UNIVERSITY OF CALIFORNIA, SAN DIEGO

Through the looking glass: population dynamics through membrane potentials.

A dissertation submitted in partial satisfaction of the
requirements for the Degree Doctor of Philosophy

in

Neurosciences

by

Krista Eva Perks

Committee in charge:

Professor Timothy Gentner, Chair
Professor Jeffry Isaacson
Professor Takaki Komiyama
Professor John Reynolds
Professor Massimo Scanziani

2016

The dissertation of Krista Eva Perks is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2016

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Below is the content of the dissertation, which is either published or submitted for publication. I was the primary investigator and author on all manuscripts in this dissertation.

Chapter 2, in full, is a reprint of materials as it appears in Perks, K., Gentner, T. (2015). "Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex" in *European Journal of Neuroscience*, v. 41, p. 725-733. DOI:10.1111/ejn.12831. The dissertation author was the primary investigator and author of this manuscript.

Chapter 3, in full, is in preparation for publication as: Perks, K. and Gentner, T. "Natural signals drive fast modulation of stimulus-specific functional networks in cortex." The dissertation author was the primary investigator and author of this manuscript.

VITAE

2005	Howard Hughes Institute Summer Fellow, Wesleyan University, CT
2005	Phi Beta Kappa, CT Gamma chapter
2005, 2006	Teaching Assistant: Introduction to Neurosciences, Wesleyan University, CT
2006	Teaching Assistant: Neurophysiology Laboratory, Wesleyan University, CT
2006	Bachelor of Arts: Neuroscience and Behavior, Wesleyan University, CT
2007	Master of Arts: Neuroscience and Behavior, Wesleyan University, CT
2010	Teaching Assistant: Cellular and Molecular Neuroscience, University of California, San Diego
2011	Teaching Assistant: Neuroanatomy (Neurosciences Program), University of California, San Diego
2012	Teaching Assistant: Sensation and Perception (Psychology Department), University of California, San Diego
2012-2013	NSF GK-12 STEM Fellow: Helix High School, La Mesa, CA
2013-2015	NIH Predoctoral Fellowship, Institute for Neural Computation, University of California, San Diego
2015	Associate-In (Instructor): Multicellular Life (Biology Department), University of California, San Diego
2016	Doctor of Philosophy: Neurosciences, University of California, San Diego

PUBLICATIONS

Perks, K., Gentner T. (2015). "Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex": *European Journal of Neuroscience*, v. 41, p. 725-733.

Perks, K. (2007). "Adaptive sensory filtering in the cerebellum-like mechanosensory nucleus of the hindbrain in Raja Erinacea." Masters Thesis, Wesleyan University, Connecticut. 66p.

ABSTRACT OF THE DISSERTATION

Through the looking glass: population dynamics through membrane potentials.

by

Krista Eva Perks

Doctor of Philosophy in Neurosciences

University of California, San Diego, 2016

Professor Timothy Q. Gentner, Chair

Relating the collective activities of neural populations to external sensory stimuli or to motor output is essential to understanding how nervous systems support behavior. Equally important is examining these processes in the context of ethological signals, which are typically high-dimensional with a wide range of co-varying features at multiple timescales. The synaptic and network mechanisms for encoding these complex signals

are largely unknown, due in part to limitations inherent in experimental design practices leveraging dimensionality reduction rather than embracing the full suite of stimulus complexity. Additional limitations arise because most studies to date examine spiking statistics in randomly-selected populations without consideration for the functional relevance of sub-network groupings. We do not know how spiking responses of individual neurons are pooled, functionally, by downstream neurons – a mechanism that could significantly alter population coding. Stimulus and behavior likely modulate the functional selection of sub-populations, which then likely exhibit different spiking statistics than the population at large.

Using a combination of intracellular and extracellular electrophysiology techniques in the caudal mesopallium (CM) and the caudal nidopallium (NCM) of the common European Starling, I examine synaptic and spiking activity driven by previously-recorded conspecific vocalizations, maintaining the full ethologically-relevant complexity contained in each signal. Uniquely, I feature the synaptic response as an independent variable in the examination of population spiking activity. The main implication of the collective results presented in this study is that, rather than a model of hierarchical processing in which stimulus-specific information is restricted to parallel circuits within each region, information about even the most complex stimuli is likely massively redundant and shared among the population at large. This scaffolds a sensory processing model in which flexible network re-organization - on short timescales and in a stimulus-specific way - support the complexities of spiking output observed in sensory cortex in response to learning, adaptation, attention, and context to meet the demands of an ever-changing environment.

CHAPTER 1

Introduction to the Dissertation

1.1 General scope and outline of the dissertation

A fundamental goal of neuroscience is to model how interactions with the environment guide animal behavior. Neurophysiological approaches to this include recording some aspect of neural activity (spike rate, spike timing, membrane potential, synaptic current, etc.) and accounting for its variance across specific dimensions of a stimulus input or a behavioral output (Werner & Mountcastle, 1963). It is very hard, however, to embrace the full suite of complexity inherent in ethological signals that drive adaptive animal behavior in unpredictable environments. Experimentalists often implement dimensionality reduction on the stimulus set or on the measured neural activity to make the problem more tractable, which then imposes inherent limitations on the generalization of results beyond experimental conditions.

The experiments of this thesis were designed to examine functional pooling and population dynamics driven by high-dimensional temporally-evolving conspecific vocal communication signals. I use a combination of intracellular and extracellular electrophysiology techniques targeting cortical auditory regions (mainly region NCM – caudomedial nidopallium, but also CM - caudal mesopallium) in the common European Starling. The Starling is a model system well-suited for experiments that examine fundamental cognitive processes like those requisite for the aspects of language common to all vocal learning species. However it is also a system generally considered inaccessible to tools most often employed (imaging, genetic manipulation, optogenetics, etc.) in models such as the mouse or fruit fly for dissecting neural circuit architecture and synaptic mechanisms of stimulus encoding.

The physiological techniques and perspectives implemented here provide alternative approaches to ask mechanistic questions about signal processing in cortical sensory systems poised to support high-level communication behavior. As a result, I obtained results that suggest interesting modifications to contemporary models of hierarchical response tuning. What I find is that, although individual neurons in regions of sensory cortex are known to have very sparsely selective spiking output, the synaptic response of an individual neuron readily reflects apparently massively redundant information about the entire range of conspecific vocalizations encountered. I also find evidence for stimulus-driven functional reorganization of neurons within NCM into sub-networks on short timescales suggesting that single downstream targets flexibly pool stimulus information within these cortical regions.

At the single-cell level, the ability for individual neurons to tap into the diversity of a massively redundant set of spiking outputs may be a fundamental pooling strategy underlying characteristically non-linear and flexible spiking output. At the population level, the ability for downstream mechanisms to implement flexible pooling could modulate the structure of population-level spiking correlations in stimulus-specific ways and play a pivotal role in supporting adaptive animal behavior.

In **Chapter 1**, we differentiate two predictions about the stimulus specificity of the subthreshold response underlying sparsely selective spiking in single neurons of CM and NCM based on distinct synaptic pooling strategies. We find that, consistent with a distributed (rather than a sparse) pooling strategy, the net stimulus-evoked synaptic input is specific (i.e. unique for different stimuli), but non-selective (i.e. driven by all stimuli). Although the modulation of subthreshold activity is not always tied to the modulation of

spike output, it is always stimulus-specific, supporting the ability for each cell to be made to spike selectively (through a variety of mechanisms) to an extremely diverse range of relevant features in conspecific vocalizations. Stimulus specificity relies on response reliability across trials, and we find that because local inhibition is necessary for its maintenance in this circuit, it is also therefore poised to modulate it. This study was published in *European Journal of Neuroscience* (Perks & Gentner, 2015).

In **Chapter 2**, we address limitations inherent in contemporary models of sensory processing, which have been supported largely by results from studies that target *randomly* selected single neurons and pairs of neurons from larger populations to estimate properties of stimulus encoding in coordinated spiking output. Instead of examining stimulus-driven structure in the population response by comparing the spiking output of one set of neurons to another, we compare the spiking output of one set of neurons to the synaptic response of a single neuron - representing a spiking population pooled by a downstream neuron. The collective results of this study suggest that sensory integration and processing are supported by a system in which information about even the most complex stimuli is likely massively redundant and shared among the population at large. This study is currently being written up into manuscript form for submission to *Journal of Neuroscience*.

To introduce the main chapters of the thesis, I provide a review of phenomena motivating those experiments and the contemporary models of sensory processing that they address, including more detailed summaries of the main experiments and results presented in each chapter as appropriate.

1.2 Variance and Representation: Mapping the behavioral environment

Variability in neural activity is a feature of neurons and neural systems that allows for information coding. When some metric of neural activity co-varies with some metric of the environment or behavior, then we say that parameter of the neural activity "codes for" that parameter of the environment or behavior. This applies in both sensory and motor systems, but for simplicity here I will focus on sensory physiology.

In sensory systems, "tuning" or "sensitivity" is the relationship between a change in some aspect of neural activity and the stimulus. The visual system is "tuned to" changes in the number and frequency of photons, the auditory system is "tuned to" changes in air pressure. A neuron is "tuned to" or "sensitive to" a particular stimulus parameter if changes in that parameter modulates its synaptic input or its spike output. Within each sensory system, different neurons are sensitive to different features of stimuli in that modality. This is easiest to characterize at the periphery where spiking in single neurons can be driven by changes in single low-dimensional stimulus parameters (such as temporal or spatial frequency). For some neurons near the sensory periphery, linear models of spiking and synaptic receptive fields (the stimuli that modulate spiking output of or synaptic input to the neuron) yield excellent predictions for the responses of that neuron to novel stimuli. For example: orientation tuning in retinal ganglion cells through the thalamo-recipient layer/region of visual cortex; frequency tuning in cochlea through to the thalamic input layer/region of auditory cortex.

The tuning of neurons further along the processing pathway within each sensory system is much harder to characterize, particularly for the complex sensory signals that are essential to many natural behaviors. Linear receptive field estimation for natural

sounds requires additional considerations such as removing signal correlations and sampling sufficient degrees of freedom (Calabrese, Schumacher, Schneider, Paninski, & Woolley, 2011; Meyer, Diepenbrock, Happel, Ohl, & Anemuller, 2014; Meyer, Diepenbrock, Ohl, & Anemuller, 2014; Nelken, 2004; Smyth, Willmore, Baker, Thompson, & Tolhurst, 2003; Touryan & Dan, 2001). For these neurons, the hassle is necessary at the very least because an estimate of the linear receptive field using low dimensional stimuli does not match the receptive field estimated under natural stimulus conditions and generates a worse prediction of responses to novel stimuli (Laudanski, Edeline, & Huetz, 2012; Theunissen, Sen, & Doupe, 2000).

The STRF method – a linear characterization of the complex stimulus-response transformations seen in sensory neurons – is a common method employed to examine responses to temporally and often spectrally complex (high-dimensional) vocal communication signals in cortical neurons of the songbird auditory system well as the auditory midbrain of bats and frogs (Theunissen et al., 2001; Woolley, Gill, Fremouw, & Theunissen, 2009; Woolley, Gill, & Theunissen, 2006). Overall, attempts to create linear models of the spiking receptive field for high-order sensory neurons yield poor predictions of responses to novel stimuli compared to the performance of these models to predict more peripheral responses (Machens, Wehr, & Zador, 2004; Theunissen et al., 2000). Several factors probably affect this. The parameterization of natural stimuli (particularly vocal communication signals) is not clear; these stimuli exist in a higher dimensional space and we do not have a clear model of the relevant dimensions nor their interactions to compare against the neural activity. Additionally, receptive fields of these neurons are likely non-stationary such that a single spike may mean different things in

different contexts. Neurons in Starling cortex exhibit flexible feature recombination at their synapses as well as composite receptive fields (Kozlov & Gentner, 2014, 2016). In A1, it is clear that choice of stimulus identity and context experiments lead to seemingly contradictory conclusions about stimulus-response relationships with the interplay of acoustic events across diverse timescales as a major culprit for causing misinterpretation (Nelken, 2004).

The temporal nature of the stimulus can be ignored as an attempt to better parameterize the stimulus-response relationship. This comes at the cost of not having control over meaningful variance in the neural activity. Baudot et al (2013) decompose the variance in both spiking output and in subthreshold input into signal and noise components(Baudot et al., 2013). After performing a frequency decomposition using wavelet analysis, they compare the spectral profile of signal and noise variance across categorically different stimulus conditions: drifting gratings, natural scenes, and dense noise. The results show that, *in-vivo*, trial-to-trial variability and temporal irregularity depend on stimulus statistics. Spiking and synaptic responses to the low-dimensional, temporally-static stimuli used to characterize receptive fields are highly unreliable with low temporal variance. This has lead to the observation of supra-Poisson spike count variability and the justification of only interpreting activity under “rate coding” (averaging across time) or “population coding” (averaging across many neurons) models. Time-varying stimuli with natural scene statistics maximize the coding capacity of synaptic and spiking activity in individual neurons by instead driving responses with simultaneously high variability over time and low variability across trials.

The mechanisms through which statistics naturally encountered in ethological contexts constrain network dynamics to increase coding efficiency are unknown. It is clear that the response profiles under low-dimensional stimulus conditions are categorically different than under natural conditions and using them to generate models of receptive fields and mechanisms of sensory processing is misleading. Additionally, further along the processing pathway, low dimensional, well-parameterized stimuli do not even do a sufficient job at eliciting spikes from single neurons. We therefore know very little about the mechanisms of response tuning and synaptic pooling strategies for time-varying natural stimuli in high-order object-selective neurons.

1.3 Pooling strategies for non-selective, stimulus-specific synaptic responses

Responses to high-dimensional stimuli are sparse (population sparse and lifetime sparse) (Vinje & Gallant, 2002)Hierarchical models of sensory processing predict that cortical circuits perform increasingly complex and integrative computations as information flows from primary to secondary regions and beyond, which could imply that information available is restricted to more and more specialized neurons with sparsely-selective input pools, all operating in parallel. However, behavior and context can quickly and drastically modulate what are already often non-linear spiking receptive fields (Fritz, David, Radtke-Schuller, Yin, & Shamma, 2010; Lee & Middlebrooks, 2011; Osmanski & Wang, 2015; Sharpee & Victor, 2009). Kozlov et al (2014,2016) demonstrate that the combination functions of inputs to single neurons can be flexible and comprise orthogonal features from the stimulus set (Kozlov & Gentner, 2014, 2016). Modulation of even more fundamental receptive field properties such as size (Pettet & Gilbert, 1992)

belie that neurons may have access to information about a much larger portion of the relevant global stimulus set than revealed by the spiking response. All of these phenomena seem to demand mechanisms of hierarchical processing that allow neurons to rapidly and drastically modulate sensory integration. By using intracellular techniques we can measure response properties of the synaptic input space directly and start to test assumptions of various synaptic pooling models.

Chapter 1 begins with an examination of the pooling principles of individual neurons nestled deep along the auditory processing pathway in Starling cortex (in areas CM and NCM). I used whole-cell patch clamp techniques to measure subthreshold changes in the cross-membrane potential driven by stimulation with conspecific vocalizations over the course of seconds. Doing so revealed that the membrane is under continuous stimulus control with specificity at extremely high temporal resolution (sub-motif at least if not sub-feature; on the order of 10s of milliseconds). Sparse spiking activity characteristic of these regions, although specific and reliable, does not belie the breadth of stimulus information potentially available to the output of individual neurons. Stimulus specificity depends on the entropy of the temporal response distribution but also on the reliability across trials given a particular stimulus. These results are consistent with results from primary auditory and visual cortex in the mouse (Baudot et al., 2013; Machens et al., 2004), which both demonstrate that sparse spike responses arise from irregular and highly reproducible membrane potential trajectories. Such pervasive stimulus specificity implies that the synaptic receptive field is extremely broad, pooling from diverse regions of the stimulus space not necessarily represented in the spiking output.

Variance in the neural activity not accounted for by variance in the stimulus is most often referred to as "noise," but a less biased term is "trial variance." Previous studies show that sensory stimulation decreases trial variance in spiking output of individual neurons throughout cortex under a variety of stimulus and behavior conditions (Churchland et al., 2010). Again, this is consistent with the decrease in trial variance in the membrane potential that I observe. To examine local inhibition as a potential mechanisms driving this decrease in trial variance I revived a previously published dataset (Thompson, Jeanne, & Gentner, 2013) to compare the trial-to-trial variability in spiking output from neurons recorded in NCM when inhibition is intact versus when inhibition is blocked locally by GABA antagonists. Blocking local inhibition increased spiking variability as measured by the FANO factor. This reveals that trial variability is potentially under circuit control –the significance of which surfaces when we consider population-level stimulus encoding.

1.4 Dissociable sources of spiking covariance modulate discriminability.

With regards to the task of the nervous system - neurons do not need to be "tuned to" a feature to contribute to its encoding - the presence of a stimulus often matters less for behavior than the ability to discriminate among a set of stimuli. The animal needs to know what to do in response to a stimulus, not what the stimulus identity is explicitly. We can think of the set of stimuli in the environment creating a map to guide behavior - stimulus discrimination increases the resolution of that map (whatever space/shape/form that map may take). Although a relationship between variance in a single-neuron response and some parameter of a stimulus set can often be defined, it is really the

coordinated activities of many individual neurons that drive behavior. Stimulus discrimination is not just a function of response strength, but of variation in the response and the coordinated variation in populations of neurons (because – with the exception of invertebrates –single neurons are rarely necessary and sufficient for supporting discrimination processes that drive behavioral output). Although individual neurons can be "tuned to" particular objects, the nervous system does not necessarily source these single-unit responses for stimulus discrimination.

If we are to understand how neural activity drives behavior, we must at some point model population dynamics. Population-level activity can be resolved by large-array extracellular recording techniques as well as imaging techniques to examine coordinated neural activity within populations of various sizes. The most heavily studied population size in which coordinated activity has been studied thus far is populations of size n = 2 (pairs of neurons). Covariance in the magnitude of spiking output between pairs of neurons is often measured by calculating the correlation coefficient between activity vectors for the two neurons (Cohen & Kohn, 2011; Lyamzin, Garcia-Lazaro, & Lesica, 2012). The total variance in spiking output can be partitioned and attributed to different sources (Goris, Movshon, & Simoncelli, 2014). As already mentioned, one major source of variance is due to stimulus identity. Among populations of neurons this type of covariance is usually referred to as the "signal correlation." "Noise correlations" (more aptly termed "trial correlations") describe the trial-to-trial covariance in spiking output – the variance unaccounted for by changes in the stimulus.

The relationship between these two major sources of variability is currently an area of intense research owing to how profoundly the structure of coordinated spiking

among neurons across stimuli and across time can affect stimulus discrimination from the population response. Orthogonal patterns of "signal" and "noise" correlation support stimulus "maps" with the highest resolution ideal for decision making processes dependent on stimulus discrimination) (Averbeck, Latham, & Pouget, 2006; Franke et al., 2016; Hu, Zylberberg, & Shea-Brown, 2014; Salinas & Sejnowski, 2000).

1.5 Flexible and stimulus-dependent spiking co-variability.

Even more critical than a specific pattern of spiking correlation is the ability to modulate these correlations and bias response discrimination in predictable ways. Jeanne et al (2013) found behavior-induced modulation in the joint statistics of signal and noise spiking correlations that benefits stimulus coding for pairwise populations in songbird auditory cortex (Jeanne, Sharpee, & Gentner, 2013). After subjects learned that only a certain subset of acoustic stimuli were informative for their reward-driven task, a linear classifier trained on stimulus-driven spike counts was able to better discriminate among those stimuli than among stimuli that were not informative for the task. Discrimination improved with increased population size. For populations of size 2, the authors calculated the pairwise signal and noise correlations and found that task-relevance increased the orthogonality of signal and noise correlation directions (positive/negative).

Mechanisms for controlling spiking covariance must exist and are likely to be accessible by top-down and bottom-up systems involved in cognitive processes such as learning and attention. Stimulus-driven decrease in trial variance seems to be a pervasive phenomenon (Churchland et al., 2010). In chapter 1 I show that in the Starling cortex, continuous acoustic stimulation is accompanied by a persistent decrease in trial variance

relative to the pre-stimulus period (silence) in both the spiking output of individual neurons as well the synaptic input to individual neurons - regardless of whether the stimulus drives spiking output or not. Blocking local inhibition, however, caused an increase in the Fano Factor of spiking output. Inhibition is therefore one circuit mechanism poised to modulate trial variance. Inhibitory circuits have been proposed as a mechanism for decorrelating spiking output among pairs of neurons (Bernacchia & Wang, 2013; Tetzlaff, Helias, Einevoll, & Diesmann, 2012). Jaime de la Rocha et al (2007) also show that pairs of neurons with higher firing rates tend to have correlations higher than expected given the increase in spike count alone (de la Rocha, Doiron, Shea-Brown, Josic, & Reyes, 2007). Schultz et al (2015) measured the dependence of noise correlations on 11 factors and found effects of firing rate and sensory tuning (including effects of distance between neurons, spike width, and spike isolation quality) (Schulz, Sahani, & Carandini, 2015). Mechanisms driving either of these physiological factors may provide points of control over variance in the system.

A change in trial-to-trial variance alone does not necessarily affect stimulus discrimination. The effect of noise correlations on stimulus encoding depends on the signal correlations in that population. Several studies have even shown that noise correlations are stimulus-dependent (Ponce-Alvarez, Thiele, Albright, Stoner, & Deco, 2013; Zylberberg, Cafaro, Turner, Shea-Brown, & Rieke, 2016). Interestingly, discrepancy among studies measuring the strength of noise correlations arises in part due to reporting the average noise correlations over the stimulus set rather than reporting the stimulus-specific values. Results describing stimulus-specific noise correlations are consistent with stimulus-specific modulation of that relationship as shown now in several

studies (Downer, Niwa, & Sutter, 2015; Jeanne et al., 2013; Ruff & Cohen, 2014a, 2014b).

Neural activity also shows higher-order statistical structure, meaning that the dynamics cannot be captured by the mean activity and pairwise statistics (the lower-order moments) alone. Since pairwise correlations do not uniquely define patterns of coactivity, we must consider higher-order correlations and what it means to have orthogonal signal and noise correlations in those higher dimensions. High-dimensional coordination among synaptic input also modulates diverse properties of spiking output in single neurons including spike timing and rate (A., 2002; Bohte, Spekreijse, & Roelfsema, 2000; Ohiorhenuan et al., 2010; Salinas & Sejnowski, 2000). Higher-order correlations are observed in neural populations from the periphery (like the retina) through to cortical areas. Higher-order correlations can modulate stimulus encoding (Ganmor, Segev, & Schneidman, 2011; Montani et al., 2009; Ohiorhenuan et al., 2010), an effect that can even depend on the stimulus-specificity of these correlations (Josic, Shea-Brown, Doiron, & de la Rocha, 2009).

Synchronous spiking in larger subgroups of cells can be infrequent, making higher-order correlations are difficult to detect because dataset size is often insufficient to capture a sufficient probability function for all event types. Because higher-order correlations can be explained geometrically as either positively or negatively skewing the distribution of the summed population activity in short time windows (Cayco-Gajic, Zylberberg, & Shea-Brown, 2015), methods have been resolved for inferring the probability distribution of higher order events from the lower order statistics of the response distribution. Staude et al (2010) leverage the first, second, and third order

cummulants of the distribution describing the population spiking response magnitude across time to estimate the probability of high-order synchoronous events (Staude, Rotter, & Grun, 2010).

Even if tools were perfected that reliably and efficiently enabled quantification of high-order covariance among population of spiking neurons, experimentors would still be faced with the issue of choosing the relevant population to study. The growth of high-throughput technologies has resulted in exponential growth in datasets with respect to both dimensionality and sample size, but bigger datasets do not necessarily result in the ability to better capture statistics of coordinated spiking output that support behavior. Co-active populations do not necessarily impact downstream neurons together (Okun et al., 2015). The existence of correlations or a particular pattern of covariance does not guarantee that those statistics are seen by downstream neurons in the processing hierarchy. For example, a randomly-recorded population may exhibit the full distribution of signal and noise correlations and their interaction, while downstream neurons only sample from the subpopulation with orthogonal correlations.

1.6 Identifying functionally-defined sub-populations.

Within a given sensory region, how do we identify "functional" sub-populations in order to examine issues of spiking covariance, like their effect on stimulus discrimination? The choice of a "relevant population" in part it depends on the question. Neurons can be subsampled based on a variety of factors including, but not limited to: cell type, anatomical location, and response tuning. Extracellular techniques such as imaging or dense electrode arrays enable the identification of sets of neurons from large-

scale population recordings that are co-active or functionally grouped by their response properties (spatial and temporal variance of population spiking) or by their co-activity. In mammalian cortex anatomically and genetically defined subpopulations can be targeted in experiments examining the specificity of correlation structure among different components of a neural circuit (like inhibitory versus excitatory neurons). Another way to define functional networks is by their connectivity (specifically, by common synaptic targets). Although histological techniques allow the identification of networks of anatomically connected neurons we lack extracellular tools to target the spiking activity of a population of neurons defined by a common synaptic target.

In Chapter 2, I examine the selection of stimulus-specific functional networks among neurons in NCM of the Starling. I begin by determining the temporal and spatial organization of spiking variance throughout NCM and throughout conspecific vocalizations as this organization necessarily constrains mechanisms for stimulus-driven pooling of the spiking activity. Evidence for "tonotopy" and "feature-topy" in songbird auditory cortex comes from results of electrophysiology and IEG and ZENK expression experiments in zebra finch (Muller & Leppelsack, 1985; Terleph, Mello, & Vicario, 2006). And STRF estimation has revealed spatial organization of receptive field profiles across regions of auditory cortex in zebra finch (Kim & Doupe, 2011; Nagel & Doupe, 2008; Woolley et al., 2009). Neurons in sensory cortices are often topographically organized according to their response preferences. Again, this has mainly been examined using static, low-dimensional stimuli (visual system: oriented bars of light, moving bars of light; auditory cortex: frequency of a sound). There are very few studies on the spatial organization of tuning to high-order objects, which Chapter 2 of this thesis examines. The

experiments of Dahl et al (2009) reveal spatial organization of modality preferences in a higher association cortex, but they did not examine responses at the resolution of stimulus-specificity within a modality (Dahl, Logothetis, & Kayser, 2009). I show that in Starling NCM, near-neighbors tend to have higher signal correlations than distant units when driven with conspecific vocalizations. Spatial clustering of response similarity constrains mechanisms that can account for the synaptic input observed in individual neurons, which broadly representing the stimulus set.

Although spiking activity is spatially distributed such that the output of near-neighbors is more highly correlated than distant units, we find that a remarkable amount of variance in the net (mean-field) NCM spiking response is reflected in the synaptic response of single neurons randomly selected from within the region. Insofar as a neuron's synaptic input and/or membrane potential is a reflection of spiking activity in the pre-synaptic network then that neuron becomes an electrode through which we can examine the activity of a functionally relevant, spatially unrestricted population of neurons un-biased by things like spike sorting or extracellular sampling techniques.

I constrain the units contributing to the mean-field NCM response by implementing regularized linear regression models to those most able to predict temporal variance in the synaptic response of single neurons from within NCM. I find that these functionally targeted networks are more accurately/reliably identified at the time-scale of single motifs (the second-long elements comprising song) than at the timescale of whole vocalization sequences (20-seconds). Across different motif stimuli, the identity of the functionally-relevant subpopulation changes such that synaptic response tuning identified a temporally heterogeneous coordination of population spiking activity. Although pairs of

neurons in these functional networks tend to have more positively correlated stimulus-driven responses than expected if they were sampled randomly from the population at large, the neurons comprising synaptically-defined populations are not spatially clustered but are flatly distributed throughout NCM.

Each functional network potentially contains highly correlated variables. I conclude the second chapter by quantifying the dimensionality reduction necessary to eliminate redundant information about the synaptic response from the identified functional networks, where the dimension of the network is defined by the number of units contributing spikes to the fit of the synaptic response. The degree of redundancy is inherently related to key aspects of population coding, where a key aspect of the problem is determining the correlations in spiking activity among the population and whether those correlations effect decoding of relevant information from population activity. The issue of whether correlations are important has been a subject of heated debate and is an area of future follow-up experiments. The mechanisms by which correlations effect decoding and the important types of correlations will take a long time to decipher, in part due to the many ways in which these phenomena are quantified (Latham & Nirenberg, 2005). However, one good start is to focus on functionally relevant populations of neurons when asking the questions, which is the major contribution to the field provided by the results from Chapter 2.

I close the dissertation with a chapter dedicated to future directions and potential caveats of the studies published from this thesis and a discussion of relevant open questions in the field.

1.7 Summary and Broader Impacts

With respect to the approach of recording some aspect of neural activity (spike rate, spike timing, membrane potential, synaptic current, etc) and accounting for its variance across specific dimensions of a stimulus input or a behavioral output, the advent of larger-scale extracellular methods at higher temporal and spatial resolutions has allowed for an explosion in the number of units among which analyses of coordinated population activity can be applied. In some ways, larger datasets of more neurons are better in that we can get a larger sampling of the full distribution of response statistics. This can be particularly important for resolving higher order correlations, which can be extremely rare but also extremely influential in models of stimulus discrimination. However, researchers are often left faced with a dimensionality reduction problem in the dataset. There are many ways to reduce the dimensionality of a set of data in order to lower computational complexity, build models, and decrease the amount of storage and working memory required for data processing. These methods are often inherent in the analyses used to examine mechanisms of sensory processing.

One way that dimensionality reduction occurs is by designing experiments that allow researchers to decipher a "population code" – the set of elements or the dimensions of a stimulus that the coordinated spiking of large populations of neurons are "encoding." This approach has traditionally been applied before the neural response is even acquired. Dimensionality reduction on the stimulus set – reducing the rich, transforms high-dimensional structure of natural signals into an arbitrary set of lower-dimensional features that are more easily parameterized and amenable to application in generating models for linear receptive fields. We can categorize these approaches as forms of *feature*

extraction – mapping the original feature space (the spiking activity of individual neurons or the full statistics of natural stimuli) to a new feature space with lower dimensions (some parameter of a stimulus). One caveat of this approach is that the resulting set of dimensions lacks any direct interpretability in terms of the original physical meaning of those elements. Another caveat is that the features over which the transformations and mappings are being applied is not necessarily the features that such operations are pooled across by individual neurons.

Another way to reduce dimensionality is to select a subset of features from the original pool without any transformation. In the context of neural circuits, this method is akin to how downstream neurons restrict their pooling operation to sample from specific subsets of a population. This method offers more direct interpretability of the original physical meaning of individual elements (like cell identity). Functionally, if selection of functional networks is restricted in this way (by some classification based on the pooling operations of individual downstream neurons) then one is left with the opportunity to examine the spiking behavior of populations directly relevant to the integration functions directing behavior in ways that studying randomly selected networks does not allow. Identifying pre-synaptic networks to test models of synaptic pooling strategies and the effects of their coordinated spiking activity presents technical challenges currently unresolved by large-scale extracellular recording techniques alone.

The synaptic response of individual neurons of sensory cortex can be very temporally specific when driven with “natural” stimuli containing the set of statistics that the network has developed to represent, providing rich information about the stimulus. Regularization methods, such as the L1 norm (or “lasso”) are one tool for implementing

feature selection on a set of dependent variables in linear regression. By setting the independent variable in a linear regression model to the synaptic response of a single neuron, one would target subsets of the spiking units among the population able to model that pre-synaptic network. Although seemingly unbelievable given prior assumptions implicit in contemporary hierarchical models of restrictive, parallel sensory processing, I have demonstrated the effort as a viable approach to targeting subnetworks within a region based on the relationship between their spiking output and some downstream synaptic response.

In this dissertation I have approached the dimensionality reduction problem from a functional perspective, embracing the complexity of natural Starling vocalizations and the diversity of spiking output maintained throughout NCM. I have instead acted creatively on the perspective lens through which I have examined the resulting neural activity. What has benefited from this has been appreciation of the rich information available in the synaptic responses underlying neural representations traditionally regarded and modeled as sparse and independent. The results have scaffolded the field to support future directions of adapting analyses targeted at examining the structure of coordinated spiking output supporting behaviorally adaptive stimulus discrimination into a framework amenable to tracking that structure through time in synaptically-defined subnetworks of populations that have traditionally been sampled at random or categorized by cell type or anatomical location.

The Starling is a prized model organism for the ethology of complex sequence processing and learning and is often used as a model for language acquisition through development and learned cognitive auditory phenomena in humans. Advances that push

the Starling forward as a model to study the neural mechanisms supporting these complex cognitive processes and downstream modulation of sensory cortical processing would yield great benefits beyond the lab in understanding these processes in healthy adults as well as developmental disorders. By combining intracellular and extracellular electrophysiological recording techniques in novel ways within this system I have contributed unique perspectives to our understanding of how cortical circuits process complex natural signals.

1.8 References

A., K. A. R. S. A. (2002). Correlated input spike trains and their effects on the response of the leaky integrate-and-fire neuron. *Neurocomputing, 44-46*, 121-126.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci, 7*(5), 358-366. doi:10.1038/nrn1888

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., & Fregnac, Y. (2013). Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. *Front Neural Circuits, 7*, 206. doi:10.3389/fncir.2013.00206

Bernacchia, A., & Wang, X. J. (2013). Decorrelation by recurrent inhibition in heterogeneous neural circuits. *Neural Comput, 25*(7), 1732-1767.

Bohte, S. M., Spekreijse, H., & Roelfsema, P. R. (2000). The effects of pair-wise and higher order correlations on the firing rate of a post-synaptic neuron. *Neural Comput, 12*(1), 153-179.

Calabrese, A., Schumacher, J. W., Schneider, D. M., Paninski, L., & Woolley, S. M. (2011). A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. *PLoS One, 6*(1), e16104. doi:10.1371/journal.pone.0016104

Cayco-Gajic, N. A., Zylberberg, J., & Shea-Brown, E. (2015). Triplet correlations among similarly tuned cells impact population coding. *Front Comput Neurosci, 9*, 57. doi:10.3389/fncom.2015.00057

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., . . . Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci, 13*(3), 369-378.

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat Neurosci, 14*(7), 811-819.

Dahl, C. D., Logothetis, N. K., & Kayser, C. (2009). Spatial organization of multisensory responses in temporal association cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(38), 11924-11932. doi:10.1523/JNEUROSCI.3437-09.2009

de la Rocha, J., Doiron, B., Shea-Brown, E., Josic, K., & Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature, 448*(7155), 802-806. doi:10.1038/nature06028

Downer, J. D., Niwa, M., & Sutter, M. L. (2015). Task engagement selectively modulates neural correlations in primary auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 35*(19), 7565-7574. doi:10.1523/JNEUROSCI.4094-14.2015

Franke, F., Fiscella, M., Sevelev, M., Roska, B., Hierlemann, A., & da Silveira, R. A. (2016). Structures of Neural Correlation and How They Favor Coding. *Neuron, 89*(2), 409-422. doi:10.1016/j.neuron.2015.12.037

Fritz, J. B., David, S. V., Radtke-Schuller, S., Yin, P., & Shamma, S. A. (2010). Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat Neurosci, 13*(8), 1011-1019. doi:10.1038/nn.2598

Ganmor, E., Segev, R., & Schneidman, E. (2011). Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proc Natl Acad Sci U S A, 108*(23), 9679-9684. doi:10.1073/pnas.1019641108

Goris, R. L., Movshon, J. A., & Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nat Neurosci, 17*(6), 858-865. doi:10.1038/nn.3711

Hu, Y., Zylberberg, J., & Shea-Brown, E. (2014). The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. *PLoS Comput Biol, 10*(2), e1003469. doi:10.1371/journal.pcbi.1003469

Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron, 78*(2), 352-363.

Josic, K., Shea-Brown, E., Doiron, B., & de la Rocha, J. (2009). Stimulus-dependent correlations and population codes. *Neural Comput, 21*(10), 2774-2804. doi:10.1162/neco.2009.10-08-879

Kim, G., & Doupe, A. (2011). Organized representation of spectrotemporal features in songbird auditory forebrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 31*(47), 16977-16990. doi:10.1523/JNEUROSCI.2003-11.2011

Kozlov, A. S., & Gentner, T. Q. (2014). Central auditory neurons display flexible feature recombination functions. *J Neurophysiol, 111*(6), 1183-1189. doi:10.1152/jn.00637.2013

Kozlov, A. S., & Gentner, T. Q. (2016). Central auditory neurons have composite receptive fields. *Proc Natl Acad Sci U S A, 113*(5), 1441-1446. doi:10.1073/pnas.1506903113

Latham, P. E., & Nirenberg, S. (2005). Synergy, redundancy, and independence in population codes, revisited. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 25*(21), 5195-5206. doi:10.1523/JNEUROSCI.5319-04.2005

Laudanski, J., Edeline, J. M., & Huetz, C. (2012). Differences between spectro-temporal receptive fields derived from artificial and natural stimuli in the auditory cortex. *PLoS One, 7*(11), e50539. doi:10.1371/journal.pone.0050539

Lee, C. C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nat Neurosci, 14*(1), 108-114. doi:10.1038/nn.2713

Lyamzin, D. R., Garcia-Lazaro, J. A., & Lesica, N. A. (2012). Analysis and modelling of variability and covariability of population spike trains across multiple time scales. *Network, 23*(1-2), 76-103. doi:10.3109/0954898X.2012.679334

Machens, C. K., Wehr, M. S., & Zador, A. M. (2004). Linearity of cortical receptive fields measured with natural sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 24*(5), 1089-1100. doi:10.1523/JNEUROSCI.4445-03.2004

Meyer, A. F., Diepenbrock, J. P., Happel, M. F., Ohl, F. W., & Anemuller, J. (2014). Discriminative learning of receptive fields from responses to non-Gaussian stimulus ensembles. *PLoS One, 9*(4), e93062. doi:10.1371/journal.pone.0093062

Meyer, A. F., Diepenbrock, J. P., Ohl, F. W., & Anemuller, J. (2014). Temporal variability of spectro-temporal receptive fields in the anesthetized auditory cortex. *Front Comput Neurosci, 8*, 165. doi:10.3389/fncom.2014.00165

Montani, F., Ince, R. A., Senatore, R., Arabzadeh, E., Diamond, M. E., & Panzeri, S. (2009). The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. *Philos Trans A Math Phys Eng Sci, 367*(1901), 3297-3310. doi:10.1098/rsta.2009.0082

Muller, C. M., & Leppelsack, H. J. (1985). Feature extraction and tonotopic organization in the avian auditory forebrain. *Exp Brain Res, 59*(3), 587-599.

Nagel, K. I., & Doupe, A. J. (2008). Organizing principles of spectro-temporal encoding in the avian primary auditory area field L. *Neuron, 58*(6), 938-955. doi:10.1016/j.neuron.2008.04.028

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol, 14*(4), 474-480. doi:10.1016/j.conb.2004.06.005

Ohiorhenuan, I. E., Mechler, F., Purpura, K. P., Schmid, A. M., Hu, Q., & Victor, J. D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. *Nature, 466*(7306), 617-621. doi:10.1038/nature09178

Okun, M., Steinmetz, N. A., Cossell, L., Iacaruso, M. F., Ko, H., Bartho, P., . . . Harris, K. D. (2015). Diverse coupling of neurons to populations in sensory cortex. *Nature, 521*(7553), 511-515. doi:10.1038/nature14273

Osmanski, M. S., & Wang, X. (2015). Behavioral dependence of auditory cortical responses. *Brain Topogr, 28*(3), 365-378. doi:10.1007/s10548-015-0428-4

Perks, K. E., & Gentner, T. Q. (2015). Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex. *Eur J Neurosci, 41*(5), 725-733. doi:10.1111/ejn.12831

Pettet, M. W., & Gilbert, C. D. (1992). Dynamic changes in receptive-field size in cat primary visual cortex. *Proc Natl Acad Sci U S A, 89*(17), 8366-8370.

Ponce-Alvarez, A., Thiele, A., Albright, T. D., Stoner, G. R., & Deco, G. (2013). Stimulus-dependent variability and noise correlations in cortical MT neurons. *Proc Natl Acad Sci U S A, 110*(32), 13162-13167. doi:10.1073/pnas.1300098110

Ruff, D. A., & Cohen, M. R. (2014a). Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci, 17*(11), 1591-1597. doi:10.1038/nn.3835

Ruff, D. A., & Cohen, M. R. (2014b). Global cognitive factors modulate correlated response variability between V4 neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(49), 16408-16416. doi:10.1523/JNEUROSCI.2750-14.2014

Salinas, E., & Sejnowski, T. J. (2000). Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 20*(16), 6193-6209.

Schulz, D. P., Sahani, M., & Carandini, M. (2015). Five key factors determining pairwise correlations in visual cortex. *J Neurophysiol, 114*(2), 1022-1033. doi:10.1152/jn.00094.2015

Sharpee, T. O., & Victor, J. D. (2009). Contextual modulation of V1 receptive fields depends on their spatial symmetry. *J Comput Neurosci, 26*(2), 203-218. doi:10.1007/s10827-008-0107-5

Smyth, D., Willmore, B., Baker, G. E., Thompson, I. D., & Tolhurst, D. J. (2003). The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 23*(11), 4746-4759.

Staude, B., Rotter, S., & Grun, S. (2010). CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains. *J Comput Neurosci, 29*(1-2), 327-350. doi:10.1007/s10827-009-0195-x

Terleph, T. A., Mello, C. V., & Vicario, D. S. (2006). Auditory topography and temporal response dynamics of canary caudal telencephalon. *J Neurobiol, 66*(3), 281-292. doi:10.1002/neu.20219

Tetzlaff, T., Helias, M., Einevoll, G. T., & Diesmann, M. (2012). Decorrelation of neural-network activity by inhibitory feedback. *PLoS Comput Biol, 8*(8).

Theunissen, F. E., David, S. V., Singh, N. C., Hsu, A., Vinje, W. E., & Gallant, J. L. (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network, 12*(3), 289-316.

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 20*(6), 2315-2331.

Thompson, J. V., Jeanne, J. M., & Gentner, T. Q. (2013). Local inhibition modulates learning-dependent song encoding in the songbird auditory cortex. *J Neurophysiol, 109*(3), 721-733.

Touryan, J., & Dan, Y. (2001). Analysis of sensory coding with complex stimuli. *Curr Opin Neurobiol, 11*(4), 443-448.

Vinje, W. E., & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 22*(7), 2904-2915.

Werner, G., & Mountcastle, V. B. (1963). The Variability of Central Neural Activity in a Sensory System, and Its Implications for the Central Reflection of Sensory Events. *J Neurophysiol, 26*, 958-977.

Woolley, S. M., Gill, P. R., Fremouw, T., & Theunissen, F. E. (2009). Functional groups in the avian auditory system. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(9), 2780-2793. doi:10.1523/JNEUROSCI.2042-08.2009

Woolley, S. M., Gill, P. R., & Theunissen, F. E. (2006). Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 26*(9), 2499-2512. doi:10.1523/JNEUROSCI.3731-05.2006

Zylberberg, J., Cafaro, J., Turner, M. H., Shea-Brown, E., & Rieke, F. (2016). Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. *Neuron, 89*(2), 369-383. doi:10.1016/j.neuron.2015.11.019

CHAPTER 2

Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex.

EJN European Journal of Neuroscience

FENS

European Journal of Neuroscience, Vol. 41, pp. 725–733, 2015 doi:10.1111/ejn.12831

Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex

Krista E. Perks[1] and Timothy Q. Gentner[1,2,3,4]
[1]Neurosciences Graduate Program, University of California San Diego, 9500 Gilman Dr, La Jolla, CA, USA
[2]Department of Psychology, University of California San Diego, La Jolla, CA, USA
[3]Neurobiology Section, Division of Biological Sciences, University of California San Diego, La Jolla, CA, USA
[4]Kavli Institute for Brain and Mind, La Jolla, CA, USA

Keywords: auditory cortex, hierarchical models, natural stimuli, object selectivity, synaptic integration

Abstract

Natural acoustic communication signals, such as speech, are typically high-dimensional with a wide range of co-varying spectro-temporal features at multiple timescales. The synaptic and network mechanisms for encoding these complex signals are largely unknown. We are investigating these mechanisms in high-level sensory regions of the songbird auditory forebrain, where single neurons show sparse, object-selective spiking responses to conspecific songs. Using whole-cell *in vivo* patch clamp techniques in the caudal mesopallium and the caudal nidopallium of starlings, we examine song-driven subthreshold and spiking activity. We find that both the subthreshold and the spiking activity are reliable (i.e. the same song drives a similar response each time it is presented) and specific (i.e. responses to different songs are distinct). Surprisingly, however, the reliability and specificity of the subthreshold response was uniformly high regardless of when the cell spiked, even for song stimuli that drove no spikes. We conclude that despite a selective and sparse spiking response, high-level auditory cortical neurons are under continuous, non-selective, stimulus-specific synaptic control. To investigate the role of local network inhibition in this synaptic control, we then recorded extracellularly while pharmacologically blocking local GABAergic transmission. This manipulation modulated the strength and the reliability of stimulus-driven spiking, consistent with a role for local inhibition in regulating the reliability of network activity and the stimulus specificity of the subthreshold response in single cells. We discuss these results in the context of underlying computations that could generate sparse, stimulus-selective spiking responses, and models for hierarchical pooling.

Introduction

Stimulus encoding – the relationship between an external event and the accompanying neural response – is the cornerstone of sensory neurophysiology (Adrian, 1926). Yet, for the complex sensory signals that are essential to many natural behaviors, our understanding of stimulus encoding is poor. In particular, we know very little about the synaptic inputs evoked by natural signals, and the operations governing their integration and transformation into spiking responses in single neurons.

Here we test specific predictions about the selectivity of stimulus-specific synaptic drive underlying sparse, selective spiking in the auditory cortex of European starlings, a species of songbird. Songbirds, in particular starlings, are well suited for these studies. Starling songs are acoustically complex and composed of very diverse, brief segments (motifs) that are perceived as distinct auditory objects (Gentner, 2008). Stimulus-driven spiking activity in the higher-order cortical regions caudal mesopallium (CM) and caudomedial nidopallium (NCM) is sparse: only a small portion of all possible motifs evoke robust spiking in single neurons (Gentner & Margoliash, 2003; Meliza *et al.*, 2010; Thompson *et al.*, 2010; Jeanne *et al.*, 2011; Meliza & Margoliash, 2012) and each motif evokes spiking from only a small number of neurons (Gentner & Margoliash, 2003). Responses to song elements are dependent on acoustic context (Jeanne *et al.*, 2011; Kozlov & Gentner, 2014), and (in zebra finches) combining song elements into longer bouts increases the sparseness of spiking responses (Schneider & Woolley, 2013).

Sparse spiking responses to natural signals appear to be a general property of sensory cortex (Vinje & Gallant, 2000, 2002; Olshausen & Field, 2004; Graham & Field, 2007b; Hromadka *et al.*, 2008; Sakata & Harris, 2009; Tolhurst *et al.*, 2009). That is, only a small proportion of all stimuli evoke spikes from any given neuron (lifetime sparseness, which we refer to as selectivity), and only a small proportion of neurons spike at any point in time (population sparseness). Sparse representations convey a range of computational benefits to downstream neurons for the classification and recognition of complex signals (Ganguli & Sompolinsky, 2012; Babadi & Sompolinsky, 2014). Models for object recognition and classification rely on feed-forward hierarchical pooling of the outputs from simpler feature detectors to build sparse, selective spiking responses to increasingly complex signals that carry behaviorally relevant information (Riesenhuber

Correspondence: Dr T. Q. Gentner, [1]Neurosciences Graduate Program, as above.
E-mail: tgentner@ucsd.edu

Received 12 September 2014, revised 7 December 2014, accepted 11 December 2014

& Poggio, 1999). The synaptic and network mechanisms underlying this pooling remain unclear.

We distinguish two potential synaptic pooling regimes: 'sparse' and 'distributed'. Each regime makes different predictions about the selectivity of the subthreshold response underlying selective spiking in single neurons. In a sparse pooling regime, neurons with selective spiking responses pool inputs that are necessarily less selective but still biased towards the features in complex natural signals that ultimately drive changes in spike rates. Sparse synaptic pooling predicts that stimulus-specific subthreshold activity will be selective (i.e. driven only by a subset of the potential stimuli), and that the degree of selectivity will vary only quantitatively from that of the spiking response. Moreover, the timing of stimulus-specific subthreshold responses should, on average, align with the stimulus-evoked changes in spike rate. We contrast sparse pooling with distributed pooling, in which neurons with selective spiking responses pool synaptic inputs that are not biased towards the features in complex natural signals that drive changes in spike rates. Distributed pooling predicts that the net stimulus-evoked synaptic input will be stimulus-specific (i.e. unique for different stimuli), but non-selective (i.e. driven by all stimuli). Because it is non-selective, the stimulus-specific subthreshold activity under distributed pooling will not necessarily be tied to stimulus-evoked spike rate changes. So long as the inputs are stimulus-specific, however, the cell could be made to spike selectively (through a variety of mechanisms) to a potentially wide range of relevant features.

The foregoing hypothetical pooling scenarios both predict that the subthreshold membrane response of neurons whose stimulus-driven spiking responses are selective will contain stimulus-specific activity. Only the sparse pooling hypothesis requires that the stimulus-specific subthreshold activity is selective for subsets of stimuli. To distinguish these two hypothetical pooling scenarios, we compare properties of the subthreshold response to spiking. Direct measurement of the subthreshold activity along with spiking was accomplished using whole-cell, *in vivo* recording techniques in NCM and CM of starlings during presentation of natural conspecific vocalizations, i.e. songs. We characterized variability across time and across trials for both subthreshold and spiking activity. In concert with the observation of very selective (sparse) stimulus-driven spiking responses, we find strong, remarkably persistent, stimulus specificity in the subthreshold response to all songs regardless of spiking. These results provide strong evidence against sparse synaptic pooling and support instead a distributed synaptic pooling regime in which, in sharp contrast to spiking output, net subthreshold responses are stimulus-specific and non-selective.

In a distributed pooling regime, stimulus-specificity throughout the subthreshold response depends on the trial-to-trial reliability of the inputs. Inhibition is a widespread feature of cortical networks (Wilent & Contreras, 2005; Ayaz & Chance, 2009; Isaacson & Scanziani, 2011) and in Starling NCM, local inhibition modulates the selectivity of spiking responses (Thompson *et al.*, 2013). We test whether inhibition is necessary for maintaining stimulus specificity in the subthreshold response throughout stimulation. If inhibition modulates the reliability (and by extension the specificity) of the synaptic input to individual neurons then it would necessarily modulate the reliability of the network spiking response, and the subthreshold and spiking responses of individual neurons. We find that transient local blockade of gamma-aminobutyric acid (GABA) receptors decreases the reliability of stimulus-driven spiking responses in NCM, consistent with a role of inhibition for supporting a non-selective, stimulus-specific distributed synaptic pooling regime underlying selective spiking in NCM.

Methods

Animal preparation

Experiments used adult European starlings (*Sturnus vulgaris*), wild-caught in southern California. We prepared the starlings for the recording session by attaching a small pin stereotaxically to the surface of the skull with dental cement (under isoflurane anesthesia). For electrophysiological recordings, we anesthetized the starlings with 20% urethane (7–8 mL/kg, in three to four intramuscular injections over ~2 h) prior to being placed in the recording chamber or with a continuous infusion of 1.3% ketamine in 5% glucose saline at 2 mL/kg/h throughout the recording. Starlings were placed in a cloth jacket and secured via the attached pin to a stereotaxic apparatus inside a sound attenuation chamber. A small craniotomy was made dorsal to the recording site (NCM: 0–300 rostral of Y sinus and 500 lateral of midline; CM: 2500 caudal of Y sinus and 500–1500 lateral of midline), the dura removed and electrodes advanced into the brain.

Ethical standards

All procedures were conducted in accordance with approved IACUC protocols and in accordance with the guidelines laid down by the NIH in the US regarding the care and use of animals for experimental procedures.

Electrophysiology

Whole cell patch current clamp recordings (MultiClamp 700B amplifier; Axon Instruments, Union City, CA, USA) of 5 to 60-min duration were made using the blind patch technique (Margrie *et al.*, 2002). Whole-cell patch pipettes (3 to 6-MΩ tip resistance) were fabricated from filament (0.25 mm) borosilicate glass (OD 2.0 mm, ID 1.5 mm; Hilgenberg, Malsfeld, Germany). The standard K^+-based internal solution was: (in mM) potassium gluconate 135, NaCl 8, HEPES 10, Mg-ATP 4, Na-GTP 0.3, EGTA 0.3 (pH 7.4, 298 osm). Recordings were obtained by slowly advancing the electrode through the region of interest (about 1500–2000 μm below the surface) while monitoring its resistance with voltage steps. During the initial descent through the hyperpallium, a large amount of positive pressure (~300 mbar) was applied to keep the electrode tip free from debris. After arriving at the depth of interest, the positive pressure was reduced to 25–35 mbar, capacitance compensation was adjusted, and the pipette was advanced in 2- to 3-μm steps until direct contact with the cell membrane was detected as an increase in resistance. Immediately upon contact, pressure was released and formation of the giga-seal between the electrode tip and the cell membrane occurred either spontaneously or after slight suction applied by mouth. After the giga-seal stabilized (typically within a few minutes) suction was used to obtain whole cell access (access resistances ranging from 5 to 90 MΩ; see 'Intrinsic physiology').

Extracellular recordings

To examine the role of local inhibition in controlling spiking reliability in the network, we used data from extracellular recordings in NCM collected as a part of previous experiments (Thompson *et al.*, 2013). Briefly, commercial multibarreled glass pipettes containing a carbon fiber electrode (5 μm diameter; 400–1200 kΩ impedance) and six attached barrels (~3 μm diameter) were used for drug microiontophoresis (Kation Scientific, Minneapolis, MN, USA). Gabazine (SR95531, 3 mM, pH 3.2; Sigma Aldrich, St Louis, MO,

USA), or a gabazine/saclofen combination, was used to inhibit GABA-mediated inhibition locally in NCM around the recording site. The combined application of gabazine and saclofen (n = 12 sites) did not elicit responses different from those during application of gabazine alone (Thompson *et al.*, 2013). Following Thompson *et al.* (2013) we considered drug delivery at a particular site to be successful if the average song-driven firing rate during iontophoresis was significantly different (either higher or lower) from the firing rate prior to iontophoresis.

Auditory stimulation

All stimuli were extracted from previously recorded song repertoires of adult European starlings. Single motifs (stereotyped multi-note elements of natural Starling song) played two or three times in succession, or longer segments (5–10 s) of continuous song were played to the anesthetized animal in an anechoic recording chamber. Auditory stimuli were presented free field from a full-range speaker mounted 30 cm from the center point of the subject's head, where the mean sound pressure level ranged from 40 to 80 dB SPL.

Data analysis

Electrical activity recorded in whole-cell configuration was low-pass filtered (10 kHz), digitally sampled (44.1 kHz), and saved for offline analysis (IGOR PRO2; WaveMetrics Software, Lake Oswego, OR, USA). For further analysis, data were down-sampled to 10 kHz and exported to a format used by custom-written MATLAB (Mathworks Software) routines. Only stimulus–response pairs for which there were at least five repeats were included in the analyses of the whole-cell data.

Intrinsic physiology

We present data recorded in whole-cell configuration from 20 single neurons in NCM and CM. As this is the first report using whole-cell *in vivo* recordings in CM and NCM of the starling, we include here a basic characterization of intrinsic properties for reference/comparison with other brain regions and species. To estimate the passive input resistance (R_{in}) and time constant (T) of the membrane, we applied a negative current pulse (−75 pA) through the recording pipette and fit the voltage response with a double exponential function to isolate the electrode artifact from the membrane response. To avoid contamination by slow, voltage-activated conductance in our estimate of the passive membrane properties (R_{in} and T), we fit only the first 100 ms of the voltage response. The median series resistance (R_s) across all 20 neurons used in this study was 32 MΩ (CI: 5–64 MΩ), which allowed sufficient isolation of the membrane response. The median R_{in} was 157 MΩ (CI: 56–309 MΩ) and the median T was 14 ms (CI: 7 and 21 ms). Independence of estimates for R_{in} and T from R_s was confirmed with a simple linear regression model.

The negative current step also regularly induced slowly depolarizing voltage sags. We estimated the time course and magnitude of this effect by removing the passive components of the voltage response (by subtracting the exponentials used to estimate R_{in} and R_s), and then re-fitting the remaining voltage response with a single exponential (Zhu *et al.*, 1999). In all cases there was a significant exponential fit with a median time constant of 184 ms (CI: 110–364 ms) that was depolarizing in 19/20 cases (across which the steady-state membrane voltages ranged from −96 to −64 mV). In a subset of neurons recorded for another experiment, we measured the dependence of the input resistance on the membrane potential and found a positive correlation. This has not been well characterized, but is consistent with hyperpolarization-induced activation of a conductance.

In 14 of 20 neurons, the series resistance was sufficiently low (< 50 MΩ) to analyse the temporal dynamics of spike shape. For these neurons, the median width at half-height was 1.0 ms (CI: 0.5–1.7 ms) and the median spike threshold was 25 mV above the resting membrane potential (CI: 19–30 mV). R_s did not correlate with spike height.

We measured a limited number of properties but the data did not suggest any clear distinction into cell types. The reported analyses therefore pool across all recorded cells in two brain regions. Sampling biases inherent in whole-cell patch techniques (relating to factors such as soma size, morphology and myelination) are probably present in our dataset; other cell types not recorded could, of course, show very different responses. Notwithstanding such issues, the consistency of our main results of the study across a potentially diverse set of cell types highlights the robustness of the effects.

Signal filtering

To detect spike times from intracellular records, the signal was high-pass filtered at 200 Hz and thresholded. For extracellularly recorded data using carbon-fiber electrodes, spike times were detected as reported by Thompson *et al.* (2013). Measures of subthreshold activity were made after the voltage record was smoothed using a one-dimensional, 18-ms median filter to suppress noise and clip spikes near threshold.

Response epoch detection

One important component of our experiment is the use of natural stimuli. As such, we did not want to arbitrarily tailor the duration of our stimulus to a duration that was experimentally convenient to analyse a 'response', as is common when using artificial stimuli. Instead, under the assumption that the activity across the entire duration of a complex stimulus does not necessarily constitute a meaningful response, we presented long bouts of song within which we defined multiple shorter responses. To do this, we separately parsed the ongoing spiking and subthreshold activity using a simple algorithm that identified localized responses (relative to the pre-stimulus trial-averaged activity), which we call 'response epochs'. To detect significant subthreshold response epochs we took the median-filtered voltage response to a song stimulus and overlaid a sliding 200-sample (20 ms)-wide analysis window broken into 20 bins. At each time-step the activity within that window was considered a response if 85% of the bins had values that exceeded the confidence bounds set by the trial-averaged pre-stimulus activity (Fig. 1A). We detected significant spiking response epochs in the same way, except that we first created a smoothed spiking probability function by converting the vector of raw spike times to a binary string where '1' indicated a spike and '0' no spike, convolving with a narrow Gaussian filter, averaging across trials, and then normalizing to max = 1. To parse spiking activity we used a 500-sample (50 ms)-wide analysis window broken into 50 bins and 95% confidence intervals on the pre-stimulus activity. To maximize the number of independent responses and minimize the duration of responses, it was necessary to use slightly different values for the window size and confidence bounds in the analysis of spiking and subthreshold activity. Although the statistics of the subthreshold activity vector necessitated a smaller analysis window than for spiking, the minimum

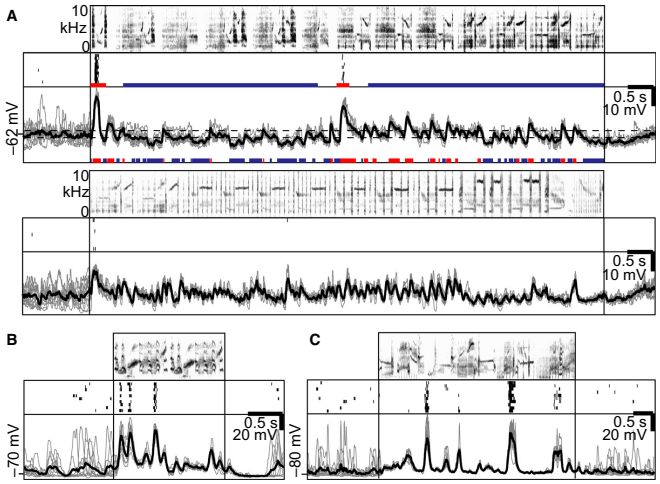

FIG. 1. Examples of spiking and subthreshold activity recorded in whole-cell configuration. (A) Neural recording from an example neuron before, during and after the presentation of two different 10-s starling song segments sampled from a longer bout. Top, spectrogram of the song. Middle, raster of spike times for each of 10 trials in which the song sample above was presented. Bottom, isolated subthreshold membrane potential recordings (gray) for the same neuron on the same 10 trials as above, overlaid with the trial-averaged voltage in black. Colored bars below spiking and subthreshold activity denote facilitatory (red) and suppressive (blue) response epochs (see Methods); mean trial-averaged pre-stimulus voltage = −62 mV. Dotted lines mark the 98% confidence bounds on the pre-stimulus range of trial-averaged subthreshold activity. (B,C) Example spiking and subthreshold activity recorded from two different neurons across trials during which song segments sampled from longer bouts were played. Neural activity and song are displayed as described in (A).

response duration was smaller than the analysis window for both spiking and subthreshold responses (62 and 22 ms, respectively) indicating that the width of the analysis window did not set the lower bound on distributions of response epoch durations in either case and that comparisons can be made between the results obtained in both activity regimes.

Stimulus specificity of subthreshold potential

We recorded the membrane potential throughout the presentation of each song stimulus across multiple trials (5–10 trials) and used a median filter to clip spikes. To assess the stimulus specificity of subthreshold activity (whether the variance in the membrane potential within a given portion of song across trials is smaller than the variance between different portions) we used k-means clustering. Offline, we segmented the full song and the corresponding time-locked subthreshold activity into evenly spaced bins of a given duration (durations used for the analysis ranged from 0.09 ms to 0.9 s). We refer to each binned song segment as a 'stimulus' for our analysis. We applied k-means clustering to a random subset of five bins and sorted all subthreshold activity (which we refer to as 'response' for our analysis) within that set of bins based on similarity in (i) the time-averaged membrane potential and (ii) the temporal pattern of membrane voltages. To measure clustering accuracy we calculated the proportion of correct response–stimulus assignments contained in the k-means result for every possible bin–cluster permutation and took the maximum proportion correct as the clustering accuracy. We iterated this process 100 times and calculated the average clustering accuracy for each cell–stimulus pair and each segment duration. The noise floor for the clustering accuracy was calculated using the subthreshold activity recorded during silence before song onset.

Fano factor analysis

We adapted routines for analysing mean spike rate and Fano factor on the extracellular data set from the 'variance toolbox' available from the Churchland lab and used in a recent report (Churchland *et al.*, 2010). The mean-matched Fano factor was computed for all stimulus conditions across the set of extracellular sites using a 50-ms sliding analysis window (in 25-ms steps) to provide a time-varying estimate of the reliability of the spike rate throughout stimulation. Mean matching equalizes the firing rate distributions across time to control for the dependence of Fano factor on firing rate.

Statistical analysis

All data were tested for normality using the Kolmogorov–Smirnov test evaluated at $P < 0.05$. When appropriate, central tendencies are reported as median ± the 95% confidence interval calculated from the cumulative distribution function unless otherwise stated. Non-parametric tests were used when data were not normal.

Results

In this study we distinguish two hypothetical synaptic pooling strategies that could both support hierarchical object selectivity, but that make distinct predictions about the selectivity of stimulus-specific subthreshold activity underlying selective spiking in single neurons in high-level auditory cortex. To examine spiking and subthreshold responses, we recorded neural activity using whole-cell patch clamp techniques in 20 single neurons in regions NCM ($n = 12$) and CM ($n = 8$) of anesthetized starlings presented with a range of conspecific songs. These songs have a wide range of natural variation in

the distribution of spectro-temporal features across time. This time-varying acoustic structure is reflected in the high temporal variability of both the spike rate and the subthreshold activity (Fig. 1). Although both the spiking and the subthreshold response vary considerably across the duration of the stimulus, as shown in the example neurons (Fig. 1), they are nonetheless quite reliable for each repetition of the same stimulus. In characterizing both the within-trial variability and the between-trial reliability over multiple timescales, we use the term 'time-averaged' to refer to the mean activity averaged over the duration of a single stimulus presentation, and the term 'trial-averaged' to refer to the time-varying activity averaged across multiple presentations of the same stimulus.

Variability in spiking and subthreshold activity across time

We first quantified the within-trial variability of both the spiking and the subthreshold activity in individual neurons. For spiking activity, we isolated epochs during stimulation for which the trial-averaged spike probability density function exceeded the trial-averaged pre-stimulus range (see Methods). We sorted these epochs into facilitating or suppressing responses depending on whether the time-averaged spike rate increased or decreased relative to the pre-stimulus period (Figs 1A and 2A; see Methods). The duration of these spiking responses varied (median 217 ms, CI = 62–1122, n = 174 facilitatory responses; median 386 ms, CI = 75–2854, n = 116 suppressive responses) as did their time-averaged spike rate (median 11.1 spikes/s, CI = 0.4–33.4, n = 174 facilitatory responses; median 0 spikes/s, CI = 0–0, n = 116 suppressive responses). On average, the facilitatory responses constituted 8% of the total stimulus duration (CI = 0–55%; n = 116 stimuli), and suppressive responses constituted 9% of the stimulus duration (CI = 0–82%; n = 116 stimuli). We note that the low spontaneous spike rates observed (median 1.0 spike/s, CI = 0.3–3.3, n = 20 neurons) are common in these regions (Gentner & Margoliash, 2003; Keller & Hahnloser, 2009; Schneider & Woolley, 2013) and can make suppression of spiking difficult to measure. Thus, the actual number of suppressive responses may be greater.

For subthreshold activity, we isolated response epochs during song presentation in which the trial-averaged voltage exceeded the trial-averaged pre-stimulus range (see Methods). We sorted these epochs into facilitating or suppressive responses based on whether the time-averaged voltage within each epoch was depolarized or hyperpolarized relative to the pre-stimulus mean (Fig. 1A; see Methods). These subthreshold responses varied in duration (median 83 ms, CI = 26–449, n = 638 facilitatory responses; median 61 ms, CI = 22–302, n = 332 suppressive responses) and the time-averaged polarization (relative to the minimum potential recorded during silence; median +11 mV, CI = +7 to +19, n = 638 facilitatory responses; median +2.8 mV, CI = +0.9 to +4.1, n = 332 suppressive responses). In a few cases where we calculated the resting membrane potential in voltage clamp (data not shown), we found that the minimum membrane potential recorded during silence approximated it. Across all neurons in our sample, the facilitatory subthreshold responses constituted on average 30% of the total stimulus duration (CI = 2–80%; n = 116 stimuli), which was a significantly larger proportion than the 8% of the stimulus that contained facilitatory spiking responses (see above; $P < 0.0001$ Kolmogorov–Smirnov test, n = 116 stimuli).

Spiking always co-occurs with subthreshold depolarization (facilitatory response), and facilitatory spiking responses are by definition a subset of facilitatory subthreshold responses. Therefore, to estimate the amount of facilitatory subthreshold responses that did not co-occur with spiking responses we take the difference between the total proportion of the stimulus in which there is facilitatory subthreshold response (30%) and subtract the total proportion in which there is a facilitatory spiking response (8%). Thus, approximately 22% of facilitatory subthreshold responses occur in the absence of spiking. Based on membrane voltage alone, only a small fraction

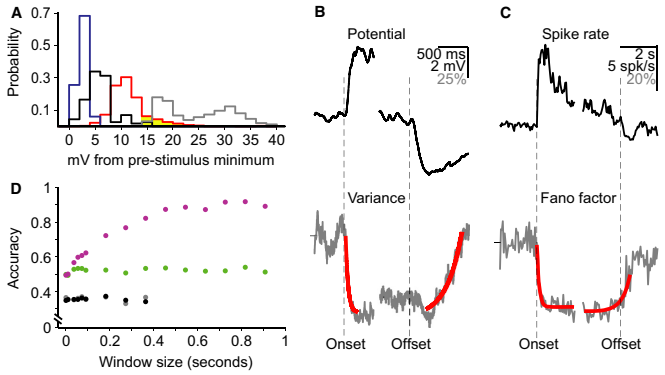

FIG. 2. Stimulus-driven response reliability and specificity. (A) Histograms of mean subthreshold activity level preceding each stimulus (black), the threshold potential for all spikes recorded from all neurons (gray), and membrane potential values contained within all facilitative (red) and suppressive (blue) response epochs. (B) The mean trial-averaged, filtered membrane potential (top) and the median trial-to-trial variance relative to pre-stimulus silence (bottom; hash = median variance before stimulus onset) across all neurons and stimulus blocks. (C) The mean spike rate (top) and the mean-matched Fano factor relative to pre-stimulus silence (bottom; hash = median Fano factor before stimulus onset) computed for all stimulus conditions across the set of extracellular sites using a 50-ms sliding window. Data in (B) and (C) are aligned to stimulus onset and offset (stippled lines). (D) Mean accuracy (from k-means clustering) with which any given segment of the membrane response can be correctly distinguished from any other segment of similar duration, for segments ranging from 0.09 ms to 0.9 s. Different colored points show the classification accuracy for the membrane voltage time series (magenta), the time-averaged membrane potential (green) during stimulus presentations, along with the same measures made during the silent interval preceding stimulus onset (gray, mean; black, time-series) for each duration of analysis time-widow.

(~17%) of the membrane potential values contained within facilitatory subthreshold responses exceeded the minimum spike threshold (16 mV above the pre-stimulus minimum voltage; Fig. 2A), and thus had the potential to elicit a spike. The actual distribution of spikes within this subset of voltages will be lower, as it depends on the refractory period and the local voltage derivative. Regardless of the exact spiking distribution, robust spiking responses are much more selective than the facilitatory subthreshold activity.

Suppressive subthreshold responses were more rare than facilitatory subthreshold responses (2% stimulus duration, CI = 0–42%; n = 116 stimuli), and more rare than suppressive spiking (9% stimulus duration, CI = 0–82%; n = 116 stimuli). Of course, spiking activity can be suppressed by subthreshold responses with a range of mean voltages anywhere below spike threshold.

Reproducibility of spiking and subthreshold activity

The two synaptic pooling strategies we address in this study are distinguished in part by the extent of stimulus control over the synaptic activity. Across-trial variability in the subthreshold and spiking activity is a reliable metric of stimulus control that we can easily measure throughout stimulation. The foregoing analyses (which showed that both spiking and subthreshold responses span a range of magnitudes and durations; Figs 1 and 2A) imply that the activity within response epochs was reproducible across trials. Reliability is not necessarily constrained to those epochs and may be measureable on a smaller timescale. We therefore measured reliability of subthreshold and spiking activity throughout stimulation by calculating the trial-to-trial variance for the subthreshold activity (single sample resolution) and the Fano factor for spiking (25-ms resolution) and characterized their temporal dynamics at stimulus onset and offset.

We observed a significant decrease in the trial-to-trial variance of the membrane potential during stimulation compared with the pre-stimulus period (Fig. 2B, bottom; $P < 0.0001$ Kolmogorov–Smirnov test; median decrease relative to baseline = 68%, CI = 84–25, n = 115 cell–stimulus pairs). The time course for the onset of this drop in the membrane potential variance was very rapid (Tau = 41 ms), but the relaxation back to the pre-stimulus levels at the offset of the stimulus was much slower (Tau = 262 ms) (Fig. 2B, bottom). The suppressed variance relative to baseline persisted throughout song stimulation, but for individual response epochs the variance depended on the mean as follows. For facilitating responses (depolarized potentials relative to pre-stimulus mean) the variance of the membrane potential decreased by 85% relative to baseline (n = 20 cells). For suppressing responses (hyperpolarized potentials relative to the pre-stimulus mean) the variance of the membrane potential decreased by 99.8% relative to baseline (n = 20 cells). We note that this difference is likely to be influenced by several factors, including voltage-dependent non-linearities that increase the spiking probability for depolarized responses, and differences in the number of synaptic events contributing to each response.

We quantified trial-to-trial variability in the spiking activity before, during, and after stimulus presentation using the Fano factor with mean-matching techniques to control for firing rate dependencies (Fig. 2C; Churchland *et al.*, 2010). This analysis is data-intensive and requires more trials than were recorded under whole-cell configuration and more neurons in order to correct for firing rate biases. To measure the between-trial reproducibility of spiking activity, we used extracellular multi-unit activity recorded in NCM to song stimuli (five adult starlings, 42 multi-unit sites; 336 site–stimulus pairs; Thompson *et al.*, 2013). At stimulus onset there was a rapid, significant decrease in the Fano factor (Tau = ~119 ms; single exponential fit), to 70% of baseline (CI = 68–81, $P < 0.0001$ Kolmogorov–Smirnov test, n = 336 stimuli). At stimulus offset, the Fano factor estimate relaxed back toward the pre-stimulus level, but with a much slower time course than the onset (Tau = ~412 ms; Fig. 2C bottom).

Stimulus specificity of subthreshold fluctuations across multiple timescales

The reproducibility of subthreshold responses suggests that neurons are under continuous synaptic control throughout the stimulus, rather than intermittently receiving input corresponding only to features most relevant to changes in output spike rates. By itself, however, trial-to-trial reliability does not confer stimulus specificity. Spiking responses in CM and NCM are both selective (show lifetime sparseness) and stimulus-specific (respond distinctly to different stimuli) (Jeanne *et al.*, 2011). Having demonstrated that the subthreshold activity is reliable, we next determined whether it is also stimulus specific throughout song presentation. Specifically, does the same portion of a song produce a unique membrane voltage response (temporal pattern or mean value) that is similar across stimulus repetitions?

To answer this question in our whole-cell data set, we applied k-means analysis to calculate an average clustering accuracy across response–stimulus pairs (see Methods: 'Stimulus specificity of subthreshold potential'). We then took this accuracy as a measure of subthreshold stimulus specificity for each cell–stimulus pair (Fig. 2D). We measured stimulus specificity using a range of stimulus durations, but always included data recorded during the entire song bout in the analysis. The noise floor for the accuracy estimate was calculated using the activity recorded during the silence before the stimulus came on. Both the mean and temporal pattern of membrane potential activity contained enough stimulus specificity to allow for accurate clustering of response–stimulus pairs (compared with the 'noise' floor) even at single-sample resolution. Beginning at relatively short timescales (~40 ms), the temporal pattern of subthreshold activity could be clustered more accurately than the mean (Fig. 2D). Thus, even very short, randomly-chosen portions of the membrane activity carry stimulus-specific information. Increasing the size of the analysis window substantially improved clustering accuracy for temporal patterns, but not for the mean, demonstrating that the time-varying membrane potential carries additional stimulus-specific information.

Inhibition modulates across-trial reproducibility of spiking in NCM

We reasoned that local inhibition might also mediate the robustness of stimulus specificity in the subthreshold activity, which depends on the reliability of spiking activity in the network from which its inputs are pooled. To test this idea, we blocked GABA receptors transiently around extracellular recording sites using iontophoretic application of gabazine (see Methods).

Local inhibition modulates the magnitude of neuronal spiking in NCM across song stimuli (Thompson *et al.*, 2013), but its effect on individuated response epochs was not previously tested. We parsed spiking response epochs (as in Fig. 1A) using the activity recorded in the gabazine condition and compared those response epochs with the corresponding stimulus–response epochs in the intact condition. Blocking local inhibition induced an increase in spiking activity during silence (median = 1.5 spikes/s, CI = 0–15.0, intact condition;

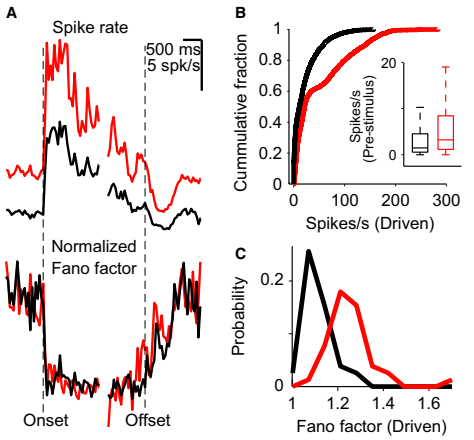

FIG. 3. Effects of gabazine. (A) The spike rate (top) and the mean-matched Fano factor (bottom) during the intact condition (black) and during gabazine iontophoresis (red), computed for all stimulus conditions across the set of extracellular sites using a 50-ms sliding analysis window (in 25-ms steps). For visualization of the relative onset and offset time course, the Fano factor trace in each condition is aligned to the median stimulus-evoked Fano factor and normalized to the maximum Fano factor before stimulus onset. (B) Cumulative distribution functions of the spike rate per response epoch (colors as in A). Inset: boxplot of spike rates during the pre-stimulus period in each condition (horizontal line shows the median; the box shows the 25th and 75th percentiles; the whiskers encompass all non-outlier data: ~99.3%). (C) Histogram of Fano factor estimates for stimulus-driven activity (colors as in A).

median = 3.2 spikes/s, CI = 0.2–24.8, gabazine condition; n = 296 stimuli) and a shift toward response epochs with higher spike rates during auditory stimulation (median = 11.6 spikes/s, CI = 0–71.6, intact condition; median = 21.5 spikes/s, CI = 1.8–162.2, gabazine condition; n = 1815 responses) consistent with previously reported effects (Thompson *et al.*, 2013; Fig. 3A and B).

The extracellularly-recorded spiking activity (from single and multi-unit sites) can serve as a reasonable proxy for estimating the reliability of the local synaptic network driving a randomly chosen cell within that network. To test whether local inhibition might contribute to the reliability of the local network activity, we compared the Fano factor computed across stimulus trials (see Methods) when local inhibition was intact and when it was blocked by gabazine. Indeed, the distribution of stimulus-driven Fano factor estimates is shifted significantly to larger values when local inhibition is blocked by gabazine than when inhibition is intact (median = 1.13 intact and 1.27 gabazine; $P < 0.0001$, Kolmogorov–Smirnov test; Fig. 3C). Notably, the Fano factor distribution across sites during the pre-stimulus period was not significantly altered by gabazine (median = 1.52 intact and 1.53 gabazine; $P = 0.7$, Kolmogorov–Smirnov test). Likewise, the time course of the change in variance at stimulus onset and offset appeared to be unaffected by blocking local inhibition with the application of gabazine (Tau_onset = 119 ms and Tau_offset = 412 ms with inhibition intact; Tau_onset = 115 ms and Tau_offset = 358 ms with inhibition blocked; Fig. 3A, bottom). Together, these results show that local inhibition mediates a stimulus-driven increase in reproducibility of spiking activity.

Discussion

We recorded the spiking and subthreshold activity of neurons in the high-level auditory regions CM and NCM, in the starling forebrain, in response to spectro-temporally diverse natural songs. We find that song stimuli drive time-varying subthreshold membrane voltage responses that are reliable across trials and stimulus-specific, but non-selective. These results are inconsistent with a model in which neurons with highly selective spiking responses pool over sets of inputs that only drive synaptic activity in service of the selective spiking response. This indicates that, at lease these and possibly other high-level sensory neurons are sampling from a much more rich and diverse stimulus space than is evinced by their output spiking. When local inhibition is blocked, spiking variability across trials increases in NCM, suggesting that inhibition plays a key role in governing the stimulus specificity of the time-varying membrane voltage activity generated in a distributed pooling regime.

Understanding how stimulus encoding supports the adaptive interaction between animals and their environment requires studying how natural stimuli are represented in the spiking activity of individual neurons and their populations. The relationship between spiking activity in single neurons and a sensory stimulus depends on both the properties of a potential set of inputs and the pooling operation across those inputs in the post-synaptic cell. In general, the complexity of stimulus selectivity increases along the sensory processing pathway with sparse-firing neurons in 'higher' regions selectively spiking in response to progressively more complex features (Sen *et al.*, 2001; Hsu *et al.*, 2004; Meliza *et al.*, 2010; Marshel *et al.*, 2011; DiCarlo *et al.*, 2012). This organization implies a functional hierarchy and presumably confers benefits for the classification and recognition of complex signals (Babadi & Sompolinsky, 2014). To explain higher-order feature selectivity, classic models first developed for primary visual cortex (Hubel & Wiesel, 1962; Movshon *et al.*, 1978) relied on the combination of inputs selective for simpler component features. This kind of feed-forward pooling forms the basis for more contemporary models of object selectivity (Riesenhuber & Poggio, 1999; Rauschecker & Scott, 2009; Lien & Scanziani, 2013), where selective convergence of inputs at each level of a hierarchical network gives rise to narrowed stimulus tuning and more complex receptive field structure. Functionally, these models capture the transformation from dense low-level feature representations to sparse encoding of high-dimensional objects, but the synaptic and network mechanisms underlying this (or other hierarchical models) is not well understood – especially in high-order neurons driven by continuous stimuli comprising diverse sets of features over long periods of time.

Natural signals are ongoing, temporally variable and diverse. We observe reliable spiking activity in response to natural song in regions CM and NCM of the starling cortex, demonstrating that the neurons are clearly auditory. In concert with the temporal variability of the signal, however, spiking is not evenly distributed throughout stimulation. In our own data, the distribution of the stimulus-evoked spike rates averaged over an entire song stimulus is not much different from the mean spontaneous spike rate (stimulus evoked: 1.0 spikes/s, CI = 0.4–5.9 spikes/s; spontaneous: 1.0 spikes/s, CI = 0.3–3.3 spikes/s), demonstrating that the mean spike rate over the song is indeed a poor measure of the cell's temporally sparse response (Fig. 1). This raises the interesting question of what inputs a cell might receive during periods of the stimulus, when there is no stimulus-specific spiking response. Drawing on the majority of physiology experiments, in which the stimulus has been abstracted from the world and arbitrarily tailored to the response, one might reason that

the inputs revert to a spontaneous state. Indeed, statements in the literature that explicitly or implicitly assume sparsely spiking neurons are not carrying out any computations for the majority of the time are not hard to find (Abeles *et al.*, 1990; Graham & Field, 2007a). But these assumptions have not been tested.

Our results address two potential synaptic pooling regimes, referred to as 'sparse' and 'distributed' pooling, that could underlie sparse, object-selective spiking to natural stimuli. Both pooling regimes can permit selective spiking output, but are distinguished by the selectivity and stimulus-specificity of subthreshold activity throughout long bouts of ongoing natural stimuli. Under sparse pooling, neurons pool inputs that are biased towards the stimulus features that ultimately drive changes in spike rates. For our analyses, this predicts that stimulus-specific activity should be largely restricted, on average, to the response epochs that inform spiking. Conversely, the subthreshold activity that is uninformative of stimulus-specific spiking responses – namely the subthreshold activity not designated as response epochs – should not be stimulus-specific. Our results do not support these predictions. Instead, we find that subthreshold activity is remarkably reliable and stimulus-specific throughout every song presented, regardless of whether there is a stimulus-driven change in the probability of spiking (Figs 1 and 2). This nearly continuous, stimulus-specific control over the subthreshold response rules out a sparse pooling regime.

Instead, the pattern of results favors a distributed pooling regime, in which neurons with selective spiking responses pool synaptic inputs that collectively produce a non-selective, but stimulus-specific, subthreshold response. Importantly, our results cannot address whether the individual pre-synaptic neurons themselves are selectively or non-selectively tuned, but two different scenarios seem most plausible to account for the observed ongoing stimulus-driven reliability and temporal specificity. One possibility is that individual CM and NCM neurons pool inputs from neurons tuned to low-dimensional features that have a relatively high probability of occurring at many points throughout song. These inputs could come from Field L, where some neurons are well-described as frequency-tuned. Their spiking responses are correspondingly non-selective throughout vocalizations and could provide continuous drive to a post-synaptic neuron because power in a given frequency channel is distributed broadly throughout the song. This scenario implies that the populations over which a given neuron is pooling have low population sparseness. Although sparseness appears to be a common feature of sensory encoding (Vinje & Gallant, 2000, 2002; Olshausen & Field, 2004; Graham & Field, 2007b; Hromadka *et al.*, 2008), it may not be the only representational scheme in place (Sakata & Harris, 2009; Tolhurst *et al.*, 2009). A second possibility is that individual NCM and CM neurons pool inputs from neurons tuned to high-dimensional features that occur with relatively low probability throughout the song. In this case inputs might come from other neurons within CM, NCM and other auditory regions that exhibit high population sparseness and whose neurons have high lifetime sparseness. Dense recurrence among (or within) these regions could allow for pooling over much larger numbers of these selective inputs and provide continuous drive to a post-synaptic neuron. In both scenarios, sparse post-synaptic spiking could emerge through the covariance of a specific set of inputs regardless of their tuning. Differentiating between these two possibilities requires a better understanding of the diversity of features to which the inputs are tuned and the combinatorial operations that govern their integration. In either case, our data are consistent with a model in which sparse spiking responses emerge through distributed pooling, which places demands on post-synaptic computational mechanisms to generate sparse spiking responses. These mechanisms, and their relationship to non-linearities imposed by the spiking threshold, will be important to investigate in future work.

Maintaining synaptic inputs for features that do not co-vary directly with selective spiking output seems somewhat counterintuitive, but it may confer computational advantages for behavior. We know that responses to elements of natural auditory signals are highly dependent on the information conveyed about behavior by particular acoustic material, so the stimulus–response relationship is heavily modulated by behavioral experience (Blake *et al.*, 2002, 2006; Gentner & Margoliash, 2003; Thompson *et al.*, 2010, 2013; Jeanne *et al.*, 2011, 2013; Meliza & Margoliash, 2012). Distributed pooling may confer individual neurons with flexible encoding across a wide diversity of stimulus features, allowing behavioral feedback to shape the pluripotency of the same inputs through synaptic plasticity, adaptation or other mechanisms (Kozlov & Gentner, 2014). Thus, a given neuron may 'represent' multiple objects based on the dynamic functional organization of the system (Kozlov *et al.*, 2013).

The subthreshold activity of all neurons in our sample is consistent with a distributed synaptic pooling architecture. This, in turn, implies that most (if not all) the neurons in these regions are highly interconnected, potentially causing correlations that could lead to deleterious effects on the encoding/decoding of spiking responses across cells (Cohen & Maunsell, 2009; Cohen & Kohn, 2011). Theoretical work has shown, however, that recurrent connectivity in cortical networks can actually decorrelate the activity patterns of neurons with shared presynaptic input (Helias *et al.*, 2014), specifically though inhibitory feedback (Tetzlaff *et al.*, 2012; Bernacchia & Wang, 2013). Correlated firing can also modulate the gain of postsynaptic cells (Salinas & Sejnowski, 2000). We find a major role for inhibition in shaping both the trial-to-trial reproducibility of the post-synaptic response to ongoing natural stimuli, and the magnitude of spike rates – effects that inhibition could manifest by altering the correlation structure of the network. The gain of responses could also be modulated directly by feed-forward inhibition (Mejias *et al.*, 2014). Input–output mappings even in single neurons are not static over time or behavioral conditions (Kozlov & Gentner, 2014), and our results demonstrate that inhibition is poised to provide flexible control over these response characteristics under natural conditions. To understand how inhibition is modulating the sensitivity of the input–output function in neurons of the Starling auditory cortex it will be useful to develop more precise ways to isolate and manipulate different sources of inhibition in the network.

The results of the current study expand our understanding of the synaptic and network mechanisms that underlie hierarchical selective representations of complex natural communication signals in the auditory system. The implications of specific types of synaptic convergence on the computations performed by object-selective neurons in sensory cortex have not been well established and they constrain models for the sparse, selective encoding of natural stimuli. Demands on synaptic plasticity to tune the output of individual neurons from broadly selective synaptic input could provide a computational advantage for increasing the amount of information a sensory signal conveys about behavior across different contexts, improving both the flexibility and the efficiency of stimulus encoding. Network organization and synaptic integration both shape the input–output relationship between the stimulus and the neural response in single cells. Knowing more about these operations in the context of naturally occurring stimuli is a critical step to improving on current models of how stimulus selectivity arises in neural networks.

Acknowledgements

We thank the members of the Gentner laboratory for comments on the manuscript. The research was funded by grant R01DC008358 from the National Institutes of Health and the Kavli Institue for Brain and Mind to T.Q.G.; and by the Institute for Neural Computation's NIMH Training Program in Cognitive Neuroscience at UCSD to K.E.P. We have no conflicts of interest to declare.

Abbreviations

CM, caudal mesopallium; GABA, gamma-aminobutyric acid; NCM, caudomedial nidopallium.

References

Abeles, M., Vaadia, E. & Bergman, H. (1990) Firing patterns of single units in the prefrontal cortex and neural network models. *Network*, **1**, 13–25.

Adrian, E.D. (1926) The impulses produced by sensory nerve-endings: Part 4. Impulses from pain receptors. *J. Physiol.*, **62**, 33–51.

Ayaz, A. & Chance, F.S. (2009) Gain modulation of neuronal responses by subtractive and divisive mechanisms of inhibition. *J. Neurophysiol.*, **101**, 958–968.

Babadi, B. & Sompolinsky, H. (2014) Sparseness and expansion in sensory representations. *Neuron*, **83**, 1213–1226.

Bernacchia, A. & Wang, X.J. (2013) Decorrelation by recurrent inhibition in heterogeneous neural circuits. *Neural Comput.*, **25**, 1732–1767.

Blake, D.T., Strata, F., Churchland, A.K. & Merzenich, M.M. (2002) Neural correlates of instrumental learning in primary auditory cortex. *Proc. Natl. Acad. Sci. USA*, **99**, 10114–10119.

Blake, D.T., Heiser, M.A., Caywood, M. & Merzenich, M.M. (2006) Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. *Neuron*, **52**, 371–381.

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S. & Shenoy, K.V. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.*, **13**, 369–378.

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

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

DiCarlo, J.J., Zoccolan, D. & Rust, N.C. (2012) How does the brain solve visual object recognition? *Neuron*, **73**, 415–434.

Ganguli, S. & Sompolinsky, H. (2012) Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Annu. Rev. Neurosci.*, **35**, 485–508.

Gentner, T.Q. (2008) Temporal scales of auditory objects underlying birdsong vocal recognition. *J. Acoust. Soc. Am.*, **124**, 1350–1359.

Gentner, T.Q. & Margoliash, D. (2003) Neuronal populations and single cells representing learned auditory objects. *Nature*, **424**, 669–674.

Graham, D.J. & Field, D.J. (2007a) Sparse coding in the neocortex. In Kaas, J.H. (Ed.), *Evolution of Nervous Systems*, *vol 3*. Academic Press, Oxford, pp. 181–187.

Graham, D.J. & Field, D.J. (2007b) Statistical regularities of art images and natural scenes: spectra, sparseness and nonlinearities. *Spatial Vision*, **21**, 149–164.

Helias, M., Tetzlaff, T. & Diesmann, M. (2014) The correlation structure of local neuronal networks intrinsically results from recurrent dynamics. *PLoS Comput. Biol.*, **10**, e1003428.

Hromadka, T., Deweese, M.R. & Zador, A.M. (2008) Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol.*, **6**, e16.

Hsu, A., Woolley, S.M., Fremouw, T.E. & Theunissen, F.E. (2004) Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *J. Neurosci.*, **24**, 9201–9211.

Hubel, D.H. & Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, **160**, 106–154.

Isaacson, J.S. & Scanziani, M. (2011) How inhibition shapes cortical activity. *Neuron*, **72**, 231–243.

Jeanne, J.M., Thompson, J.V., Sharpee, T.O. & Gentner, T.Q. (2011) Emergence of learned categorical representations within an auditory forebrain circuit. *J. Neurosci.*, **31**, 2595–2606.

Jeanne, J.M., Sharpee, T.O. & Gentner, T.Q. (2013) Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron*, **78**, 352–363.

Keller, G.B. & Hahnloser, R.H. (2009) Neural processing of auditory feedback during vocal practice in a songbird. *Nature*, **457**, 187–190.

Kozlov, A.S. & Gentner, T.Q. (2014) Central auditory neurons display flexible feature recombination functions. *J. Neurophysiol.*, **111**, 1183–1189.

Kozlov, A.S., Briggs, A., Gentner, T.Q. & Sharpee, T.O. (2013) *Mosaic receptive fields of central auditory neurons optimized to the natural sound statistics*. Paper presented at the Society for Neuroscience, San Diego.

Lien, A.D. & Scanziani, M. (2013) Tuned thalamic excitation is amplified by visual cortical circuits. *Nat. Neurosci.*, **16**, 1315–1323.

Margrie, T.W., Brecht, M. & Sakmann, B. (2002) In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflug. Arch.*, **444**, 491–498.

Marshel, J.H., Garrett, M.E., Nauhaus, I. & Callaway, E.M. (2011) Functional specialization of seven mouse visual cortical areas. *Neuron*, **72**, 1040–1054.

Mejias, J.F., Payeur, A., Selin, E., Maler, L. & Longtin, A. (2014) Subtractive, divisive and non-monotonic gain control in feedforward nets linearized by noise and delays. *Front. Comput. Neurosci.*, **8**, 19.

Meliza, C.D. & Margoliash, D. (2012) Emergence of selectivity and tolerance in the avian auditory cortex. *J. Neurosci.*, **32**, 15158–15168.

Meliza, C.D., Chi, Z. & Margoliash, D. (2010) Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. *J. Neurophysiol.*, **103**, 1195–1208.

Movshon, J.A., Thompson, I.D. & Tolhurst, D.J. (1978) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.*, **283**, 53–77.

Olshausen, B.A. & Field, D.J. (2004) Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.*, **14**, 481–487.

Rauschecker, J.P. & Scott, S.K. (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.*, **12**, 718–724.

Riesenhuber, M. & Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nat. Neurosci.*, **2**, 1019–1025.

Sakata, S. & Harris, K.D. (2009) Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. *Neuron*, **64**, 404–418.

Salinas, E. & Sejnowski, T.J. (2000) Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J. Neurosci.*, **20**, 6193–6209.

Schneider, D.M. & Woolley, S.M. (2013) Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron*, **79**, 141–152.

Sen, K., Theunissen, F.E. & Doupe, A.J. (2001) Feature analysis of natural sounds in the songbird auditory forebrain. *J. Neurophysiol.*, **86**, 1445–1458.

Tetzlaff, T., Helias, M., Einevoll, G.T. & Diesmann, M. (2012) Decorrelation of neural-network activity by inhibitory feedback. *PLoS Comput. Biol.*, **8**, e1002596.

Thompson, J.V., Jeanne, J. & Gentner, T.Q. (2010) *Local inhibition shapes the learned responses to song in NCM*. Paper presented at the Soc. Neurosci Absracts.

Thompson, J.V., Jeanne, J.M. & Gentner, T.Q. (2013) Local inhibition modulates learning-dependent song encoding in the songbird auditory cortex. *J. Neurophysiol.*, **109**, 721–733.

Tolhurst, D.J., Smyth, D. & Thompson, I.D. (2009) The sparseness of neuronal responses in ferret primary visual cortex. *J. Neurosci.*, **29**, 2355–2370.

Vinje, W.E. & Gallant, J.L. (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, **287**, 1273–1276.

Vinje, W.E. & Gallant, J.L. (2002) Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *J. Neurosci.*, **22**, 2904–2915.

Wilent, W.B. & Contreras, D. (2005) Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. *Nat. Neurosci.*, **8**, 1364–1370.

Zhu, J.J., Uhlrich, D.J. & Lytton, W.W. (1999) Properties of a hyperpolarization-activated cation current in interneurons in the rat lateral geniculate nucleus. *Neuroscience*, **92**, 445–457.

2.9 Acknowledgements

Chapter 2, in full, is a reprint of materials as it appears in Perks, K., Gentner, T. (2015). “Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex” in *European Journal of Neuroscience*, v. 41, p. 725-733. DOI:10.1111/ejn.12831. The dissertation author was the primary investigator and author of this manuscript.

CHAPTER 3

Natural signals drive fast modulation of stimulus-specific functional networks in cortex.

3.1 Abstract

Relating the collective activities of neural populations to external sensory stimuli or to motor output is essential to understanding how nervous systems support behavior. Equally important is examining the covariance among spiking outputs of many neurons. What we know about the mechanisms of these processes comes primarily from examining the spiking output of randomly selected neurons from within larger populations, and we do not know how those spiking responses are pooled, functionally, by downstream neurons – a mechanism that could significantly alter population coding. In the current study, we find that not only does the spiking activity across the "NCM" region of Starling auditory cortex exhibit temporal and spatial variance, but that a remarkable amount of variance in net spiking response is reflected in the synaptic response of single neurons within NCM. We further leverage the information in the continuous, time-varying synaptic response of individual neurons to refine the selection of functional networks within NCM identified by their ability to predict the synaptic response of a single neuron by implementing linear regression techniques. By examining the stimulus-driven properties of these identified functional networks we determine how spiking responses in NCM are pooled across time and across space to reflect synaptic representations of conspecific vocalizations. The main implication of the collective results presented in this study is that, rather than a model of hierarchical processing in which stimulus-specific information is restricted to parallel circuits within each region, sensory integration and processing are supported by a system in which information about even the most complex stimuli is likely massively redundant and shared among the population at large.

3.2 Introduction

Relating the collective activities of neural populations to external sensory stimuli or to motor output is essential to understanding how nervous systems support behavior. Equally important is examining the covariance among spiking outputs of many neurons. What we know about the mechanisms of these processes comes primarily from examining the spiking output of randomly selected neurons and pairs of neurons from within larger populations. We find that the modulation of spiking covariance can interact with stimulus tuning to bias response discriminability in predictable ways. Yet we do not know how those spiking responses are pooled, functionally, by downstream neurons – a mechanism that could significantly alter population coding if the spiking statistics of the pooled/integrated neurons differ from that of the population at large. In the current study, we aim to address the limitations inherent to examining spiking statistics of randomly selected neural populations by leveraging the synaptic response of an individual neuron embedded among that population to identify functional subsets of units whose co-activity reflects the pre-synaptic input of that single neuron.

Studies from the literature do establish that the synaptic response of individual neurons provides a continuous, time-varying signal containing a breadth of stimulus-specific information (Baudot et al., 2013; Machens, Wehr, & Zador, 2004; Perks & Gentner, 2015). One possibility is that the synaptic input of a single neuron represents only a tiny fraction of the total population in any given region. Insofar as the spiking pattern among the pre-synaptic population for each neuron is unique and independent, predicting the synaptic response from the spiking pattern of a randomly-recorded sample of neurons from the total population would be difficult. In line with this, contemporary

models of sensory processing invoke hierarchical, progressively restrictive feature pooling like that observed in the retina, thalamus, and even V1 to shape progressively complex stimulus-specific receptive fields (Lien & Scanziani, 2013; Rauschecker & Scott, 2009; Riesenhuber & Poggio, 1999). One might imagine that to predict the sub-threshold response of a given neuron, you need to record from a population of neurons in which the inputs to that cell are strongly represented.

On the other hand, if the synaptic responses of different neurons carry redundant information about the stimulus, and this shared information is reflected in the population, then the spiking activity of the total population may actually provide a very good predictor of the synaptic response in any randomly chosen neuron. We do know that, although spiking responses from individual neurons nestled deep within sensory cortex are sparsely stimulus selective, sensory processing must meet the demands of behavior in flexible environments. In order to support flexible spiking output, individual neurons may be driven by multiple subnetworks that are differently engaged by different stimuli and can be modulated independently.

The Starling auditory cortex offers an excellent model for examining these processes and is the target of the current study. Starling songs are composed of many acoustically complex and distinct sound patterns, called motifs, that are perceived as different auditory objects (Gentner, 2008). In higher-order regions of the auditory pathway (the caudal mesopallium, CM, and the caudomedial nidopallium, NCM, in particular) spiking responses are sparse - any given neuron is driven strongly by only a small portion of all possible motifs (Gentner & Margoliash, 2003; Jeanne, Thompson, Sharpee, & Gentner, 2011; Meliza, Chi, & Margoliash, 2010; Thompson, Jeanne, &

Gentner, 2010) and each motif evokes spiking from only a small number of neurons (Gentner & Margoliash, 2003). Pervasive stimulus specificity in the subthreshold membrane potential of individual neurons within NCM implies that the synaptic receptive field of these neurons are extremely broad, pooling from spiking responses corresponding to a diverse range of stimulus information not necessarily represented in the spiking output (Perks & Gentner, 2015). In this paper, we first quantify the stimulus-driven temporal and spatial variance of spiking activity across NCM, which could support the prediction of highly temporally-specific synaptic responses recorded in individual neurons among that population. The population response vector summed across all spiking units was also temporally specific and consistent across trials much like the synaptic response of individual neurons within that region. We find that a remarkable amount of variance in the population activity is reflected in the synaptic response of a single randomly selected neuron from within the region.

We then constrain the population by identifying sets of units selected for their ability to sufficiently predict temporal variance in the synaptic response, and refer to this set as the "*functional network.*" The identity of a functional network was more accurately/reliably identified at the time-scale of single motifs (1 second) than at the timescale of whole vocalization sequences (2-20 seconds), and in response to changes in motif identity, the identity of units contained in the functional network changed - suggesting that synaptic responses may be tuned from a temporally heterogeneous coordination of population spiking activity. We also found that the units comprising each functional network were spatially arranged such that they carpet NCM flatly, yet maintain a bias toward more positive pairwise signal correlations than expected if the

population was sampled at random, suggesting significant stimulus-driven structure in the pooling strategies integrating covarying spiking responses. Each functional network potentially contains highly correlated variables and we conclude by quantifying the dimensionality reduction necessary on each functional network to eliminate redundant information about the synaptic response.

The degree of redundancy is inherently related to key aspects of population coding, such as determining spiking correlations among the population and whether those correlations effect information decoding - aspects that will require additional computational tools to decipher (Latham & Nirenberg, 2005). However, it will be important to examine functionally relevant populations of neurons when asking these questions, which the results of this study directly address. The main implication of the collective results presented in this study is that, rather than a model of hierarchical processing in which stimulus-specific information is restricted to parallel circuits within each region, sensory integration and processing are supported by a system in which information about even the most complex stimuli is likely massively redundant and shared among the population at large.

3.3 Methods

3.3.1 Animal Preparation

Experiments used adult European starlings (Sturnus vulgaris), wild-caught in southern California. We prepare the starlings for the recording session by attaching a small pin stereotaxically to the surface of the skull with dental cement (under isoflurane

anaesthesia). For electrophysiolgical recordings, we anesthetized the starlings with 20% urethane (7–8 ml/kg, in three to four intramuscular injections over ~2 h) prior to being placed in the recording chamber. Starlings were placed in a cloth jacket and secured via the attached pin to a stereotaxic apparatus inside of a sound attenuation chamber. A small craniotomy was made just dorsal to the recording site (NCM), the dura removed, and electrodes advanced into the brain.

3.3.2 Electrophysiology

Recordings of the cross-membrane potential were obtained using standard intracellular sharp electrode techniques under current clamp configuration (MultiClamp 700B amplifier, Axon Instruments). *Sharp* pipettes (70–100-MΩ tip resistance) were fabricated from filament (0.25 mm) borosilicate glass (OD 1 mm, ID 0.5 mm, Hilgenberg, Malsfeld, Germany). The standard electrode fill solution was 3M Potassium Acetate, pH 7.4). Recordings were obtained by slowly advancing the electrode in 2- to 3-μm steps through the region of interest (starting around 1500-2000 μm below the surface) while monitoring its resistance with current steps. Once direct contact with cell membrane was detected as an increase in electrode tip resistance, electrical "buzz" was created with the amplifier's Buzz and (+)Clear functions to break through the cell membrane. A large (~350pA) hyperpolarizing current was used initially to stabilize the recording but subsequently slowly released. In some cases a smaller hyperpolarizing current (<100pA) was used when a cell's feverish spiking activity obscured recovery of reliable subthreshold potential from the recording. Intracellular quality was assessed by spike height and all intracellular data included in these analyses maintained >40mV

spikes throughout the recording. Electrical activity recorded in the intracellular configuration was low-pass filtered (10kHz), digitally sampled (44.1 kHz), and saved offline analysis (IgorPro2; WaveMetrics Software). For further analysis, data was down-sampled to 10kHz and exported to a format used by custom-written MATLAB (Mathworks Software) routines.

To examine the relationship between synaptic input to single cells and spiking responses in larger local neuron populations, we used data from extracellular recordings in NCM. Briefly, commercial 32-channel NeuroNexus probes (32-channel Edge and Poly3) were used to obtain analog voltage signals that were then amplified (5,000 x gain; AM Systems), filtered (high pass, 300 Hz; low pass, 3–5 kHz), sampled (20 kHz) and digitized using Spike2 software and then stored for offline analysis. Using custom routines, the data was then exported to Matlab-compatible files for further processing. For spike sorting, we used open-source automated routines from Klusta team (github.com/kwikteam) to isolate spiking events from the raw voltage waveforms on multiple neighboring channels, cluster these events, and the associate clusters (with varying degrees of certainty) to unique neuronal sources. Manual sorting using the phy package and klustaviewer (phy.cluster.manual) followed the automated sorting process to omit noisy and non-neural event clusters.

3.3.3 Auditory stimulation

All stimuli were extracted from previously recorded song repertoires of adult European starlings. Single motifs (stereotyped multi-note elements of natural starling song; selected at 1 second long each) were concatenated together into continuous 5-

second pseudo-song sequences – with each sequence containing a unique set of motifs. These pseudo-songs were played to the anesthetized animal in an anechoic recording chamber. To target the analysis at "ongoing" auditory processing and avoid analyzing onset responses, the same 1-second-long "intro" motif was appended to the onset of every pseudo song. The response to this intro motif was excluded from all analyses since it was over-represented and identical across sequences. Auditory stimuli were presented free field from a full-range speaker mounted 30 cm from the center point of the subject's head, where the sound pressure level ranged from 40 - 80 dB SPL.

3.3.4 Data Analysis

Extracellular Spike waveforms: Single neurons isolated from extracellularly-recorded dataset were classified as either "regular-spiking" or "fast-spiking" based on the temporal profile of the average spike waveform (aligned to the trough of the biphasic signal). Threshold for fast-spiking classification was <160μsec duration at half-height and <300μsec duration from trough to peak. For each subject, there were about 10 times as many regular spiking units designated than fast-spiking units. Some of the regular spiking units likely comprised more than one neuron (multi-unit activity).

Trial-averaged pseudo-synaptic activity: After spiking events were extracted and attributed to isolated units, they were filtered to create a continuous time series for all following analysis. First, each spike event was assigned an exponential kernel (tau = 10msec). For each unit under a given stimulus condition, the pseudo-synaptic input was averaged across trials for use in the analyses of signal correlation and as input variables in each linear regression model.

Regularized Least Squares Linear Regression Model: At each time point, we obtain an estimate (average across trials) of the membrane potential of a single neuron isolated using an intracellular sharp electrode, and an estimate (average across trials) of the pseudo-synaptic output from each isolated unit along the extracellular electrode array. For the set of temporal observations under each stimulus condition, we implemented least squares regression to determine the identity of independent variables (spiking output of neurons) that predicted the membrane potential response (dependent variable) of each single neuron. The identity of the unit associated with each independent spiking response variable was preserved throughout the analysis and linked a specific physical location along the recording penetration (relative to dorsal surface; obtained from the electrode geometry).

We implemented two different regularization algorithms - lasso and ridge - that constrain the degrees of freedom in the least squares regression according to the L1 or squared L2 norm, respectively. By regularizing with the L1 norm of the coefficients for the set of predictors, the lasso algorithm identifies and removes redundant predictors from the regression regularized with the squared L2 norm of the coefficients. To toggle between these two regularization algorithms we use Matlab to implement an elastic net model: the lasso algorithm dominates when the alpha parameter is set to 1, while the ridge algorithm dominates when the alpha parameter set to 0.01. The number of non-zero coefficients depends on the gain of the regularization term - when the regularization gain is as small as possible, the model generates the best prediction possible (minimum MSE between the prediction and the trial-average membrane potential), but it also uses the maximum number of non-zero coefficients to do so.

In pre-processing for the regression analysis, we maintain the relative response tuning across the stimulus set within each neuron by normalizing all trial-averaged responses for a neuron to its maximum trial-averaged response. This normalization creates a pre-training condition in which all predictors are equipotential. The trial-averaged membrane potential response that the model is trained to predict is also normalized to the maximum response across stimuli.

As the gain on the regularization parameter is decreased, the reliability and accuracy of the prediction increases, but the number of non-zero coefficients also increases. For each regularization gain (lambda value) used in the regression model, a different set of coefficients is generated and the number of non-zero coefficients decreases as the regularization gain increases. To determine the model that would allow for a minimal set of non-zero coefficients while still sufficiently predicting the membrane potential, we implemented a physiologically relevant benchmark of sufficiency for the model's prediction. We defined a "sufficiency threshold" for each model: [the mean squared error (MSE) between the membrane potential and the prediction produced by the model with the smallest regularization gain] + [a standard prediction error in the trial-averaged membrane potential estimate] (Figure 3B). Where the Prediction Error = 3*[MSE between pairs of trial-averaged membrane potential estimates from random subsamples of 75% of the trials] (n = 50 subsample estimates) (Figure 3A). All signals were z-scored before calculating the MSE between them. For each value of lambda a new model was generated and we calculated the Mean Squared Error between the model prediction and the trial averaged membrane potential (Figure 3B). Once the lambda value was low enough that the prediction of the model crossed this threshold, we considered it

sufficient and use that model for all further analyses as described in the main text (Figure 3C).

Unless otherwise specified, results are reported as median value with the 95% confidence interval following in brackets.

3.4 Results

Examining population statistics and the selection of functionally-relevant subpopulations driven by temporally evolving natural stimuli presents challenges that this study addresses. The work focuses in the auditory cortex of the Starling, a model system that enables experiments targeting the neural mechanisms underlying the fundamental cognitive processes like those requisite for language processes common to all vocal learning species.

3.4.1 Temporal and Spatial Variance in Population Spiking

In the songbird cortex, NCM is a region - poised to support sensory-driven behavioral decisions – that demonstrates flexible high-dimensional structure in single neuron and population activity critical for adaptive spike-based stimulus classification. Presumably, downstream processes pool this activity into functional networks comprised of subpopulations within NCM, each with their own specific stimulus-response tuning function and unique co-variability structure that is modulated - potentially independently - to support downstream tasks. The spatial and temporal organization of spiking responses within regions that individual neurons pool from constrains the mechanisms that enable those neurons to achieve a breadth of stimulus-specific synaptic input.

As a first step to determining how vocal communication signals organize functional spiking networks within NCM, we examine the structure of stimulus-driven spiking activity through space and time among large populations (200-300 units of varying isolation quality) randomly selected from within NCM while presenting the animals with long sequences of conspecific vocalizations (40-80seconds of acoustic material in 6second segments). Although many studies have contributed to our knowledge of how regions comprising the ascending auditory pathway in songbird cortex are grossly connected, the local circuit architecture within each region is unparsed (Saini & Leppelsack, 1977, 1981; Wang, Brzozowska-Prechtl, & Karten, 2010). Even for model systems where we know more about local neural architecture, little is known about the distribution of temporal and spatial response similarity to time-varying, complex natural stimuli (but see (Woolley, Gill, Fremouw, & Theunissen, 2009)).

In Figure 1, we have plotted, for each NCM unit in a single subject, the normalized trial-averaged spiking response to a 6-second sequence of Starling vocalizations. Each unit's response is plotted spatially according to the electrode location at which it was recorded (along the dorsal-ventral axis in NCM). At any given moment in time, we observe variation across the population in the identity of active units, and the pattern of this activation across the population varies through time (stimuli). Additionally, it appears that near-neighbors tend to have more similar patterns of activation across time than units further from each other. To quantify these effects we measure the temporal and spatial correlations in the population spiking activity.

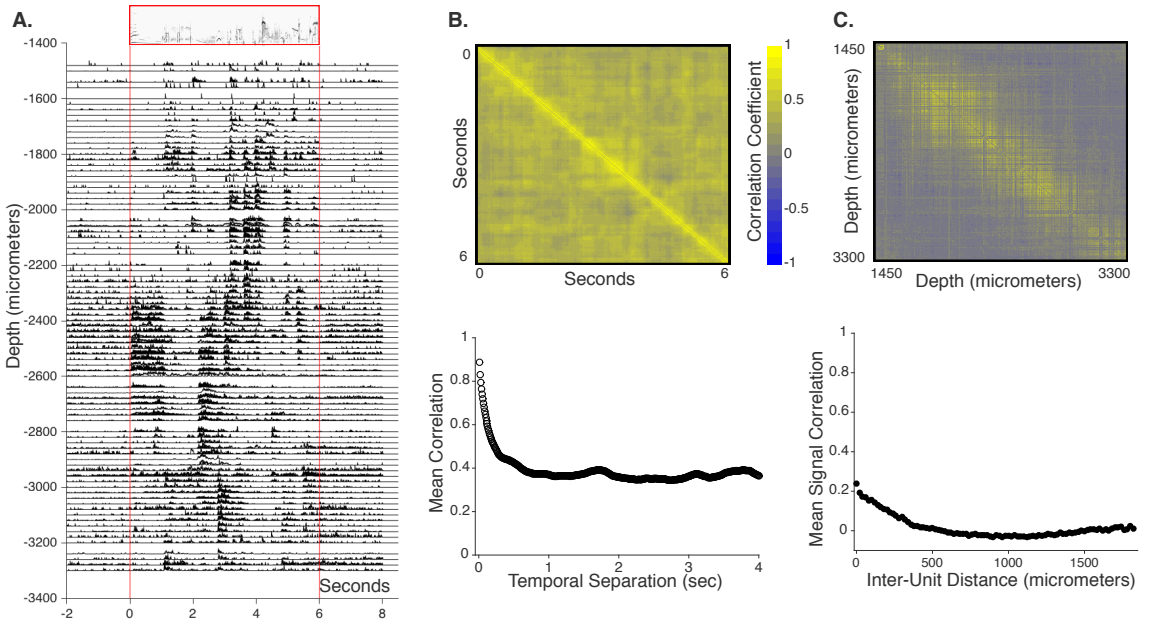

Figure 3-1: Spatial and temporal variance of spiking activity in NCM.
(A) Example of the trial-averaged spiking response (spike events filtered with a synaptic kernel (tau = 10ms) of each unit recorded in a single subject before, during, and after a 6-second vocalization sequence (spectrogram:top). Each trial-averaged response is plotted corresponding to the electrode location at which their spike events were isolated. (B) top: Example 2D image plot from a single subject of the correlation coefficient between the population activity vector at each pair of time points during the presentation of a 6-second sequence of vocalizations (colorbar right). Image plot averaged across 6 vocalization sequences for this subject. Bottom: average correlation for each temporal distance. (C) top: Example 2D image plot from the same subject shown in (B) of the correlation coefficient between the trial-averaged activity of pairs of unit (indexed by the electrode location at which their spike events were isolated) (colorbar left). Image plot averaged across 6 vocalization sequences for this subject. Bottom: average correlation for each temporal distance.

First, we calculate the correlation coefficient between the population activity vectors (spiking response strength across units) at two time points and compare that to their temporal separation (Figure 1B). We find that for time-points close to each other, the population spiking output is highly correlated (mean = 0.68 [max 0.75, min 0.54]; n = 4 subjects), and across time (stimuli) the correlation decays such that the log of the correlation scales with the log of the temporal distance by a factor of -0.21 (mean across

4 subjects, [-0.14 min, -0.28 max]). Spiking is highly reliable across trials (Churchland et al., 2010; Perks & Gentner, 2015), meaning that the observed temporal variance can provide rich stimulus-specific information. Instability in the population response across time could enable neurons pooling inputs from this region to achieve high temporal variability in their synaptic input simply by sampling consistently from one set of neurons. Alone, this would not fully constrain mechanisms of synaptic pooling that maintain stimulus-specificity. A combination of the spatial organization of spiking response tuning and the spatial organization of synaptic sampling across that spiking population also contributes to synaptic tuning.

In this dataset we can explicitly ask how the spiking response tuning is spatially organized in NCM - without needing to model the receptive fields of each neuron, which is otherwise a notoriously difficult solution (Kozlov & Gentner, 2014, 2016; Sharpee, 2013; Theunissen, Sen, & Doupe, 2000). We quantify the spatial structure of stimulus tuning among spiking responses through NCM by calculating the correlation coefficient between trial-averaged spiking responses of isolated units – the signal correlation – and comparing that to the spatial distance between each pair of units (Figure 1C). We find that near-neighbors tend to have higher signal correlation (median 0.21 [0.15min, 0.25max] for units within 200microns; n = 4 subjects) and that along the dorsal-ventral axis in NCM, this correlation had a constant exponential decay rate (space constant) of 0.0060 micrometer^{-1} [0.0035 min, 0.0074 max](n = 4 subjects).

For the representation of temporally-varying high-dimensional stimuli in sensory cortex, "*object-topic*" mapping like that shown in Figure 1 has not previously been parsed. The anatomical arrangement of feature selectivity has been heavily studied in the visual

cortex, but mainly limited to well-parameterized low-dimensional stimulus sets – for example, defining orientation columns in V1. Combinatorial explosion in stimulus set size combined with receptive field non-linearity makes any spatial characterization of object -selectivity challenging. Vocal communication signals used by species such as humans and songbirds provide a continuously time-varying sample of a high-dimensional diverse feature space that lends itself well to characterization of response relationships among sparsely selective neuron populations. Demonstrating the spatial organization of response similarity to temporally varying complex objects is non-trivial result that constrains mechanistic models of information processing through NCM. Several mechanisms could account for such a result - including massive recurrent local pooling and/or spatially restricted shared input from upstream areas.

Presumably, the identity of sets of units comprising functional networks pooled by downstream neurons NCM ultimately determines the temporal and spatial variance among the synaptic population shaping stimulus-specific temporal variance in the synaptic input to individual neurons. Although our extracellular recordings demonstrate strong variations in the population response across space and time (Fig 1), the functional relationships between neurons (and sets of neurons) within this population remains obscure.

3.4.2 Synaptic response reveals coordinated population dynamics.

While extracellular techniques such as imaging or dense electrode arrays enable the identification of neuron populations that are co-active or functionally grouped by their response properties, these co-active populations do not necessarily impact downstream

neurons together. Identifying the actual sub-networks within this large population that provide synaptic input to individual neurons and isolating their spiking activity presents technical challenges – we lack tools like molecular markers that target neurons flexibly recruited by synaptic processing in cortex. Insofar as a neuron's synaptic input and/or membrane potential is a reflection of spiking activity in the pre-synaptic network, then that neuron becomes an electrode through which we can examine the activity of functionally-relevant, spatially-unrestricted populations of neurons unharried by common biases of spike sorting or extracellular signal sampling.

To probe the population in this way, we must measure synaptic activity from individual neurons whose synaptic input reflects the statistics of the population. The ability of the population spiking activity to predict the membrane potential of a single neuron would provide confirmation of this, but presents a potential problem - presumably the synaptic input of a single neuron represents only a tiny fraction of the total population in any given region. Insofar as the spiking pattern among the pre-synaptic population for each neuron is unique and independent, predicting the synaptic response from the spiking pattern of a randomly-recorded sample of neurons from the total population would be difficult. One might imagine that to predict the sub-threshold response of a given neuron, you need to record from a population of neurons in which the inputs to that cell are strongly represented. On the other hand, if the synaptic responses of different neurons carry redundant information about the stimulus, and this shared information is reflected in the population, then the spiking activity of the total population may provide a very good predictor of the synaptic response in any randomly chosen neuron.

We know that individual NCM neurons receive highly structured non-specific synaptic input even when that neuron's spiking response is very selective (Perks & Gentner, 2015). Here we use intracellular sharp electrodes to measure, in the same bird and within the same area of NCM where spiking responses were collected extracellularly, the synaptic response (mV) of individual neurons and examine how its temporal variance can be explained by the temporal and spatial variance in NCM population spiking activity. We first convolved each spike event with a synaptic kernel (tau = 10msec) to generate a pseudo-synaptic response for each unit. At each time point we then sum the pseudo-synaptic response across all NCM units to generate a *net* pseudo-synaptic response vector through the entire stimulus set. We calculated the correlation coefficient between the trial-averaged synaptic response and the net pseudo-synaptic response using a sliding 500msec window. To account for potential temporal offset in stimulus drive for these two signals we took the maximum cross-correlation value within each window (limited by a maximum lag of 50 msec – reflecting the onset of stimulus-driven activity generally observed in NCM). The median average correlation was 0.93 (Figure 2D; (n = 8 datasets from 4 subjects). As a comparison for the magnitude of this correlation, we calculated the mean correlation between the membrane potential response within each window to 10 other randomly-selected (temporally mis-aligned) net pseudo-synaptic responses. The aligned condition had a significantly higher correlation than the mis-aligned conditions for each stimulus (Figure 2,E; $p<0.001$ for all datasets except one with $p = 0.007$; median average misaligned correlation = 0.91).

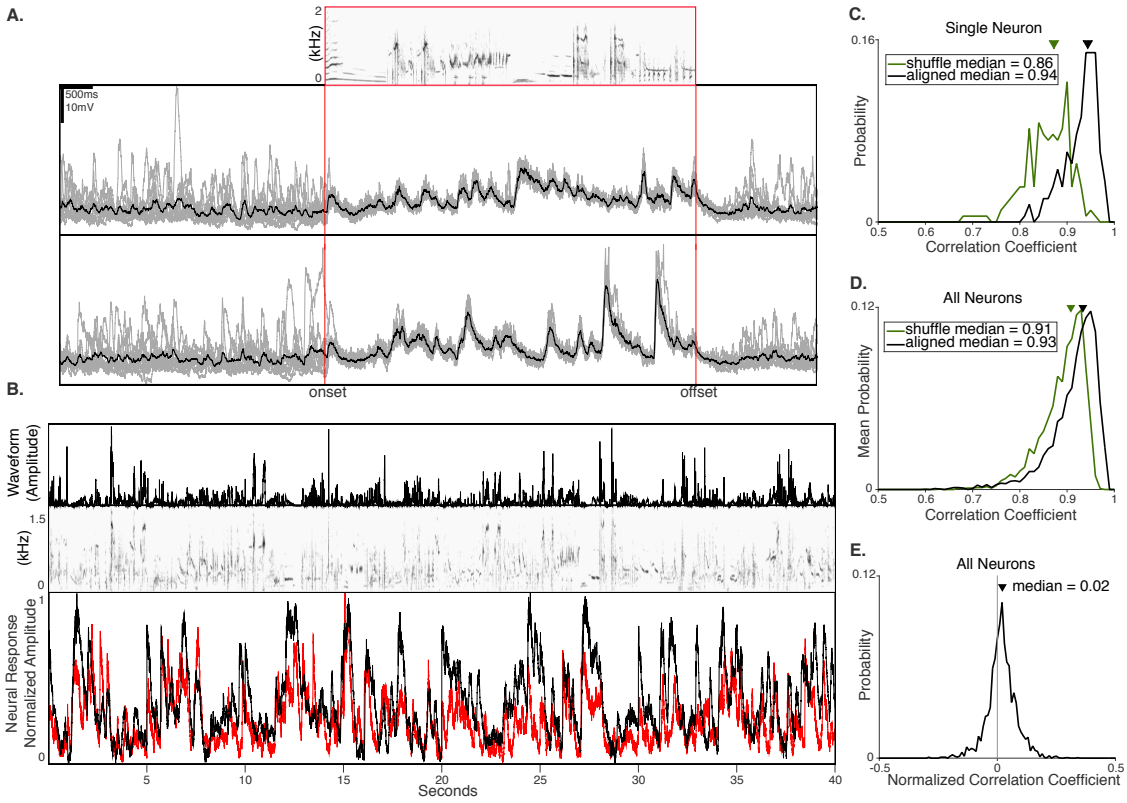

Figure 3-2: The synaptic response reflects net population spiking activity.
(A) Example membrane potential of one neuron before, during, and after the presentation of a vocalization sequence (top and bottom panels correspond to two unique vocalization sequences). Gray: single trials; black: trial-averaged. Responses in top panel correspond to the vocalization sequence plotted as a spectrogram (top). (B) top: Amplitude waveform concatenated across all vocalization sequences presented to a single subject. Middle: Same vocalization sequence as in the top panel; plotted as a spectrogram. Bottom: Trial-averaged membrane potential (black) recorded from a single neuron in NCM and trial-averaged net pseudo-synaptic response (population spiking response) from 279 units isolated from the entire dorsal-ventral extent of NCM sampled in the same subject. Both signals normalized to max. (C) Probability distribution of the binned correlation coefficient between the membrane potential of one neuron and the population spiking response calculated within a 500 msec window stepped every 100msec through the response. black: temporal alignment between the membrane potential and the population response maintained, green: windows shuffled randomly without replacement before calculating correlation coefficients (corresponds to the data plotted in 'B: bottom'). (D) Same as in C, Probability distribution of correlation coefficients between membrane potential and net pseudo-synaptic response across all single-neuron:population pairs (n = 8) (E) Same data as in D, Probability distribution of difference between aligned and shuffled conditions (all response windows, all neuron:population pairs).

Together, this demonstrates that temporal variance in the net population spiking response across NCM throughout the stimulus can explain a lot of the variance in the subthreshold membrane potential of individual neurons within NCM.

Spiking responses from some subset of the NCM population could provide more explanatory power than from other subsets of the population. As an initial test of the ability to identify smaller functional networks within NCM whose spiking optimally co-varies with the synaptic representation of the stimulus (measured as the subthreshold membrane potential), we employ a simple multiple linear regression model in which the independent (predictor) variables are given by the trial-averaged pseudo-synaptic response of isolated spiking units, and the dependent variable is the trial-averaged synaptic membrane potential of a single neuron recorded *in-vivo* from the same population (Figure 3A,B; see Methods). We implemented an elastic net regularization parameter biased heavily toward the ridge algorithm (0.01 Lasso:Ridge). To restrict the regression model to the pseudo-synaptic responses that were most informative of the synaptic input (mV) of a single neuron, we chose the regularization gain that generated a model with the smallest number of non-zero coefficients able to generate a fit on the membrane potential exceeding the "sufficiency threshold" (Figure 3C-E; see Methods for more detail). We defined the "functional network" as the set of units whose spiking activity contributes to sufficient fits of the subthreshold membrane potential variance (the predictors with non-zero coefficients; Figure 3C-D, 4A).

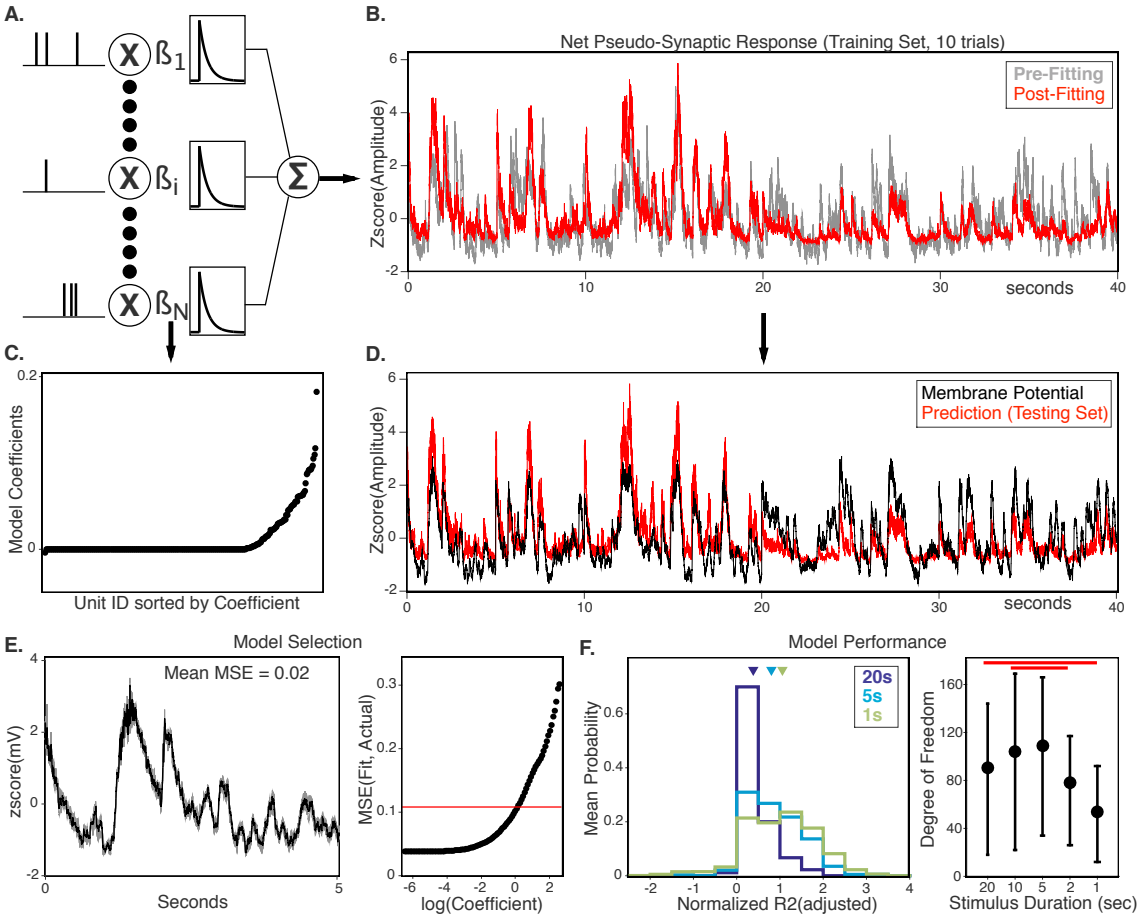

Figure 3-3: Cross-validated linear regression performs best with shorter stimuli. (A) Each spike event is convoluted with an exponential kernel (tau=10ms) and weighted by a coefficient (Bi). (B) Pseudo-synaptic responses summed across all units. Pre-fitting (gray) and post-fitting (red) for the training set of 10/20 trials. (C) Unit ID is sorted in ascending order corresponding to the magnitude of its coefficient assigned by the linear regression (D) After fitting; Prediction (red; on 10/20 withheld trials) of the synaptic response (black). (E) For the example dataset depicted in B, left: (gray) 100 estimates of the trial-averaged membrane potential using 75% of recorded trials on each iteration and (black) the inclusive trial-averaged membrane potential for a sample of the response plotted in black in D. The average Mean Squared Error (MSE) has been calcualted for the entire response; right: (red) Sufficiency threhold for the model fit. (scatter) MSE between the membrane potential and the model fit for each regularization coefficient value. (F) For all datasets (n = 8) left: Histogram of adjusted-Rsquare values between model prediction and the membrane potential under three segment duration conditions (1s, 5s, 20s). The distributions were significantly different for all of these conditions. right: For each segment duration condition, the median (+/- 95% confidence intervalse) on the distribution of degrees of freedom for each model generated by the ridge regularized regression fit. Significant comparisons denoted by red lines.

To determine whether each model fit was a meaningful representation of the membrane potential with predictive power, we assessed model performance by calculating an estimate of the adjusted R-square value between the membrane potential and predictions of the membrane potential from the model. We built a cross-validation with 50% of the trials (randomly chosen) and testing with the other 50% to calculate the adjusted R-square value on each of 5 iterations and averaging across cross-validations. Although the synaptic input to individual neurons was highly correlated with the net population spiking response throughout the duration of the entire stimulus set (40-80seconds), there is evidence from the literature for stimulus-specific control over the representation of these vocalization signals on smaller timescales of single motifs and features – on the order of seconds, and even tens of milliseconds – (Jeanne, Sharpee, & Gentner, 2013; Kozlov & Gentner, 2014).

Accordingly, we binned the responses by segment durations of 20, 10, 5, 2, or 1 seconds, generated a model for each segment, and quantified model performance for each. To combine the results across datasets, for each stimulus we calculated the difference between the adjusted R-square value under normal conditions and spike shuffled conditions. Across subjects, the estimate (median) of this normalized adjusted R-square for 20-second segments was 0.38 [0,1.20], while for 1-second segments the median was 1.06 [-0.22,2.31] indicating a trend for increased model performance with decreased stimulus duration (Figure 2F; 10s = median 0.51 [0,1.58]; 5s = median 0.80 [0,1.93]; 2s = median 1.01 [0,2.17]). We conducted a Kruskal–Wallis one-way analysis of variance to compare the effect of segment duration on model performance, which indicated a significant effect at the $p<0.05$ level for the five conditions [Chi2 (4, 5845) = 151.38,

$p<0.001$]. Post-hoc comparisons using the Tukey HSD test indicated that all comparisons were significant with $p<0.001$; except between the 5 and 2 second groups $p = 0.003$, between the 20 and 10 second groups $p = 0.117$, and between the 2 and 1 second groups $p = 0.388$. We then conducted a Kruskal-Wallis test to compare the effect of segment duration on the degree of freedom in each model (Figure 2F; 20s = median 91 [18,144]; 10s = median 104 [22,169]; 5s = median 110 [34,166]; 2s = median 78 [26,117]; 1s = median 54 [12,92]). There was a significant effect at the $p<0.05$ level for the 5 conditions [$Chi2(4, 2863) = 345.77$, $p<0.001$). Post-hoc comparisons using the Tukey HSD test indicated a significant different between the 1-second group and all other groups (each comparison $p<0.001$) and, additionally, a significant difference between the 2-second group and the 10 and 5 second groups ($p<0.001$ for each). Overall, decreasing segment duration tended to decrease the degree of freedom for each models. Combined, these results imply that with decreased stimulus duration, the linear regression trained on a specific stimulus segment is able to hone in on a refined representation of the functional network associated with each synaptic representation of that stimulus. One significant interpretation of this result is that it could suggest mechanisms of stimulus-specific modulation in the covariance structure of the population.

For all following analyses, we segmented the response into bins corresponding to the 1-second long acoustic elements comprising the vocalization sequences – referred to as motifs. In addition to this being the group with the highest model performance requiring the fewest degrees of freedom (Figure 3F), this segment duration is consistent with ethologically-relevant segmentation of starling vocalizations (Gentner, 2008).

The models' performance could be a trivial result of over-fitting due to the large number of predictor variables (units isolated extracellularly). To address this, we created a noisy condition in which we permuted the population spike events to destroy any stimulus-specific structure in the pattern of activity across the population at any given point in time. In this condition each spike time was preserved, but its occurrence in a given neuron in response to a given stimulus on a given trial was randomly shuffled before performing the linear regression as before. We found that the functional networks identified by linear regression under each stimulus condition were able to predict the dependent neuron's synaptic input (mV) under the normal condition better than under this shuffle condition (Figure 3F, distributions of the difference between R2 in the normal condition and R2 in the shuffle condition were all centered above zero). Additionally, the degree of freedom necessary for a sufficient fit was higher under the shuffle condition (for the 1-second segment group, median 57 [4,113] units under the normal condition and median 80 [11,165] units under spike shuffled condition) – a trend that was consistent and significant across all stimulus durations tested. Together, these results support the use of synaptic input recorded from single neurons (measured here as subthreshold membrane potential) to obtain a snapshot of the functional pooling of spiking activity across NCM, which we then capitalized on to examine the temporal and spatial structure of stimulus-specific functional networks in NCM.

3.4.3 Heterogeneous pooling of spiking supports stimulus-specific synaptic responses.

Across datasets, we identified a trend for model performance to increase with decreasing response segment duration (Figure 3F and associated statistics), suggesting

that with smaller chunks of vocalizations we are better able to identify the units whose spiking are unique to each synaptic representation of that vocalization. To directly test the stimulus-specificity of each functional network, we compared the predictive ability of the functional network under each motif condition on all other motif conditions.

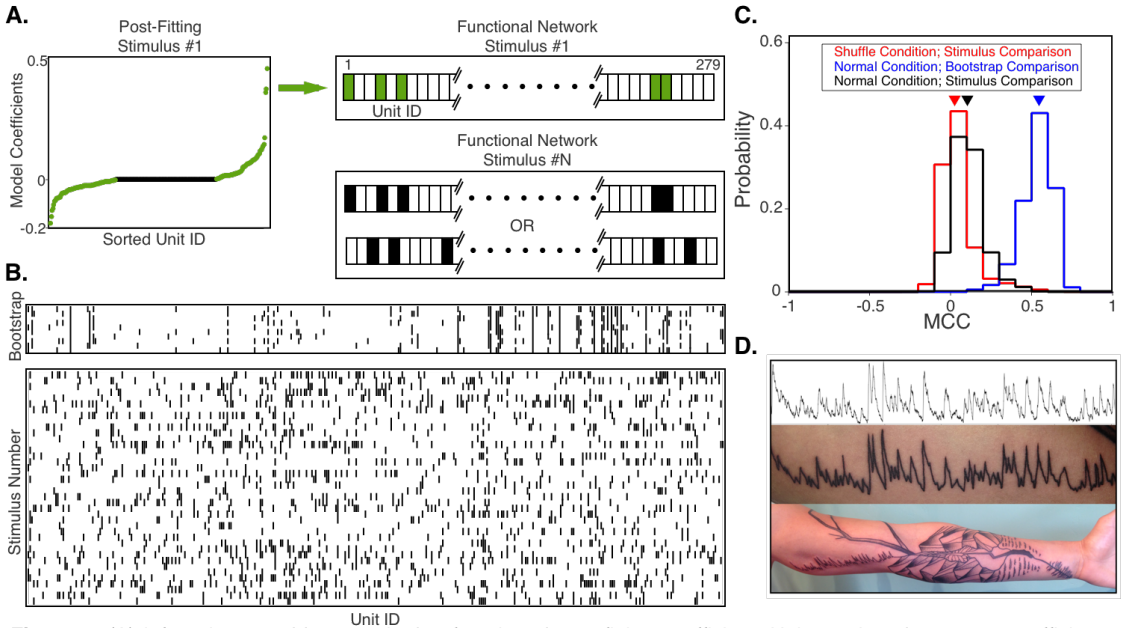

Figure 3-4: Functional network specificity.
(A) left: units sorted by magnitude of assigned post-fitting coefficient. Units assigned non-zero coefficients highlighted in green. right, top: The functional network is a binarized vector in which each unit with a non-zero coefficient is assigned a "1" and all other units are assigned a "0". right, bottom: Hypotheses under a different stimulus condition. Two potential functional networks are depicted: one is the same as for stimulus #1 and the other is different. (B) For an example dataset (single-neuron:population pair), top: funcitonal network identified for each of 10 bootstrap iterations under the same stimulus condition. bottom: functional network identified for each of 40 motif stimuli under the same bootstrap iteration (C) Probability distribution of Matthews Correlation Coefficient (MCC) values calculated between functional networks under 3 conditions: (red) comparison across stimuli under the spikes shuffled condition, (blue) comparison across bootstrap iterations under the non-shuffled condition, (balck) comparison across stimului under the non-shuffled condition. Difference between distributions is significant for all comparisons. (D) Placeholder for balanced graphical design, but don't have another graph that seems necessary to show here. Instead, some images from the thesis-inspired tattoo. Pink Floyd's "Another Brick in the Wall" synaptically filtered through the starling's auditory system and depicted as a dendrite in the artistic tattoo neuron.

After generating a linear regression model under each motif condition we constructed a binary population vector of 1's and 0's where each entry identifies whether a unit was assigned to the functional network or not, respectively (Figure 4A). We compare the functional network under each stimulus condition using the "Matthews Correlation Coefficient" (MCC), which is a measure of how well one binary vector predicts another. Sampling the spiking activity introduces inconsistency that might account for any observed differences between the functional networks identified for different motifs. To address this we implement a bootstrap: subsampling the extracellular dataset (10 out of 20 trials chosen randomly without replacement) and generating a model for each bootstrapped dataset under the same stimulus condition (n = 10 iterations). We obtained an estimate of the maximum MCC expected across stimuli if the functional network was consistent by calculating the MCC across bootstrapped datasets under each stimulus condition (median 0.54 [0.35, 0.66]). We obtained an estimate of the minimum MCC value expected given no constancy between functional networks representing different motifs by calculating the median MCC across stimuli for the spike-shuffled datasets used to test model performance originally, (median 0.03 [-0.08, 0.28]. Finally, we calculated the MCC between different stimuli in each (non-shuffled) bootstrapped dataset (median 0.10 [-0.02,0.29]) to compare amongst the distribution of MCC values in the maximum-overlap and maximum-noise conditions (bootstrap comparison and spike-shuffled condition respectively). We conducted a Kruskal–Wallis one-way analysis of variance to compare the effect of group on the MCC, which indicated a significant effect at the $p<0.05$ level for the 3 conditions [Chi2 (2, 16639) = 3398.49, p=0]. Post-hoc

comparisons using the Tukey HSD test indicated that all comparisons were significant with $p<0.001$. Although there was some correlation in the functional network identified for each motif, motif identity modulated the identity of the functional network more than expected due to sampling differences (Figure 4C). Instead of the same subset of the population being sampled continuously throughout the stimulus set, we found that the functional networks were more likely to be closer to heterogeneous (non-overlapping across motifs). This result supports the possibility for significant reorganization of downstream pooling across stimuli that could allow for, among other changes in population dynamics, shifts in the correlation structure among neurons comprising functional subnetworks for different stimuli.

3.4.4 Spatially flat distribution of functional networks.

In the first section (Figure 1) we examined the spatial structure of spiking response similarity among the units isolated throughout NCM. The units comprising functional networks identified by linear regression correspond in some way to the structure of spiking response tuning of units identified anatomically. The functional network identified for each stimulus may be comprised of units clustered with similar responses or comprised of units distributed throughout the response space. We characterized the flatness of the spatial distribution of coefficients (corresponding to individual units linearly arranged throughout the dorsal-ventral extent of NCM) by creating a probability distribution of coefficients across recording locations (Figure 5A). For each functional network identified we generated probability distributions for a maximally-flat and a maximally-peaked model to compare the empirical data against. To

generate the flat model, the regression coefficients were randomly (without replacement) re-assigned to each unit before calculating the location probability distribution. To generate the maximally peaked model, the regression coefficients were sorted in descending order and re-assigned sequentially to each unit according to its ranked location (so the strongest weights are assigned to the units at location 1 and the weakest weights are assigned to the unit at location N, where locations 1 through N are linearly arranged from most dorsal to most ventral). We then calculated the KL divergence between these location probability distributions (Figure 5A,B). The median KL divergence between the extreme models (maximally-peaked and maximally-flat) was 0.44 [0.01, 2.08]. The median KL divergence between the functional network and the maximally flat model was 0.03 [0, 0.15]. The median KL divergence between the functional network and the maximally peaked model was 0.36 [0.01, 1.96] (Figure 5B). The distribution of coefficients in the functional network were more like the flat model than the peaked model. We conducted a Kruskal–Wallis one-way analysis of variance to compare the effect of group on the KL divergence, which indicated a significant effect at the $p<0.05$ level for the 3 conditions [$Chi2\ (2, 1077) = 423.73$, $p<0.001$]. Post-hoc comparisons using the Tukey HSD test indicated that the only comparison that was not significant at $p<0.001$ was the KL divergence between the two extreme models and the KL divergence between the functional network and the maximally peaked model ($p = 0.26$) indicating that the functional network and the maximally-flat model were the same distance from the maximally-peaked model.

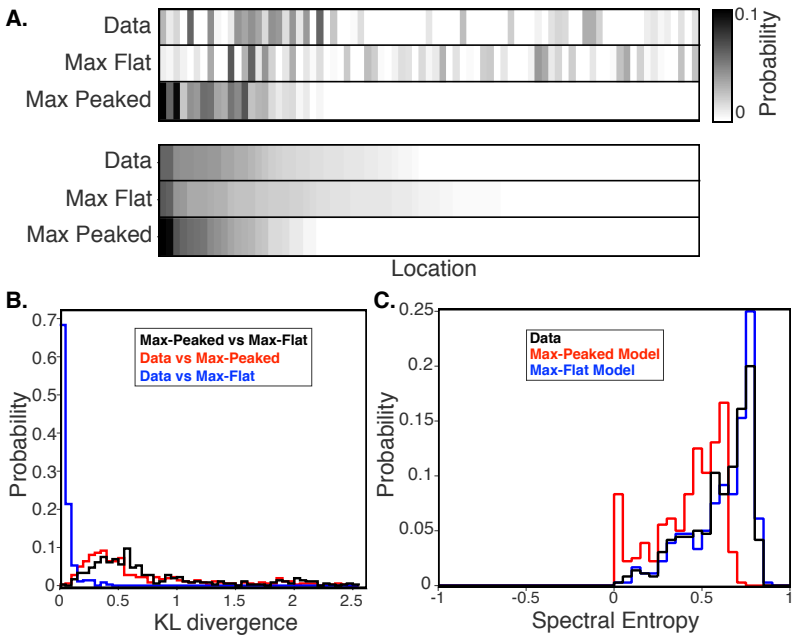

Figure 3-5: Functional network spatial distribution.
(A) top: Example from one dataset of the relative weight assigned to each location on the recording electrode (quanitified as the probability) by linear regression under each model ("Maximally Peaked" and "Maximally Flat" as described in the text) and as calcualted from the unmanipulated data. bottom: Locations sorted by their relative assigned magnitudes. (B) Probability distribution of the KL divervence for each of three comparisons: between the data and the maximally peaked model, between the data and the maximally flat model, and between the two extreme models (flat and peaked) (n = 8 datsets) (C) Probabilty distribution of the Spectral Entropy calculated on the coefficient probability distrubution for the unmanipulated data and under the two extreme models (peaked and flat) (n = 8 datasets).

One way to quantify the flatness of a distribution (in this case across physical space) is the Wiener entropy in which values approaching 1 indicate that each location has a similar amount of predictive power whereas values approaching 0 indicate that the predictive power is clustered around a single location. We found that the maximally peaked model distribution had the lowest entropy (0.46 [0.01, 0.65]), the maximally flat model distribution had the highest entropy (0.64 [0.26, 0.80]), and the functional network distribution had a median entropy of 0.62 [0.25, 0.78]. We conducted a Kruskal–Wallis one-way analysis of variance to compare the effect of group on the entropy, which

indicated a significant effect at the $p<0.05$ level for the 3 conditions [Chi2 (2, 822) = 158.31, $p<0.001$]. Post-hoc comparisons using the Tukey HSD test indicated that the only comparison that was not significant was between the maximally-flat model and the functional network ($p = 0.30$) consistent with the results obtained from the analysis of KL divergence among the models. (Figure 5C).

We find that spiking activity in NCM, although spatially organized by stimulus-driven response probabilities, can be pooled heterogeneously across space to support encoding of conspecific vocalizations. Given that the functional networks do not just sample units clustered together anatomically, and signal correlation decreases with physical distance in NCM, one might expect that the input space samples from neurons with diverse spike tuning. We examined this by comparing the distribution of pairwise signal correlations among each functional network to the distribution of pairwise signal correlations among all pairs recorded in NCM. We found that functional networks are actually biased to contain units with more positive mean pairwise signal correlations [re-running quantification of median & ci]. (Figure 6A).

3.4.5 Redundancy.

Given that pairwise signal correlations among functional networks identified for each stimulus were higher than expected if sampling evenly from population distribution, we next examined redundancy of the information carried in spiking response of individual units more explicitly by implementing lasso regularization. Lasso emphasizes sparseness in the regression model and eliminates redundancy by penalizing the error minimization routine by the L1 norm of the coefficients. If there is a group of highly

correlated variables, then the lasso tends to select one variable from a group and ignore the others. Thus the Lasso regularizaton yields the sparsest possible input space that predicts the mV with a given degree of accuracy, while the Ridge regularization yields the largest/densest possible input space that predicts the membrane potential to the same level of accuracy, including all the units that carry redundant information. Accordingly, the population redundancy can be quantified directly as the as the ratio of the size of the input space defined by these two models (Severson et al 2015), where size is the number of units assigned non-zero coefficients (i.e., degrees of freedom). We express this ratio as a Redundancy Index (RI; methods), where RI equal to 1 indicates full redundancy (all units perfectly correlated) and RI = 0 indicates no redundancy (all units are independent). The median RI across datasets was 0.64 [0.41, 0.85] (Figure 6).

To quantify the constancy with which lasso and ridge regularization identified stimulus-specific functional networks we calculated the MCC between the functional networks generated by each model for a given stimulus in a given dataset (median 0.78 [0.66, 0.86]; Figure 6D). As expected given that a functional network identified under ridge strongly predicts the functional network identified under lasso, we find that the sparse functional networks identified under lasso were also stimulus specific and distributed flatly across NCM.

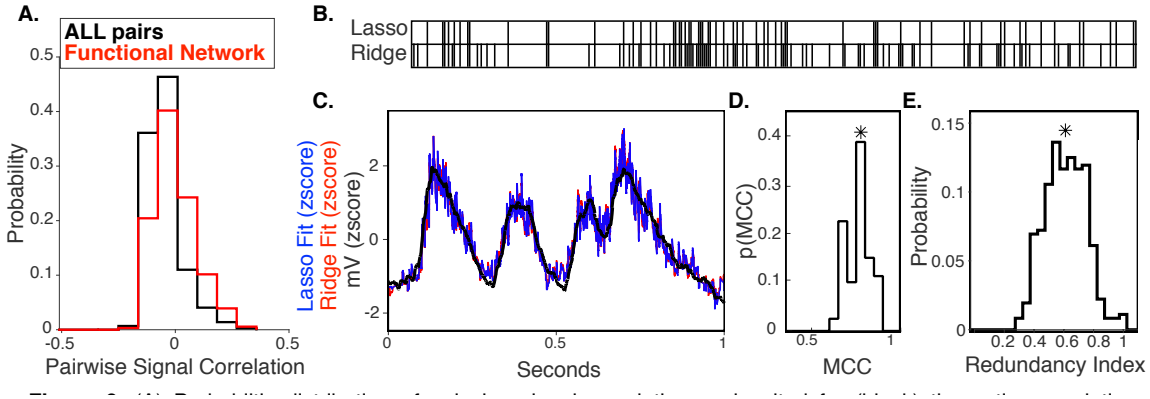

Figure 3-6: Redundancy.
(A) Probabiltiy distribution of pairwise signal correlations calcualted for (black) the entire population recorded extracellularly and (red) for units within identified functional networks. (n = 8 datasets). (B) Functional Network for the same motif stimulus under the same bootstrap iteration under Lasso (top) or Ridge (bottom) regularization. (C) Example membrane potential (black) recorded in a single neuron in response to a single motif stimulus overlaid with the model fit of that membrane potential from the population spiking response under Lasso (blue) or Ridge (red) regularization. All waveforms plotted as z-score to normalize. (D) Probability distribution of MCC values calculated between the Functional Network identified under Lasso and Ridge for the same motif stimulus under the same bootstrap iteration. (n = 8 datasets) (E) Probability distribution of the Redundancy Index comparing the Degrees of Freedom for the regression model under Lasso and under Ridge regularization for the same motif stimulus under the same bootstrap iteration.

To quantify the stimulus specificity we calculated the MCC between the functional networks across stimuli and across bootstrapped sub-samples, generated under both normal and shuffle conditions. As before, we implemented a bootstrap by subsampling the extracellular dataset (10 out of 20 trials chosen randomly without replacement) and generating a model for each bootstrapped dataset under the same stimulus condition (n = 10 iterations). We obtained an estimate of the maximum MCC expected across stimuli if the functional network was consistent by calculating the MCC across bootstrapped datasets under each stimulus condition (median 0.45 [0.28, 0.60]).

We obtained an estimate of the minimum MCC value expected given no constancy between functional networks representing different motifs by calculating the median MCC across stimuli for the spike-shuffled datasets used to test model performance originally, (median 0.02 [-0.08, 0.24]. Finally, we calculated the MCC between different stimuli in each (non-shuffled) bootstrapped dataset (median 0.07 [-0.02,0.24]) to compare amongst the distribution of MCC values in the maximum-overlap and maximum-noise conditions (bootstrap comparison and spike-shuffled condition respectively). We conducted a Kruskal–Wallis one-way analysis of variance to compare the effect of group on the MCC, which indicated a significant effect at the $p<0.05$ level for the 3 conditions [Chi2 (2, 16997) = 2501.31, p=0]. Post-hoc comparisons using the Tukey HSD test indicated that all comparisons were significant with $p<0.001$.

To quantify the spatial distribution of regularization coefficients, we generated a maximally flat and a maximally peaked model to compare against the functional networks using the entropy of the probability distirbutions and the KL divergence between them. The maximally peaked model distribution had the lowest entropy (0.37[0, 0.56]), the maximally flat model distribution had the highest entropy (0.0.55 [0.19, 0.73]), and the functional network distribution had a median entropy of 0.54 [0.18, 0.72]. We conducted a Kruskal–Wallis one-way analysis of variance to compare the effect of group on the entropy, which indicated a significant effect at the $p<0.05$ level for the 3 conditions [Chi2 (2, 831) = 189.8, $p<0.001$]. Post-hoc comparisons using the Tukey HSD test indicated that the only comparison that was not significant was between the maximally-flat model and the functional network ($p = 0.60$) consistent with the results obtained from the analysis of KL divergence among the models.

The degree of redundancy is inherently related to key aspects of population coding, such as spiking correlations among the population, which we address in the discussion. Together, these results suggest that sensory integration and processing are supported by a system in which information about even the most complex stimuli is likely massively redundant and shared among the population at large.

3.5 Discussion

For the complex sensory signals that guide many behaviors, our understanding of stimulus encoding is especially poor. Synaptic pooling strategies allowing individual neurons to flexibly integrate across diverse spiking responses are unknown, as most studies examining population coding utilize randomly-sampled spiking output from among anatomically-defined regions under low-dimension stimulus conditions. In the current study, we embrace the complexity of a rich stimulus set (the conspecific vocalizations of the Starling) and utilize the synaptic response of individual neurons to study the stimulus-driven dynamics of population spiking activity across space and time. What we find is that much of the temporal variability in the population spiking response (across a large spatial extent of tissue) is reflected in the synaptic response of individual neurons embedded anywhere in that region. The subset of the population that best describes the synaptic response is stimulus-specific on short timescales, and the population must be sampled heterogeneously through space and time to reflect the synaptic representation of the stimulus. These results bear directly on contemporary models of hierarchical sensory processing supporting flexible behavior in natural environments.

Songbirds, in particular starlings, are excellent models for examining neural mechanisms supporting cognitive processes such as those requisite for language and vocal communication in many species. In high-order auditory cortex of the Starling, spiking responses are very sparse and stimulus-specific (Jeanne et al., 2011) but not well-described by linear receptive fields (Theunissen et al., 2000), with individual neurons demonstrating flexible input recombination functions driven by diverse sets of independent (orthogonal) features contained in conspecific vocalizations (Kozlov & Gentner, 2014, 2016). Additionally, these already complex spiking responses are modulated by task-relevant information, acoustic context, and attention (Jeanne et al., 2011; Kozlov & Gentner, 2014)(Caporello & Gentner, unpublished). The ability for individual neurons and populations of neurons within high-order auditory cortex of the Starling (and other animals) to demonstrate complex and flexible spiking behavior is critical for successfully processing of incoming sensory information to guide behavior.

At a population level, task-relevant stimulus information modulates the spiking covariance patterns among neurons in high-order sensory cortex, which improves stimulus discrimination for informative stimuli in neural populations poised to support sensory-driven behavioral decisions (Downer, Niwa, & Sutter, 2015; Gu et al., 2011; Jeanne et al., 2013). Modulation of spiking covariance among neural populations seems to be an important component for success in neural processing in general. Although beyond the scope of this study, experiments designed to examine the effects of spiking correlations on population coding - and subsequently behavior - will benefit from identifying the specific sets of neurons whose spiking output is integrated by downstream targets. Covariance may be distributed among these functionally-relevant populations

differently than the population at large. We find that pairs of neurons within synaptically-identified functional networks have a higher probability of positive signal correlations than pairs of neurons randomly selected from the population. This suggests that the synaptic integration process is selective for sets of neurons with specific correlation structure. To examine this directly using the techniques implemented in the current study would require simultaneous recording of the single-neuron membrane potential and the population spiking to estimate the covariance structure on a single-trials. But the results would help resolve heated debates in the literature about the ways in which correlated spiking activity effects population coding.

Most of previous studies examining the affect of correlations on population coding have focused on calculating pairwise correlations, which are easier to measure with limited amounts of data than higher-order correlations (Averbeck, Latham, & Pouget, 2006; Franke et al., 2016). This limitation could be addressed by using the synaptic response of single neurons to estimate the correlation structure of a population since the statistics of the subthreshold membrane potential offer a complete image of the statistics of the pre-synaptic network's spiking ensemble. Mathematical tools to accomplish this are readily available in the literature (Benucci, Verschure, & Konig, 2007; Bohte, Spekreijse, & Roelfsema, 2000; DeWeese & Zador, 2006; Kuhn, Aertsen, & Rotter, 2003; Levitan, Segundo, Moore, & Perkel, 1968; Renart et al., 2010; Rudolph & Destexhe, 2003; Tan, Chen, Scholl, Seidemann, & Priebe, 2014). Additionally, the results of the current study reveal the existence of stimulus-specific populations, each potentially exhibiting a unique correlation structure, which could be modulated independently by behavior, context, or attention. We are left with an open question as to

whether the correlations among the pre-synaptic network of each individual neuron reflect the correlations among its synaptically-identified functional networks. This could be directly addressed by obtaining simultaneous recordings of these signals combined with an estimation of the correlation structure in the pre-synaptic population and the randomly-sampled population. Implementing these techniques in future studies will enable us to resolve more mechanistic detail bearing on how representations of stimuli are modulated to meat the behavioral demands of ever-changing environments.

3.6 Acknowledgements

Chapter 3, in full, is in preparation for publication as: Perks, K. and Gentner, T. "Natural signals drive fast modulation of stimulus-specific functional networks in cortex." The dissertation author was the primary investigator and author of this manuscript.

3.7 References

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci, 7*(5), 358-366. doi:10.1038/nrn1888

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., & Fregnac, Y. (2013). Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. *Front Neural Circuits, 7*, 206. doi:10.3389/fncir.2013.00206

Benucci, A., Verschure, P. F., & Konig, P. (2007). Dynamical features of higher-order correlation events: impact on cortical cells. *Cogn Neurodyn, 1*(1), 53-69. doi:10.1007/s11571-006-9000-y

Bohte, S. M., Spekreijse, H., & Roelfsema, P. R. (2000). The effects of pair-wise and higher order correlations on the firing rate of a post-synaptic neuron. *Neural Comput, 12*(1), 153-179.

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., . . . Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci, 13*(3), 369-378.

DeWeese, M. R., & Zador, A. M. (2006). Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 26*(47), 12206-12218. doi:10.1523/JNEUROSCI.2813-06.2006

Downer, J. D., Niwa, M., & Sutter, M. L. (2015). Task engagement selectively modulates neural correlations in primary auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 35*(19), 7565-7574. doi:10.1523/JNEUROSCI.4094-14.2015

Franke, F., Fiscella, M., Sevelev, M., Roska, B., Hierlemann, A., & da Silveira, R. A. (2016). Structures of Neural Correlation and How They Favor Coding. *Neuron, 89*(2), 409-422. doi:10.1016/j.neuron.2015.12.037

Gentner, T. Q. (2008). Temporal scales of auditory objects underlying birdsong vocal recognition. *J Acoust Soc Am, 124*(2), 1350-1359.

Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature, 424*(6949), 669-674.

Gu, Y., Liu, S., Fetsch, C. R., Yang, Y., Fok, S., Sunkara, A., . . . Angelaki, D. E. (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron, 71*(4), 750-761. doi:10.1016/j.neuron.2011.06.015

Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron, 78*(2), 352-363.

Jeanne, J. M., Thompson, J. V., Sharpee, T. O., & Gentner, T. Q. (2011). Emergence of learned categorical representations within an auditory forebrain circuit. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 31*(7), 2595-2606.

Kozlov, A. S., & Gentner, T. Q. (2014). Central auditory neurons display flexible feature recombination functions. *J Neurophysiol, 111*(6), 1183-1189.

Kozlov, A. S., & Gentner, T. Q. (2016). Central auditory neurons have composite receptive fields. *Proc Natl Acad Sci U S A, 113*(5), 1441-1446. doi:10.1073/pnas.1506903113

Kuhn, A., Aertsen, A., & Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput, 15*(1), 67-101. doi:10.1162/089976603321043702

Latham, P. E., & Nirenberg, S. (2005). Synergy, redundancy, and independence in population codes, revisited. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 25*(21), 5195-5206. doi:10.1523/JNEUROSCI.5319-04.2005

Levitan, H., Segundo, J. P., Moore, G. P., & Perkel, D. H. (1968). Statistical analysis of membrane potential fluctuations. Relation with presynaptic spike train. *Biophys J, 8*(11), 1256-1274. doi:10.1016/S0006-3495(68)86554-3

Lien, A. D., & Scanziani, M. (2013). Tuned thalamic excitation is amplified by visual cortical circuits. *Nat Neurosci, 16*(9), 1315-1323.

Machens, C. K., Wehr, M. S., & Zador, A. M. (2004). Linearity of cortical receptive fields measured with natural sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 24*(5), 1089-1100. doi:10.1523/JNEUROSCI.4445-03.2004

Meliza, C. D., Chi, Z., & Margoliash, D. (2010). Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. *J Neurophysiol, 103*(3), 1195-1208.

Perks, K. E., & Gentner, T. Q. (2015). Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex. *Eur J Neurosci, 41*(5), 725-733. doi:10.1111/ejn.12831

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci, 12*(6), 718-724. doi:10.1038/nn.2331

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., & Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science, 327*(5965), 587-590. doi:10.1126/science.1179850

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat Neurosci, 2*(11), 1019-1025.

Rudolph, M., & Destexhe, A. (2003). Characterization of subthreshold voltage fluctuations in neuronal membranes. *Neural Comput, 15*(11), 2577-2618. doi:10.1162/089976603322385081

Saini, K. D., & Leppelsack, H. J. (1977). Neuronal arrangement in the auditory field L of the neostriatum of the starling. *Cell Tissue Res, 176*(3), 309-316.

Saini, K. D., & Leppelsack, H. J. (1981). Cell types of the auditory caudomedial neostriatum of the starling (Sturnus vulgaris). *J Comp Neurol, 198*(2), 209-229. doi:10.1002/cne.901980203

Sharpee, T. O. (2013). Computational identification of receptive fields. *Annu Rev Neurosci, 36*, 103-120. doi:10.1146/annurev-neuro-062012-170253

Tan, A. Y., Chen, Y., Scholl, B., Seidemann, E., & Priebe, N. J. (2014). Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature, 509*(7499), 226-229. doi:10.1038/nature13159

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 20*(6), 2315-2331.

Thompson, J. V., Jeanne, J., & Gentner, T. Q. (2010). *Local inhibition shapes the learned responses to song in NCM.* Paper presented at the Soc. Neurosci Absracts.

Wang, Y., Brzozowska-Prechtl, A., & Karten, H. J. (2010). Laminar and columnar auditory cortex in avian brain. *Proc Natl Acad Sci U S A, 107*(28), 12676-12681. doi:10.1073/pnas.1006645107

Woolley, S. M., Gill, P. R., Fremouw, T., & Theunissen, F. E. (2009). Functional groups in the avian auditory system. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(9), 2780-2793. doi:10.1523/JNEUROSCI.2042-08.2009

CHAPTER 4

Discussion to the Dissertation

Most studies to date that examine population spiking statistics do so without consideration for a selection process that would restrict the identity of neural populations in which spiking is measured. The experiments presented in this thesis address this limitation and lead to new perspectives on neural coding by featuring the synaptic response as a dependent variable in the examination of population spiking activity. Most generally, the implication that information about even the most complex stimuli is likely massively redundant and shared among the population at large scaffolds a sensory processing model in which flexible network re-organization - on short timescales and in a stimulus-specific way - support the complexities of spiking output observed in sensory cortex in response to learning, adaptation, attention, and context to meet the demands of an ever-changing environment. Accordingly, the results raise many more important questions. For example, although there is ample evidence in the literature that spiking correlations among randomly selected neurons affect stimulus discrimination, are these effects "seen by" downstream targets? Do mechanisms for flexible synaptic pooling exist that modulate the correlation structure of a neuron's inputs in a task-specific manner by stimulus context, attention, or goal? These questions remain largely unexplored, but broader applications of the approach taken in the preceding experiments to the mechanistic study of sensory processing and population coding are discussed further.

4.1 Correlation structures effecting population coding and behavior.

Measurements of covariance among spiking outputs seem to dominate much of the contemporary literature on population coding and sensory discrimination. Measured covariance is further dissociated and attributed to different sources: covariance occurring

over the stimulus set is calculated as “signal correlations”, while covariance unexplained by the stimulus is calculated as “noise correlations.” The distinction is germane because it is the relationship among different sources of co-variance –not covariance alone – that has been proposed to strongly influence population coding performance (Averbeck, Latham, & Pouget, 2006). Importantly, spiking correlations can be modulated by context (Kohn, Zandvakili, & Smith, 2009; Ruff & Cohen, 2014a, 2014b), attention (Cohen & Maunsell, 2009; Downer, Niwa, & Sutter, 2015; Mitchell, Sundberg, & Reynolds, 2009), learning (Gu et al., 2011; Jeanne, Sharpee, & Gentner, 2013) and adaptation (Adibi, McDonald, Clifford, & Arabzadeh, 2013; Gutnisky & Dragoi, 2008), effects that suggest a significant role in behavior. However the issue of whether correlations are actually behaviorally important for sensory discrimination has become a subject of heated debate – a debate that the approach taken in this thesis offers unique perspectives on (Averbeck et al., 2006; Averbeck & Lee, 2003; Dan, Alonso, Usrey, & Reid, 1998; deCharms & Merzenich, 1996; Eckhorn et al., 1988; Golledge et al., 2003; Gray, Konig, Engel, & Singer, 1989; Gray & Singer, 1989; Levine, 2004; Meister, Lagnado, & Baylor, 1995; Nirenberg, Carcieri, Jacobs, & Latham, 2001; Oram, Hatsopoulos, Richmond, & Donoghue, 2001; Panzeri, Pola, Petroni, Young, & Petersen, 2002; Petersen, Panzeri, & Diamond, 2001; Steinmetz et al., 2000; Vaadia et al., 1995).

First, the effects of correlations on population coding have predominantly been studied through calculating pairwise correlations. The spiking output of single neurons and pairs of neurons offers only a low-resolution sample of the population’s activity space, whereas subthreshold changes in membrane potential offer a more complete image of a network’s spiking ensemble. Accordingly, resolving the covariance structure of a

pre-synaptic population based on the synaptic response would provide significant advances in our understanding of mechanisms of coordinated activity that support sensory processing and behavior.

Second, spiking correlation alone is not a key indicator of connectivity among populations and does not imply a shared downstream target, yet models of stimulus encoding and discrimination processes are supported largely by results from studies examining the spiking of neurons *randomly* selected from large populations. Insofar as correlation structure effects stimulus discrimination it becomes even more relevant to study these relationships in neural populations whose spiking outputs are pooled by downstream targets.

The results of this thesis demonstrate that it is tractable to design experiments that address limitations inherent to current studies and potentially resolve conflicting results in the literature. By experimentally utilizing the synaptic response of individual neurons to target sets of neurons whose spiking activity reflected the inputs actually being pooled by downstream processes (represented in the synaptic response of a single neuron), we were able to examine the stimulus-driven dynamics of functional networks driven by complex, natural auditory signals (Chapter 2). A major challenge of such experiments in the future would be dissociating independent components of the pre-synaptic population covariance structure (such as signal and noise correlations) from the continuous one-dimensional synaptic response (either current or voltage over time).

4.2 Resolving correlation structure form the synaptic response

Ideally, we could estimate the correlation structure of each neuron's actual pre-synaptic population from its synaptic response and determine whether the same correlation structure was reflected among spiking outputs within the functional networks identified using that synaptic response (Chapter 2). Such a result would provide empirical support for using the synaptic response to make more direct hypotheses about mechanisms of synaptic integration shaping the correlation structure of pooled neural populations to support adaptive sensory-driven behaviors.

The ability to infer the statistics of the pre-synaptic population's coordinated spiking dynamics from a single neuron's synaptic response is well-supported, but the development of tools to do so is still an active area of research. The idea of examining the statistics of the membrane potential is not a new one (Levitan, Segundo, Moore, & Perkel, 1968; Rudolph & Destexhe, 2003). The spiking output (rate and timing) of a neuron is not just sensitive to the strength of its synaptic inputs but also to correlations among its inputs (Benucci, Verschure, & Konig, 2007; Bohte, Spekreijse, & Roelfsema, 2000; Kuhn, Aertsen, & Rotter, 2003) – evidence that input correlations modulate statistics of the synaptic response. Interaction among neurons in a pre-synaptic population affect the statistics of the membrane potential in predictable ways (Renart et al., 2010; Rudolph & Destexhe, 2003), and the application of mathematic techniques designed specifically to decipher spiking correlations from membrane potential signals have been implemented in several nervous systems (Tan and Priebe 2014, visual cortex; DeWeese and Zador 2006, auditory cortex; Renart et al 2010, cortical circuits) (DeWeese & Zador, 2006; Renart et al., 2010; Tan, Chen, Scholl, Seidemann, & Priebe, 2014).

In order to examine the relationship between signal and noise correlations, however, they must be dissociated from single membrane potential measurements. The approach used by Baudot et al (Baudot et al., 2013) seems to provide some means to this end. The component of the membrane potential under stimulus (the "signal" component) can be measured as the trial-averaged activity, and the component of the membrane potential unaccounted for by stimulus identity (the "noise" component) can be measured as the trial residuals (the differences between the membrane potential on each trial and the trial-averaged estimate). Insofar as the statistics of the membrane potential reflect correlations in the pre-synaptic population, then the statistics of the signal and noise component of the membrane potential should then be dissociable. These methods have not been used/tested directly to model pre-synaptic correlation structure, but the application seems tractable.

So far we have focused on the discussion of pairwise statistics of spiking covariance. High-order correlations are likely at least as influential as pairwise correlations in mechanisms of stimulus discrimination, but they have been notoriously difficult to capture with limited amounts of spiking data. Since the synaptic response of a neuron necessarily represents a complete sample of the pre-synaptic population, it provides a great low-dimensional representation of that high-dimensional space to enable less data-intensive methods for estimating high-order spiking covariance statistics. Higher order correlations have a profound impact on the spiking output of post-synaptic neurons (Benucci et al., 2007; Bohte et al., 2000; Kuhn et al., 2003), which provides evidence that input correlations modulate statistics of the synaptic response and that the latter could be used to estimate the former. Reimer et al 2013 establish one method for

inferring higher-order spike correlations from the measurement of subthreshold membrane potential (Reimer, Staude, Boucsein, & Rotter, 2013). This method is an extension of previous work establishing a cummulant-based approach for inferring the high-order spike correlations from the probability distribution of the net population spiking response (Staude, Rotter, & Grun, 2010).

Established mathematical tools are readily available for estimating the *magnitude* of correlations from the synaptic response, and it even seems reasonably tractable to dissociate signal and noise components from this estimate. One major caveat of this approach is that not all correlations created equal when it comes to enhancing stimulus discrimination in a population spike code (Moreno-Bote et al., 2014). The effect of correlations on stimulus discrimination depends on the *direction* of correlations, not their magnitude alone. Hu et al (Hu, Zylberberg, & Shea-Brown, 2014) demonstrate that coding performance depends on whether signal and noise correlations are both positive, both negative, or have opposite signs. Zylberberg et al (Zylberberg, Cafaro, Turner, Shea-Brown, & Rieke, 2016) demonstrate this principle in the retina where stimulus-specific noise correlations improve the decoding accuracy for oriented bars of light based on their relationship to the shape of the response space due to signal correlations. These findings have been extended to higher-order neural populations using more generalized language applicable beyond 2-dimensional stimulus space by establishing that stimulus decoding is enhanced specifically by noise correlations orthogonal to the surface of a stimulus space defined by the signal correlations (Franke et al., 2016). It is less clear how to estimate the "direction" of the noise correlations and the shape of the stimulus space from a synaptic response, but this ability to do so would drastically advance this field of study.

Performing experiments that provide empirical proof of some set of mathematical tools able to quantitatively capture that property of the correlations from the statistics of a synaptic response would be a useful project for someone to undertake.

4.3 Resolving cell-type specific contributions to the synaptic response.

Dissociating different presynaptic cell types (EX: inhibition or excitation) from the synaptic response is currently more tractable than dissociating distinct components of the measured correlation (EX: signal or noise) or the orthogonality between those components. Inhibitory and excitatory pre-synaptic populations may have different correlation structures, and examining this would yield mechanistic inferences about the integration of information-limiting correlations in downstream targets.

Inhibitory and excitatory currents can be measured directly using whole-cell voltage clamp techniques, or indirectly by manipulating current clamp across subsequent trials (Wilent & Contreras, 2005). In-vitro, the pharmalogical isolation of specific synaptic currents yields great mechanistic insights, however these pharmacological techniques are less applicable to the in-vivo experiments required to examine how sensory processing supports behavior. Another way to dissociate specific subpopulations from a single synaptic response (such as the membrane potential) is to leverage the varying kinetics among those specific types of synaptic inputs (Yasar, Wright, & Wessel, 2016). Both Kostuk et al (Kostuk, Toth, Meliza, Margoliash, & Abarbanel, 2012) and Meliza et al (Meliza et al., 2014) have demonstrated the feasibility of estimating numerous state variables and parameters of single neurons from their membrane potential

based on models under the Hudgkin-Huxley framework and with the injection of specific dynamical currents targeted at resolving these statistics.

Once the component of the synaptic response associated with a specific sub-population was isolated, one could estimate the correlation structure of that population from the statistics of the input using the same methods established to make these estimates from the probability distribution of the membrane potential (Berg & Ditlevsen, 2013; Borg-Graham, Monier, & Fregnac, 1996; Rudolph & Destexhe, 2003; Tan et al., 2014; Wilent & Contreras, 2005).

4.4 Real-time intracellular and extracellular interactions.

In the literature, one finds that trial-to-trial fluctuations of synaptic inputs significantly affect population coding, especially when considering the covariance of spiking underlying these fluctuations are considered. In the experiments of Chapter 2 we calculated the trial-averaged activity of single neurons and populations and therefore were restricted to examining stimulus-driven population dynamics and network re-organization. In order to resolve specificity in the fluctuation of functional networks across trials under the same stimulus conditions, one would need to simultaneously record population spiking and the single-neuron synaptic response. Single-trial analyses of trial correlations is mathematically tractable and would allow for experimental designs that leverage single-trial behavioral parameters (like attention or outcome choice) as independent variables in these analyses (Snyder, Morais, & Smith, 2013).

4.5 Population coding from the perspective of single spikes.

Contemporary sensory and motor system physiology has seen the advent of diversity in the perspectives with which to understand neural representation and neural coding. Particularly attractive given the collective results of this thesis are models that account for the spiking of each neuron in relation to the spiking of each other neuron.

In the end, an organism's decision to choose between one of several behavioral outcomes does not depend on the identity of a particular driving stimulus, per se, but rather on the specific emergent trajectory of coordinated population spiking activity relative to other possible trajectories. Several labs have made substantial progress in understanding the meaning of individual spikes relative to each other and the application of such models to the representation of a target stimulus (as opposed to the spike rate from each neuron reflecting some underlying probability distribution representing the target stimulus). In the model of Boerlin et al (Boerlin & Deneve, 2011; Boerlin, Machens, & Deneve, 2013) a spike may vary on a given trial relative to a specific external sensory stimulus or a specific behavioral output parameter, but instead of such variation being "noise," it is simply a reflection of an underlying network process that maintains a constant representation of a target stimulus without regard for absolute spike timing from each component neuron. Specifically, the subthreshold potential is actually an error function between the network spiking output and the target representation such that when that error function reaches threshold, a spike is elicited in that neuron that reduces its own error function (membrane potential) and the error function represented in the membrane potential of all connected neurons. The theoretical studies of Romaine Brette (Brette, 2012) propose the concept of a *synchrony receptive field*, in which

stimulus features are extracted by synchronous population events rather than by spike rates of individual neurons or populations of neurons.

From these perspectives it is particularly interesting to think about how the results from Chapter2 would differ if a single-trial version of the analysis were implemented. Recall that when correlations between the *average* stimulus responses were calculated, synchronous activity seemed to dominate locally (in time and in space). On a smaller timescale of a single trial, does synchronous activity dominate or do neurons contribute "for each other" so that on single trials one would not observe synchrony like we observed among the trial-averaged responses? The redundancy observed among the identified functional networks in Chapter 2 could support these mechanisms. In the current dataset, I do observe that on each single trial, the net population response is similar to the trial-averaged population response, but this does not mean that the same cell is contributing on every trial. Additionally, one might imagine that when comparing the synaptic response with the mean-field population spiking response that one could look at the correlation between these signals (as in Chapter 2) and identify moments when the synaptic response deviates from the mean-field response. These may signify the moments when that cell is contributing a unique component to the representation of the stimulus, aspects of which could be resolved by examining the network behavior at those time points relative to all other time points when the two signals are redundant.

To examine the mechanisms governing the relationship among individual spikes in the population and the synaptic response of a single neuron would require additional effort to accomplish simultaneous recordings of these signals, but would yield results that

contribute significantly to our understanding of how neural circuits represent sensory stimuli in guiding behavioral outputs.

4.6 Caveats

One of the main caveats of the current study, and many of the proposed future directions implementing the same methods, is error in estimates of redundancy and correlation due to the spike sorting process. This is a known limitation and not one specific to the current study (Schulz, Sahani, & Carandini, 2015). Reporting spike isolation quality metrics (such as ISI violations and measures of the shape of distributions generated by the spike sorting algorithm implemented) is one way to at least provide that information as reference for future studies able to resolve these issues with advancements in recording and/or spike isolation techniques. Another way to address this before publication of the manuscript from Chapter 2 would be to compare (on a trial by trial basis) the difference in correlation between silence and stimulus among simultaneously recorded units.

4.7 Précis

A traditional approach in sensory physiology is to leverage the parameterization of stimulus sets amenable to dimension reduction, rather than embracing the full suite of complexity inherent to ethological signals supporting adaptive animal behavior. The limitations of this are apparent in the difference between the spiking and synaptic behavior under stimulus conditions with different statistics (Baudot et al., 2013; Theunissen, Sen, & Doupe, 2000). In this study, I have embraced the complexity of

natural Starling vocalizations and instead acted creatively on the perspective lens through which I have examined the resulting neural activity. What has benefited from this has been appreciation of the rich information available in the synaptic responses underlying neural representations traditionally regarded and modeled as sparse and independent, and the implementation of the synaptic response as a lens through which to examine population spiking dynamics.

4.8 References

Adibi, M., McDonald, J. S., Clifford, C. W., & Arabzadeh, E. (2013). Adaptation improves neural coding efficiency despite increasing correlations in variability. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 33*(5), 2108-2120. doi:10.1523/JNEUROSCI.3449-12.2013

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci, 7*(5), 358-366. doi:10.1038/nrn1888

Averbeck, B. B., & Lee, D. (2003). Neural noise and movement-related codes in the macaque supplementary motor area. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 23*(20), 7630-7641.

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., & Fregnac, Y. (2013). Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. *Front Neural Circuits, 7*, 206. doi:10.3389/fncir.2013.00206

Benucci, A., Verschure, P. F., & Konig, P. (2007). Dynamical features of higher-order correlation events: impact on cortical cells. *Cogn Neurodyn, 1*(1), 53-69. doi:10.1007/s11571-006-9000-y

Berg, R. W., & Ditlevsen, S. (2013). Synaptic inhibition and excitation estimated via the time constant of membrane potential fluctuations. *J Neurophysiol, 110*(4), 1021-1034. doi:10.1152/jn.00006.2013

Boerlin, M., & Deneve, S. (2011). Spike-based population coding and working memory. *PLoS Comput Biol, 7*(2), e1001080. doi:10.1371/journal.pcbi.1001080

Boerlin, M., Machens, C. K., & Deneve, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput Biol, 9*(11), e1003258. doi:10.1371/journal.pcbi.1003258

Bohte, S. M., Spekreijse, H., & Roelfsema, P. R. (2000). The effects of pair-wise and higher order correlations on the firing rate of a post-synaptic neuron. *Neural Comput, 12*(1), 153-179.

Borg-Graham, L., Monier, C., & Fregnac, Y. (1996). Voltage-clamp measurement of visually-evoked conductances with whole-cell patch recordings in primary visual cortex. *J Physiol Paris, 90*(3-4), 185-188.

Brette, R. (2012). Computing with neural synchrony. *PLoS Comput Biol, 8*(6), e1002561. doi:10.1371/journal.pcbi.1002561

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

Dan, Y., Alonso, J. M., Usrey, W. M., & Reid, R. C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat Neurosci, 1*(6), 501-507. doi:10.1038/2217

deCharms, R. C., & Merzenich, M. M. (1996). Primary cortical representation of sounds by the coordination of action-potential timing. *Nature, 381*(6583), 610-613. doi:10.1038/381610a0

DeWeese, M. R., & Zador, A. M. (2006). Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 26*(47), 12206-12218. doi:10.1523/JNEUROSCI.2813-06.2006

Downer, J. D., Niwa, M., & Sutter, M. L. (2015). Task engagement selectively modulates neural correlations in primary auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 35*(19), 7565-7574. doi:10.1523/JNEUROSCI.4094-14.2015

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biol Cybern, 60*(2), 121-130.

Franke, F., Fiscella, M., Sevelev, M., Roska, B., Hierlemann, A., & da Silveira, R. A. (2016). Structures of Neural Correlation and How They Favor Coding. *Neuron, 89*(2), 409-422. doi:10.1016/j.neuron.2015.12.037

Golledge, H. D., Panzeri, S., Zheng, F., Pola, G., Scannell, J. W., Giannikopoulos, D. V., . . . Young, M. P. (2003). Correlations, feature-binding and population coding in primary visual cortex. *Neuroreport, 14*(7), 1045-1050. doi:10.1097/01.wnr.0000073681.00308.9c

Gray, C. M., Konig, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338*(6213), 334-337. doi:10.1038/338334a0

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci U S A, 86*(5), 1698-1702.

Gu, Y., Liu, S., Fetsch, C. R., Yang, Y., Fok, S., Sunkara, A., . . . Angelaki, D. E. (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron, 71*(4), 750-761. doi:10.1016/j.neuron.2011.06.015

Gutnisky, D. A., & Dragoi, V. (2008). Adaptive coding of visual information in neural populations. *Nature, 452*(7184), 220-224. doi:10.1038/nature06563

Hu, Y., Zylberberg, J., & Shea-Brown, E. (2014). The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. *PLoS Comput Biol, 10*(2), e1003469. doi:10.1371/journal.pcbi.1003469

Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron, 78*(2), 352-363.

Kohn, A., Zandvakili, A., & Smith, M. A. (2009). Correlations and brain states: from electrophysiology to functional imaging. *Curr Opin Neurobiol, 19*(4), 434-438. doi:10.1016/j.conb.2009.06.007

Kostuk, M., Toth, B. A., Meliza, C. D., Margoliash, D., & Abarbanel, H. D. (2012). Dynamical estimation of neuron and network properties II: Path integral Monte Carlo methods. *Biol Cybern, 106*(3), 155-167. doi:10.1007/s00422-012-0487-5

Kuhn, A., Aertsen, A., & Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput, 15*(1), 67-101. doi:10.1162/089976603321043702

Levine, M. W. (2004). The potential coding utility of intercell cross-correlations in the retina. *Biol Cybern, 91*(3), 182-187. doi:10.1007/s00422-004-0492-4

Levitan, H., Segundo, J. P., Moore, G. P., & Perkel, D. H. (1968). Statistical analysis of membrane potential fluctuations. Relation with presynaptic spike train. *Biophys J, 8*(11), 1256-1274. doi:10.1016/S0006-3495(68)86554-3

Meister, M., Lagnado, L., & Baylor, D. A. (1995). Concerted signaling by retinal ganglion cells. *Science, 270*(5239), 1207-1210.

Meliza, C. D., Kostuk, M., Huang, H., Nogaret, A., Margoliash, D., & Abarbanel, H. D. (2014). Estimating parameters and predicting membrane voltages with conductance-based neuron models. *Biol Cybern, 108*(4), 495-516. doi:10.1007/s00422-014-0615-5

Mitchell, J. F., Sundberg, K. A., & Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron, 63*(6), 879-888. doi:10.1016/j.neuron.2009.09.013

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., & Pouget, A. (2014). Information-limiting correlations. *Nat Neurosci, 17*(10), 1410-1417. doi:10.1038/nn.3807

Nirenberg, S., Carcieri, S. M., Jacobs, A. L., & Latham, P. E. (2001). Retinal ganglion cells act largely as independent encoders. *Nature, 411*(6838), 698-701. doi:10.1038/35079612

Oram, M. W., Hatsopoulos, N. G., Richmond, B. J., & Donoghue, J. P. (2001). Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J Neurophysiol, 86*(4), 1700-1716.

Panzeri, S., Pola, G., Petroni, F., Young, M. P., & Petersen, R. S. (2002). A critical assessment of different measures of the information carried by correlated neuronal firing. *Biosystems, 67*(1-3), 177-185.

Petersen, R. S., Panzeri, S., & Diamond, M. E. (2001). Population coding of stimulus location in rat somatosensory cortex. *Neuron, 32*(3), 503-514.

Reimer, I. C., Staude, B., Boucsein, C., & Rotter, S. (2013). A new method to infer higher-order spike correlations from membrane potentials. *J Comput Neurosci, 35*(2), 169-186. doi:10.1007/s10827-013-0446-8

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., & Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science, 327*(5965), 587-590. doi:10.1126/science.1179850

Rudolph, M., & Destexhe, A. (2003). Characterization of subthreshold voltage fluctuations in neuronal membranes. *Neural Comput, 15*(11), 2577-2618. doi:10.1162/089976603322385081

Ruff, D. A., & Cohen, M. R. (2014a). Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci, 17*(11), 1591-1597. doi:10.1038/nn.3835

Ruff, D. A., & Cohen, M. R. (2014b). Global cognitive factors modulate correlated response variability between V4 neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(49), 16408-16416. doi:10.1523/JNEUROSCI.2750-14.2014

Schulz, D. P., Sahani, M., & Carandini, M. (2015). Five key factors determining pairwise correlations in visual cortex. *J Neurophysiol, 114*(2), 1022-1033. doi:10.1152/jn.00094.2015

Snyder, A. C., Morais, M. J., & Smith, M. A. (2013). Variance in population firing rate as a measure of slow time-scale correlation. *Front Comput Neurosci, 7*, 176. doi:10.3389/fncom.2013.00176

Staude, B., Rotter, S., & Grun, S. (2010). CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains. *J Comput Neurosci, 29*(1-2), 327-350. doi:10.1007/s10827-009-0195-x

Steinmetz, P. N., Roy, A., Fitzgerald, P. J., Hsiao, S. S., Johnson, K. O., & Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature, 404*(6774), 187-190. doi:10.1038/35004588

Tan, A. Y., Chen, Y., Scholl, B., Seidemann, E., & Priebe, N. J. (2014). Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature, 509*(7499), 226-229. doi:10.1038/nature13159

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 20*(6), 2315-2331.

Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., & Aertsen, A. (1995). Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature, 373*(6514), 515-518. doi:10.1038/373515a0

Wilent, W. B., & Contreras, D. (2005). Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. *Nat Neurosci, 8*(10), 1364-1370.

Yasar, T. B., Wright, N. C., & Wessel, R. (2016). Inferring presynaptic population spiking from single-trial membrane potential recordings. *J Neurosci Methods, 259*, 13-21. doi:10.1016/j.jneumeth.2015.11.019

Zylberberg, J., Cafaro, J., Turner, M. H., Shea-Brown, E., & Rieke, F. (2016). Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. *Neuron, 89*(2), 369-383. doi:10.1016/j.neuron.2015.11.019