in red. **(C)** Gap statistic (Tibshirani et al., 2001) for clustering, which
determines the ideal number of clusters from the data points in two-dimensional PC space. The
gap statistic significantly favors a single cluster across the entire population of neurons that
belong to at least one cNE. Neurons that had a higher gap statistic for 2 clusters (~5%) were
omitted from the analyses in Figures 3.5 – 3.7.

A

cNE1-a spikes   cNE5-a spikes

Frequency (kHz)

Time before spike (ms)

Real STRF similarity

1. Combine spikes from both groups

2. Randomly sort them into two groups and calculate STRFs

3. Get STRF similarity between these two random groups

4. Iterate over 1000 times

B

- ■ Simulated STRF similarity, n = 1000
- ── Real STRF similarity
- ╌╌ Median simulated STRF similarity

Uniqueness index = 0.005

Ratio

STRF similarity

C

Ratio

n = 343

Uniqueness index

Figure 3.S5. Statistical method for quantifying the difference between STRFs generated by cNE-a subsets of spikes from the same neuron.
(A) Sample of two cNE-a subset STRFs from a neuron with significant STRF that is a member of two significant cNEs. The method for generating a null distribution of STRF similarity is described in dark green. **(B)** Histogram of simulated STRF similarity, with the median represented by the dark green dashed line. The STRF similarity between the two STRFs in **(A)** is represented by the bright green solid line. The uniqueness index is the proportion of the entire distribution (dark green bars and bright green solid line; n = 1001) that is smaller than or equal to the real STRF similarity value (bright green solid line). Hence, the minimum possible uniqueness index ≈ 0.001. **(C)** Histogram of uniqueness index values across all neurons shared between multiple significant cNEs. Almost 50% of such neurons had uniqueness index values of < 0.05.
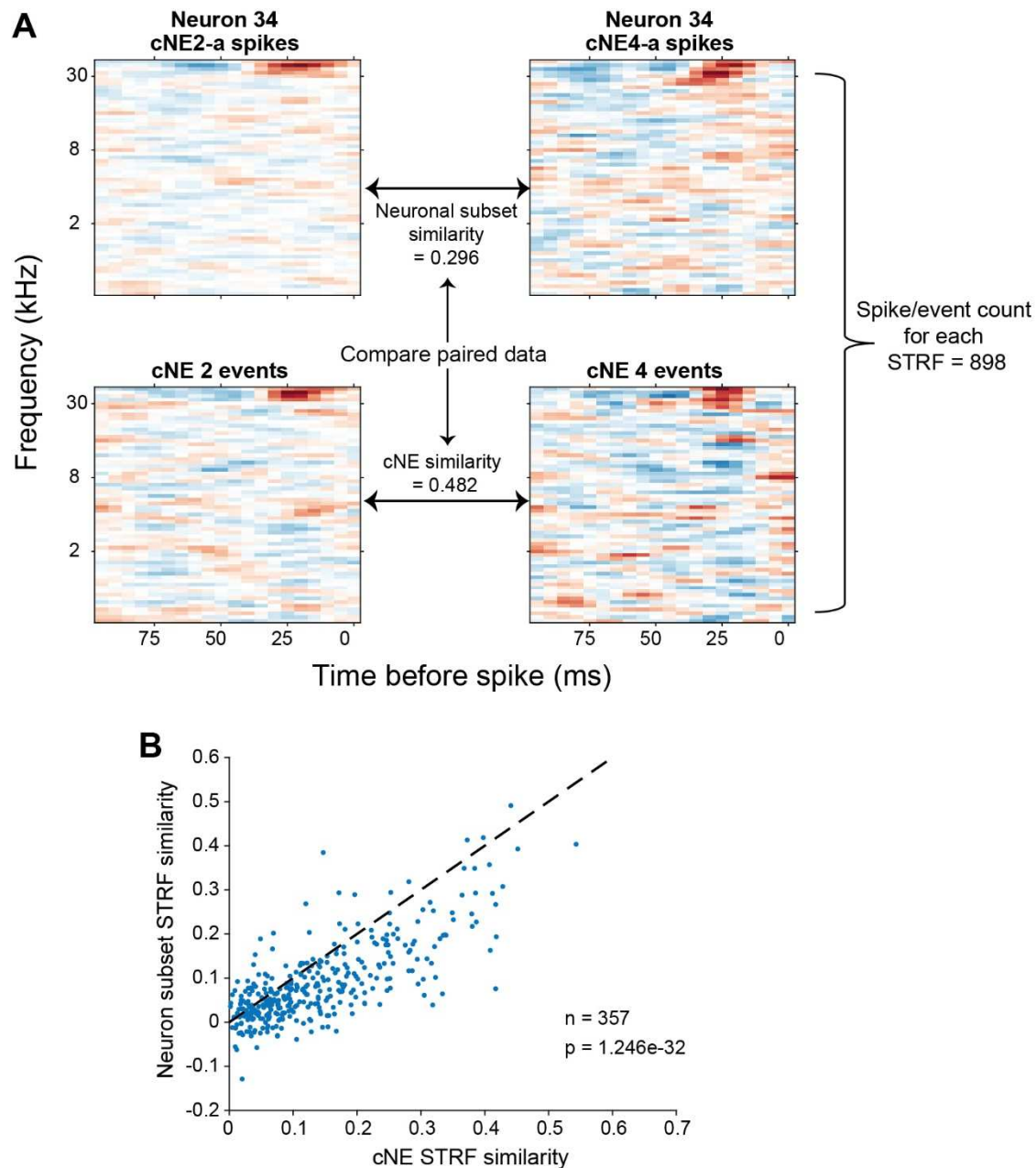
Figure 3.S6. Difference in STRFs between neuronal subsets determined by two cNEs cannot be trivially explained by the difference in STRFs between the same cNEs. (A) Illustration of the comparison between neuronal subset STRF similarity and its corresponding cNE STRF similarity. Spike/event counts for each spike/event train that generated the STRFs were normalized to have the same number of spikes/events. In this example, the STRF similarity between the neuronal subsets is lower than that between the cNEs. (B) Comparison of STRF similarity across the whole population of shared neurons. cNE STRF similarity was significantly higher than that of neuronal subset STRF similarity (≈80% below black dashed unity line). Wilcoxon signed-rank test was used in (B).

CHAPTER 4

Conclusion and future directions

4.1 Conclusion

The results presented in this dissertation support the hypothesis that cNEs should be considered the fundamental unit of information processing in the brain (for review, see Buzsáki, 2010; Harris and Mrsic-Flogel, 2013; Yuste, 2015), instead of single neurons. Specifically, I showed that cNEs recorded in AI encoded auditory information more reliably than single neurons, and more stably represented different types of information, auditory or not, in AI.

In chapter 2, I established that cNEs in AI are meaningful biological constructs by rigorously examining alternative hypotheses for their generation (e.g. stimulus synchronization, second-order correlations). I also showed that they were more reliable in encoding repeated stimuli, and had STRFs that provided more information about their activity than single neurons or random groups of simultaneously recorded neurons of the same size. Importantly, I verified that statistically similar cNEs were present and active in both spontaneous and evoked activity, strongly suggesting that cNEs are "biological infrastructure" that are present regardless of the source of neuronal information, and are biologically meaningful, supporting the idea that they are important and fundamental units of information processing in AI.

In chapter 3, I classified cNEs into different functional categories based on their STRFs and showed that they processed different kinds of information, and had statistically different physical and functional properties. Furthermore, I showed evidence of AI neurons multiplexing both auditory and non-auditory information based on their instantaneous association with different cNEs. Meanwhile, information processing by each cNE was stable, by the type of information that each cNE processed and potentially transmitted. This further reinforced the idea

that cNEs, with their stable representation of information, should be considered the elementary

units of information processing in AI, instead of single neurons.

4.2 Future directions

Other than providing strong evidence for coordinated groups of neurons as the fundamental units of information processing in AI, this dissertation has also provided a strong foundation and a useful toolbox for future research in the analysis and investigation of cNEs or cNE-like constructs in AI and other brain areas. In this final section, I will summarize some interesting future directions that the results from my work will support.

cNEs in awake, behaving animals

In chapter 3, I discussed the existence of cNEs that do not encode spectro-temporal features of the auditory stimulus (spectrotemporally unspecific cNEs), and speculated that these cNEs might encode for auditory behaviors like choice in an auditory discrimination task (Francis et al., 2018; Rodgers and DeWeese, 2014). Hence, a very important next step is to verify the existence of these "unspecific" cNEs in awake animals, and show that they are relevant to an animal's behavior. Given previous studies on the rapid changes of AI neuronal receptive fields during an auditory detection task (Fritz et al., 2003, 2005b; Yin et al., 2014) and the decreased noise correlations between neurons associated with an increase in arousal in animals doing a behavioral task (Ni et al., 2018; Stringer et al., 2016), I postulate that my observation that neurons can multiplex information based on instantaneous association with different cNEs ties these ideas together. When an animal performs an auditory behavioral task, top-down modulation biases neurons to fire in smaller but more reliable cNEs that are more optimized for performing the task, and changes each neuron's receptive field based on the biased association with the optimized cNEs. These ideas are more thoroughly discussed in chapter 3.4. In summary, one of the top priorities for future work is to record from animals trained to perform an auditory

detection or discrimination task, and to detect cNEs when an animal is passively listening or engaged in the task, and test the ideas discussed above.

cNEs in auditory areas upstream of AI

Another important idea discussed in this dissertation is that coordinated activity is vital for information transmission downstream, because of the probabilistic nature of synaptic transmission (Matsumura et al., 1996; Stevens and Zador, 1998). This idea is also related to a part of Hebb's cell assembly hypothesis, that activation of a few members of a cell assembly would be sufficient to cause the activation of the whole cell assembly reliably (Hebb, 1949). More recently, others have also proposed that cell assemblies or synchronous cNE-like constructs are only meaningful from the perspective of their downstream neurons that integrate the information from upstream neurons (Buzsáki, 2010). Hence, another important future direction of this work is to record and isolate cNEs from two highly interconnected areas, and show that cNEs in the upstream area has a much higher probability of activating single neurons or cNEs in the downstream area. In the lemniscal auditory pathway, it is well-established the ventral area of the medial geniculate body (MGBv) has strong functional inputs into AI (Lee and Winer, 2011). Therefore, investigating the functional connectivity between cNEs in MGBv and cNEs in AI is also an important next step, and this could also be integrated with auditory behavior to test the difference in cNE dynamics between those two areas in engaged vs passively listening animals.

Longitudinal tracking of cNEs during changes in neuronal connectivity

While this dissertation has shown that cNEs are relatively stable over short periods of recording time of up to 2.5 hours (Figures 2.S8, 2.10A and B), it is highly unlikely that cNEs are

stable over longer periods of time, especially when plasticity and changes in neuronal connectivity are actively occurring. One such paradigm, in which there are large changes in neuronal connectivity in AI, is during auditory behavior training. Previous studies have shown that auditory perceptual learning could lead to changes in cortical stimulus representations (Polley et al., 2006; Schnupp et al., 2006). Meanwhile, long-term exposure to different types of sound during developmental "critical periods" of animals have also been shown to influence the receptive field properties of auditory neurons in those animals (Insanally et al., 2009; Pysanenko et al., 2018). On a shorter timescale, activation of neuromodulatory systems like the cholinergic system, when paired with exposure to stimuli, has also been shown to cause reorganization of synaptic tuning curves in AI (Froemke et al., 2007). In any of these paradigms, longitudinally isolating cNEs and recording their activity over time would provide a macroscopic view of long-term plasticity in AI or other auditory areas, and help us understand, on a functional level, how synaptic reorganization might occur. With the advent of flexible electrodes that allow for long-term stable and chronic monitoring of neurons (Chung et al., 2019; Fu et al., 2016), another interesting future direction would be to densely record AI neurons over time and monitor the evolution of cNE membership and activity as synaptic reorganization occurs.

Generalizability of the cNE framework

While I have discussed the more immediate future directions of my work above as related to the auditory system, the cNE framework could easily be generalized to other systems. Indeed, studies have shown the importance of population activity in the hippocampus (Harris et al., 2003), the cerebellum (Herzfeld et al., 2015), the motor cortex (Gulati et al., 2014), the amygdala (Gründemann et al., 2019), the prefrontal cortex (Kiani et al., 2014), and other sensory areas (Choi et al., 2011; Miller et al., 2014; Nicolelis et al., 1995), just to name a few. These studies

have generally described the improved encoding of behaviors in the activity of indiscriminately pooled populations of neurons than the activity of single neurons. With the cNE validation framework and the idea of neuronal multiplexing that I have developed on top of the cNE detection algorithm (Lopes-dos-Santos et al., 2013), neuroscientists could investigate sub-populations of synchronous neurons that potentially more saliently encode behaviors or the interaction of organisms with their environments. It is my hope that my dissertation work will provide a foundation for systems neuroscientists to consider the basic unit of information processing in the brain as coordinated groups of neurons.

References

Abrams, D.A., Nicol, T., White-Schwoch, T., Zecker, S., and Kraus, N. (2017). Population responses in primary auditory cortex simultaneously represent the temporal envelope and periodicity features in natural speech. Hear. Res. *348*, 31–43.

Aertsen, A.M.H.J., and Johannesma, P.I.M. (1981). The Spectro-Temporal Receptive Field. Biol. Cybern. *42*, 133–143.

Ainsworth, M., Lee, S., Cunningham, M.O., Traub, R.D., Kopell, N.J., and Whittington, M.A. (2012). Rates and Rhythms: A Synergistic View of Frequency and Temporal Coding in Neuronal Networks. Neuron *75*, 572–583.

Almeida-Filho, D.G., Lopes-dos-Santos, V., Vasconcelos, N.A.P., Miranda, J.G. V, Tort, A.B.L., and Ribeiro, S. (2014). An investigation of Hebbian phase sequences as assembly graphs. Front. Neural Circuits *8*, 34.

Atencio, C.A., and Schreiner, C.E. (2010a). Columnar connectivity and laminar processing in cat primary auditory cortex. PLoS One *5*, e9521.

Atencio, C.A., and Schreiner, C.E. (2010b). Laminar diversity of dynamic sound processing in cat primary auditory cortex. J. Neurophysiol. *103*, 192–205.

Atencio, C.A., and Schreiner, C.E. (2012). Spectrotemporal processing in spectral tuning modules of cat primary auditory cortex. PLoS One *7*, e31537.

Atencio, C.A., and Schreiner, C.E. (2013). Auditory Cortical Local Subnetworks Are Characterized by Sharply Synchronous Activity. J. Neurosci. *33*, 18503–18514.

Atencio, C.A., and Schreiner, C.E. (2016). Functional congruity in local auditory cortical microcircuits. Neuroscience *316*, 402–419.

Atencio, C.A., Sharpee, T.O., and Schreiner, C.E. (2008). Cooperative Nonlinearities in Auditory Cortical Neurons. Neuron *58*, 956–966.

Atencio, C.A., Sharpee, T.O., and Schreiner, C.E. (2012). Receptive field dimensionality increases from the auditory midbrain to cortex. J Neurophysiol *107*, 2594–2603.

Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nat. Rev. Neurosci. *7*, 358–366.

Balaguer-Ballester, E., Lapish, C.C., Seamans, J.K., and Durstewitz, D. (2011). Attracting Dynamics of Frontal Cortex Ensembles during Memory-Guided Decision-Making. PLoS Comput. Biol. *7*, e1002057.

Bathellier, B., Ushakova, L., and Rumpel, S. (2012). Discrete Neocortical Dynamics Predict Behavioral Categorization of Sounds. Neuron *76*, 435–449.

Belitski, A., Gretton, A., Magri, C., Murayama, Y., Montemurro, M.A., Logothetis, N.K., and Panzeri, S. (2008). Low-Frequency Local Field Potentials and Spikes in Primary Visual Cortex Convey Independent Visual Information. J. Neurosci. *28*, 5696–5709.

Bell, A.H., Summerfield, C., Morin, E.L., Malecek, N.J., and Ungerleider, L.G. (2016). Encoding of Stimulus Probability in Macaque Inferior Temporal Cortex. Curr. Biol. *26*, 2280–2290.

Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. Nature *436*, 1161–1165.

Berens, P., Ecker, A.S., Cotton, R.J., Ma, W.J., Bethge, M., and Tolias, A.S. (2012). A fast and simple population code for orientation in primate V1. J. Neurosci. *32*, 10618–10626.

Bermudez Contreras, E.J., Schjetnan, A.G.P., Muhammad, A., Bartho, P., McNaughton, B.L., Kolb, B., Gruber, A.J., and Luczak, A. (2013). Formation and Reverberation of Sequential Neural Activity Patterns Evoked by Sensory Stimulation Are Enhanced during Cortical Desynchronization. Neuron *79*, 555–566.

Bharmauria, V., Bachatene, L., Cattan, S., Brodeur, S., Chanauria, N., Rouat, J., and Molotchnikoff, S. (2016). Network-selectivity and stimulus-discrimination in the primary visual cortex: Cell-assembly dynamics. Eur. J. Neurosci. *43*, 204–219.

Bi, G.Q., and Poo, M.M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J. Neurosci. *18*, 10464–10472.

Billeh, Y.N., Schaub, M.T., Anastassiou, C.A., Barahona, M., and Koch, C. (2014). Revealing cell assemblies at multiple levels of granularity. J. Neurosci. Methods *236*, 92–106.

Bizley, J.K., Walker, K.M.M., King, A.J., and Schnupp, J.W.H. (2010). Neural ensemble codes for stimulus periodicity in auditory cortex. J. Neurosci. *30*, 5078–5091.

Blanche, T.J., Spacek, M.A., Hetke, J.F., and Swindale, N. V. (2005). Polytrodes: High-Density Silicon Electrode Arrays for Large-Scale Multiunit Recording. J. Neurophysiol. *93*, 2987–3000.

Braitenberg, V., and Schüz, A. (1991). Anatomy of the Cortex (Berlin, Heidelberg: Springer Berlin Heidelberg).

Branco, T., and Häusser, M. (2011). Synaptic Integration Gradients in Single Cortical Pyramidal Cell Dendrites. Neuron *69*, 885–892.

Brasselet, R., Panzeri, S., Logothetis, N.K., and Kayser, C. (2012). Neurons with stereotyped and rapid responses provide a reference frame for relative temporal coding in primate auditory cortex. J. Neurosci. *32*, 2998–3008.

Brenner, N., Strong, S.P., Koberle, R., Bialek, W., and de Ruyter van Steveninck, R.R. (2000). Synergy in a neural code. Neural Comput. *12*, 1531–1552.

Britten, K., Newsome, W., Shadlen, M.N., Celebrini, S., and Movshon, J. (1996). A relationship between behavoiral choice and the visual responses of neurons in macaque MT. Vis. Neurosci. *13*, 87–100.

Brosch, M., and Schreiner, C.E. (1999). Correlations between neural discharges are related to receptive field properties in cat primary auditory cortex. Eur. J. Neurosci. *11*, 3517–3530.

Brosch, M., Selezneva, E., and Scheich, H. (2015). Neuronal activity in primate auditory cortex during the performance of audiovisual tasks. Eur. J. Neurosci. *41*, 603–614.

Buzsáki, G. (2010). Neural Syntax: Cell Assemblies, Synapsembles, and Readers. Neuron *68*, 362–385.

Calabrese, A., Schumacher, J.W., Schneider, D.M., Paninski, L., and Woolley, S.M.N. (2011). A Generalized Linear Model for Estimating Spectrotemporal Receptive Fields from Responses to Natural Sounds. PLoS One *6*, e16104.

Carrillo-Reid, L., Hernández-López, S., Tapia, D., Galarraga, E., and Bargas, J. (2011). Dopaminergic modulation of the striatal microcircuit: receptor-specific configuration of cell assemblies. J. Neurosci. *31*, 14972–14983.

Carrillo-Reid, L., Lopez-Huerta, V.G., Garcia-Munoz, M., Theiss, S., and Arbuthnott, G.W. (2015). Cell Assembly Signatures Defined by Short-Term Synaptic Plasticity in Cortical Networks. Int. J. Neural Syst. *25*, 1550026.

Celikel, T., Szostak, V.A., and Feldman, D.E. (2004). Modulation of spike timing by sensory deprivation during induction of cortical map plasticity. Nat. Neurosci. *7*, 534–541.

Chakraborty, S., Sandberg, A., and Greenfield, S.A. (2007). Differential dynamics of transient neuronal assemblies in visual compared to auditory cortex. Exp. Brain Res. *182*, 491–498.

Chambers, A.R., and Rumpel, S. (2017). A stable brain from unstable components: Emerging concepts and implications for neural computation. Neuroscience *357*, 172–184.

Chelaru, M.I., and Dragoi, V. (2008). Efficient coding in heterogeneous neuronal populations. Proc. Natl. Acad. Sci. *105*, 16344–16349.

Choi, G.B., Stettler, D.D., Kallman, B.R., Bhaskar, S.T., Fleischmann, A., and Axel, R. (2011). Driving Opposing Behaviors with Ensembles of Piriform Neurons. Cell *146*, 1004–1015.

Chung, J.E., Magland, J.F., Barnett, A.H., Tolosa, V.M., Tooker, A.C., Lee, K.Y., Shah, K.G., Felix, S.H., Frank, L.M., and Greengard, L.F. (2017). A Fully Automated Approach to Spike Sorting. Neuron *95*, 1381-1394.e6.

Chung, J.E., Joo, H.R., Fan, J.L., Liu, D.F., Barnett, A.H., Chen, S., Geaghan-Breiner, C., Karlsson, M.P., Karlsson, M., Lee, K.Y., et al. (2019). High-Density, Long-Lasting, and Multi-region Electrophysiological Recordings Using Polymer Electrode Arrays. Neuron.

Clement, E.A., Richard, A., Thwaites, M., Ailon, J., Peters, S., and Dickson, C.T. (2008). Cyclic and Sleep-Like Spontaneous Alternations of Brain State Under Urethane Anaesthesia. PLoS One *3*, e2004.

Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. Nat. Neurosci. *14*, 811–819.

Cohen, M.R., and Maunsell, J.H.R. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. *12*, 1594–1600.

David, S.V., and Shamma, S.A. (2013). Integration over Multiple Timescales in Primary Auditory Cortex. J. Neurosci. *33*, 19154–19166.

DeNardo, L.A., Liu, C.D., Allen, W.E., Adams, E.L., Friedmann, D., Fu, L., Guenthner, C.J., Tessier-Lavigne, M., and Luo, L. (2019). Temporal evolution of cortical ensembles promoting remote memory retrieval. Nat. Neurosci. *22*, 460–469.

Depireux, D. a, Simon, J.Z., Klein, D.J., and Shamma, S. a (2001). Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. J. Neurophysiol. *85*, 1220–1234.

Dragoi, G., and Tonegawa, S. (2013). Distinct preplay of multiple novel spatial experiences in the rat. Proc. Natl. Acad. Sci. *110*, 9100–9105.

Du, J., Blanche, T.J., Harrison, R.R., Lester, H.A., and Masmanidis, S.C. (2011). Multiplexed, High Density Electrophysiology with Nanofabricated Neural Probes. PLoS One *6*, e26204.

Eagleman, S.L., and Dragoi, V. (2012). Image sequence reactivation in awake V4 networks. Proc. Natl. Acad. Sci. *109*, 19450–19455.

Eggermont, J.J. (1992). Neural interaction in cat primary auditory cortex. Dependence on recording depth, electrode separation, and age. J. Neurophysiol. *68*, 1216–1228.

Eggermont, J.J. (2006). Properties of correlated neural activity clusters in cat auditory cortex resemble those of neural assemblies. J. Neurophysiol. *96*, 746–764.

Eggermont, J.J., Munguia, R., and Shaw, G. (2013). Cross-correlations between three units in cat primary auditory cortex. Hear. Res. *304*, 179–187.

Eldawlatly, S., and Oweiss, K.G. (2011). Millisecond-Timescale Local Network Coding in the Rat Primary Somatosensory Cortex. PLoS One *6*, e21649.

Engineer, C.T., Perez, C.A., Chen, Y.H., Carraway, R.S., Reed, A.C., Shetake, J.A., Jakkamsetti, V., Chang, K.Q., and Kilgard, M.P. (2008). Cortical activity patterns predict speech discrimination ability. Nat. Neurosci. *11*, 603–608.

Escabi, M.A., and Schreiner, C.E. (2002). Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. J. Neurosci. *22*, 4114–4131.

Escabí, M.A., Read, H.L., Viventi, J., Kim, D.-H., Higgins, N.C., Storace, D.A., Liu, A.S.K., Gifford, A.M., Burke, J.F., Campisi, M., et al. (2014). A high-density, high-channel count, multiplexed µECoG array for auditory-cortex recordings. J. Neurophysiol. *112*, 1566–1583.

Francis, N.A., Winkowski, D.E., Sheikhattar, A., Armengol, K., Babadi, B., and Kanold, P.O. (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. Neuron *97*, 885-897.e6.

Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat. Neurosci. *6*, 1216–1223.

Fritz, J., Elhilali, M., and Shamma, S. (2005a). Active listening: Task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. Hear. Res. *206*, 159–176.

Fritz, J.B., Elhilali, M., and Shamma, S.A. (2005b). Differential Dynamic Plasticity of A1 Receptive Fields during Multiple Spectral Tasks. J. Neurosci. *25*, 7623–7635.

Froemke, R.C., Merzenich, M.M., and Schreiner, C.E. (2007). A synaptic memory trace for cortical receptive field plasticity. Nature *450*, 425–429.

Fu, T.-M., Hong, G., Zhou, T., Schuhmann, T.G., Viveros, R.D., and Lieber, C.M. (2016). Stable long-term chronic brain mapping at the single-neuron level. Nat. Methods *13*, 875–882.

Furukawa, S., Xu, L., and Middlebrooks, J.C. (2000). Coding of sound-source location by ensembles of cortical neurons. J Neurosci *20*, 1216–1228.

Georgopoulos, A., Schwartz, A., and Kettner, R. (1986). Neuronal population coding of movement direction. Science (80-.). *233*, 1416–1419.

Goris, R.L.T., Movshon, J.A., and Simoncelli, E.P. (2014). Partitioning neuronal variability. Nat. Neurosci. *17*, 858–865.

Gourévitch, B., and Eggermont, J.J. (2010). Maximum decoding abilities of temporal patterns and synchronized firings: application to auditory neurons responding to click trains and amplitude modulated white noise. J. Comput. Neurosci. *29*, 253–277.

Granot-Atedgi, E., Tkačik, G., Segev, R., and Schneidman, E. (2013). Stimulus-dependent Maximum Entropy Models of Neural Population Codes. PLoS Comput. Biol. *9*, e1002922.

Gründemann, J., Bitterman, Y., Lu, T., Krabbe, S., Grewe, B.F., Schnitzer, M.J., and Lüthi, A. (2019).

Amygdala ensembles encode behavioral states. Science (80-.). *364*, eaav8736.

Gulati, T., Ramanathan, D.S., Wong, C.C., and Ganguly, K. (2014). Reactivation of emergent task-related ensembles during slow-wave sleep after neuroprosthetic learning. Nat. Neurosci. *17*, 1107–1113.

Guo, W., Chambers, A.R., Darrow, K.N., Hancock, K.E., Shinn-Cunningham, B.G., and Polley, D.B. (2012). Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. J. Neurosci. *32*, 9159–9172.

Gururangan, S.S., Sadovsky, A.J., and MacLean, J.N. (2014). Analysis of Graph Invariants in Functional Neocortical Circuitry Reveals Generalized Features Common to Three Areas of Sensory Cortex. PLoS Comput. Biol. *10*.

Hamel, E.J.O., Grewe, B.F., Parker, J.G., and Schnitzer, M.J. (2015). Cellular Level Brain Imaging in Behaving Mammals: An Engineering Approach. Neuron *86*, 140–159.

Han, F., Caporale, N., and Dan, Y. (2008). Reverberation of Recent Visual Experience in Spontaneous Cortical Waves. Neuron *60*, 321–327.

Harper, N.S., Schoppe, O., Willmore, B.D.B., Cui, Z., Schnupp, J.W.H., and King, A.J. (2016). Network Receptive Field Modeling Reveals Extensive Integration and Multi-feature Selectivity in Auditory Cortical Neurons. PLOS Comput. Biol. *12*, e1005113.

Harris, K.D. (2005). Neural signatures of cell assembly organization. Nat. Rev. Neurosci. *6*, 399–407.

Harris, K.D., and Mrsic-Flogel, T.D. (2013). Cortical connectivity and sensory coding. Nature *503*, 51–58.

Harris, K.D., and Thiele, A. (2011). Cortical state and attention. Nat. Rev. Neurosci. *12*, 509–523.

Harris, K., Csicsvari, J., Hirase, H., Dragoi, G., and Buzsáki, G. (2003). Organization of cell assemblies in the hippocampus. Nature *424*, 552–556.

Harvey, M.A., Saal, H.P., Dammann, J.F., and Bensmaia, S.J. (2013). Multiplexing Stimulus Information through Rate and Temporal Codes in Primate Somatosensory Cortex. PLoS Biol. *11*.

Hebb, D.O. (1949). The Organization of Behavior.

Herzfeld, D.J., Kojima, Y., Soetedjo, R., and Shadmehr, R. (2015). Encoding of action by the Purkinje cells of the cerebellum. Nature *526*, 439–442.

Higgins, I., Stringer, S., and Schnupp, J. (2017). Unsupervised learning of temporal features for word categorization in a spiking neural network model of the auditory brain. PLoS One *12*, e0180174.

Hopfield, J.J., and Brody, C.D. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. Proc. Natl. Acad. Sci. *98*, 1282–1287.

Hromádka, T., DeWeese, M.R., and Zador, A.M. (2008). Sparse Representation of Sounds in the Unanesthetized Auditory Cortex. PLoS Biol. *6*, e16.

Hyvärinen, A., and Oja, E. (1997). A Fast Fixed-Point Algorithm for Independent Component Analysis. Neural Comput. *9*, 1483–1492.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., and Yuste, R. (2004). Synfire chains and cortical songs: temporal modules of cortical activity. Science *304*, 559–564.

Ince, R.A.A., Panzeri, S., and Kayser, C. (2013). Neural Codes Formed by Small and Temporally Precise Populations in Auditory Cortex. J. Neurosci. *33*, 18277–18287.

Insanally, M.N., Kover, H., Kim, H., and Bao, S. (2009). Feature-Dependent Sensitive Periods in the Development of Complex Sound Representation. J. Neurosci. *29*, 5456–5462.

Jacobs, N.S., Chen-Bee, C.H., and Frostig, R.D. (2015). Emergence of spatiotemporal invariance in large neuronal ensembles in rat barrel cortex. Front. Neural Circuits *9*, 1–16.

Jermakowicz, W.J., Chen, X., Khaytin, I., Bonds, A.B., and Casagrande, V.A. (2008). Relationship

Between Spontaneous and Evoked Spike-Time Correlations in Primate Visual Cortex. J. Neurophysiol. *101*, 2279–2289.

Jun, J.J., Steinmetz, N.A., Siegle, J.H., Denman, D.J., Bauza, M., Barbarits, B., Lee, A.K., Anastassiou, C.A., Andrei, A., Aydın, Ç., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature *551*, 232–236.

Kanitscheider, I., Coen-Cagli, R., and Pouget, A. (2015). Origin of information-limiting noise correlations. Proc. Natl. Acad. Sci. *112*, E6973–E6982.

Kaur, S., Rose, H.J., Lazar, R., Liang, K., and Metherate, R. (2005). Spectral integration in primary auditory cortex: laminar processing of afferent input, in vivo and in vitro. Neuroscience *134*, 1033–1045.

Kayser, C., Montemurro, M.A., Logothetis, N.K., and Panzeri, S. (2009). Spike-Phase Coding Boosts and Stabilizes Information Carried by Spatial and Temporal Spike Patterns. Neuron *61*, 597–608.

Kayser, C., Logothetis, N.K., and Panzeri, S. (2010). Millisecond encoding precision of auditory cortex neurons. Proc. Natl. Acad. Sci. *107*, 16976–16981.

Kiani, R., Cueva, C.J., Reppas, J.B., and Newsome, W.T. (2014). Dynamics of Neural Population Responses in Prefrontal Cortex Indicate Changes of Mind on Single Trials. Curr. Biol. *24*, 1542–1547.

Kohn, A., and Smith, M.A. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J. Neurosci. *25*, 3661–3673.

Kohn, A., Coen-cagli, R., Kanitscheider, I., and Pouget, A. (2016). Correlations and Neuronal Population Information. Annu. Rev. Neurosci. *39*, 237–256.

Krause, B.M., Raz, A., Uhlrich, D.J., Smith, P.H., and Banks, M.I. (2014). Spiking in auditory cortex following thalamic stimulation is dominated by cortical network activity. Front. Syst. Neurosci. *8*, 1–24.

Kubota, M., Sugimoto, S., Horikawa, J., Nasu, M., and Taniguchi, I. (1997). Optical imaging of dynamic

horizontal spread of excitation in rat auditory cortex slices. Neurosci. Lett. *237*, 77–80.

Kudrimoti, H.S., Barnes, C.A., and McNaughton, B.L. (1999). Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. J. Neurosci. *19*, 4090–4101.

Laje, R., and Buonomano, D. V. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat. Neurosci. *16*, 925–933.

Lankarany, M., Al-Basha, D., Ratté, S., and Prescott, S.A. (2019). Differentially synchronized spiking enables multiplexed neural coding. Proc. Natl. Acad. Sci. *116*, 201812171.

Laubach, M., Wessberg, J., and Nicolelis, M.A.L. (2000). Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. Nature *405*, 567–571.

Lee, C.C., and Winer, J. a. (2011). Convergence of thalamic and cortical pathways in cat auditory cortex. Hear. Res. *274*, 85–94.

Lewicki, M.S. (1994). Bayesian Modeling and Classification of Neural Signals. Neural Comput. *6*, 1005–1030.

Lopes-dos-Santos, V., Ribeiro, S., and Tort, A.B.L. (2013). Detecting cell assemblies in large neuronal populations. J. Neurosci. Methods *220*, 149–166.

Luczak, A., and MacLean, J.N. (2012). Default activity patterns at the neocortical microcircuit level. Front. Integr. Neurosci. *6*, 1–6.

Luczak, A., Barthó, P., and Harris, K.D. (2009). Spontaneous Events Outline the Realm of Possible Sensory Responses in Neocortical Populations. Neuron *62*, 413–425.

Luczak, A., Mcnaughton, B.L., and Harris, K.D. (2015). Packet-based communication in the cortex. Nat. Publ. Gr. *16*, 1–12.

Macke, J.H., Berens, P., Ecker, A.S., Tolias, A.S., and Bethge, M. (2009). Generating Spike Trains with

Specified Correlation Coefficients. Neural Comput. *21*, 397–423.

Macke, J.H., Opper, M., and Bethge, M. (2011). Common Input Explains Higher-Order Correlations and Entropy in a Simple Model of Neural Population Activity. Phys. Rev. Lett. *106*, 208102.

Malvache, A., Reichinnek, S., Villette, V., Haimerl, C., and Cossart, R. (2016). Awake hippocampal reactivations project onto orthogonal neuronal assemblies. Science *353*, 1280–1283.

Marčenko, V.A., and Pastur, L.A. (1967). Distribution of Eigenvalues for Some Sets of Random Matrices. Math. USSR-Sbornik *1*, 457–483.

Marguet, S.L., and Harris, K.D. (2011). State-dependent representation of amplitude-modulated noise stimuli in rat auditory cortex. J. Neurosci. *31*, 6414–6420.

Markram, H. (1997). Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs. Science (80-.). *275*, 213–215.

Marsat, G., and Maler, L. (2010). Neural Heterogeneity and Efficient Population Codes for Communication Signals. J. Neurophysiol. *104*, 2543–2555.

Matsumura, M., Chen, D., Sawaguchi, T., Kubota, K., and Fetz, E.E. (1996). Synaptic interactions between primate precentral cortex neurons revealed by spike-triggered averaging of intracellular membrane potentials in vivo. J. Neurosci. *16*, 7757–7767.

McGinley, M.J., David, S.V., and McCormick, D.A. (2015). Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron *87*, 179–192.

Miller, L.M., and Recanzone, G.H. (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. Proc. Natl. Acad. Sci. U. S. A. *106*, 5931–5935.

Miller, J.-E.K., Ayzenshtat, I., Carrillo-Reid, L., and Yuste, R. (2014). Visual stimuli recruit intrinsically generated cortical ensembles. Proc. Natl. Acad. Sci. *111*, E4053–E4061.

Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2009). Spatial Attention Decorrelates Intrinsic Activity Fluctuations in Macaque Area V4. Neuron *63*, 879–888.

Montani, F., Ince, R.A.A., Senatore, R., Arabzadeh, E., Diamond, M.E., and Panzeri, S. (2009). The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. *367*, 3297–3310.

Morrill, R.J., and Hasenstaub, A.R. (2018). Visual Information Present in Infragranular Layers of Mouse Auditory Cortex. J. Neurosci. *38*, 2854–2862.

Ni, A.M., Ruff, D.A., Alberts, J.J., Symmonds, J., and Cohen, M.R. (2018). Learning and attention reveal a general relationship between neuronal variability and perception. Science (80-.). *465*, 1–28.

Nicolelis, M., Baccala, L., Lin, R., and Chapin, J. (1995). Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system. Science (80-.). *268*, 1353–1358.

Ohiorhenuan, I.E., Mechler, F., Purpura, K.P., Schmid, A.M., Hu, Q., and Victor, J.D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. Nature *466*, 617–621.

Okun, M., Naim, A., and Lampl, I. (2010). The Subthreshold Relation between Cortical Local Field Potential and Neuronal Firing Unveiled by Intracellular Recordings in Awake Rats. J. Neurosci. *30*, 4440–4448.

Osborne, L.C., Palmer, S.E., Lisberger, S.G., and Bialek, W. (2008). The Neural Basis for Combinatorial Coding in a Cortical Population Response. J. Neurosci. *28*, 13522–13531.

Pagliardini, S., Gosgnach, S., and Dickson, C.T. (2013). Spontaneous Sleep-Like Brain State Alternations and Breathing Characteristics in Urethane Anesthetized Mice. PLoS One *8*, 1–10.

Pakkenberg, B., and Gundersen, H.J. (1997). Neocortical neuron number in humans: effect of sex and age. J. Comp. Neurol. *384*, 312–320.

Paninski, L., Shoham, S., Fellows, M.R., Hatsopoulos, N.G., and Donoghue, J.P. (2004). Superlinear Population Encoding of Dynamic Hand Trajectory in Primary Motor Cortex. J. Neurosci. *24*, 8551–8561.

Panzeri, S., Petroni, F., Petersen, R.S., and Diamond, M.E. (2003). Decoding Neuronal Population Activity in Rat Somatosensory Cortex: Role of Columnar Organization. Cereb. Cortex *13*, 45–52.

Perel, S., Sadtler, P.T., Oby, E.R., Ryu, S.I., Tyler-Kabara, E.C., Batista, A.P., and Chase, S.M. (2015). Single-unit activity, threshold crossings, and local field potentials in motor cortex differentially encode reach kinematics. J. Neurophysiol. *114*, 1500–1512.

Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S.I., and Battaglia, F.P. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. Nat. Neurosci. *12*, 919–926.

Peyrache, A., Benchenane, K., Khamassi, M., Wiener, S.I., and Battaglia, F.P. (2010). Principal component analysis of ensemble recordings reveals cell assemblies at high temporal resolution. J. Comput. Neurosci. *29*, 309–325.

Pipa, G., Wheeler, D.W., Singer, W., and Nikolić, D. (2008). NeuroXidence: Reliable and efficient analysis of an excess or deficiency of joint-spike events. J. Comput. Neurosci. *25*, 64–88.

Polley, D., Read, H., Storace, D., and Merzenich, M. (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. J. … *97*, 3621–3638.

Polley, D.B., Steinberg, E.E., and Merzenich, M.M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. J. Neurosci. *26*, 4970–4982.

Polsky, A., Mel, B.W., and Schiller, J. (2004). Computational subunits in thin dendrites of pyramidal cells. Nat. Neurosci. *7*, 621–627.

Poulet, J.F.A., and Petersen, C.C.H. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. Nature *454*, 881–885.

Pysanenko, K., Bureš, Z., Lindovský, J., and Syka, J. (2018). The Effect of Complex Acoustic Environment during Early Development on the Responses of Auditory Cortex Neurons in Rats. Neuroscience *371*, 221–228.

Reimann, M.W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., and Markram, H. (2017). Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function. Front. Comput. Neurosci. *11*, 48.

Rios, G., Lubenov, E. V., Chi, D., Roukes, M.L., and Siapas, A.G. (2016). Nanofabricated Neural Probes for Dense 3-D Recordings of Brain Activity. Nano Lett. *16*, 6857–6862.

Rodgers, C.C., and DeWeese, M.R. (2014). Neural Correlates of Task Switching in Prefrontal Cortex and Primary Auditory Cortex in a Novel Stimulus Selection Task for Rodents. Neuron *82*, 1157–1170.

Roland, B., Deneux, T., Franks, K.M., Bathellier, B., and Fleischmann, A. (2017). Odor identity coding by distributed ensembles of neurons in the mouse olfactory cortex. Elife *6*, 1–26.

Russo, E., and Durstewitz, D. (2017). Cell assemblies at multiple time scales with arbitrary lag constellations. Elife *6*, 1–39.

de Ruyter van Steveninck, R.R., Lewen, G.D., Strong, S.P., Koberle, R., and Bialek, W. (1997). Reproducibility and variability in neural spike trains. Science *275*, 1805–1808.

Schnupp, J.W.H., Hall, T.M., Kokelaar, R.F., and Ahmed, B. (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. J. Neurosci. *26*, 4785–4795.

Schölvinck, M.L., Saleem, A.B., Benucci, A., Harris, K.D., and Carandini, M. (2015). Cortical State Determines Global Variability and Correlations in Visual Cortex. J. Neurosci. *35*, 170–178.

See, J.Z., Atencio, C.A., Sohal, V.S., and Schreiner, C.E. (2018). Coordinated neuronal ensembles in primary auditory cortical columns. Elife *7*, 1–33.

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

Shih, J.Y., Atencio, C.A., and Schreiner, C.E. (2011). Improved stimulus representation by short interspike intervals in primary auditory cortex. J. Neurophysiol. *105*, 1908–1917.

Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. PLoS Biol. *3*, e68.

Stevens, C.F., and Zador, A.M. (1998). Input synchrony and the irregular firing of cortical neurons. Nat. Neurosci. *1*, 210–217.

Stevenson, I.H., and Kording, K.P. (2011). How advances in neural recording affect data analysis. Nat. Neurosci. *14*, 139–142.

Stringer, C., Pachitariu, M., Steinmetz, N.A., Okun, M., Bartho, P., Harris, K.D., Sahani, M., and Lesica, N.A. (2016). Inhibitory control of correlated intrinsic variability in cortical networks. Elife *5*, 1–33.

Strong, S.P., Köberle, R., de Ruyter van Steveninck, R.R., and Bialek, W. (1998). Entropy and Information in Neural Spike Trains. Phys. Rev. Lett. *80*, 197–200.

Szymanski, F.D., Garcia-Lazaro, J.A., and Schnupp, J.W.H. (2009). Current source density profiles of stimulus-specific adaptation in rat auditory cortex. J. Neurophysiol. *102*, 1483.

Teufel, C., and Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. Conscious. Cogn. *47*, 17–25.

Thorson, I.L., Liénard, J., and David, S. V. (2015). The Essential Complexity of Auditory Receptive Fields. PLoS Comput. Biol. *11*, 1–33.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. Ser. B (Statistical Methodol. *63*, 411–423.

Tingley, D., Alexander, A.S., Quinn, L.K., Chiba, A.A., and Nitz, D.A. (2015). Cell Assemblies of the Basal Forebrain. J. Neurosci. *35*, 2992–3000.

Tremblay, S., Pieper, F., Sachs, A., and Martinez-Trujillo, J. (2015). Attentional Filtering of Visual Information by Neuronal Ensembles in the Primate Lateral Prefrontal Cortex. Neuron *85*, 202–215.

Walker, K.M.M., Bizley, J.K., King, A.J., and Schnupp, J.W.H. (2011). Multiplexed and Robust Representations of Sound Features in Auditory Cortex. J. Neurosci. *31*, 14565–14576.

Wallace, M.N., and Palmer, A.R. (2007). Laminar differences in the response properties of cells in the primary auditory cortex. Exp. Brain Res. *184*, 179–191.

Wessberg, J., Stambaugh, C.R., Kralik, J.D., Beck, P.D., Laubach, M., Chapin, J.K., Kim, J., Biggs, S.J., Srinivasan, M.A., and Nicolelis, M.A. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature *408*, 361–365.

Williams, S.R., and Stuart, G.J. (2002). Dependence of EPSP efficacy on synapse location in neocortical pyramidal neurons. Science *295*, 1907–1910.

Winkowski, D.E., and Kanold, P.O. (2013). Laminar Transformation of Frequency Organization in Auditory Cortex. J. Neurosci. *33*, 1498–1508.

Yin, P., Fritz, J.B., and Shamma, S. a. (2014). Rapid Spectrotemporal Plasticity in Primary Auditory Cortex during Behavior. J. Neurosci. *34*, 4396–4408.

Yoshimura, Y., Dantzker, J.L.M., and Callaway, E.M. (2005). Excitatory cortical neurons form fine-scale functional networks. Nature *433*, 868–873.

Yu, S., Yang, H., Nakahara, H., Santos, G.S., Nikolic, D., and Plenz, D. (2011). Higher-Order Interactions Characterized in Cortical Activity. J. Neurosci. *31*, 17514–17526.

Yuste, R. (2015). From the neuron doctrine to neural networks. Nat. Rev. Neurosci. *16*, 487–497.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

_____   8/20/19
Author Signature   _____
                                          Date

151