CORRELATIONS IN POPULATIONS OF SENSORY NEURONS

Dmitry Lyamzin

Dissertation at the Graduate School of Systemic Neurosciences at the

Ludwig-Maximilians-Universität München

Graduate School of
Systemic Neurosciences

LMU Munich

Supervisor: Dr Nicholas A Lesica

2nd reviewer: Dr Christian Leibold

Datum mündlichen Prüfung: 25.09.2014

Table of Contents

I Summary

This thesis is composed of three completed research projects, two of which have been published, and the third one is currently being submitted for publication.

The Introduction (Chapter II) gives an overview of the role of correlations in sensory processing, and shows the need for a framework that would be able to fit, analyze, and simulate the activity of populations of sensory neurons. At the end of the introduction we formulate the requirements to the model.

In Chapter III "Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations" we develop a basic model based on dichotomized Gaussian, that allows us independent manipulation of all the mentioned parameters of neuronal activity, including independent manipulation of signal and noise correlations.

In Chapter IV "Analysis and modeling of variability and covariability of population spike trains across multiple timescales" we investigate the ability of the model to reproduce and manipulate temporal correlations, which we introduce by Gaussian conditioning (MacKay, 2003). This paper also builds a link between our framework and articles that quantify cell responses, variability, and signal and noise correlations based on cell firing rates rather than probability of their spiking at a certain time during the presentation of the stimulus.

In Chapter V "Nonlinear transfer of signal and noise correlations in cortical networks" we use the framework to determine the origin of the correlations observed in extracellular recordings in vivo. We show that the transfer of correlations happening in pairs of cells, i.e. the change in the magnitude of correlations between input currents and output spike trains is nonlinear, with these nonlinearities defined by input current variances and correlations.

We show on whole-cell recordings in slices that dichotomized Gaussian framework predicts input signal to noise ratio and input signal and noise correlations with high precision based only on the observed spike trains. Using this fact we infer input signal to noise ratios and input correlations in the inputs to the pairs of cells in the recordings from primary auditory cortex of Mongolian gerbils. Finally, we look at the correlation structure observed in responses to the sounds with varying location.

All chapters are given a brief introduction, and the contributions to each of the chapters are stated.

II Introduction

With the development of new methods of simultaneous data acquisition from populations of neurons (O'Keefe & Recce, 1993; Wilson & McNaughton, 1993), and with the advances in mathematical and software tools that allow separating responses of individual neurons from a recorded population response (K D Harris, Henze, Csicsvari, Hirase, & Buzsáki, 2000), there is a known gap in the body of papers on the statistical analysis and modeling of population spike trains.

Despite the fact that a large number of works that discuss the implications of population coding relies on data from sensory systems (Ohiorhenuan et al., 2010; Pillow et al., 2008), few research groups make an accent on the modeling of spiking patterns of specifically sensory neuronal populations.

The main feature that distinguishes the activity of populations of sensory neurons in the brain is that during active sensory processing it carries information about the external stimulus (Abbott & Dayan, 2001)d, whereas the activity of other neuronal populations is to a larger extent endogenous.

Because of that, the statistics of the external stimulus is partially reflected in the statistics of the sensory neuronal response, and a share of correlations observed in neuronal spiking arises as the result of the structure of the stimulus that has correlations in time and feature space.

The part of correlations in neural activity that arises due to the sensory drive is called *signal correlations*.

Correlations in spiking that are caused by any other factors, not related to the stimulus, are called *noise correlations*.

There are numerous sources of such stimulus-unrelated correlations, such as the activity of an upstream neuron projecting to a pair of the observed cells (Moore, Segundo, Perkel, & Levitan, 1970), top-down firing rate modulation (Moran & Desimone, 1985), changes in brain state (Poulet & Petersen, 2008) or behavior (Bair, Zohary, & Newsome, 2001) and anesthesia (Marguet & Harris, 2011). Generally, the relationship between signal and noise correlations (that we will be referring to as *correlation structure*) can be indicative of a particular behavioral state or can characterize the ongoing sensory processing.

Currently there is not always a clear understanding of 1) why the changes in correlation structure occur in a given behavioral condition, and 2) whether these changes at all matter, despite the significant progress in both directions.

Both of these questions are extremely broad. The first one may require understanding the mechanisms of sensory processing in all modalities as well as the general principles of correlated spiking. The second question requires estimating the efficiency of the neural code in behaving animals or animals engaged in sensory processing.

In the next two sections we will briefly touch upon works that analyze the role of signal and noise correlations in conjunction with changes in other properties of neural response (firing rates, variability), as well as high-level changes, e.g. behavior and learning, and formulate the information-theoretic approach that can be used as a zero-assumption framework for the analysis of neural code.

At the end of the introduction we will set the goals of current thesis given the scope of these two approaches.

Significance of signal and noise correlations from the information-theoretic perspective

The distinction between signal and noise correlations is central from the viewpoint of the information theoretic analysis. The relationship between these two kinds of correlations ultimately determines the *mutual information* in the observed population spike trains (Abbott & Dayan, 2001), an objective measure of information carried by the responses of any neuronal population:

$$I(R,S) = \sum_{R,S} P(r,s) log\left(\frac{P(r,s)}{P(r)P(s)}\right)$$

Specifically, an expression for mutual information contained in the responses of a neuronal population to a set of stimuli $I(R,S)$ can be written out as a sum of four terms (Pola, Thiele, Hoffmann, & Panzeri, 2003).

$$I(R,S) = I_{lin} + I_{sig-sim} + I_{cor-ind} + I_{cor-dep}$$

The linear term I_{lin} gives the amount of information conveyed by all cells had they all been independent. The latter three terms depend on the signal and noise correlations: the term $I_{sig-sim}$ arises due to the similarity of signal responses between different cells and is always non-positive, the stimulus-independent correlation component $I_{cor-ind}$ depends on the signal similarity and existing correlations when the latter are not modulated by the stimuli, and the term $I_{cor-dep}$ that is always non-negative arises due to the stimulus-dependency of noise correlations.

The three latter terms can increase or decrease mutual information depending on the correlation structure of the network. Moreover, the sign of the contribution of the third component is not defined a priori and will depend on the signs of the signal and noise

correlations: both positive or both negative signal and noise correlations for a given pair will lead to redundancy and to the loss of information, while the opposite signs of signal and noise correlations will lead to synergy and to the increase of information.

Information theoretic analysis requires calculating exact probabilities of spiking in responses to a stimulus $P(r|s)$, which is feasible when responses to a range of stimuli are sampled extensively. Moreover, some approaches like GLM allow one to predict responses to unobserved stimuli given that the parameters of GLM for each neuron are estimated accurately (Chornoboy, Schramm, & Karr, 1988; Pillow et al., 2008). Even despite possible inconsistencies due to often highly nonlinear behavior of neurons and the necessity of sampling responses to a wide range of stimuli, the ability to predict responses to novel stimuli is an undisputable advantage of GLM.

What GLM is not capable of doing is to directly control the parameters of population spiking, including PSTHs of individual cells and signal and noise correlations. Therefore it would be impossible to sample spike trains of a pair of cells that would have for example the same PSTHs as a pair of cells observed experimentally but different noise correlations.
However such flexibility in manipulating the statistics of neuronal responses would be desirable for a systematic study of the factors affecting mutual information.

In this thesis we will provide the tools for sampling spike trains with desired statistics and for modifying the statistics of the spike trains observed experimentally.

Although primary research has been done towards the use of our framework in MI analysis (see Discussion), we did not get novel results from that analysis, most likely due to the specifics of the data we got from the local populations in A1.

Changes in the structure of correlations and neural coding

Currently there is an extensive body of works that regard correlations as signatures of particular modes of sensory processing and behavior, directed toward the inference of coding principles rather than the estimation of total information content in the activity of a neuronal population.

Noise correlations have been shown to change as the result of learning, either reducing overall (Gu et al., 2011), or being a target for learning themselves (Jeanne, Sharpee, & Gentner, 2013). The balance between signal and noise correlations shifts with attention (Cohen & Maunsell, 2009; Mitchell, Sundberg, & Reynolds, 2009), as a result of adaptation (Gutnisky & Dragoi, 2008), and with changes in brain state (Marguet & Harris, 2011).

Significant progress has been made in finding the potential mechanisms of top-down attention and associating them with the emergence of a synchronized state in a local neuronal population (Kenneth D Harris & Thiele, 2011; Zagha, Casale, Sachdev, McGinley, & McCormick, 2013).

The study of these top-down and intra-cortical mechanisms of shaping correlation structure lies largely beyond the scope of current work, although we discuss them briefly in the context of the results of Chapter V.

Another source of neural noise is periphery. Such noise can carry information about certain features of external stimulus and is also important for understanding basic principles of sensory coding.

For example decreasing contrast of a visual stimulus results in temporally imprecise firing of LGN and V1 neurons (Desbordes et al., 2008; Sadagopan & Ferster, 2012). Because the relative timing of spikes remains relatively precise across the population, such reduced precision of absolute timing results in increased noise correlations.

An even more powerful example is presented in Chapter V where we show a change in spike time reliability of A1 neurons without changes in firing rate in case when the position of the sound source is altered. In that case again we see simultaneous changes in noise correlations.

It is not quite clear however if such changes in the variability of individual neurons would *necessarily* result in changes in correlated noise.

Importantly for the analysis that we carry out in this Thesis, it has been shown that the correlations observed also depend on the firing rates of a pair of neurons (de la Rocha, Doiron, Shea-Brown, Josić, & Reyes, 2007), even when they are corrected for the change in chance correlations. Thus it is not only the activity of the local network, or top-down modulations, but also the individual properties of neurons that have to be taken into account when explaining the experimentally observed correlations.

Finally, several authors observe both top-down modulations of noise correlations due to attention and changes in individual cell variability occurring simultaneously (Cohen & Maunsell, 2009), and it is not well understood how these two factors interact and shape the observed noise correlations in population spike trains.

From these examples it becomes apparent that temporal precision of spiking, or indeed lack thereof should be reflected in the measure of neuronal noise.

Thesis objectives

The main goal of this thesis is to develop a model of a population of sensory neurons, and to apply this model to the analysis of the role of pairwise signal and noise correlations in sensory coding.

For an effective use of our model, including information theoretic analysis, we should be able to sample population spike trains with given parameters and arbitrarily manipulate these parameters if necessary.

In order to study coding of sensory features we need to be able to fit model to experimentally obtained population spike trains, matching PSTH and trial-to-trial variability of individual cells and signal and noise correlations for pairs of cells. As noted before, the degree of variability should be estimated based on timing precision.

That way we will cover parameters of population activity that reflect its properties fully excluding high order correlations and up or down states.

As pointed out above (Cohen & Maunsell, 2009; de la Rocha et al., 2007), correlations are found to depend on or change simultaneously with single cell parameters. Noise correlations are often shown to be positively related to signal correlations (Averbeck & Lee, 2003; Bair et al., 2001; Jung, Qin, Lee, & Mook-Jung, 2000; Kohn & Smith, 2005; Liu, Gu, DeAngelis, & Angelaki, 2013; Martin & Schröder, 2013). Using our model we would like to look into these dependencies and find their mechanisms.

III Modeling Population Spike Trains with Specified Time-Varying Spike Rates, Trial-to-Trial Variability, and Pairwise Signal and Noise Correlations

In this chapter, we formulate the dichotomized Gaussian (DG) model as a means for the generation and analysis of spike trains with particular single cell and pairwise properties. We show that PSTHs and firing rates for single cells, and signal and noise correlations for pairs of cells are matched exactly to the precision that is allowed by the initial data.

The model can generate population activity if no experimental data is given a priori, using the desired parameters of spike trains: firing rates and trial-to-trial variability of individual cells, and signal and noise correlations of pairs of cells. Additionally, we demonstrate how matching temporal cross- and auto-correlations is possible in the DG framework.

Sampling multivariate binary vectors with known joint probabilities can be done algorithmically using several different methods (e.g. (Devroye, 1986) Inversion method; (Headrick, 2002)). An overview of methods for generating correlated binary variables with certain means and correlation coefficients is given by (Qaqish, 2003).

One of these methods, called Dichotomized Gaussian, is of particular interest to us due to its analytical tractability (Cox & Wermuth, 2002; Emrich & Piedmonte, 2013; Gutnisky & Josić, 2010; Macke, Berens, Ecker, Tolias, & Bethge, 2009; Pearson, 1909). Dichotomized Gaussian is a generative model that is based on a thresholded (dichotomized) latent Gaussian. By manipulating means, variances, and correlations between the components of a multivariate Gaussian, as well as thresholds for each component, it is possible to obtain the required values of mean and pairwise correlations for the corresponding components of the observed multivariate binary variable.

Authors' contributions

The work was carried out under the supervision of Nicholas Lesica. The research was conceived and designed by NL, JM, and DL; NL and DL performed research; NL and DL wrote the paper. Research results were presented at *Mathematical Neuroscience* in April 2010 in Edinburgh, UK.

The paper has been accepted in *Frontiers in Computational Neuroscience* under the following reference:

"Modeling Population Spike Trains with Specified Time-Varying Spike Rates, Trial-to-Trial Variability, and Pairwise Signal and Noise Correlations", Dmitry R Lyamzin, Jakob H Macke, Nicholas A Lesica, *Front Comput Neurosci*. 2010; 4: 144. Accepted October 5, 2010.

frontiers in
COMPUTATIONAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 15 November 2010
doi: 10.3389/fncom.2010.00144

Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations

Dmitry R. Lyamzin[1†], Jakob H. Macke[2,3†] and Nicholas A. Lesica[1*†]

[1] *Division of Neurobiology, Department of Biology II, Ludwig-Maximilians-University Munich, Martinsried, Germany*
[2] *Computational Vision and Neuroscience Group, Max Planck Institute for Biological Cybernetics, Tübingen, Germany*
[3] *Werner Reichhardt Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany*

Edited by:
Klaus R. Pawelzik, University of Bremen, Germany

Reviewed by:
Udo Ernst, University of Bremen, Germany
Kresimir Josic, University of Houston, USA

****Correspondence:***
Nicholas A. Lesica, Ear Institute, University College London, 332 Gray's Inn Rd., London WC1X 8EE, UK.
e-mail: n.lesica@ucl.ac.uk

†Current address:
Dmitry R. Lyamzin and Nicholas A. Lesica, Ear Institute, University College London, London, UK.
Jakob H. Macke, Gatsby Computational Neuroscience Unit, University College London, London, UK.

As multi-electrode and imaging technology begin to provide us with simultaneous recordings of large neuronal populations, new methods for modeling such data must also be developed. Here, we present a model for the type of data commonly recorded in early sensory pathways: responses to repeated trials of a sensory stimulus in which each neuron has it own time-varying spike rate (as described by its PSTH) and the dependencies between cells are characterized by both signal and noise correlations. This model is an extension of previous attempts to model population spike trains designed to control only the total correlation between cells. In our model, the response of each cell is represented as a binary vector given by the dichotomized sum of a deterministic "signal" that is repeated on each trial and a Gaussian random "noise" that is different on each trial. This model allows the simulation of population spike trains with PSTHs, trial-to-trial variability, and pairwise correlations that match those measured experimentally. Furthermore, the model also allows the noise correlations in the spike trains to be manipulated independently of the signal correlations and single-cell properties. To demonstrate the utility of the model, we use it to simulate and manipulate experimental responses from the mammalian auditory and visual systems. We also present a general form of the model in which both the signal and noise are Gaussian random processes, allowing the mean spike rate, trial-to-trial variability, and pairwise signal and noise correlations to be specified independently. Together, these methods for modeling spike trains comprise a potentially powerful set of tools for both theorists and experimentalists studying population responses in sensory systems.

Keywords: population, correlation, noise correlation, simulation, model

INTRODUCTION

Correlated spiking activity in neuronal populations has been a subject of intense theoretical and experimental research over the past several decades, and the importance of correlations has been demonstrated in a number of contexts, including plasticity and information processing (for a recent review, see Averbeck et al., 2006). Recent advances in experimental technology have finally made it possible to observe the activity of large neuronal populations simultaneously. In order to take full advantage of these advances, new methods for the analysis and modeling of population activity must also be developed.

A number of methods exist for modeling correlated population spike trains in which some fraction of the input driving the activity of each neuron is shared with other neurons, including integrate-and-fire models and other spiking models with correlated input currents or synaptic conductances (Destexhe and Pare, 1999; Feng and Brown, 2000; Song and Abbott, 2001; Stroeve and Gilen, 2001; Salinas and Sejnowski, 2002; Dorn and Ringach, 2003; Gutig et al., 2003; Galan et al., 2006; De La Rocha et al., 2007; Shea-Brown et al., 2008; Tchumatchenko et al., 2008), stochastic spiking models with correlated rate functions (Galan et al., 2006; Brette, 2009; Krumin and Shoham, 2009), and models based on a dichotomized Gaussian (DG) framework (Macke et al., 2009; Gutnisky and Josic, 2010). There are also a variety of methods for capturing precise synchrony between neurons through explicit sharing of spikes (Kuhn et al., 2003; Galan et al., 2006; Niebur, 2007; Brette, 2009) and several models based on statistical frameworks such as maximum entropy (Schneidman et al., 2006; Shlens et al., 2006; Roudi et al., 2009).

All of the approaches described above are designed to capture and/or control the total correlation between spike trains and, as such, are of limited utility in the context of early sensory systems where it is important to separate internal network correlations from those due to the external stimulus. In this paper, we propose a framework designed specifically to model spike trains in which the total correlation can be separated into signal and noise components. If responses to repeated trials of an identical sensory stimulus are observed, the signal correlation, which reflects both correlation in the stimulus itself and similarities in neurons' preferred stimulus features, will be evident in the fraction of the response that is repeatable from trial-to-trial. Noise correlation, which results from the activity of network and intrinsic cellular mechanisms, will be evident in the fraction of the response that is variable from trial-to-trial (note that the term noise correlation is not meant to imply that the activity underlying this correlation is unimportant, but simply that is not directly dependent on the stimulus).

For modeling the population spike trains of early sensory neurons, another class of methods based on generalized linear models (GLMs) has been developed (Chornoboy et al., 1988; Paninski, 2004; Kulkarni and Paninski, 2007; Paninski et al., 2007; Pillow et al., 2008). In its typical formulation, the GLM is parameterized by a series of filters that relate the time-varying spike rate in one neuron to the sensory stimulus and the responses of other neurons. This formulation has the great strength that once the filter parameters have been estimated, the model can be used not only to simulate responses that match those measured experimentally, but also to simulate responses to novel stimuli. However, this generality comes at a cost: specifying the filters requires the estimation of a large number of parameters and, thus, a large amount of experimental data – much more than is necessary for a model designed only to simulate responses to the same stimuli that have been tested experimentally. It is possible to formulate alternatives to the typical GLM that require less experimental data by forgoing the ability to predict responses to novel stimuli and parameterizing the time-varying firing rate in response to a particular stimulus directly. However, even in this formulation, the GLM lacks a critical property: it does not enable straightforward specification or manipulation of one response property independent of the others (Krumin and Shoham, 2009; Toyoizumi et al., 2009).

In the model we present below, the time-varying spike rate, trial-to-trial variability, and pairwise signal and noise correlations can be matched to those measured experimentally, and the noise correlations can be manipulated without changes in the signal correlations or the single-cell properties. The model is an extension of previous attempts to model population spike trains as DGs (Emrich and Piedmonte, 1991; Cox and Wermuth, 2002; Macke et al., 2009; Gutnisky and Josic, 2010). In our model, the response of each cell is a binary vector determined by the thresholded sum of two inputs: a signal, which is the same for each trial of a given stimulus, and a noise, which is different for each trial, both of which can be correlated across neurons. In the first part of the paper, we show how the model parameters can be estimated from experimental data and used to simulate spike trains with properties that match those measured experimentally. We also demonstrate how the model parameters can be manipulated to obtain spike trains with arbitrary pairwise noise correlations without changes in single-cell properties. In the second part of the paper, we describe a general form of the model that can be used model spike trains with arbitrary single-cell properties and pairwise correlations.

All of the Matlab code required to perform the analyses described in this paper is available for download at http://www.ucl.ac.uk/ear/research/lesicalab.

A MODEL FOR SIMULATING AND MANIPULATING EXPERIMENTALLY RECORDED POPULATION SPIKE TRAINS

SINGLE-CELL RESPONSES

To represent a set of spike times from a single cell on a single trial $i \in \{1,2,\ldots,I\}$ of a particular stimulus, we discretize time into $n \in \{1,2,\ldots,N\}$ bins of length Δ and set $r_i[n] = 1$ if a spike occurs in bin n on trial i, and $r_i[n] = 0$ otherwise. In general, we assume that Δ is small enough that no more than one spike occurs in any given bin. Based on the responses to all trials r (an $N \times I$ binary matrix), we can define several quantities of interest:

Mean spike rate	$r_0 = \langle r \rangle_{n,i}$	(scalar)
Time-varying spike rate (PSTH)	$\bar{r} = \langle r \rangle_i$	(N-dimensional vector)
Spike train signal to noise ratio	$SNR = \dfrac{\mathrm{var}(\bar{r})}{\langle \mathrm{var}(\xi_i) \rangle_i}$	(scalar)

where $\xi_i = \bar{r} - r_i$ is the residual on trial i

Note that we use the notation $\langle \cdot \rangle_x$ to represent the expectation over all possible values of x, $\langle \cdot \rangle_{x,y}$ to represent the expectation over all possible values of x followed by the expectation of all possible values of y, and $\langle \cdot \rangle_{x \neq y}$ to represent the expectation over all possible combinations of x and y in which their values are not equal. We chose to use the above definition of SNR as the measure of trial-to-trial variability because is it commonly used in early sensory systems (Borst and Theunissen, 1999). One important property of this measure that should be noted is that its value is dependent on the bin size Δ. Thus, all of the computations described below for fitting model parameters must be repeated if the bin size is changed.

We model the response as a dichotomized sum of a deterministic "signal" and Gaussian "noise"

$$r_i[n] = \begin{cases} 1, & (s[n] + z_i[n]) > 0 \\ 0, & (s[n] + z_i[n]) \leq 0 \end{cases} \quad (1)$$

Where $r_i[n]$ is the response in time bin N on trial i, s (an N-dimensional vector) is the same on every trial and $z \sim \mathcal{N}(0, 1)$ (an N-dimensional vector) is different on every trial [note that neither s nor z are intended to correspond directly to any intracellular quantities]. Given the experimentally recorded responses of a cell, we wish to simulate responses with the same PSTH $\bar{r}$. This can be done by solving

$$\bar{r}[n] = \langle r_i[n] \rangle_i = \Phi(s[n], 1) \quad (2)$$

for $s[n]$ in each bin, where $\Phi(x, \sigma^2)$ is the CDF for a Gaussian with zero mean and variance σ^2 evaluated at x. Equation 2 is easily solved numerically, as the function is monotonic and has unique level crossings. It is clear from Eq. 2 that the choice of one for the variance of z is somewhat arbitrary; for any finite value of the variance of z, an $s[n]$ can be found to achieve any desired value of $\bar{r}[n]$. Note that if $\bar{r}[n] = 0$ or 1, then $s[n]$ must be either $-\infty$ or $+\infty$. If finite values of $s[n]$ are desired, then $\bar{r}[n]$ can be constrained to the interval $[1/I, 1 - 1/I]$ before solving Eq. 2.

Importantly, since this approach matches $\bar{r}$ exactly, it will also match the mean spike probability r_0 and the spike train signal to noise ratio SNR, as both can be uniquely defined in terms of the PSTH $\bar{r}$:

$$r_0 = \langle \bar{r} \rangle_n$$

$$\mathrm{SNR} = \frac{\mathrm{var}(\bar{r})}{\langle \mathrm{var}(\xi_i) \rangle_i} = \frac{\mathrm{var}(\bar{r})}{\langle \mathrm{var}(\bar{r} - r_i) \rangle_i} = \frac{\mathrm{var}(\bar{r})}{\mathrm{var}(\bar{r}) + \langle \mathrm{var}(r_i) \rangle_i - 2\langle \mathrm{cov}(\bar{r}, r_i) \rangle_i}$$

where, because r_i is binary,

$$\langle \mathrm{var}(r_i) \rangle_i = r_0(1 - r_0) \quad \text{and} \quad \langle \mathrm{cov}(\bar{r}, r_i) \rangle_i = \langle \bar{r}^2 \rangle_n - r_0^2.$$

Matching $\bar{r}$ exactly will also match the mutual information transmitted by single spikes (Brenner et al., 2000). Note that if it is not necessary to match the bin to bin spike probabilities of the experimental response, but only the distribution of overall spike counts, a reduced model can be used (Macke et al., 2009).

To demonstrate the utility of this model, we first generated responses using Eq. 1 with a variety of different signals, and then attempted to reproduce the model responses after estimating s using Eq. 2. Typical results are shown in **Figure 1A**. For uniform random, sine wave, and square wave signals, the PSTH and, consequently, r_0 and SNR of the responses simulated with the estimated s closely match those of the original model generated data.

Next, we tested the model's ability to reproduce the single-cell properties of experimentally recorded responses. **Figure 1B** shows the responses of neurons in the gerbil inferior colliculus to repeated presentations of a variety of sounds. In each case, we estimated s from the experimental data using Eq. 2 and were able to simulate new responses with PSTH, r_0, and SNR that match those measured experimentally.

POPULATION RESPONSES

As described in the Introduction, correlations between cells in early sensory systems can have both signal and noise components: signal correlations arise from correlations in the stimulus itself and/or similarity in preferred stimulus features (frequency, orientation, etc.), while noise correlations arise from shared inputs that contribute to the trial-to-trial variability in responses. In our model, we adopt the most common definition of noise correlation, where ρ^{pq}_{noise}, the noise correlation between cells p and q, is given by the difference between the total correlation and the signal correlation, $\rho^{pq}_{\text{noise}} = \rho^{pq}_{\text{total}} - \rho^{pq}_{\text{signal}}$, and ρ^{pq}_{total} and $\rho^{pq}_{\text{signal}}$ are the correlation coefficients between the responses of cells p and q before and after the trial order has been shuffled. The model described above for a single cell is easily extended to capture the pairwise signal and noise correlations in a population, where the response of cell $p \in \{1, 2,\ldots,P\}$ is given by

$$r_i^p[n] = \begin{cases} 1, & \left(s^p[n] + z_i^p[n]\right) > 0 \\ 0, & \left(s^p[n] + z_i^p[n]\right) \le 0 \end{cases} \tag{3}$$

where each cell has its own s^p that is the same on every trial and z^p that is different on every trial. In this population model, $z \sim \mathcal{N}(0, \Sigma_z)$ is a multivariate (P-dimensional) Gaussian random process with covariance matrix

$$\Sigma_z = \begin{bmatrix} 1 & \rho_z^{12} & \cdots & \rho_z^{1P} \\ \rho_z^{21} & 1 & & \\ \vdots & & \ddots & \\ \rho_z^{P1} & & & 1 \end{bmatrix},$$

where ρ_z^{pq}, which is assumed to be constant across time bins and trials, is the pairwise correlation coefficient between z^p and z^q and $\rho_z^{pq} = \rho_z^{qp}$. Assuming we have the responses of a population to repeated trials of a particular stimulus, we can estimate each s^p separately to match the single-cell properties as described above. Because the response is binary and this approach matches $\bar{r}$ exactly for each cell, it will also match the signal correlation between cells. To match the noise correlation, it is necessary to find the appropriate covariance matrix Σ_z. This can be done by solving the equation that relates ρ_z^{pq} to the spike train noise correlation ρ^{pq}_{noise} numerically for each pair of cells (again, the function is monotonic and, because z is Gaussian, each ρ_z^{pq} can be solved for independently).

Thus, ρ^{pq}_{noise} can be written as

$$\begin{aligned} \rho^{pq}_{\text{noise}} &= \rho^{pq}_{\text{total}} - \rho^{pq}_{\text{signal}} \\ &= \frac{\left\langle \text{cov}\left(r_i^p, r_i^q\right)\right\rangle_i}{\sqrt{\left\langle \text{var}\left(r_i^p\right)\right\rangle_i \left\langle \text{var}\left(r_i^q\right)\right\rangle_i}} - \frac{\left\langle \text{cov}\left(r_i^p, r_j^q\right)\right\rangle_{i\neq j}}{\sqrt{\left\langle \text{var}\left(r_i^p\right)\right\rangle_i \left\langle \text{var}\left(r_i^q\right)\right\rangle_i}} \end{aligned} \tag{4}$$

where, because r_i is binary,

$$\left\langle \text{var}\left(r_i^p\right)\right\rangle_i = r_0{}^p\left(1 - r_0{}^p\right)$$

and, because z is Gaussian,

$$\left\langle \text{cov}\left(r_i^p, r_i^q\right)\right\rangle_i = \left\langle \Phi_2\left(\begin{bmatrix} s^p \\ s^q \end{bmatrix}, \begin{bmatrix} 1 & \rho_z^{pq} \\ \rho_z^{pq} & 1 \end{bmatrix}\right)\right\rangle_n - r_0{}^p r_0{}^q \quad \text{and}$$

$$\left\langle \text{cov}\left(r_i^p, r_j^q\right)\right\rangle_{i\neq j} = \left\langle \Phi\left(s^p, 1\right)\Phi\left(s^q, 1\right)\right\rangle_n - r_0{}^p r_0{}^q$$

where $\Phi_2(\vec{x}, \Sigma)$ is the CDF for a two-dimensional Gaussian with zero mean and covariance Σ evaluated at $\vec{x}$.

To demonstrate the utility of this approach, we first attempted to reproduce the single-cell properties and pairwise correlations recorded experimentally from a population of 10 cells in the cat lateral geniculate nucleus in response to repeated presentations of a natural scene movie. **Figure 2A** shows the experimental and simulated responses for two cells. As expected, the PSTH, r_0, and SNR of the experimental and simulated responses are closely matched. As shown in **Figure 2B**, the measured and simulated pairwise noise correlations in the population are also closely matched.

In addition to matching the experimentally observed responses, this approach can also be used to manipulate pairwise correlations without disturbing single-cell properties by changing the value of ρ^{pq}_{noise} on the left side of Eq. 4 before solving for ρ_z^{pq} (note that there are a number of constraints on the realizable values of ρ^{pq}_{noise} – for example, because the covariance matrix Σ_z must be positive semi-definite, it may be difficult to obtain strong negative correlations; see Macke et al., 2009 for a detailed discussion). As a demonstration, we attempted to simulate population spike trains in which the noise correlations were twice as large as those observed experimentally. As shown in **Figure 2C**, the noise correlations in the simulated data match those desired.

EVALUATING GOODNESS OF FIT

Our model is not fit directly to observed spike trains, but rather to the PSTHs and pairwise noise correlations that are extracted from them. In our framework, any set of PSTHs and noise correlations can be fit with a unique set of model parameters, but that does not mean, of course, that the model is a good description of the original spike trains. The actual goodness of fit of the model is determined by two factors: the measurement noise in the PSTHs and noise correlations and the validity of the assumption that the spike trains can be described by our model

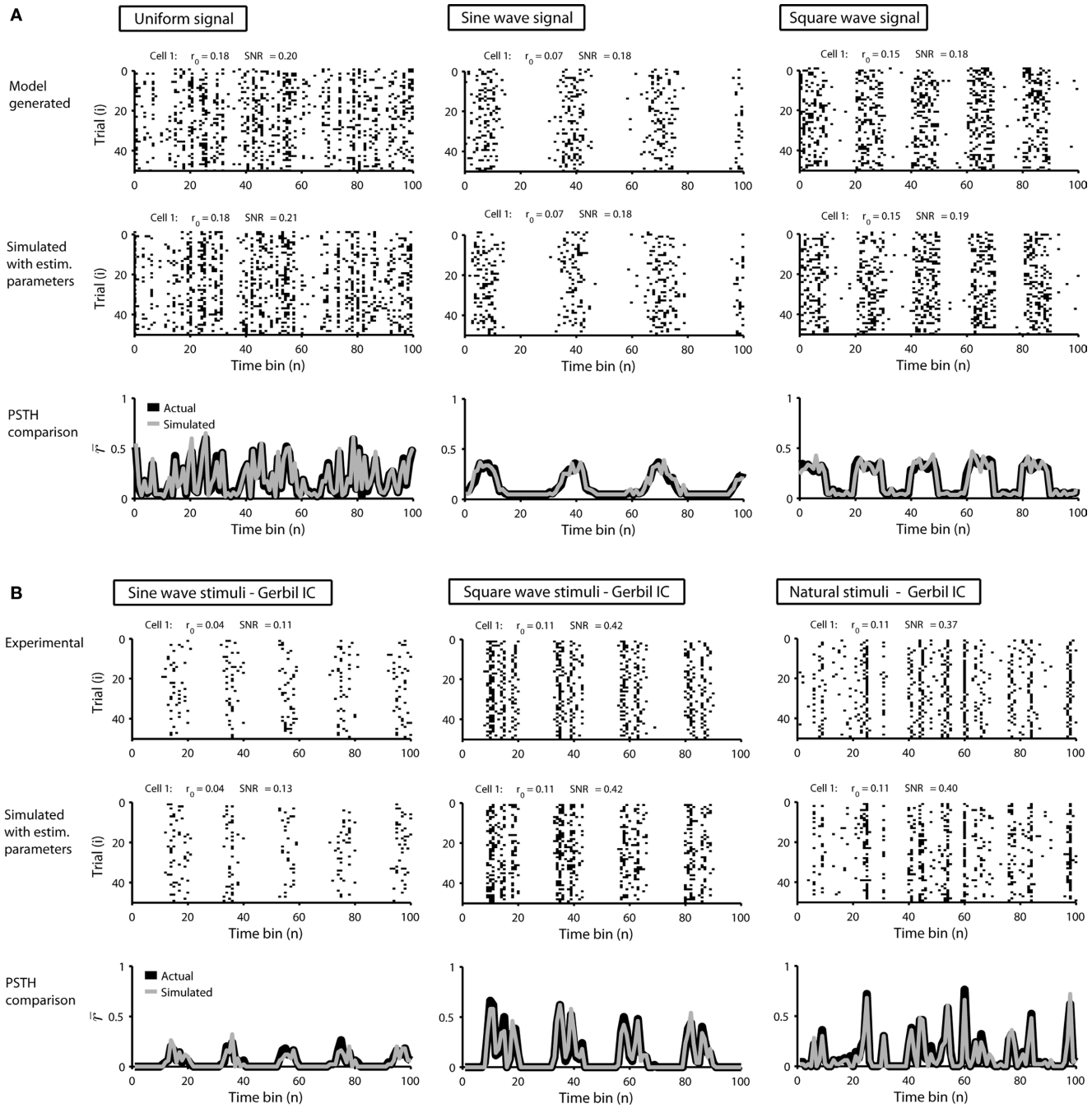


FIGURE 1 | Simulated responses match the single-cell properties of model generated and experimental data. (A) The top row shows model generated responses to repeated trials with a variety of waveforms for *s*: a uniform random noise signal in the range [–0.5 to 0.5]; a sinusoidal signal 0.75 sin(*n*/5) – 1; a square wave signal with a 50% duty cycle, a period of 20 samples, and the same mean value and peak-to-peak amplitude as the sinusoid. The middle row shows simulated responses with *s* estimated from the responses in the top row. The bottom row shows the PSTHs of the original model generated (black) and simulated (gray) responses. **(B)** The top row shows experimental responses of a cell recorded in the inferior colliculus of an anesthetized gerbil to repeated trials of a variety of sounds (experimental methods are described in Lesica and Grothe, 2008a,b): a sinusoidally amplitude modulated (SAM) tone with a carrier frequency of 6 KHz, a modulation depth of 100%, an intensity of 70 dB, and a modulation frequency of 50 Hz; a square wave modulated tone with the same carrier frequency, modulation depth, and intensity, a 50% duty cycle, and a period of 25 ms; a tone with the same carrier frequency and intensity modulated by a signal with power spectra matched to that measured from a series of animal vocalizations. The bin size Δ for these responses was 1 ms. The middle and bottom rows show the simulated responses simulated and PSTHs presented as in **(A)**.

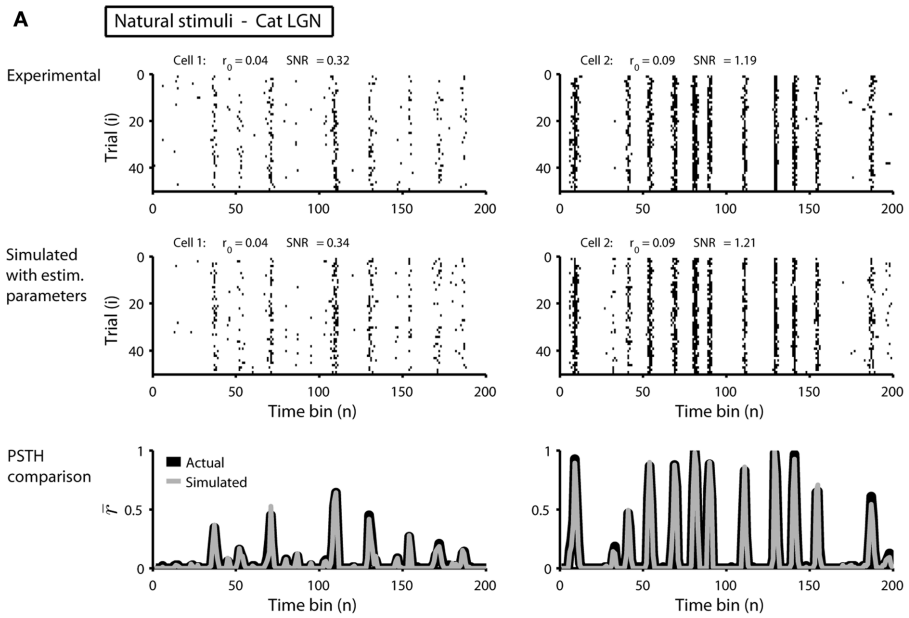


FIGURE 2 | Simulated responses match the single-cell properties and pairwise correlations in experimental data. (A) The top row shows the experimental responses of two cells recorded simultaneously in the lateral geniculate nucleus of an anesthetized cat to repeated trials of a natural scene movie (experimental methods are described in Lesica et al., 2007) with an RMS contrast of 0.4. The bin size Δ for these responses was 4 ms. The middle and bottom rows show the simulated responses and PSTHs presented as in **Figure 1**. **(B)** The image shows the pairwise noise correlations in the experimental (lower triangular portion) and simulated (upper triangular portion) responses of a population of 10 neurons recorded in the cat LGN. The responses of all 10 cells were recorded simultaneously. **(C)** The image shows the desired pairwise noise correlations (lower triangular portion; values are double those measured in the original data) and the pairwise noise correlations realized in the simulated responses (upper triangular portion).

framework. The goodness of fit can be measured by separating the available spike trains into "training" and "testing" sets, fitting the model parameters on the training spike trains, and calculating the (log) likelihood of the testing spike trains from the resulting model. The absolute likelihood may be difficult to interpret, but the ratio of the likelihoods from two different models can give be an informative measure.

To demonstrate the use of likelihood as a measure of goodness of fit, we simulated population spike trains from a known model (see figure legend for model parameters), split the spike trains into training and testing sets, and estimated the model parameters from the training spike trains. To determine whether including noise correlations in the estimated model improved the goodness of fit, we then compared the likelihood of the testing spike trains from the estimated model with and without noise correlations (i.e., with Σ_z estimated as described above or set to the identity matrix) for different numbers of training trials. The likelihood of a given testing spike train was computed as

$$L(\boldsymbol{r}_i) = \sum_{n=1}^{N} \log \Phi_P\left((-2\boldsymbol{r}_i[n]+1)\cdot \boldsymbol{s}[n], \Sigma_z\right) \quad (5)$$

where $\boldsymbol{r}_i$ is the $N \times P$ binary matrix of the responses of a population of P cells on a given trial i, $\boldsymbol{r}_i[n]$ is the vector of the responses $r_i^p[n]$ for each cell, $\boldsymbol{s}[n]$ is the vector of the signals $s^p[n]$ for each cell, $\boldsymbol{\Phi}_P(\vec{x}, \Sigma)$ is the CDF for a P-dimensional Gaussian with zero mean and covariance Σ evaluated at $\vec{x}$, and $\cdot$ denotes a point-by-point vector product. To isolate the effects of the noise correlations on the goodness of fit, we set the PSTHs in the estimated model to be the same as those in the actual model and used only the estimated noise correlations. As shown in **Figure 3**, as the number of training trials increased, the measurement noise in noise correlations

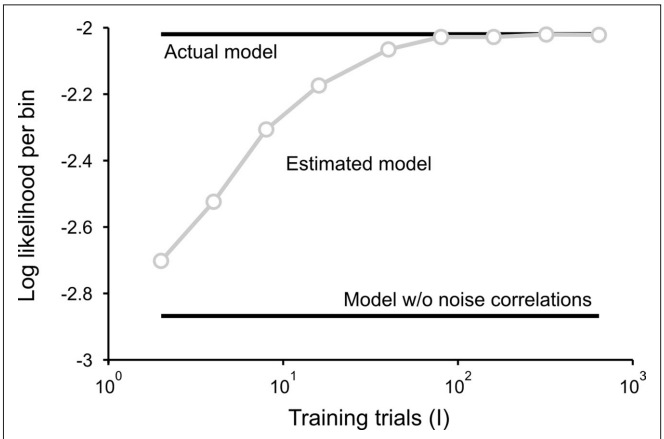


FIGURE 3 | Model goodness of fit increases with increasing trials. The gray circles show the log likelihood (per bin) for the estimated model as a function of the number of trials used for fitting Σ_z. The likelihood was computed for 100 trials not used for fitting the model. The likelihoods for the actual model with and without noise correlations are also shown. Spike trains were generated using the model described in Eq. 10 with the following parameters: $N = 500$, $P = 10$, σ_s^2 and θ were chosen so that $r_0 \sim \mathcal{M}(0.5, 0.35)$, and ρ_s^{pq} and ρ_z^{pq} were chosen so that $\rho_{\text{signal}}^{pq} \sim \mathcal{N}(0.12, 0.03)$ and $\rho_{\text{noise}}^{pq} \sim \mathcal{N}(0.07, 0.02)$.

decreased, and the likelihood from the estimated model with noise correlations approached that of the actual model, reaching the same value with $I = 80$ trials.

NOISE WITH TEMPORAL CORRELATIONS

The model as described above captures both the instantaneous and long-term signal correlations between cells by matching their individual PSTHs, but captures only instantaneous noise correlations because z is uncorrelated in time. While instantaneous noise correlations are likely to be sufficient to describe population spike trains in early sensory systems, the model can also be extended to capture long-term noise correlations if necessary, for example, to capture the high level of trial-to-trial variability in higher cortical areas. Long-term noise correlations can be captured by adding temporal correlations to z via Gaussian conditioning (MacKay, 2003; Macke et al., 2009) so that z in each time bin is drawn from a distribution with mean and covariance dependent on the values of z in the preceding time bins:

$$z[n+1 \mid n-K+1,...,n] \sim \mathcal{N}\left(CB^{-1}z'[n-K+1,...,n], \Sigma_0 - CB^{-1}C^T\right)$$

$$\text{where} \quad B = \begin{bmatrix} \Sigma_0 & \Sigma_1 & \cdots & \Sigma_{K-1} \\ \Sigma_1 & \Sigma_0 & \cdots & \Sigma_{K-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{K-1} & \Sigma_{K-2} & \cdots & \Sigma_0 \end{bmatrix} \tag{6}$$

$$C = \begin{bmatrix} \Sigma_1 & \Sigma_2 & \cdots & \Sigma_K \end{bmatrix}$$

$$\Sigma_k = \begin{bmatrix} \rho_z^{pp}[k] & \rho_z^{pq}[k] \\ \rho_z^{qp}[k] & \rho_z^{qq}[k] \end{bmatrix}$$

$$z'[n-K+1,...,n] = \left[z^p[n-K+1] \quad ,z^q[n-K+1],\cdots,z^p[n], \quad z^q[n]\right]$$

and K is the number of preceding time bins to condition on. An example of such conditioning is shown in **Figure 4**. We used a known model to generate population spike trains with the noise auto- and cross-correlation functions shown by the thick black lines (see figure legend for model parameters). We then estimated model parameters from those spike trains, including those required for conditioning z on the preceding $K = 4$ time bins. The noise auto- and cross-correlation functions for the spike trains simulated from the estimated model (shown by the gray circles) closely match those of the original spike trains.

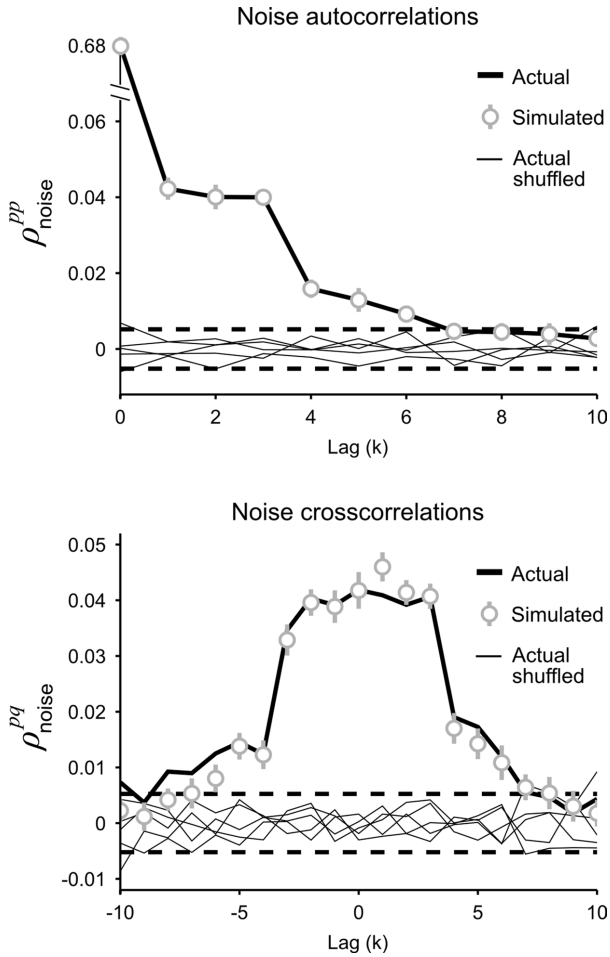


FIGURE 4 | Gaussian conditioning captures long-term noise correlations. The gray circles show the noise auto- and cross-correlation functions for spike trains simulated from the estimated model. Error bars represent the standard deviation in the correlation values computed from independent realizations of the model. The model parameters were estimated from spike trains simulated by the model in Eq. 10, including Gaussian conditioning of z, with the following parameters: $N = 500$, $P = 10$, σ_s^2 and θ were chosen so that $r_0 \sim \mathcal{N}(0.16, 0.04)$ and SNR $\sim \mathcal{N}(0.5, 0.35)$, $\rho_s^{pq} = 0.5$, and $\rho_z^{pp}[k]$ and $\rho_z^{pq}[k] = 0.12$ for $k \leq 4$ and zero thereafter. The noise auto- and cross-correlation functions of the original spike trains are shown by the thick black lines. Examples of the noise auto- and cross-correlation functions of the original spike trains after trial shuffling are shown by the thin black lines. The dashed lines indicate the mean ±2 standard deviations of the distribution of the correlation values after shuffling.

A GENERAL MODEL FOR POPULATION SPIKE TRAINS

SINGLE-CELL RESPONSES

When modeling experimental spike trains as described above, the noise correlations can be chosen arbitrarily, but the mean spike rate, trial-to-trial variability, and signal correlations are dependent on the PSTH. It may also be useful to have a general model for population spike trains in which all of the response properties can be specified independently. For a single cell, this is achieved by replacing the deterministic signal s in the model framework described above with a Gaussian random process:

$$r_i[n] = \begin{cases} 1, & (s[n] + z_i[n]) > \theta \\ 0, & (s[n] + z_i[n]) \leq \theta \end{cases} \tag{7}$$

where $z \sim \mathcal{N}(0, 1)$ is again a Gaussian random process that is different on every trial, s is a Gaussian random process $s \sim \mathcal{N}(0, \sigma_s^2)$ that is the same on every trial, and the threshold θ is allowed to take on any value (note that in this case, θ cannot simply be set to an arbitrary value; in order to achieve any combination of r_0 and SNR, 2 degrees of freedom are required). Such a model could be used, for example, to simulate spike trains with any mean spike rate and trial-to-trial variability. Furthermore, because the model is based on Gaussian processes, it may enable certain population response properties to be investigated analytically or numerically directly from the model parameters, without the need for simulations.

To specify the model parameters, the equations for r_0 and SNR can be written in terms of σ_s^2 and θ and solved numerically to obtain the appropriate values:

$$r_0 = \Phi(-\theta, \sigma_s^2 + 1) \tag{8}$$

$$\mathrm{SNR} = \frac{\mathrm{var}(\bar{r})}{\left\langle \mathrm{var}(\xi_i) \right\rangle_i} \tag{9}$$

While Eq. 8 is already written in terms of σ_s^2 and θ, Eq. 9 requires some manipulation. The numerator can be written as

$$\mathrm{var}(\bar{r}) = \mathrm{var}\left(\frac{1}{I}\sum_{i=1}^{I} r_i\right) = \frac{1}{I^2}\left(I\left\langle \mathrm{var}(r_i)\right\rangle_i + 2I(I-1)\left\langle \mathrm{cov}(r_i, r_j)\right\rangle_{i\neq j}\right)$$

where, because r_i is binary,

$$\left\langle \mathrm{var}(r_i)\right\rangle_i = r_0(1-r_0)$$

and, because s and z are Gaussian,

$$\left\langle \mathrm{cov}(r_i, r_j)\right\rangle_{i\neq j} = \Phi_2\left(-\theta, \begin{bmatrix} \sigma_s^2+1 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2+1 \end{bmatrix}\right) - r_0^2.$$

The denominator can be written as

$$\left\langle \mathrm{var}(\xi)\right\rangle_i = \left\langle \mathrm{var}(\bar{r} - r_i)\right\rangle_i = \mathrm{var}(\bar{r}) + \left\langle \mathrm{var}(r_i)\right\rangle_i - 2\left\langle \mathrm{cov}(\bar{r}, r_i)\right\rangle_i$$

where, because r_i is binary and s and z are Gaussian,

$$\begin{aligned}\left\langle \mathrm{cov}(\bar{r}, r_i)\right\rangle_i &= \left\langle \bar{r}\cdot r_i\right\rangle_{n,i} - r_0^2 = \left\langle \bar{r}^2\right\rangle_n - r_0^2 = \left\langle \left(\frac{1}{I}\sum_{i=1}^{I} r_i\right)^2\right\rangle_n - r_0^2 \\ &= \frac{1}{I^2}\left(I\left\langle r_i^2\right\rangle_{n,i} + I(I-1)\left\langle r_i\cdot r_j\right\rangle_{n,i\neq j}\right) - r_0^2 \\ &= \frac{1}{I^2}\left(Ir_0^2 + I(I-1)\Phi_2\left(-\theta, \begin{bmatrix} \sigma_s^2+1 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2+1 \end{bmatrix}\right)\right) - r_0^2\end{aligned}$$

Note that in these equations, $\bar{r}\cdot r_i$ and $r_i \cdot r_j$ denotes point-by-point vector products, and $\bar{r}^2$ denotes point-by-point squaring. Thus, for any realizable combination of r_0 and SNR, appropriate σ_s^2 and θ can be found (the minimum realizable SNR depends on the number of trials, see Appendix). To demonstrate this approach, we randomly chose a variety of values for r_0 and SNR, estimated the corresponding values of σ_s^2 and θ, and generated responses using the estimated values. As shown in **Figure 5A**, r_0 and SNR of the simulated responses closely match the desired values.

POPULATION RESPONSES

The model described above for a single cell is easily extended to a population, where the response of cell $p \in \{1, 2, \ldots, P\}$ is given by

$$r_i^p[n] = \begin{cases} 1, & \left(s^p[n] + z_i^p[n]\right) > \theta^p \\ 0, & \left(s^p[n] + z_i^p[n]\right) \leq \theta^p \end{cases} \tag{10}$$

where both s and z are multivariate Gaussian random process $s \sim \mathcal{N}(0, \Sigma_s)$ and $z \sim \mathcal{N}(0, \Sigma_z)$ with covariance matrices

$$\Sigma_s = \begin{bmatrix} \sigma_{s^1}^2 & \rho_s^{12}\sigma_{s^2}\sigma_{s^1} & \cdots & \rho_s^{1P}\sigma_{s^P}\sigma_{s^1} \\ \rho_s^{21}\sigma_{s^1}\sigma_{s^2} & \sigma_{s^2}^2 & & \\ \vdots & & \ddots & \\ \rho_s^{P1}\sigma_{s^1}\sigma_{s^P} & & & \sigma_{s^P}^2 \end{bmatrix} \text{ and}$$

$$\Sigma_z = \begin{bmatrix} 1 & \rho_z^{12} & \cdots & \rho_z^{1P} \\ \rho_z^{21} & 1 & & \\ \vdots & & \ddots & \\ \rho_z^{P1} & & & 1 \end{bmatrix}.$$

After determining σ_s^2 and θ for each cell based on the desired r_0 and SNR as described above, the pairwise correlation coefficients ρ_s^{pq} and ρ_z^{pq} required to obtain the desired spike train signal and noise correlations $\rho_{\mathrm{signal}}^{pq}$ and $\rho_{\mathrm{noise}}^{pq}$ can be found by solving the following equations numerically:

$$\rho_{\mathrm{signal}}^{pq} = \frac{\left\langle \mathrm{cov}\left(r_i^p, r_j^q\right)\right\rangle_{i\neq j}}{\sqrt{\left\langle \mathrm{var}\left(r_i^p\right)\right\rangle_i \left\langle \mathrm{var}\left(r_i^q\right)\right\rangle_i}} \quad \text{and}$$

$$\rho_{\mathrm{noise}}^{pq} = \rho_{\mathrm{total}}^{pq} - \rho_{\mathrm{signal}}^{pq} = \frac{\left\langle \mathrm{cov}\left(r_i^p, r_i^q\right)\right\rangle_i}{\sqrt{\left\langle \mathrm{var}\left(r_i^p\right)\right\rangle_i \left\langle \mathrm{var}\left(r_i^q\right)\right\rangle_i}} - \frac{\left\langle \mathrm{cov}\left(r_i^p, r_j^q\right)\right\rangle_{i\neq j}}{\sqrt{\left\langle \mathrm{var}\left(r_i^p\right)\right\rangle_i \left\langle \mathrm{var}\left(r_i^q\right)\right\rangle_i}}$$

where, because r_i is binary,

$$\left\langle \mathrm{var}\left(r_i^p\right)\right\rangle_i = r_0^p\left(1 - r_0^p\right)$$

and, because s and z are Gaussian,

$$\left\langle \mathrm{cov}\left(r_i^p, r_j^q\right)\right\rangle_{i\neq j} = \Phi_2\left(-\begin{bmatrix}\theta^p \\ \theta^q\end{bmatrix}, \begin{bmatrix} \sigma_{s^p}^2+1 & \rho_s^{pq}\sigma_{s^p}\sigma_{s^q} \\ \rho_s^{pq}\sigma_{s^p}\sigma_{s^q} & \sigma_{s^q}^2+1 \end{bmatrix}\right) - r_0^p r_0^q$$

and

$$\left\langle \mathrm{cov}\left(r_i^p, r_i^q\right)\right\rangle_i = \Phi_2\left(-\begin{bmatrix}\theta^p \\ \theta^q\end{bmatrix}, \begin{bmatrix} \sigma_{s^p}^2+1 & \rho_s^{pq}\sigma_{s^p}\sigma_{s^q} + \rho_z^{pq} \\ \rho_s^{pq}\sigma_{s^p}\sigma_{s^q} + \rho_z^{pq} & \sigma_{s^q}^2+1 \end{bmatrix}\right) - r_0^p r_0^q$$

Again, the functions are monotonic and each ρ_s^{pq} and ρ_z^{pq} can be solved for independently.

To demonstrate this approach, we generated a random set of pairwise signal and noise correlation coefficients $\rho_{\mathrm{signal}}^{pq}$ and $\rho_{\mathrm{noise}}^{pq}$, estimated the corresponding values of ρ_s^{pq} and ρ_z^{pq}, and simulated population spike trains with these values. As shown in **Figures 5B,C**, the correlations in the simulated spike trains closely matched the desired values.

DISCUSSION

We have described a model for simulating population spike trains typical of early sensory systems. The model has two forms: the first requires the specification of PSTHs and noise correlations and

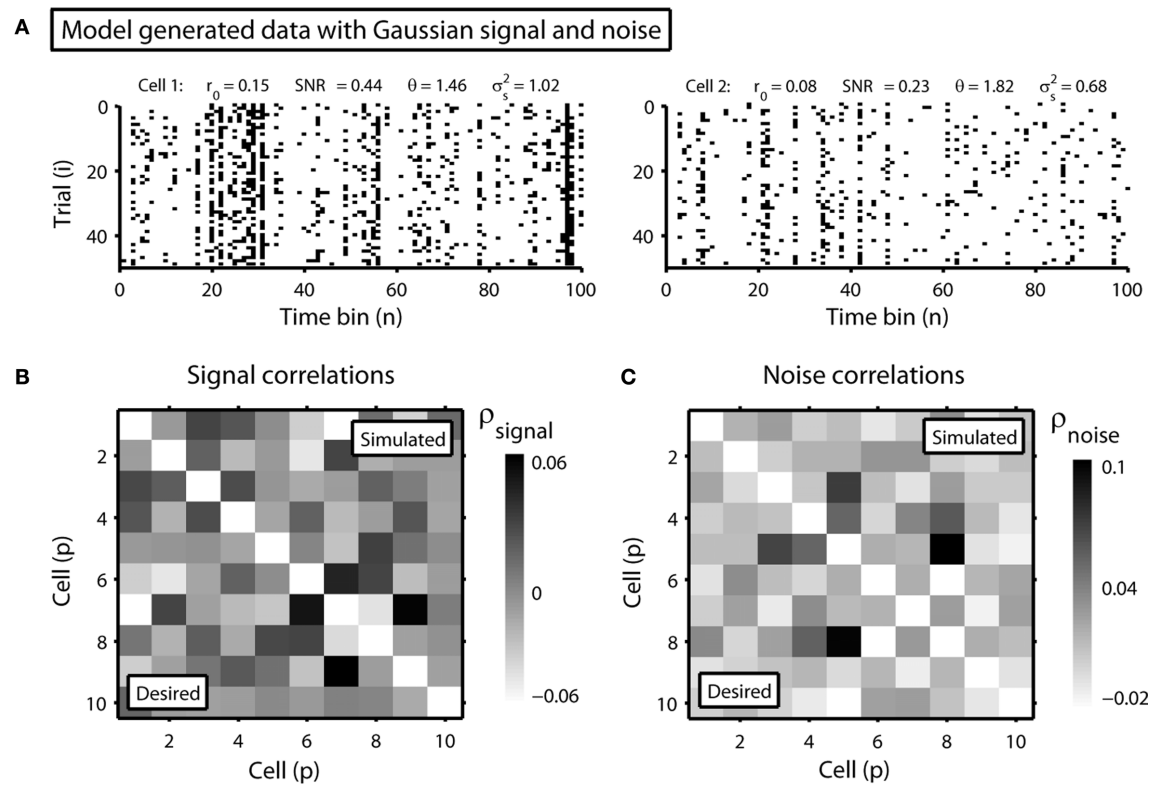


FIGURE 5 | Simulated population responses with specified single-cell properties and pairwise correlations. (A) The top row shows the simulated responses for two cells with specified r_0 and SNR. **(B)** The image shows the desired pairwise signal correlations (lower triangular portion) and the pairwise signal correlations realized in the simulated responses (upper triangular portion) for a population of 10 cells. **(C)** The desired and realized pairwise noise correlations for a population of 10 cells presented as in **(B)**.

can be used to match and manipulate experimental data, and the second is more general and allows for population spike trains with any mean spike rates, trial-to-trial variabilities, signal correlations, and noise correlations. Both forms of the model are easily implemented as parameter fitting requires simply finding the level crossings of monotonic functions and correlations can be determined independently for each pair of cells. The Matlab code required to fit the model parameters is available for download at http://www.ucl.ac.uk/ear/research/lesicalab.

Our model improves on the existing methods for generating population spike trains described in the Introduction in several important ways. First, the model framework is explicitly designed around the response properties that are important for early sensory neurons: time-varying spike rate (PSTH), trial-to-trial variability, and signal and noise correlations. Second, the model allows independent and straightforward manipulation of one response property without changes in the other properties. One can imagine a number of potential uses for a model with these properties. The fact that the model matches the single-cell properties and correlations observed experimentally is in itself of some utility, such as providing a simple framework for computing the likelihood of observed spike trains given only pairwise interactions. These likelihoods could be used to, for example, test how important noise correlations are in determining population spike patterns by comparing models with and without noise correlations. The model also provides the ability to manipulate noise correlations without affecting the signal correlations or single-cell properties. In the brain, these properties are coupled to each other – for example, one can decrease the spike rate of visual neurons by decreasing the contrast of the stimulus, but this will also likely change the trial-to-trial variability and the correlations. Thus, a question such as whether or not changes in correlations with changes in contrast are detrimental or beneficial to a population code is impossible to answer experimentally. With our model, one could compare simulated populations with high contrast single-cell properties and correlations to simulated populations with high contrast single-cell properties and low contrast correlations to directly test whether or not the change in correlations is important. A similar example can be used to illustrate the utility of the general form of the model: Because the general form of the model allows for spike trains with any mean spike rate, trial-to-trial variability, and pairwise signal and noise correlations (within statistical constraints), it could be used to perform a systematic investigation of the effects of noise correlations on populations with different levels of signal correlations that would be impossible to conduct experimentally.

There are several ways in which the formulation of our model described here could potentially be improved. For example, the assumption that no more than one spike can occur in any time

bin could potentially be relaxed, using a formulation analogous to that described in (Macke et al., 2009). Also, the current formulation of the model cannot be directly connected to biophysical quantities, so its parameters cannot be used, for example, to determine the current necessary to achieve a desired set of correlations in a stimulation experiment. With further effort, however, it may be possible to extend some of the desirable properties of our model into a more realistic framework such as an integrate-and-fire model (Paninski et al., 2004). Another limitation of the model that may be difficult to overcome is that in estimating a single value for each parameter across all trials, it implicitly assumes that the population is stationary. Certain types of non-stationarities, such as trial-to-trial fluctuations in the PSTH, may be captured by introducing long-term noise correlations via Gaussian conditioning, but others may require integrating the model into an adaptive framework (Eden et al., 2004).

There are also certain limitations of the model that are inherent in its underlying statistical framework. For example, it is difficult to generate spike trains with very low SNRs with a small number of trials – even if the variance of the internal signal in the model $\sigma_s^2 = 0$, because the spikes are generated stochastically, the PSTH computed from a small number of trials will not have zero variance, and, thus, the SNR will not be zero. There are also constraints on the values of signal and noise correlations that can be achieved. Because the model is based on multivariate Gaussian processes which must have positive semi-definite covariance matrices, it is difficult, for example, to generate spike trains with strong negative correlations. Finally, while the model can be specified and fit to a population of any size, it is only guaranteed to capture second-order correlations. As with any model that includes only pairwise interactions, our model's ability to capture the higher-order structure in population spike trains will depend on the nature of the population (Schneidman et al., 2006; Shlens et al., 2006; Roudi et al., 2009; Ohiorhenuan et al., 2010).

ACKNOWLEDGMENTS

We thank Benedikt Grothe for the use of the gerbil data and Chong Weng, Jianzhong Jin, Chun-I Yeh, Dan Butts, Garrett Stanley, and Jose-Manuel Alonso for the use of the cat data. Dmitry R. Lyamzin and Nicholas A. Lesica were supported by the Deutsche Forschungsgemeinschaft (LE2522/1-1) and the Wellcome Trust. Jakob H. Macke was supported by the German Ministry of Education, Science, Research and Technology through the Bernstein award to Matthias Bethge (BMBF; FKZ: 01GQ0601).

REFERENCES

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

Borst, A., and Theunissen, F. E. (1999). Information theory and neural coding. *Nat. Neurosci.* 2, 947–957.

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

Brette, R. (2009). Generation of correlated spike trains. *Neural Comput.* 21, 188–215.

Chornoboy, E., Schramm, L., and Karr, A. (1988). Maximum likelihood identification of neural point process systems. *Biol. Cybern.* 59, 265–275.

Cox, D. R., and Wermuth, N. (2002). On some models for multivariate binary variables parallel in complexity with the multivariate Gaussian distribution. *Biometrika* 89, 462–469.

De La Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., and Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature* 448, 802–806.

Destexhe, A., and Pare, D. (1999). Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *J. Neurophysiol.* 81, 1531.

Dorn, J. D., and Ringach, D. L. (2003). Estimating membrane voltage correlations from extracellular spike trains. *J. Neurophysiol.* 89, 2271–2278.

Eden, U. T., Frank, L. M., Barbieri, R., Solo, V., and Brown, E. N. (2004). Dynamic analysis of neural encoding by point process adaptive filtering. *Neural Comput.* 16, 971–998.

Emrich, L. J., and Piedmonte, M. R. (1991). A method for generating high-dimensional multivariate binary variates. *Am. Stat.* 45, 302–304.

Feng, J., and Brown, D. (2000). Impact of correlated inputs on the output of the integrate-and-fire model. *Neural Comput.* 12, 671–692.

Galan, R. F., Fourcaud-Trocme, N., Ermentrout, G. B., and Urban, N. N. (2006). Correlation-induced synchronization of oscillations in olfactory bulb neurons. *J. Neurosci.* 26, 3646.

Gutig, R., Aharonov, R., Rotter, S., and Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *J. Neurosci.* 23, 3697.

Gutnisky, D. A., and Josic, K. (2010). Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process. *J. Neurophysiol.* 103, 2912–2930.

Krumin, M., and Shoham, S. (2009). Generation of spike trains with controlled auto-and cross-correlation functions. *Neural Comput.* 21, 1642–1664.

Kuhn, A., Aertsen, A., and Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput.* 15, 67–101.

Kulkarni, J. E., and Paninski, L. (2007). Common-input models for multiple neural spike-train data. *Network* 18, 375–407.

Lesica, N. A., and Grothe, B. (2008a). Efficient temporal processing of naturalistic sounds. *PLoS ONE* 3, e1655. doi: 10.1371/journal.pone.0001655.

Lesica, N. A., and Grothe, B. (2008b). Dynamic spectrotemporal feature selectivity in the auditory midbrain. *J. Neurosci.* 28, 5412–5421.

Lesica, N. A., Jin, J., Weng, C., Yeh, C. I., Butts, D. A., Stanley, G. B., and Alonso, J. M. (2007). Adaptation to stimulus contrast and correlations during natural visual stimulation. *Neuron* 55, 479–491.

MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press.

Macke, J. H., Berens, P., Ecker, A. S., Tolias, A. S., and Bethge, M. (2009). Generating spike trains with specified correlation coefficients. *Neural Comput.* 21, 397–423.

Niebur, E. (2007). Generation of synthetic spike trains with defined pairwise correlations. *Neural Comput.* 19, 1720–1738.

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

Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network* 15, 243–262.

Paninski, L., Pillow, J., and Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. *Prog. Brain Res.* 165, 493–507.

Paninski, L., Pillow, J. W., and Simoncelli, E. P. (2004). Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural Comput.* 16, 2533–2561.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E., and Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454, 995.

Roudi, Y., Nirenberg, S., and Latham, P. E. (2009). Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. *PLoS Comput. Biol.* 5, e1000380. doi: 10.1371/journal.pcbi.1000380.

Salinas, E., and Sejnowski, T. J. (2002). Integrate-and-fire neurons driven by correlated stochastic input. *Neural Comput.* 14, 2111–2155.

Schneidman, E., Berry, M. J. II, R. S., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 1007.

Shea-Brown, E., Josic, K., De La Rocha, J., and Doiron, B. (2008). Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding. *Phys. Rev. Lett.* 100, 108102.

Shlens, J., Field, G. D., Gauthier, J. L., Grivich, M. I., Petrusca, D., Sher, A., Litke, A. M., and Chichilnisky, E. J. (2006). The structure of multi-neuron firing patterns in primate retina. *J. Neurosci.* 26, 8254–8266.

Song, S., and Abbott, L. (2001). Cortical development and remapping through spike timing-dependent plasticity. *Neuron* 32, 339–350.

Stroeve, S., and Gilen, S. (2001). Correlation between uncoupled conductance-based integrate-and-fire neurons due to common and synchronous presynaptic firing. *Nat. Neurosci.* 13, 2005–2029.

Tchumatchenko, T., Malyshev, A., Geisel, T., Volgushev, M., and Wolf, F. (2008). Correlations and synchrony in threshold neuron models. *Quant. Biol. arXiv* 810, 2–6.

Toyoizumi, T., Rad, K. R., and Paninski, L. (2009). Mean-field approximations for coupled populations of generalized linear model spiking neurons with Markov refractoriness. *Neural Comput.* 21, 1203–1243.

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

Received: 15 November 2009; paper pending published: 19 January 2010; accepted: 05 October 2010; published online: 15 November 2010.

Citation: Lyamzin DR, Macke JH and Lesica NA (2010) Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations. Front. Comput. Neurosci. ***4****:144. doi: 10.3389/fncom.2010.00144*

APPENDIX

MINIMUM ACHIEVABLE SNR

For any binary response r with N time bins and I trials, the minimum realizable SNR depends only on the number of trials. Recalling the definition of SNR

$$\mathrm{SNR} = \frac{\mathrm{var}(\bar{r})}{\left\langle \mathrm{var}(\xi_i) \right\rangle_i}$$

the minimum value within the context of our model framework is clearly achieved when the variance of the signal $\sigma_s^2 = 0$ (in the case of $I = \infty$, this results in $\mathrm{var}(\bar{r}) = 0$). Recalling the expansion of $\mathrm{var}(\bar{r})$ in section "Single Cell Responses" and the fact that $r_0 = \Phi(-\theta, \sigma_s^2 + 1)$ within our framework, we find that when $\sigma_s^2 = 0$,

$$\begin{aligned}
\mathrm{var}(\bar{r}) &= \frac{1}{I^2}\left(I\left\langle \mathrm{var}\left(r_i\right)\right\rangle_i + 2I(I-1)\left\langle \mathrm{cov}\left(r_i, r_j\right)\right\rangle_{i\neq j}\right) \\
&= \frac{1}{I^2}\left(Ir_0(1-r_0) + 2I(I-1)\Phi_2\left(-\theta, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) - r_0^2\right) \\
&= \frac{1}{I^2}\left(Ir_0(1-r_0)\right)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left\langle \mathrm{var}(\xi)\right\rangle_i &= \left\langle \mathrm{var}\left(\bar{r} - r_i\right)\right\rangle_i \\
&= \mathrm{var}(\bar{r}) + \left\langle \mathrm{var}\left(r_i\right)\right\rangle_i - 2\left\langle \mathrm{cov}\left(\bar{r}, r_i\right)\right\rangle_i \\
&= \frac{1}{I^2}\left(Ir_0\left(1-r_0\right)\right) + r_0(1-r_0) \\
&\quad - 2\left(\frac{1}{I^2}\left(Ir_0^2 + I(I-1)\Phi_2\left(-\theta, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\right) - r_0^2\right) \\
&= \frac{1}{I^2}\left(Ir_0\left(1-r_0\right)\right) + r_0\left(1-r_0\right)
\end{aligned}$$

Simplifying the resulting expression gives the minimum realizable SNR:

$$\mathrm{SNR} = \frac{\frac{1}{I^2}\left(Ir_0\left(1-r_0\right)\right)}{\frac{1}{I^2}\left(Ir_0\left(1-r_0\right)\right) + r_0\left(1-r_0\right)} = \frac{1}{(I+1)}$$

IV Analysis and modelling of variability and covariability of population spike trains across multiple time scales

In this chapter we further explore the capability of DG model to fit experimental data and modify its properties. In particular we add auto- and cross-correlations between neurons in the time domain. For this purpose we use Gaussian conditioning algorithm (MacKay, 2003; Macke et al., 2009) to create correlations in the underlying Gaussian processes, and then find expressions for the correlations in the observed binary spike trains.

We show that controlling noise correlations this way allows one to arbitrarily change the Fano factors, ISI distributions of individual cells, and count correlations between pairs of cells.

Authors' contributions

The work was carried out under the supervision of Nicholas Lesica. The research was conceived and designed by NL and DL; NL and JGL collected experimental data; DL performed analysis; NL and DL wrote the paper. Preliminary research results were presented at *Mathematical Neuroscience* in April 2011 in Edinburgh, UK.

The paper has been accepted in *Network: Computation in Neural Systems* under the following reference:

"Analysis and modelling of variability and covariability of population spike trains across multiple time scales", Dmitry R Lyamzin, Jose A Garcia-Lazaro, Nicholas A Lesica, *Network: Computation in Neural Systems* 2012, Vol. 23, No. 1-2, Pages 76-103.

Title: **Analysis and modeling of variability and covariability of population spike trains across multiple time scales**

Authors: Dmitry R. Lyamzin[1, 2], Jose A. Garcia-Lazaro[1], Nicholas A. Lesica[1*]

Affiliations: [1] Ear Institute, University College London, London, UK

[2] Division of Neurobiology, Department of Biology II, Ludwig-Maximilians-University Munich, Martinsried, Germany

* Correspondence: Nicholas A. Lesica, Ear Institute, University College London, 332 Gray's Inn Rd., London WC1X 8EE, UK. E-mail: n.lesica@ucl.ac.uk

Abbreviated title: Modeling covariability across multiple time scales

Acknowledgements: The authors thank Jakob Macke and Maneesh Sahani for helpful discussions. This work was supported by the German Research Foundation (DFG) and the Wellcome Trust.

Abstract:

As multi-electrode and imaging technology begin to provide us with simultaneous recordings of large neuronal populations, new methods for modeling such data must also be developed. We present a model of responses to repeated trials of a sensory stimulus based on thresholded Gaussian processes that allows for analysis and modeling of variability and covariability of population spike trains across multiple time scales. The model framework can be used to specify the values of many different variability measures including spike timing precision across trials, coefficient of variation of the interspike interval distribution, and Fano factor of spike counts for individual neurons, as well as signal and noise correlations and correlations of spike counts across multiple neurons. Using both simulated data and data from different stages of the mammalian auditory pathway, we demonstrate the range of possible independent manipulations of different variability measures, and explore how this range depends on the sensory stimulus. The model provides a powerful framework for the study of experimental and surrogate data and for analyzing dependencies between different statistical properties of neuronal populations.

Keywords: dichotomized Gaussian, noise correlations, population coding, Fano factor

1. Introduction

In sensory systems repeated presentations of identical stimuli can evoke different spike patterns, with variability typically increasing from the periphery to more central brain areas (Kara et al. 2000). The importance of considering variability and covariability (or correlation) of spiking on different time scales has been stressed in a number of studies (Vaadia et al. 1995; Ratnam & Nelson 2000; Bair et al. 2001; Smith & Kohn 2008). For single neurons, the impact of variable spiking on stimulus coding has been studied on several time scales, from millisecond spike time jitter and phase-locking in early sensory pathways, to the scale of inter-spike intervals (ISIs) and spike counts across seconds in cortical areas (Softky & Koch 1993; Shadlen & Newsome 1998). Whether such variability is beneficial or detrimental for coding depends on the assumptions about the nature of the neural code and its significant parameters (Mainen & Sejnowski 1995; Schneidman et al. 1998; Manwani et al. 2002). When neuronal populations are considered, the same question arises with respect to the covariability of spiking of individual neurons (Zohary et al. 1994; Abbott & Dayan 1999; Romo et al. 2003; Schneidman et al. 2003; Latham & Nirenberg 2005; Averbeck et al. 2006). For populations of neurons, both spike synchrony on fine time scales and correlations in spike counts across longer time scales can result in synergistic or redundant coding depending, for example, on the similarity of the tuning curves of the neurons in the population (Zohary et al. 1994; Dan et al. 1998; Puchalla et al. 2005; Smith & Kohn 2008).

Taking this background into consideration, it is clear that tools enabling the analysis and modeling of the different types of variability in population spike trains should be developed. Previous research has shown that certain types of variability and covariability can be directly related to the spike train auto- and crosscorrelation functions (Bair et al. 2001; Nawrot 2010; Tchumatchenko et al. 2010). While existing models for the analysis of population spike trains allow specification of auto- and crosscorrelation functions (Krumin & Shoham 2009; Macke et al. 2009; Lyamzin et al. 2010; Gutnisky & Josić 2010) the parameters of these models have not yet been explicitly linked to the different measures of variability and covariability described above. Modeling spike trains with specified variability and covariability across multiple time scales, each of which could be independently manipulated, would provide a systematic way to study their impact on coding.

In this paper, we present further analysis of a model that we previously developed to capture instantaneous response properties of neuronal populations (Lyamzin et al. 2010). We show how the model can capture variability and covariability across a range of time scales, including signal to noise ratio, coefficient of variation (CV) of the ISI distribution, and Fano factor of spike counts for single neurons, and spike synchrony and correlation coefficient of spike counts for populations. Within the framework of our model, each of the variability and covariability measures can be expressed in terms of internal model parameters in closed form, and in particular, in terms of the auto- and crosscorrelation functions. Given experimental data, the model enables the simulation of spike trains with the same statistics, as well as manipulation of individual response properties. Furthermore, given a set of desired statistics, the model can be used to generate population spike trains *de novo*. Importantly, because the model is based on a dichotomized Gaussian framework, it enables not only simulation of spike trains, but

also direct analysis of the relationship between different types of variability and covariability, and the impact of different response properties on population coding. Matlab code implementing the model will be made available at www.ucl.ac.uk/ear/research/lesicalab.

2. Modeling experimental spike trains

2.1 Model framework

The basic framework of our model is described in our previous paper (Lyamzin et al. 2010). We represent spike trains as binary vectors of length *N* where each element corresponds to a time bin with a value of 1 if there is a spike and 0 otherwise. We assume the width of a time bin is small enough so that no more than one spike occurs in any time bin. Each stimulus is presented *I* times, so that for each cell a response to a single stimulus is a binary *N* x *I* matrix. We model a population of *P* cells, with binary vectors r_i^p for each cell *p* on each trial *i* obtained by thresholding a sum of Gaussian noise z_i^p that is different on each trial and a deterministic component s^p that is the same on each trial. Thresholding is done in each time bin $[n]$.

$$r_i^p[n] = \begin{cases} 1, & (s^p[n] + z_i^p[n]) > 0 \\ 0, & (s^p[n] + z_i^p[n]) \le 0 \end{cases}$$

where $i = 1 \ldots I$, $n = 1 \ldots N$, and $p = 1 \ldots P$. A schematic illustration of the model for single cells is shown in Fig. 1.

The free parameters of this model are the values of s^p, which must be fit for each cell in each time bin, and the correlations in the Gaussian process *z* across both cells and time (the component of *z* corresponding to each cell is constrained to have unit variance across time).

The values of s^p are determined by the PSTH of each cell. From experimental data, one can calculate the PSTH ($\bar{r}^p$) and then find $s^p[n]$ by solving equation (1) numerically in every time bin

$$\bar{r}^p[n] = \Phi(s^p[n], 1) \qquad (1)$$

where $\Phi(x, \sigma^2)$ is the Gaussian cumulative distribution function with zero mean and variance σ^2 evaluated at x.

The correlations in *z* are determined by measuring the noise correlations in the spike trains. Correlations in the spike trains of sensory neurons have both signal and noise components, with signal correlations reflecting covariability in the stimulus-driven component of the response that is repeatable from trial to trial, and noise correlations reflecting covariability in the component of the response that is different from trial to trial. We assume that signal and noise correlations are additive, and thus define noise correlations between two spike trains as a difference between total correlations and

signal correlations. Note that, although the dichotomized Gaussian framework can capture certain higher order statistical properties of spike trains (Macke et al. 2011, Yu et al. 2012), we assume for our purposes that only second order correlations are matched.

To set the correlations in the components of *z* corresponding to cells *p* and *q* at time lag *k*, $\rho_z^{pq}[k]$, one first measures the corresponding noise correlations from the given set of spike trains:

$$\rho_{noise}^{pq}[k] = \rho_{total}^{pq}[k] - \rho_{signal}^{pq}[k]$$

which can be rewritten as

$$\rho_{noise}^{pq}[k] = \frac{\left\langle \mathrm{cov}(r_i^p, r_i^q(k)) \right\rangle_i}{\sqrt{\left\langle \mathrm{var}(r_i^p) \right\rangle_i \left\langle \mathrm{var}(r_i^q) \right\rangle_i}} - \frac{\left\langle \mathrm{cov}(r_i^p, r_j^q(k)) \right\rangle_{i \neq j}}{\sqrt{\left\langle \mathrm{var}(r_i^p) \right\rangle_i \left\langle \mathrm{var}(r_i^q) \right\rangle_i}} \qquad (2)$$

where $\langle . \rangle_i$ denotes the expected value of the covariance of two quantities on the same trial, $\langle . \rangle_{i \neq j}$ denotes the expected value of the covariance of two quantities on different trials, and $r_i^p(k)$ is the vector r_i^p shifted by *k* time bins. Then one finds the corresponding $\rho_z^{pq}[k]$ by substituting (3) into (2) and solving numerically

$$\left\langle \mathrm{cov}(r_i^p, r_i^q(k)) \right\rangle_i = \left\langle \Phi_2 \left(\begin{bmatrix} s^p[n] \\ s^q[n+k] \end{bmatrix}, \begin{bmatrix} 1 & \rho_z^{pq}[k] \\ \rho_z^{qp}[k] & 1 \end{bmatrix} \right) \right\rangle_n - r_0^p r_0^q \qquad (3)$$

2.2 Variability of single cells on short time scales: Signal to noise ratio

In our previous paper, we specified the trial-to-trial variability of each cell as the signal to noise ratio

$$SNR = \frac{\mathrm{var}(\bar{r})}{\left\langle \mathrm{var}(\bar{r} - r_i) \right\rangle_i}$$

Because SNR is calculated for a single neuron, we omit cell indices for simplicity. We showed that, because the response is binary, if the PSTH is known, it uniquely defines SNR. This measure of variability is commonly used in the study of early sensory systems; it describes the reliability and precision of spiking across trials with respect to the stimulus. A cell with no noise, SNR = ∞, would spike on every trial in some bins, and never spike in the others. As the cell becomes less reliable, i.e. spikes are missed on some trials, or less precise, i.e. jitter moves spikes from one bin to the next, the SNR decreases. For a cell driven entirely by noise, i.e. a cell with a flat PSTH, the SNR is zero.

2.3 Variability of populations on short time scales: instantaneous signal and noise correlations

As described above, correlations in the spike trains of sensory neurons have both signal and noise components. In our previous paper, we specified the instantaneous signal and noise correlations, i.e. the correlations on the scale of a single time bin with zero lag $\rho_{signal}^{pq}[0]$ and $\rho_{noise}^{pq}[0]$. We showed that, because the responses are binary and the underlying noise process is Gaussian, matching the PSTH and signal to noise ratio of individual cells in the population will also match the instantaneous signal correlations between each pair. To match the instantaneous noise correlations between each pair of cells, the parameters $\rho_z^{pq}[0]$ that define the instantaneous correlations of *z* must be fit as described above.

2.4 Variability of single cells on long time scales: Fano Factor

The purpose of this paper is to analyze the relationship between the parameters of the Gaussian processes that underlie our model framework and experimentally relevant properties of spike trains that depend on temporal correlations – CV of the ISI distribution, Fano factor of spike counts, and correlation coefficient of spike counts for populations. For this purpose, one follows the fitting process described in Section 2.1 to find the values of the covariance matrix of *z* not only with zero lag, but at all lags *k* that correspond to significant noise correlations. To sample from the model and simulate spike trains, the correlated process *z* can be generated by using Gaussian conditioning (MacKay 2003; Macke et al. 2009; Lyamzin et al. 2010), or by using other methods, such as the multivariate autoregressive process described in (Gutnisky & Josić 2010).

Fano factor, the ratio of spike count variance to mean spike count across trials, has often been used as a measure of variability of spiking (Softky & Koch 1993; Berry et al. 1997; Shadlen & Newsome 1998; Kara et al. 2000). Fano factor is an obvious choice when neuronal activity is considered from the perspective of rate coding, and it has been used extensively to quantify variability in cortical cells, and, less commonly, cells of early sensory pathways. (Tolhurst et al. 1983; Sestokas & Lehmkuhle 1988; Hartveit & Heggelund 1994; Berry et al. 1997; Gur et al. 1997; Gershon et al. 1998; Buracas et al. 1998; Oram et al. 1999; Kara et al. 2000). A spike count distribution with a Fano factor equal to one has variability that is similar to that of a Poisson process, while neurons with Fano factors smaller than one are more reliable and those with high Fano factors are more variable.

Within the framework of our model, Fano factor can be expressed in terms of the correlation coefficients of *z*. Previous studies have shown that Fano factor depends on the integral of the spike train autocorrelation function, with Fano factor being bigger than one if the integral is positive, and less than one if the integral is negative (Macke et al. 2009). In our model, Fano factor depends not only on the integral of the autocorrelation function, but also, because the spike rate is not assumed to be stationary, on the value of *s* in every time bin.

The expression for Fano factor in terms of autocorrelations of the Gaussian process and signal can be written as (intermediate steps can be found in Appendix A):

$$FF = \frac{\sum_{n=1}^{N}\Phi\left(s[n],1\right) + 2\sum_{m=n+1}^{N}\sum_{n=1}^{N}\Phi_2\left(\begin{matrix}s[n]\\ s[m]\end{matrix}, \Sigma_z[m-n]\right) - \left(\sum_{n=1}^{N}\Phi\left(s[n],1\right)\right)^2}{\sum_{n=1}^{N}\Phi\left(s[n],1\right)} \quad (4)$$

Thus, given *s* in every time bin, when the autocorrelation function of a spike train is matched during the model fitting process, Fano factor is also matched. Modifying Fano factor will require changing the spike train correlation structure. Generally speaking, it can be changed arbitrarily, and in many cases simple scaling of correlation coefficients in $\Sigma_z[k]$ for all *k* is a valid method (note that not all Fano factors are realizable, as the covariance matrix of the underlying Gaussian process must remain valid; this limitation is discussed in detail in Section 4.1). In the case of scaling, (4) can be converted into an equation in terms of the unknown scaling factor and solved numerically. The correlation matrix will then be parameterized by the scaling factor α,

$$\Sigma_z^{\alpha}[m-n] = \begin{bmatrix} 1 & \alpha\rho_z[m-n] \\ \alpha\rho_z[m-n] & 1 \end{bmatrix}$$

so that (4) can be rewritten as

$$FF\left(\alpha\right) = \frac{\sum_{n=1}^{N}\Phi\left(s[n],1\right) + 2\sum_{m=n+1}^{N}\sum_{n=1}^{N}\Phi_2\left(\begin{matrix}s[n]\\ s[m]\end{matrix}, \Sigma_z^{\alpha}[m-n]\right) - \left(\sum_{n=1}^{N}\Phi\left(s[n],1\right)\right)^2}{\sum_{n=1}^{N}\Phi\left(s[n],1\right)} \quad (5)$$

Thus, as long as the necessary scaling factor corresponds to a realizable autocorrelation function, one can create a new Gaussian random process that would produce spike trains with the desired Fano factor.

To demonstrate the ability of our model to both match and manipulate the Fano factor in experimental spike trains, we fit our model to responses from a single neuron in mouse auditory cortex. Spike trains were recorded during the presentation of a 100 repeated trials of a 1 second segment of an amplitude modulated broadband noise sound. Fig. 2A shows a raster plot of the experimental responses.

Sampling spike trains (with the same trial length) from the fitted model resulted in responses with a Fano factor that matched the Fano factor in the original data, as shown in red in Fig. 2B (dashed line is the line of equality, error bars denote two standard deviations of the values obtained by repeated bootstrap sampling). To demonstrate the ability of the model to manipulate the Fano factor in experimental data, we used equation (5) to calculate a scaling factor for the autocorrelations of *z* such that the spike trains sampled from the model would have a different, predefined Fano factor. The responses of the model with scaled correlations accurately matched seven arbitrary values of Fano factor, including bigger and smaller values, as shown in black in Fig. 2B.

Fig. 2C shows histograms of spike counts for the two extreme cases of Fano factor modification shown in Fig. 2B. The distribution of spike counts for *FF* = 0.8, the least variable example, is shown in blue bars. Spike counts are relatively tightly clustered around the mean count. The distribution of spike counts for the most variable example, *FF* = 2.0, is shown in white bars. The distribution is considerably wider and contains many more extreme values.

The effects of the change in Fano factor can be seen by eye in the corresponding raster plots. Figs. 2D and 2E show the model responses with *FF* = 0.8 and *FF* = 2.0, respectively. The spike trains shown in Fig. 2D are much more regular than those shown in Fig. 2E, where one can see that a considerable proportion of spikes comes in bursts. Note that by design, instantaneous variability as measured by SNR is not affected by the manipulation of Fano factor: the SNR of the original data was 0.036, whereas the SNR of the manipulated data was 0.039 for *FF* =0.8 and 0.032 for *FF* =2.0. Similarly, the mean spike rate and PSTH are also unaffected.

2.5 Variability of single cells on short time scales: ISI distributions

Another important statistical property of the spike trains of individual cells is the distribution of ISIs. For the case of stationary spike rates, which is most commonly considered in the studies of ISI measures, ISI distributions are regularly shaped and are approximated well by gamma distributions (Nawrot et al. 2008; Shimokawa et al. 2010). The parameters of the gamma distribution in this case are defined by the autocorrelation function of the cell. Positive autocorrelations result in less regular spiking with ISI distributions that decrease monotonically with time. Negative autocorrelations result in more regular spiking, with the ISI distribution increasing with time for short ISIs, and then monotonically decreasing one once the peak has been passed.

A useful landmark in the family of ISI distributions of cells with stationary spike rates is the distribution associated with Poisson spiking. In this case the distribution of ISIs is exponential with decay speed depending on the spike rate. In terms of coefficient of variation (*CV*), which is the ratio of standard deviation of a distribution to its mean, these spike trains, with *CV* equal to one, lie between positively autocorrelated, irregular, spike trains (*CV*>1), and negatively correlated, regular, spike trains (*CV*<1).

When spike rate is non-stationary, as generally the case for sensory neurons, ISI distributions can be quite complex. An obvious example would be phase-locked spiking, where an ISI distribution can have several peaks at the ISIs that are multiples of the period of stimulus modulation. However, autocorrelations still have their own contribution in addition to the shape of the ISI distributions defined by the stimulus modulation. For example, as in the stationary spike rate case, negative autocorrelations due to refractoriness decrease the probability of short ISIs while increasing the probability of long ISIs, thus resulting in more regular spiking patterns (Berry et al. 1997).

For sensory systems, it is important to model spike trains with both refractoriness and flexible ISI variability in general. To characterize the regularity of spiking, we measure it

with the squared coefficient of variation (CV^2) of the ISI distribution. Because of the generally non-stationary spike rate, CV^2 cannot be used to compare spiking regularity of different cells, but it is useful to measure the changes in spiking regularity of the same cell. The absolute value of CV^2 can also provide some guidelines on the statistics of a spike train, and choosing this measure is partially motivated by the ability to reference its values in limiting cases. For example, in addition to the Poisson case noted above, CV^2 is also a useful measure for comparing spike trains to renewal processes, for which CV^2 equals the Fano factor of spike counts.

In our model, we can derive explicit expressions for the probability of an ISI given s in every time bin and the autocorrelation function of the spike trains (assuming that the ISI distribution is completely determined by second order statistics). The probability of an ISI of length m time bins is the mean probability of this ISI length over all time bins n.

$$P(ISI = m) = \left\langle P(ISI[n] = m) \right\rangle_n$$

These probabilities are calculated as follows:

$$\begin{aligned}
&P\left(ISI = m\right) = \left\langle P\left(ISI[n] = m\right)\right\rangle_n \\
&P\left(ISI[n] = m\right) = P\left(r[n+1, n+2, ..., n+m] = [0...1] \,|\, r[n] = 1\right) = \\
&= \frac{1}{P\left(r[n] = 1\right)} P\left(r[n, n+1, ..., n+m] = [10...01]\right) = \\
&= \frac{1}{\bar{r}[n]} \int_{-\infty}^{s[n]} \int_{s[n+1]}^{+\infty} \cdots \int_{m-1}^{s[n+m]} \int_{-\infty} N\ (0, \Sigma) dx
\end{aligned}$$

$$\text{where } \Sigma = \begin{bmatrix} 1 & \rho_z[1] & \cdots & \rho_z[m] \\ \rho_z[1] & 1 & & \rho_z[m-1] \\ \vdots & & \ddots & \\ \rho_z[m] & \rho_z[m-1] & & 1 \end{bmatrix} \qquad (6)$$

Thus, the probability distribution of ISIs, and hence its coefficient of variation, is matched as long as the autocorrelation function of the spike trains is matched, to the extent that the ISI distribution depends only on the second order statistics of the intrinsic noise.

Modifying the CV^2 of an ISI distribution is done by manipulating the correlation coefficients of z. As in the case of Fano factor, this can be done by arbitrary changes of correlation structure, but often works with scaling: all correlation coefficients in the covariance matrix of z are multiplied by a parameter.

$$\Sigma^{\alpha} = \begin{bmatrix} 1 & \alpha\rho_z[1] & \cdots & \alpha\rho_z[m] \\ \alpha\rho_z[1] & 1 & & \alpha\rho_z[m-1] \\ \vdots & & \ddots & \\ \alpha\rho_z[m] & \alpha\rho_z[m-1] & & 1 \end{bmatrix}$$

and the probability of an ISI of length *m* at a given time bin *n* with a spike is

$$P(\alpha) = P(ISI[n] = m) = \frac{1}{\bar{r}[n]} \int_{-\infty}^{s[n]} \int_{s[n+1]}^{+\infty} \cdots \int_{-\infty}^{s[n+m]} N \ (0, \Sigma^{\alpha}) dx$$
$$m-1$$

Once the desired value of CV^2 is specified, the value of the scaling factor can be found numerically from the following formula.

$$CV^2(\alpha) = \frac{\mathrm{var}(P(\alpha))}{\langle P(\alpha) \rangle^2} \qquad (7)$$

Where $P(\alpha)$ is the probability distribution of ISIs as a function of the scaling parameter α.

Obviously, for independent manipulation of CV^2 and *FF*, modification of the correlations of *z* via simple scaling is not suitable. In this case, one has to use modifications of the correlations of *z* that have two free parameters. One such example is discussed in Section 3.5.

To demonstrate the ability of our model to match and manipulate the ISI distribution of experimental data, we fit our model to a set of spike trains from a single cell in mouse inferior colliculus (IC). Spike trains were recorded during the presentation of 100 repeated trials of a 1 second segment of amplitude modulated broadband noise sounds (see raster in Fig. 3A).

With the assumption of ISI distribution being defined by second order statistics, we match the signal component and autocorrelation function and thus the ISI distribution. The ISI distribution of the original data and the data sampled from the fitted model are shown in Fig. 3B. The original distribution of ISIs is shown in blue bars, and the distribution of ISIs sampled from the fitted model is shown by the black and white line. The CV^2 of the original data is 0.40 and CV^2 of the data sampled from the fitted model is 0.38.

Fig. 3C demonstrates how scaling of the autocorrelations can change the CV^2 of the spike trains. We found the scaling factor of the autocorrelation function that would produce spike trains with CV^2 of 0.3 and 1.0, thus increasing and decreasing spiking regularity respectively. The resulting distributions of ISIs are compared with the original in Fig. 3C. (Note that depending on the length of a trial in time bins, the use of sampling in the numerical search for the parameter value in the equation (7) can be computationally more efficient than calculation of the ISI probability distribution by using the equations

(6). The results shown in Fig. 3C were obtained by sampling the spike trains with scaled correlation coefficients and calculating the ISI probability distribution directly from these spike trains.)

The effects of changes in the CV^2 of the ISI distribution can be seen on the raster plots in Fig. 3D and 3E. In the case of CV^2 of 1.0, there is no observed refractoriness and spiking often occurs in bursts due to strong positive autocorrelations, as can be seen in the inset in Fig. 3D. For CV^2=0.3, spiking becomes more refractory, as can be seen in the inset in Fig. 3E; almost every spike has a period of silence of one or more time bins, which is the result of increased negative noise autocorrelations.

2.6 Variability of populations on long time scales: Count correlations

The effects of noise correlations on population coding can be studied across multiple time scales (Bair et al. 2001; Averbeck et al. 2006; Smith & Kohn 2008). In our previous work (Lyamzin et al. 2010), we mainly addressed instantaneous correlations that are related to spiking synchrony between two cells. One can also consider correlations that are calculated from larger time bins or whole trials. Correlations in spike counts across seconds have the capacity to both increase and decrease the information carried by a population depending on the correlation structure across the population, and the prospective from which their effect is studied (Abbott & Dayan 1999).

As with the variability measures for individual cells described above, the pairwise count correlations for a population of cells can be related to the covariance matrix of the noise process *z*.

Count correlations are defined as

$$CC = \frac{\operatorname{cov}\left(\sum_{n=1}^{N} r^p[n], \sum_{m=1}^{N} r^q[m]\right)}{\sqrt{\operatorname{var}\left(\sum_{n=1}^{N} r^p[n]\right)\operatorname{var}\left(\sum_{m=1}^{N} r^q[m]\right)}} = \frac{\left\langle \left(\sum_{n=1}^{N} r^p[n]\right)\left(\sum_{m=1}^{N} r^q[m]\right)\right\rangle - \left\langle \sum_{n=1}^{N} r^p[n]\right\rangle \left\langle \sum_{m=1}^{N} r^q[m]\right\rangle}{\sqrt{\operatorname{var}\left(\sum_{n=1}^{N} r^p[n]\right)\operatorname{var}\left(\sum_{m=1}^{N} r^q[m]\right)}}$$

First, we expand the product of spike counts:

$$\left\langle \sum_{n=1}^{N} r^p[n] \sum_{m=1}^{N} r^q[m] \right\rangle = \left\langle r^p[1]r^q[1] + r^p[1]r^q[2] + \ldots + r^p[n]r^q[m] + \ldots + r^p[N]r^q[N] \right\rangle$$

and express each of the $r^p[n]r^q[m]$ products in terms of cumulative distributions:

$$\left\langle \sum_{n=1}^{N} r^p[n] \right\rangle = \sum_{n=1}^{N} \Phi\left(s^p[n], 1\right)$$

$$\left\langle \sum_{n=1}^{N} r^p[n] \sum_{m=1}^{N} r^q[m] \right\rangle = \sum_{m=1}^{N}\sum_{n=1}^{N} \Phi_2\left(\begin{matrix} s^p[n] \\ s^q[m] \end{matrix}, \Sigma_z[m-n]\right)$$

The formula for count correlations can then be rewritten as

$$CC = \frac{\sum_{m=1}^{N}\sum_{n=1}^{N}\Phi_2\left(\begin{matrix}s^p[n]\\ s^q[m]\end{matrix}, \Sigma_z[m-n]\right) - \sum_{n=1}^{N}\Phi\left(s^p[n],1\right)\sum_{m=1}^{N}\Phi\left(s^q[m],1\right)}{\sqrt{\mathrm{var}\left(\sum_{n=1}^{N} r^p[n]\right)\mathrm{var}\left(\sum_{m=1}^{N} r^q[m]\right)}} \quad (8)$$

where assuming symmetry of correlations

$$\Sigma_z[m-n] = \begin{bmatrix} 1 & \rho_z^{pq}[m-n] \\ \rho_z^{pq}[m-n] & 1 \end{bmatrix}$$

and

$$\mathrm{var}\left(\sum_{n=1}^{N} r^p[n]\right) = \sum_{n=1}^{N}\Phi\left(s^p[n],1\right) + \sum_{n=1}^{N}\sum_{m=1}^{N}\Phi_2\left(\begin{matrix}s^p[n]\\ s^q[m]\end{matrix}, \Sigma_z[m-n]\right) - \left(\sum_{n=1}^{N}\Phi\left(s^p[n],1\right)\right)^2$$

As you can see, count correlations depend only on *s* and the crosscorrelation function between two cells. Thus, when the crosscorrelation function between two neurons is matched during the model fitting process, the count correlations are also matched.

As in the case of the single cell properties described above, manipulation of count correlations can often be achieved by finding a scaling factor for the crosscorrelations of *z* that will result in spike trains with the desired count correlations. The covariance matrix in this case is rewritten as

$$\Sigma_z^{\alpha}[m-n] = \begin{bmatrix} 1 & \alpha\rho_z^{pq}[m-n] \\ \alpha\rho_z^{pq}[m-n] & 1 \end{bmatrix}$$

where α is the scaling factor, and the variance of spike counts can be rewritten as

$$\mathrm{var}\left(\sum_{n=1}^{N} r^p[n]\right) = \sum_{n=1}^{N}\Phi\left(s^p[n],1\right) + \sum_{n=1}^{N}\sum_{m=1}^{N}\Phi_2\left(\begin{matrix}s^p[n]\\ s^q[m]\end{matrix}, \Sigma_z^{\alpha}[m-n]\right) - \left(\sum_{n=1}^{N}\Phi\left(s^p[n],1\right)\right)^2$$

so that count correlation value from (8) now depends on the scaling factor

$$CC(\alpha) = \frac{\sum_{m=1}^{N}\sum_{n=1}^{N}\Phi_2\left(\begin{matrix}s^p[n]\\ s^q[m]\end{matrix}, \Sigma_z^{\alpha}[m-n]\right) - \sum_{n=1}^{N}\Phi\left(s^p[n],1\right)\sum_{m=1}^{N}\Phi\left(s^p[m],1\right)}{\sqrt{\mathrm{var}\left(\sum_{n=1}^{N} r^p[n]\right)\mathrm{var}\left(\sum_{m=1}^{N} r^q[m]\right)}} \quad (9)$$

Equation (9) can be solved for α numerically.

When the scaling factor of the crosscorrelation function is found, one can create a new Gaussian random process *z* that would produce spike trains with the required value of count correlations.

To demonstrate the ability of our model to both match and manipulate pairwise count correlations in experimental spike trains, we fit our model to responses from a pair of neurons in gerbil auditory cortex. Spike trains were recorded during the presentation of a 100 repeated trials of a 200 ms segment of an amplitude modulated broadband noise sound (see raster plots in Fig. 4A).

Sampling spike trains (with the same trial length) from the fitted model resulted in responses with a count correlation coefficient that matched the count correlation coefficient in the original data, as shown in red in Fig. 4B. We also used equation (9) to calculate the scaling factor for the autocorrelations of *z* such that the result of sampling from the model would have a different, predefined count correlation coefficient. The responses of the model with scaled correlations accurately matched four arbitrary values of correlation coefficient, including bigger and smaller values, as shown in black in Fig. 4B.

The difference between the two extreme simulated cases (count correlation of 0.4 and 0.7) can be seen from the plot of the joint spike count distributions in Figs. 4C and 4D (colored contours represent levels of equal probability, points within red contours correspond to the most frequently occurring pairs of spike count values, and points within blue contours correspond to less frequently occurring pairs of spike count values). The distribution of spike counts for spike trains with CC = 0.7 shown in Fig. 4C is more elongated along the diagonal than the distribution for spike trains with CC = 0.4 shown in Fig. 4D.

3. Simulating spike trains *de novo*

3.1 Model framework

In addition to matching and manipulating a particular set of experimental data, the model can also be used to generate a new spike trains with a desired set of statistics. Unlike the data-based model in which the PSTH of each neuron and several associated quantities are constrained to match those observed experimentally, the *de novo* version of the model provides a fully flexible framework in which to investigate the interaction of variability and covariability on different time scales.

In recent studies (De La Rocha et al. 2007) it has been shown, both theoretically and experimentally, that correlations between two cells depend on their spike rates. Within our the dichotomized Gaussian framework, this dependency also holds (Macke et al. 2011). The advantage of our model is that similar relationships can be observed for a wider range of parameters, e.g. SNR, signal correlations, Fano factors etc. Additionally, our model can also be used to study how values of a given parameter constrain the possible range of the others.

The model used for simulation of spike trains *de novo* is a modification of the one described above. A version of this model with specified instantaneous variability and covariability was described in detail in our previous work (Lyamzin et al. 2010). As before, spike trains are generated as binary matrices by thresholding a sum of a signal *s* that is the same on every trial and a noise *z* that is different on every trial.

$$r_i^p[n] = \begin{cases} 1, & (s^p[n] + z_i^p[n]) > \theta^p \\ 0, & (s^p[n] + z_i^p[n]) \le \theta^p \end{cases}$$

where $i = 1 \ldots I$, $n = 1 \ldots N$, and $p = 1 \ldots P$.

The PSTH of each cell is no longer constrained by the experimental responses, but must be defined in another way. One can simply choose an arbitrary deterministic signal to create a pseudo-experimental situation, or choose a parametric stochastic signal to allow precise control of the relevant spike train parameters. For the latter case, we set *s* to be a Gaussian process with variance $\sigma_{s^p}^2$ that is a free parameter for each cell. In this case, as both the mean spike rate r_0 and the SNR are no longer constrained by the PSTH, a second free parameter is required in the form of a variable threshold θ^p. Both *SNR* and r_0 can be expressed in terms of $\sigma_{s^p}^2$ and θ^p (Lyamzin et al. 2010) and the required values solved for numerically.

Because instantaneous signal correlations are also no longer defined by the PSTHs of the two cells, ρ_{signal}^{pq} becomes a free parameter of the spike trains, which can be manipulated by changing the corresponding ρ_s^{pq}, the correlation coefficient between the different components of *s*. In the previous paper, we showed that for Gaussian *s*, ρ_{signal}^{pq} and ρ_{noise}^{pq} can be expressed in terms of ρ_s^{pq} and ρ_z^{pq}. Hence, given the desired values of ρ_{signal}^{pq} and ρ_{noise}^{pq}, one can generate the correlated *s* and *z* processes that would give spike trains with desired correlations after thresholding.

3.2 Fano factor of *de novo* spike trains

If we assume the signal *s* is Gaussian, we can write closed form expressions for the same measures of variability and covariability presented in the section 2. Fano factor can be written out as

$$FF = \frac{\sum_{n=1}^{N} \Phi(-\theta, \sigma_s^2 + 1) + \sum_{n \neq m}^{N} \Phi_2 \left(\begin{matrix} -\theta \\ -\theta \end{matrix}, \Sigma_{s+z}[m-n] \right) - \left(\sum_{n=1}^{N} \Phi(-\theta, \sigma_s^2 + 1) \right)^2}{\left(\sum_{n=1}^{N} \Phi(-\theta, \sigma_s^2 + 1) \right)}$$

The correlation matrix of the two-dimensional Gaussian now contains correlation coefficients of *s* and *z*, as well as the variance of *s*. For derivation and details, see Appendix B. Using this formula for cases when s can be assumed Gaussian can dramatically decrease calculation times.

3.3 ISI distribution of *de novo* spike trains

In case of *de novo* generation of spike trains with Gaussian *s*, we can write out the probability of finding an ISI of a certain length *m* analogous to Section 3.2 (again assuming that the ISI distribution is determined only by second order statistics)

$$P\left(ISI=m\right)=\left\langle P\left(ISI[n]=m\right)\right\rangle_n$$

where

$$\begin{aligned}
&P\left(ISI=m\right)=\left\langle P\left(ISI[n]=m\right)\right\rangle_n \\
&P\left(ISI[n]=m\right)=P\left(r[n+1,n+2,\ldots,n+m]=[0\ldots1]\,|\,r[n]=1\right)= \\
&=\frac{1}{P\left(r[n]=1\right)}P\left(r[n,n+1,...,n+m]=[10...01]\right)= \\
&=\frac{1}{r_0}\int_{-\infty}^{-\theta}\int_{-\theta}^{+\infty}\ldots\int_{-\infty}^{-\theta}N\ \ (0,\Sigma_{s+z})dx \\
&\qquad\qquad m-1
\end{aligned}$$

and

$$\Sigma_{s+z}=\begin{bmatrix}
1+\sigma_s^2 & \rho_z[1]+\sigma_s^2\rho_s[1] & \cdots & \rho_z[k]+\sigma_s^2\rho_s[k] \\
\rho_z[1]+\sigma_s^2\rho_s[1] & 1+\sigma_s^2 & & \rho_z[k-1]+\sigma_s^2\rho_s[k-1] \\
\vdots & & \ddots & \\
\rho_z[k]+\sigma_s^2\rho_s[k] & \rho_z[k-1]+\sigma_s^2\rho_s[k-1] & & 1+\sigma_s^2
\end{bmatrix}$$

When the probability distribution is found for all ISI lengths, its coefficient of variation is calculated as $CV^2=\dfrac{\mathrm{var}(P)}{\left\langle P\right\rangle^2}$

3.4. Count correlations of *de novo* spike trains

Spike count correlation coefficient by definition is

$$CC=\frac{\mathrm{cov}\left(\sum_{n=1}^{N}r^p[n],\sum_{m=1}^{N}r^q[m]\right)}{\sqrt{\mathrm{var}\left(\sum_{n=1}^{N}r^p[n]\right)\mathrm{var}\left(\sum_{m=1}^{N}r^q[m]\right)}}=\frac{\left\langle\left(\sum_{n=1}^{N}r^p[n]\right)\left(\sum_{m=1}^{N}r^q[m]\right)\right\rangle-\left\langle\sum_{n=1}^{N}r^p[n]\right\rangle\left\langle\sum_{m=1}^{N}r^q[m]\right\rangle}{\sqrt{\mathrm{var}\left(\sum_{n=1}^{N}r^p[n]\right)\mathrm{var}\left(\sum_{m=1}^{N}r^q[m]\right)}}$$

where similarly to the analogous formula for Fano factor

$$\left\langle\sum_{n=1}^{N}r^p[n]\sum_{m=1}^{N}r^q[m]\right\rangle=\sum_{m=1}^{N}\sum_{n=1}^{N}\Phi_2\begin{pmatrix}-\theta^p\\-\theta^q\end{pmatrix},\Sigma_s[m-n]+\Sigma_z[m-n]\Bigg)$$

and

$$\mathrm{var}\left(\sum_{n=1}^{N} r^p[n]\right) =$$

$$N\Phi\left(-\theta^p, \sigma_{s^p}^2 + 1\right) + \sum_{n=1}^{N}\sum_{m=1}^{N} \Phi_2\left(\begin{matrix}-\theta^p \\ -\theta^p\end{matrix}, \begin{bmatrix} \sigma_{s^p}^2 + 1 & \rho_z^{pp}[m-n] \\ \rho_z^{pp}[m-n] & \sigma_{s^p}^2 + 1 \end{bmatrix}\right) - \left(N\Phi\left(-\theta^p, \sigma_{s^p}^2 + 1\right)\right)^2$$

where, in turn, assuming symmetry of correlations

$$\Sigma_s[m-n] = \begin{bmatrix} \sigma_{s^p}^2 & \sigma_{s^p}\sigma_{s^q}\rho_s^{pq}[m-n] \\ \sigma_{s^p}\sigma_{s^q}\rho_s^{pq}[m-n] & \sigma_{s^p}^2 \end{bmatrix}$$

$$\Sigma_z[m-n] = \begin{bmatrix} 1 & \rho_z^{pq}[m-n] \\ \rho_z^{pq}[m-n] & 1 \end{bmatrix}$$

This formula can be used to analyze the dependencies of count correlation on single cell and pairwise parameters. For example, because the variances of spike counts in the denominator of the expression for count correlations depend only on autocorrelations, but not on cross correlations, one can see that an increase in autocorrelations (which leads to an increase in Fano factor) decreases count correlations when crosscorrelations and other parameters are fixed.

3.5 Independent manipulation of *CV²* and *FF*

In section 2, we showed how, by using the version of the dichotomized Gaussian model with temporal correlations, one can match and manipulate different parameters of variability and covariability in experimental responses. In those examples, the different measures of variability and covariability were considered independently of each other, though all may be affected by changes in the auto- and cross correlation functions of the spike trains. However, independent manipulation of different types of variability is also possible if the correlation functions are parameterized by multiple parameters.

For example, both Fano factor and *CV²* are influenced by autocorrelations, but do not necessarily determine each other. For example, in the case of a homogeneous binary process (the case of a constant *s* in our model), the relationship between Fano factor and *CV²* of the ISI distribution is expressed as (Cox & Lewis 1966).

$$FF = CV^2\left(1 + \sum_{k=1}^{\infty} \rho_{serial}(k)\right)$$

Serial correlations $\rho_{serial}(k)$ are, in turn, functions of autocorrelations of *z* and their sum will depend on the shape of the autocorrelation functions. Hence, in this simple case one can potentially manipulate *CV²* and Fano factor independently of one another by

changing the sum of the serial correlation coefficients by changing the autocorrelation functions.

To provide a more complex example of the independent manipulation of Fano factor and CV^2 of the ISI distribution in our model, we parameterized autocorrelation functions with a family of functions with two free parameters (α, β) that model both refractory spiking and burstiness:

$$r_z[k] = \begin{cases} 1 & k = 0 \\ e^{-ak} - 2e^{-bk} & k = 1 \end{cases},$$

The sum of exponents allows for different positive overshoots after a refractory period, which in turn will make it possible to have several different values of Fano factor for a given value of CV^2.

First, we investigated how *FF* and CV^2 change with the changing parameters of the autocorrelation functions, and then investigated how *FF* changes along the lines on the (α, β) plane where the CV^2 stays constant. This analysis demonstrates the degree to which Fano factor and CV^2 can be manipulated independently. For this purpose, we created two examples – one in which *s* is a Gaussian process as in the *de novo* framework described above, and another one in which *s* is an arbitrary deterministic signal in the form of a sine wave (see figure legend for details).

Fig. 5A shows the values of CV^2 for a range of parameters of the described family of functions in the case of sine wave *s*. For certain combinations of parameters (α, β), the autocorrelation functions result in covariance matrices of *z* that are not positive-definite, and thus the process is not realizable. This region of parameter space is shown in white. For all the other possible combinations of parameter values within the range selected for this illustration (α between 0 and 2, and beta between 0 and 5), different values of CV^2 are shown in different colors from blue for the lowest values to red for the highest. We chose three values of interest: $CV^2 = 5.5$, $CV^2 = 5.7$, $CV^2 = 5.9$ and plotted the lines along which the values of CV^2 were equal to these values (shown in red). At the ends of the lines of equality we plotted the autocorrelations of *z* that correspond to the combination of parameters at the given end. From these plots, one can see that the same values of CV^2 can be achieved for different autocorrelation functions. Next, as shown in Fig. 5B, we calculated the Fano factors along each iso-CV^2 curve. The insets show autocorrelation functions for the corresponding set of parameters at both ends of the curves (marked with black circles). Fig. 5B shows that Fano factor changes considerably along the iso-CV^2 curves and, thus, there is a room for independent manipulation of CV^2 and Fano factor. However, when we chose *s* to be a Gaussian random process rather than a sine wave and performed the same analysis as above, the results were somewhat different. As shown in Fig. 5C, even though the autocorrelation functions change along the lines of equality for CV^2, the range of Fano factors is less that in the case of sinusoidal *s*. Thus, in this case, there is less room for independent manipulation of CV^2 and Fano factor.

The results in figure 5 suggest a strong influence of the signal itself on the statistics of the resulting spike trains. To investigate this issue in more detail, we chose one

particular correlation function from the family of functions described above (with α = 0.3 and β = 1.5), and varied the frequency of the sinusoidal signal and its amplitude relative to the intrinsic noise.

In the stimulus driven case (high signal amplitude), the CV^2 values changed substantially with changing frequency. As expected, in the noise driven case (with low signal amplitude), changes in frequency had no significant effect on the ISI distributions and hence the CV^2 values. The figure shows that similar values of CV^2 can be achieved through different combinations of signal and noise parameters, and through different ISI distributions. We did the same analysis for several other, qualitatively different functions from the same family, and found that the behavior of the curves is similar for all of them.

These examples demonstrate that, although our framework can be used to vary different measures of variability and covariability independently, the extent to which this is possible can depend strongly on the stimulus.

4. Limitations of the model

We have shown how our model can be applied to spike trains with a range of different statistical properties. However the dichotomized Gaussian framework has limitations as well, as have been noted before (Macke et al. 2009; Gutnisky & Josić 2010), most of which stem from the fact that the covariance matrices of Gaussian processes must be positive-definite. This constraint imposes limitations on the allowed auto- and crosscorrelation functions that can be used in our model.

In particular, one strong limitation of the model is that only weak negative correlations can be reproduced (Macke et al. 2009; Gutnisky & Josić 2010). The ability of our model to capture weak negative correlations has been illustrated in section 2.5, where we reproduced the refractory period of experimental spike trains, and then increased it, such that the spiking became even more regular. However, as shown in Fig. 7A, the model is incapable of generating spike trains with strong negative correlations.

The red line shows the values of the autocorrelations of z that would be required to match the autocorrelations in a set of spike trains with a strong refractory period recorded from the IC. The covariance matrix corresponding to these values is not positive definite, and thus the process is not realizable. We scaled the values and weakened the negative autocorrelations until we obtained a realizable process, resulting in the values shown in black.

When a process is close to unrealizable, the strongest realizable correlations, and the behavior of the process that is generated, depend on the maximum lag at which the correlations are specified. In the example in Fig. 7A, we only specified the autocorrelation function of z at lags 1 to 4. However, when we then used our model to generate a set of spike trains, the resulting spike trains contained significant positive autocorrelations at lag 5, as shown in black in Fig. 7B. If one deliberately sets a zero correlation coefficient for z at lag 5, the positive autocorrelations in the spike trains are simply delayed to lag 6. This procedure, of explicitly setting zero correlation coefficients

at successively larger lags, ultimately leads to a non-positive definite covariance matrix for z, and thus an unrealizable process (note that this artifact is present not only in the spike train correlations, but also in the z process itself). This example illustrates the importance of choosing the range of autocorrelations carefully, and verifying that the corresponding model is realizable and can generate spike trains with the desired statistics.

Because of the limitations on the negative autocorrelation functions, values of Fano factor cannot be arbitrarily small. Similarly, ISI distribution that require a strong refractory period may not be realizable.

5. Discussion

The variability of single neurons and the covariability of neuronal populations can be measured on multiple time scales from synchrony in precise spike timing to correlations in spike count over a period of seconds. In previous work, we demonstrated the ability of a dichotomized Gaussian framework to reproduce the instantaneous variability in single neurons, and the instantaneous signal and noise correlations (i.e. synchrony) in neuronal populations (Lyamzin et al. 2010). In this study, we have shown how the dichotomized Gaussian model can capture variability and covariability on multiple time scales, and demonstrated its success in capturing the Fano factor, ISI distributions, and count correlations in experimental spike trains.

Our main motivation for developing the model was to enable the systematic study of different response properties of neuronal populations that cannot be easily decoupled experimentally. For example, to determine the time scales at which variability is important for coding a particular stimulus parameter, it is helpful to be able to manipulate the variability on different time scales independently. While this manipulation may be difficult to achieve experimentally (any change in the spike train autocorrelation function is likely to affect variability on multiple timescales), it may be achievable within our model framework with appropriate parameterization. For example, as shown in section 3.5, Fano factor and CV^2 of the interspike interval distribution can be manipulated independently under certain conditions.

Our model can also be used to systematically investigate the dependencies between different properties of spike trains. For example, our model framework captures the dependency of correlations on spike rate demonstrated in (De La Rocha et al. 2007) and analyzed in-depth in (Macke et al. 2011). Because our model captures the full second order structure of population spike trains, it should prove to be a useful tool for studying a wide range of such dependencies, as well as for determining how the value of a given response property constrains the possible range of possible values of others. For example, if one assumes that only second order correlations and stimulus affect the values of Fano factor and CV^2, our model provides a way to investigate the dependencies between them. Another example would be the inverse dependency of count correlations on the temporal autocorrelations of individual neurons, when pairwise correlations remain fixed, as noted in Section 3.4.

We have also shown that the framework has significant limitations, the most important of which is the inability to generate or model spike trains with strong negative correlations. This limitation arises from the requirement that the covariance matrix of the multivariate Gaussian process z is positive definite. Note that this limitation is inherent to the model framework, and is independent of the specific method used to generate the Gaussian process when simulating spike trains (Gutnisky & Josić 2010). Because of this limitation, the model is better suited to modeling spike trains from cortical areas where correlations are typically positive (Bair et al. 2001; Smith & Kohn 2008) than subcortical areas where strong negative correlations due to refractory effects may be prominent.

Another limitation of our model is that it is only designed to capture the first and second order statistical properties of spike trains and may not accurately reflect higher order properties, such as the dependencies between successive ISIs or correlations among groups of neurons that are not predictable from second order statistics. This limitation may be important for certain brain areas where higher order correlations carry significant information (Ohiorhenuan et al. 2010, Farkhooi et al. 2009, Nesse et al. 2010). One should note that even though the dichotomized Gaussian does not allow for explicit control of higher-order correlations, it still produces them as a result of the thresholding process (Macke et al., 2011). Interestingly, the higher-order correlations produced by the Gaussian model correspond well to the higher-order correlations found in cortex (Yu et al. 2012). These results suggest that models based on the dichotomized Gaussian framework are potentially promising for research into the effects of higher-order correlations on coding, although further experimental study of this issue is needed.

References

Abbott LF & Dayan P. 1999. The effect of correlated variability on the accuracy of a population code. Neural Computation, 11(1):91-101.

Averbeck BB, Latham PE & Pouget A. 2006. Neural correlations, population coding and computation. Nature Reviews Neuroscience, 7(5):358-66.

Bair W, Zohary E & Newsome WT. 2001. Correlated firing in macaque visual area MT: time scales and relationship to behavior. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 21(5):1676-1697.

Berry MJ, Warland DK & Meister M. 1997. The structure and precision of retinal spike trains. Proceedings of the National Academy of Sciences of the United States of America, 94(10):5411-5416.

Buracas GT, Zador A, DeWeese M & Albright TD. 1998. Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. Neuron, 20(5):959-969.

Cox DR & Lewis PAW. 1966. The statistical analysis of series of events, London, Methuen.

Dan Y, Alonso JM, Usrey MW & Reid CR. 1998. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. Nature Neuroscience, 1(6):501-507.

Farkhooi F, Strube-Bloss MF, Nawrot MP. 2009. Serial correlation in neural spike trains: experimental evidence, stochastic modeling, and single neuron variability. Physical Review E, 79(2 Pt 1):021905

Gershon ED, Wiener MC, Latham PE & Richmond BJ. 1998. Coding strategies in monkey V1 and inferior temporal cortices. Journal of Neurophysiology, 79(3):1135-1144.

Gur M, Beylin A & Snodderly DM. 1997. Response variability of neurons in primary visual cortex (V1) of alert monkeys. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 17(8):2914-2920.

Gutnisky DA & Josić K. 2010. Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process. Journal of Neurophysiology, 103(5):2912-2930.

Hartveit E & Heggelund P. 1994. Response variability of single cells in the dorsal lateral geniculate nucleus of the cat. Comparison with retinal input and effect of brain stem stimulation. Journal of Neurophysiology, 72(3):1278-1289.

Kara P, Reinagel P & Reid RC. 2000. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. Neuron, 27(3):635-646.

Krumin M & Shoham S. 2009. Generation of spike trains with controlled auto-and cross-correlation functions. Neural Computation, 21:1642–1664.

De La Rocha J, Doiron B, Shea-Brown E, Josić K, & Reyes A. 2007. Correlation between neural spike trains increases with firing rate. Nature, 448:802–806.

Latham PE & 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.

Lyamzin DR, Macke JH & Lesica NA. 2010. Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations. Frontiers in Computational Neuroscience, 4:144.

MacKay D. 2003. Information theory, inference, and learning algorithms, Cambridge, UK; New York: Cambridge University Press.

Macke JH, Opper M & Bethge M. 2011. Common input explains higher-order correlations and entropy in a simple model of neural population activity. Physical Review Letters. 106(20):208102.

Macke JH, Berens P, Ecker AS, Tolias AS & Bethge M. 2009. Generating spike trains with specified correlation coefficients. Neural Computation, 21(2):397-423.

Mainen ZF & Sejnowski TJ. 1995. Reliability of spike timing in neocortical neurons. Science, 268(5216):1503-1506.

Manwani A, Steinmetz PN & Koch C. 2002. The impact of spike timing variability on the signal-encoding performance of neural spiking models. Neural Computation, 14(2):347-367.

Nawrot MP, Boucsein C, Molina VR, Riehle A, Aertsen A & Rotter S. 2008. Measurement of variability dynamics in cortical spike trains. Journal of Neuroscience Methods, 169(2):374-390.

Nawrot MP. 2010. Analysis and Interpretation of Interval and Count Variability in Neural Spike Trains. In S. Grün & S. Rotter, eds. Analysis of Parallel Spike Trains. Boston, MA: Springer US: 37-58

Nesse WH, Maler L & Longtin A. 2010. Biophysical information representation in temporally correlated spike trains. Proceedings of the National Academy of Sciences of the United States of America, 107(51):21973-21978.

Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM. & Victor JD. 2010. Sparse coding and high-order correlations in fine-scale cortical networks. Nature, 466:617–621.

Oram MW, Wiener MC, Lestienne R & Richmond BJ. 1999. Stochastic nature of precisely timed spike patterns in visual system neuronal responses. Journal of Neurophysiology, 81(6):3021-3033.

Puchalla JL, Schneidman E, Harris RA & Berry MJ 2nd. 2005. Redundancy in the population code of the retina. Neuron, 46(3):493-504.

Ratnam R & Nelson ME. 2000. Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 20(17):6672-6683.

Romo R, Hernandez A, Zainos A & Salinas E. 2003. Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. Neuron, 38(4):649–657.

Schneidman E, Freedman B & Segev I. 1998. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. Neural Computation, 10(7):1679-1703.

Schneidman E, Bialek W & Berry MJ 2nd. 2003. Synergy, redundancy, and independence in population codes. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 23(37):11539-11553.

Sestokas AK & Lehmkuhle S. 1988. Response variability of X- and Y-cells in the dorsal lateral geniculate nucleus of the cat. Journal of Neurophysiology, 59(2):317-325.

Shadlen MN & Newsome WT. 1998. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 18(10):3870-3896.

Shimokawa T, Koyama S & Shinomoto S. 2010. A characterization of the time-rescaled gamma process as a model for spike trains. Journal of Computational Neuroscience, 29(1-2):183-191.

Smith MA & Kohn A, 2008. Spatial and temporal scales of neuronal correlation in primary visual cortex. Journal of Neuroscience, 28:12591.

Softky WR & Koch C. 1993. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 13(1):334-350.

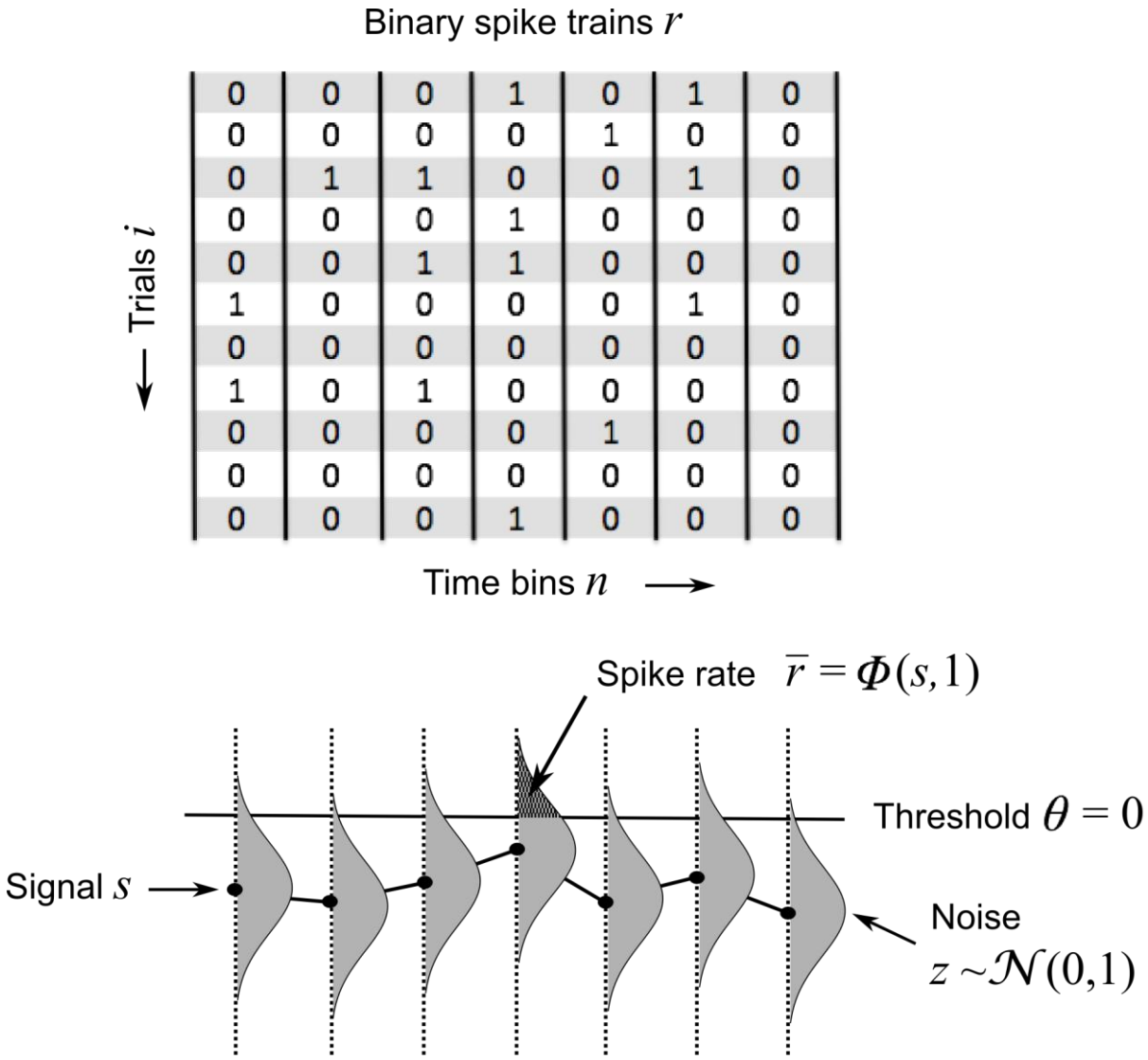
Tchumatchenko T, Giesel T, Volgushev M & Wolf F, 2010. Signatures of synchrony in pairwise count correlations. Frontiers in Computational Neuroscience, 4:1.

Tolhurst DJ, Movshon JA & Dean AF. 1983. The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Research, 23(8):775-785.

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.

Yu S, Yang H, Nakahara H, Santos GS, Nicolić D & Plenz D. 2012. Higher-order interactions characterized in cortical activity. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 31(48):17514-26.

Zohary E, Shadlen MN & Newsome WT. 1994. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature, 370(6485):140-143.



DG Model:

$$r_i[n] = \begin{cases} (s[n]+z_i[n]) > 0 \\ (s[n]+z_i[n]) \leq 0 \end{cases}$$

Figure 1. Schematic diagram of dichotomized Gaussian model for a single cell. The response of a cell to a single stimulus presentation is assumed to be represented by a binary vector *r* of length *n*, in which 1 corresponds to a time bin with a spike, and 0 corresponds to a time bin without a spike. The same stimulus is presented *I* times and the response of a cell on a particular trial in a particular time bin is denoted $r_i[n]$. Spikes are generated based on a Gaussian random 'noise' process *z* with unit variance, whose mean is different from time bin to time bin as specified by the value of a 'signal' *s.* In each time bin *n*, the probability of observing a spike across all trials (i.e. the PSTH) is given by the cumulative Gaussian distribution function evaluated at $s[n]$.

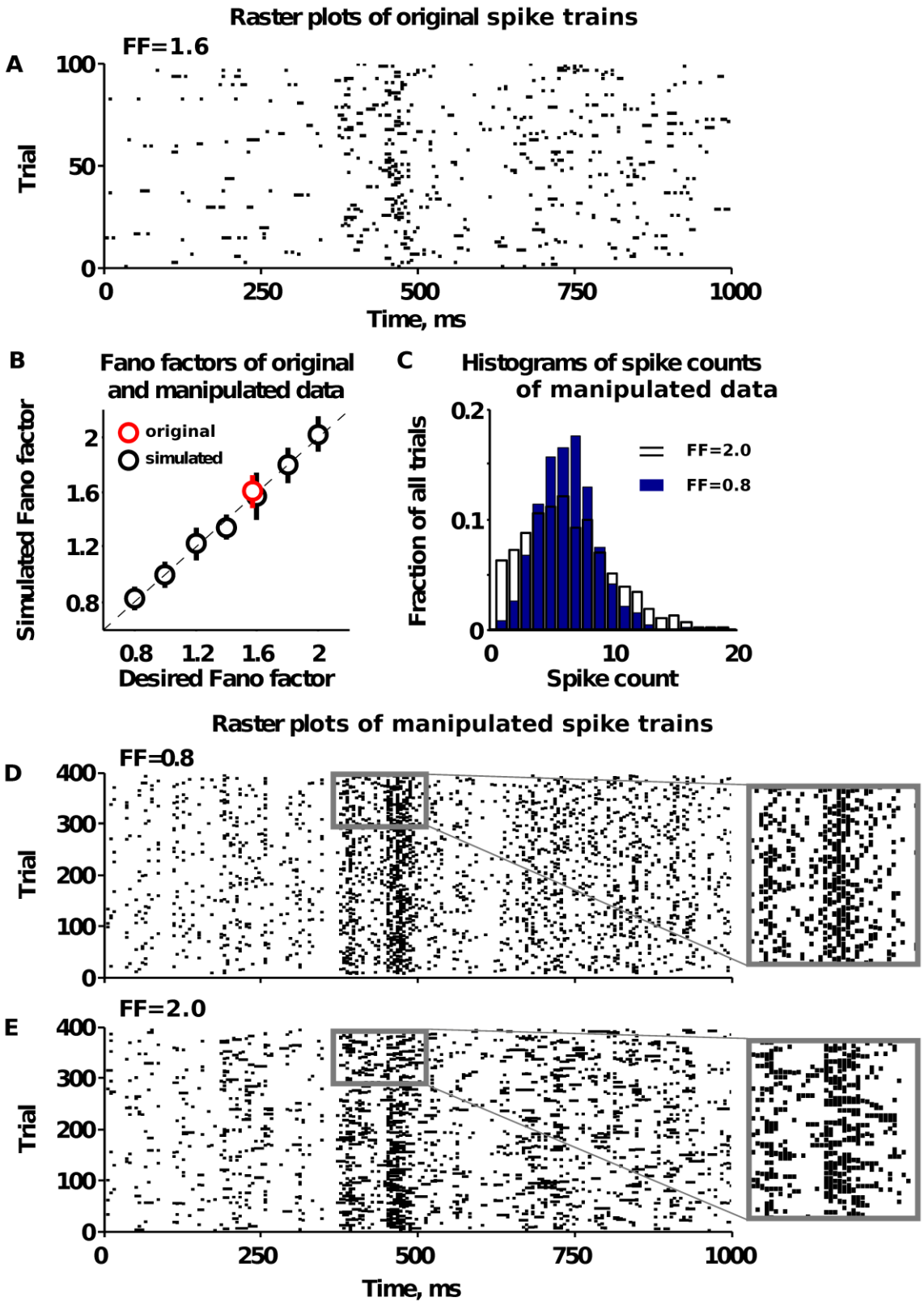


Figure 2. A. Raster plot of spike trains in mouse auditory cortex. The vertical axis specifies trial number, the horizontal axis specifies time since stimulus onset, and each

dot on the plot represents a spike. The data were collected during the presentation of a 100 repeated trials of a 1 second segment of a sinusoidally modulated broadband noise sound. For model fitting, the data were binned in 5 ms time bins. **B.** Values of Fano factor from a simulated dataset versus desired value of Fano factor. We fitted our model to experimental data shown in Fig. 2A, simulated spike trains with the same properties, and calculated Fano factor of the original and simulated data. The corresponding point is plotted in red. We set 7 other arbitrary values (from 0.8 to 2.0), solved for the scaling factor of each, simulated spike trains for every autocorrelation found, and plotted the Fano factor calculated from the simulated spike trains versus the desired Fano factor (in black). **C.** Spread of spike counts for the simulated datasets with *FF* = 0.8 and *FF* = 2.0. Blue bars show the distribution of spike counts for the simulated data with *FF* = 0.8, transparent bars with black edges show the distribution of spike counts for the simulated data with *FF* = 2.0. **D.** Raster plot of the simulated dataset where th e autocorrelation function was scaled to match Fano factor value of 0.8. A magnified part of this raster plot is shown to the right. **E.** Raster plot of the simulated dataset where autocorrelation function was scaled to match Fano factor value of 2.0, presented as in D.

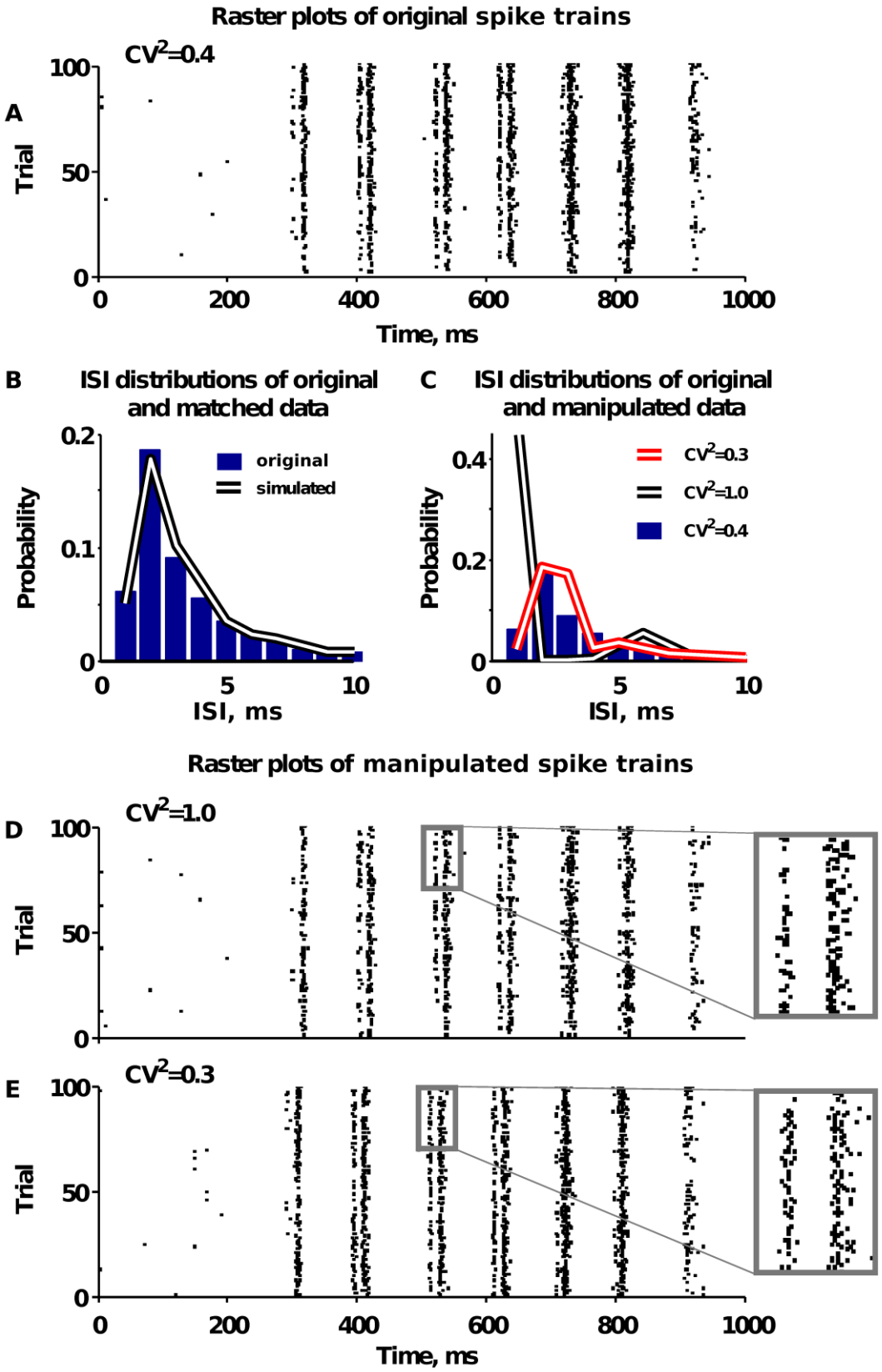


Figure 3. A. Raster plot of spike trains in mouse inferior colliculus. The data were collected during the presentation of a 100 repeated trials of a 1 second segment of a sinusoidally modulated broadband noise sound. For model fitting, the data were binned

in 1 ms time bins. **B.** ISI distribution of original data (blue bars), and ISI distribution of spike trains sampled from the model fitted to the original data (black and white line). **C.** ISI distribution of original data (blue bars), and ISI distribution of manipulated data (black and white, red and white lines). Given the parameters of the fitted model, we solved for the scaling coefficient of autocorrelation function that would give us the desired value of CV^2 of the ISI distribution: 0.3 and 1.0, with CV^2 calculated for the same range of time bins where the autocorrelations are significant (k = 10). **D.** Raster plot of the simulated dataset where autocorrelation function was scaled to match CV^2 value of 1.0. **E.** Raster plot of the simulated dataset where autocorrelation function was scaled to match CV^2 value of 0.3.

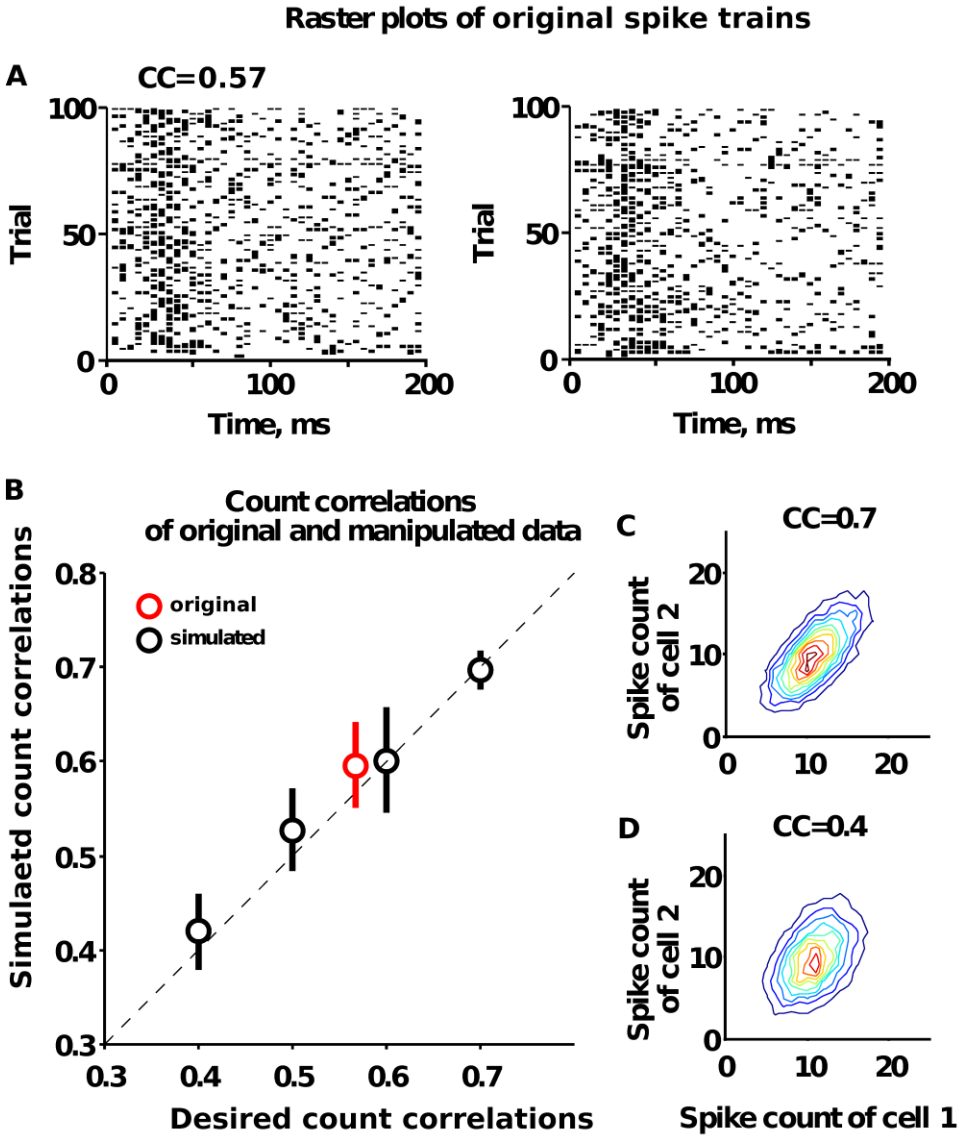


Figure 4. A Raster plot of the spike trains from a pair of cells in gerbil auditory cortex. The data were collected during the presentation of 100 repeated trials of a 200 second segment of a sinusoidally modulated broadband noise sound. The data were binned in 5 ms time bins. **B.** Matching the desired values of count correlations. The horizontal axis shows the desired value of count correlation coefficient and vertical axis shows the count correlation coefficient of the simulated spike trains. We fitted our model to the original data and sampled spike trains from the fitted model. The count correlation of the simulated spike trains is plotted versus the count correlation coefficient of the original spike train in red. We set four other arbitrary values from 0.4 to 0.7 and solved for the scaling coefficient of the autocorrelation function that would yield the desired count correlation in the simulated spike trains. The count correlations calculated from the simulated spike trains are plotted versus the desired values of correlations in black. Dashed line is the line of equality **C,D.** Contour plots of the distributions of spike counts of two cells. For the count correlations of 0.7 (C) and 0.4 (D) we created two-dimensional histograms of the spike counts of both cells with spike count of cell one on

the horizontal axis, and spike count of cell 2 on the vertical axis. C and D show these histograms as contour plots, where each line denotes a level of constant probability of falling inside the contour.

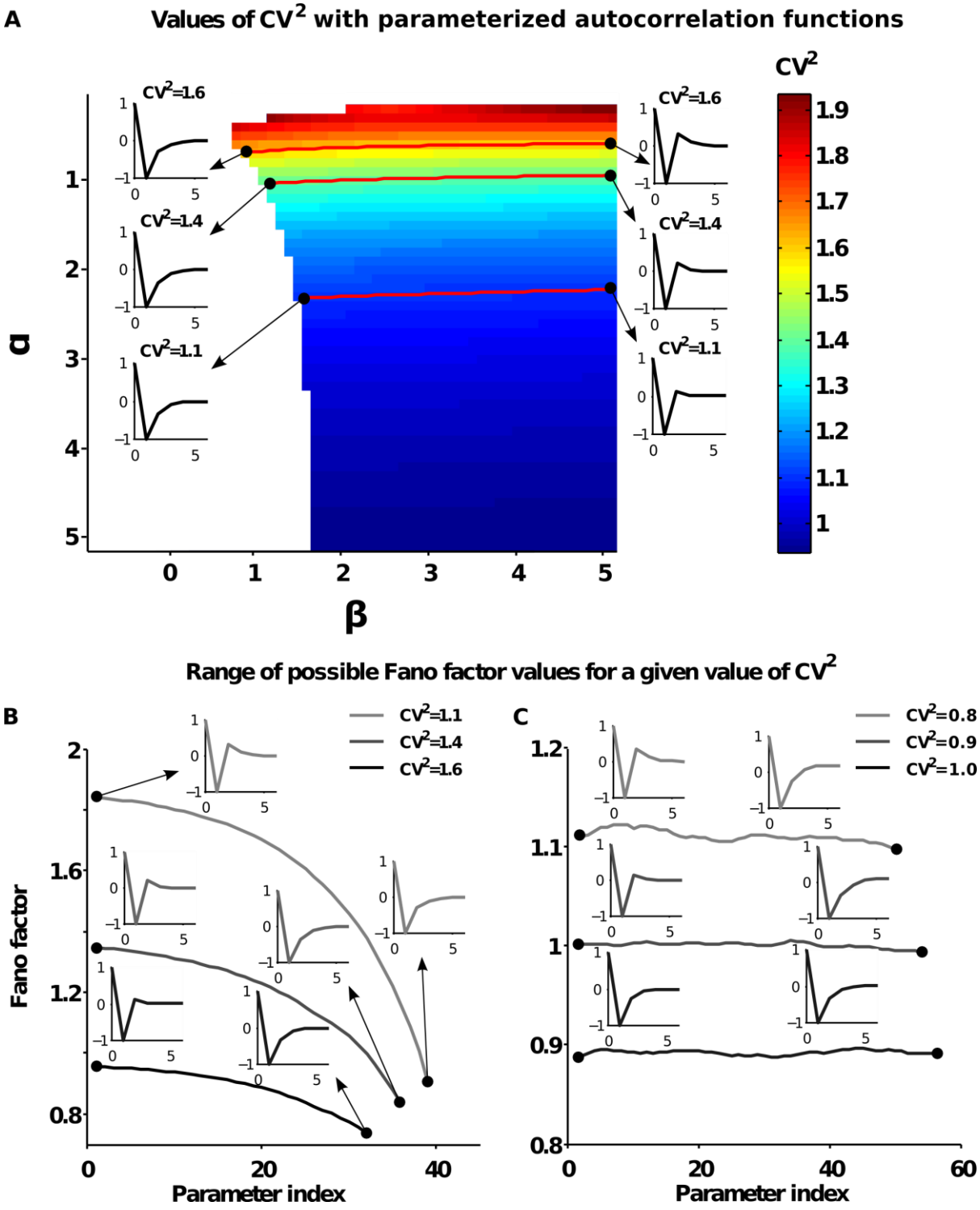


Figure 5. Map of values of CV^2 in space of parameters of autocorrelation function of z. The family of functions that we model autocorrelation functions with is described in section 5. The parameters α and β were varied between 0 and 2, and 0 and 5 correspondingly. Colorbar to the right of the map shows the actual values of CV^2 for the corresponding color. Red lines on the map show lines of equality for CV^2 of 5.5, 5.7, and 5.9. Sample autocorrelation functions are plotted next to the ends of isolines as insets (the corresponding ends are marked with black dots and arrows). The map is created by sampling from a model in which s is a sine wave with frequency of 1.6 Hz, amplitude of

1, and zero mean. The threshold was chosen such that the average firing rate was 60 spikes per second. The simulations were 1 second long, binned on 1000 time bins each and had 1000 trials for every autocorrelation function. **B.** Range of changes in Fano factor along the lines of equality of CV^2. Using the map of CV^2 values in two-parameter space (A), we calculated Fano factor for all the pairs of parameters (in the given range) where CV^2 is constant and equal to 5.5, 5.7, and 5.9. The horizontal axis shows the index of the pair of parameters (that increases as the values of parameters α and β decrease), and the vertical axis shows the value of Fano factor for this combination of parameters. Insets show plots of the autocorrelation functions on the corresponding ends of curves (marked with dots and arrows). **C.** Range of changes in Fano factor along the lines of equality of CV^2. We created a map analogous to the one shown on A (not shown here) but with the *s* component in the generative process being a Gaussian process without temporal correlations and with variance of 0.4. We chose to consider the range of changes in Fano factor for the isolines with CV^2 equal to 1.5, 1.6, and 1.7. The plot shows three corresponding curves with increasing values CV^2 of from bottom to top. As in B, the horizontal axis shows index of the pair of parameters (that increases as the values of parameters α and β decrease), and the vertical axis shows the value of Fano factor for this combination of parameters. Insets show plots of the autocorrelation functions on the corresponding ends of curves (marked with dots and arrows).

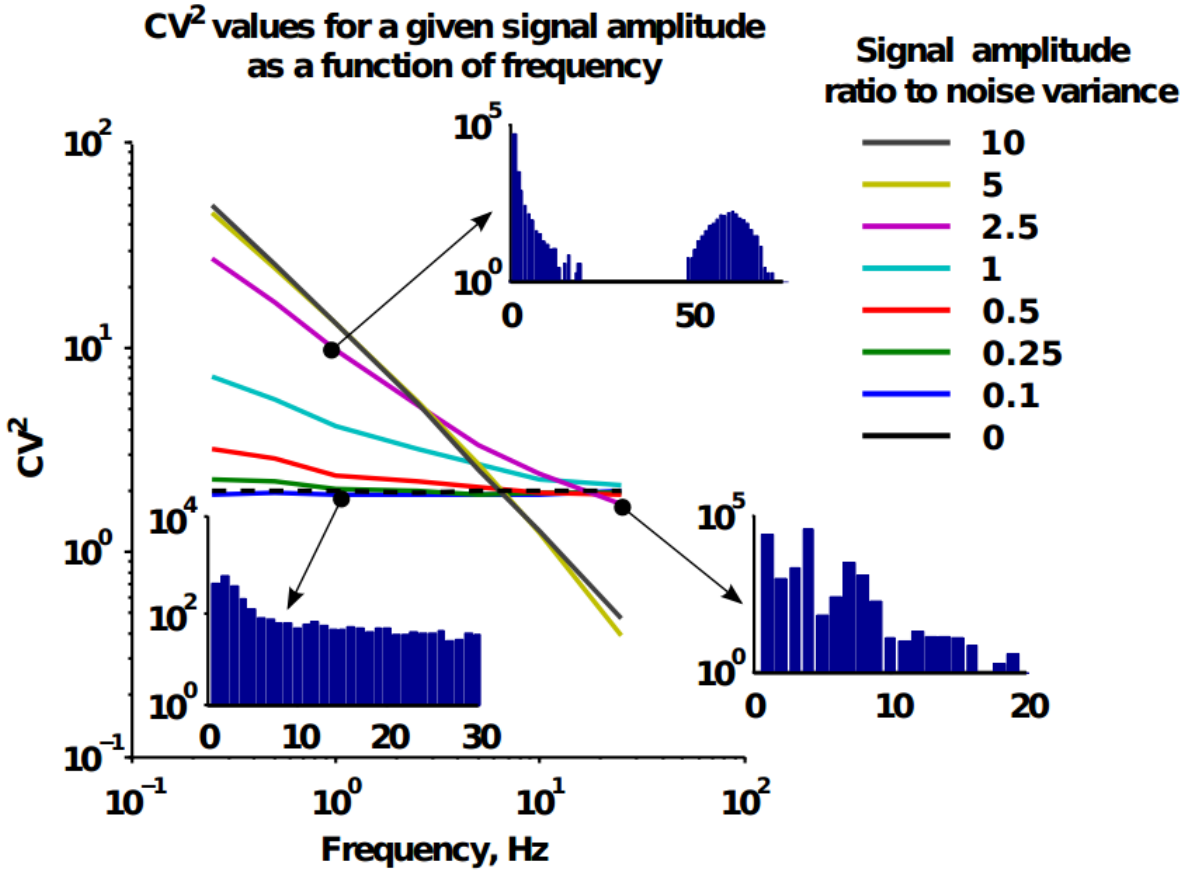


Figure 6. *CV²* of the ISI distribution as a function of amplitude and frequency of the sinusoidal signal. We varied the amplitude of signal from 0.1 to 10 with the amplitude of noise always being 1. The frequency was varied between 0.16 and 40 Hz (note that the frequency axis has logarithmic scale). Simulations were done in 200 trials for each point, were 1 second long and binned on 1000 time bins. All the curves are plotted for the same autocorrelation function chosen from the family described in Section 5 with $\alpha = 0.3$ and $\beta = 1.5$, which corresponds to a positive non-monotonically decreasing function. Dashed line shows the case when signal is completely absent. Histograms in the smaller insets show the ISI distributions for the marked points on the curves. The horizontal axis shows the length of ISIs in ms, and the vertical axis shows the number of occurrences of an ISI of a given length. With the wide range of relative amplitudes of signal one can see the behavior of the model in signal driven (high amplitude) and noise driven (low amplitude) cases. In the signal driven cases ISI histograms are defined by the structure of the signal (see the histograms on top and on the right having the periodic features of the corresponding signals), and in the noise driven cases, changes in signal frequency do not affect the ISI distribution in any significant way. The figure also illustrates that similar *CV²* values can be achieved in both signal and noise driven regimes and have qualitatively different underlying ISI distributions (compare the two lower histograms). Note that curves are plotted in logarithmic axes.

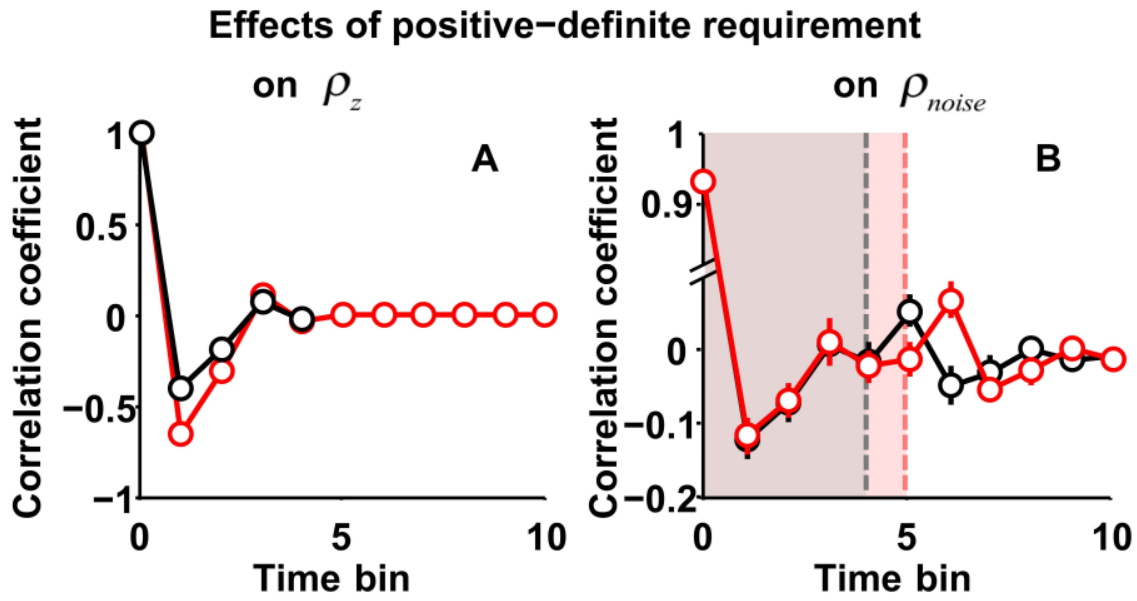


Figure 7. Effects of positive definite requirement on the covariance matrix of *z*. **A.** Example of the autocorrelation function of *z* calculated from the experimental data (red), and the closest realizable autocorrelation function (black). **B.** The attempt to reproduce strong negative correlations also leads to an artifact in the resulting spike trains. We simulated spike trains with correlations in time extending to the time bin 4 (no significant correlations beyond that limit in the original data), the simulated dataset had an artifact of significant positive correlations in time bin 5. When zero-padded (we deliberately added zero correlations to the time bin 5 in the correlation function of *z*), the artifact peak shifted to the next time bin. Further zero-padding resulted in a process that was unrealizable.

Appendix A

By definition, Fano factor is the ratio of spike count variance and mean spike count across the trials.

$$FF = \frac{\mathrm{var}\left(\sum_{n=1}^{N} r[n]\right)}{\left\langle \sum_{n=1}^{N} r[n] \right\rangle} = \frac{\left\langle \left(\sum_{n=1}^{N} r[n]\right)^2 \right\rangle - \left\langle \sum_{n=1}^{N} r[n] \right\rangle^2}{\left\langle \sum_{n=1}^{N} r[n] \right\rangle}$$

(because Fano factor is a single cell quantity, cell indices are omitted for simplicity). We expand the square of the spike count to deal with the expected values of the elements of this expansion.

$$\left(\sum_{n=1}^{N} r[n]\right)^2 = r[1]r[1] + r[1]r[2] + \ldots + r[n]r[m] + \ldots + r[N]r[N]$$

Since *r* is a binary vector, $r[n]r[n] = r[n]$ and thus $\langle r[n]r[n] \rangle = \Phi(s[n],1)$.

We define correlation matrix $\Sigma_z[k] = \begin{bmatrix} 1 & \rho_z[k] \\ \rho_z[k] & 1 \end{bmatrix}$ then rewrite the expression for the square of the spike count as follows:

$$\left\langle \left(\sum_{n=1}^{N} r[n]\right)^2 \right\rangle = \left\langle \sum_{n=1}^{N} r[n]r[n] + 2\sum_{m=n+1}^{N}\sum_{n=1}^{N} r[n]r[m] \right\rangle =$$

$$= \sum_{n=1}^{N} \Phi\left(s[n],1\right) + 2\sum_{m=n+1}^{N}\sum_{n=1}^{N} \Phi_2\left(\begin{matrix} s[n] \\ s[m] \end{matrix}, \Sigma_z[m-n]\right)$$

and the full expression for Fano factor is

$$FF = \frac{\sum_{n=1}^{N} \Phi\left(s[n],1\right) + 2\sum_{m=n+1}^{N}\sum_{n=1}^{N} \Phi_2\left(\begin{matrix} s[n] \\ s[m] \end{matrix}, \Sigma_z[m-n]\right) - \left(\sum_{n=1}^{N} \Phi\left(s[n],1\right)\right)^2}{\sum_{n=1}^{N} \Phi\left(s[n],1\right)}.$$

Appendix B

Fano factor by definition is

$$FF = \frac{\mathrm{var}\left(\sum_{n=1}^{N} r[n]\right)}{\left\langle \sum_{n=1}^{N} r[n] \right\rangle} = \frac{\left\langle \left(\sum_{n=1}^{N} r[n]\right)^2 \right\rangle - \left\langle \sum_{n=1}^{N} r[n] \right\rangle^2}{\left\langle \sum_{n=1}^{N} r[n] \right\rangle}$$

(because Fano factor is a single cell quantity, cell indices are omitted for simplicity). Expanding the first term in the numerator we get

$$\left(\sum_{n=1}^{N} r[n]\right)^2 = r[1]r[1] + \ldots + r[m]r[n] + \ldots + r[N]r[N]$$

Where each $r[n]r[n]$ term equals $r[n]$ because r is a binary vector.

$$\langle r[n]r[n]\rangle = \langle r[n]\rangle = \langle (s[n] + z[n]) > \theta\rangle = \Phi(-\theta, \sigma_s^2 + 1)$$

For all the other products $r[n]r[m]$ where $n \neq m$

$$\langle r[n]r[m]\rangle = \Phi_2\begin{pmatrix} -\theta \\ -\theta \end{pmatrix}, \Sigma_{s+z}[m-n]\Bigg)$$

Where Φ_2 is a two-dimensional Gaussian cumulative distribution function, and we define Σ_{s+z} as

$$\Sigma_{s+z}[m-n] = \begin{bmatrix} \sigma_s^2 + 1 & \rho_z[m-n] + \sigma_s^2\rho_s[m-n] \\ \rho_z[m-n] + \sigma_s^2\rho_s[m-n] & \sigma_s^2 + 1 \end{bmatrix}$$

Hence the final formula for Fano factor is (assuming symmetry of correlations)

$$FF = \frac{\sum_{n=1}^{N}\Phi(-\theta, \sigma_s^2 + 1) + \sum_{n\neq m}^{N}\Phi_2\begin{pmatrix} -\theta \\ -\theta \end{pmatrix}, \Sigma_{s+z}[m-n]\Bigg) - \left(\sum_{n=1}^{N}\Phi(-\theta, \sigma_s^2 + 1)\right)^2}{\left(\sum_{n=1}^{N}\Phi(-\theta, \sigma_s^2 + 1)\right)}$$

V Nonlinear transfer of signal and noise correlations in cortical networks

In this chapter we use the DG framework to study the mechanisms by which the signal and noise correlations observed *in vivo* are created.

We find that thresholding as a model for spike generation imposes nonlinear transfer functions between input current signal and noise correlations and spike train correlations. In vitro recordings from neurons in mouse V1 show that these nonlinearities indeed exist in pyramidal cells.

Using DG we find to what extent correlation transfer nonlinearities are responsible for the correlation structure observed in vivo. We analyze spike trains obtained *in vivo* in response to diotically presented sounds, and sounds that are presented ipsilaterally to the recording site and contralaterally.

Using the results of our analysis we show how different relationships between signal and noise correlations can arise due to the effects of correlation transfer nonlinearities or dependencies existing in input currents.

Note that in this chapter instead of fixing the variance of noise at one, we let it take on arbitrary values and fix threshold at one instead.

Authors' contributions

The work was carried out under the supervision of Nicholas Lesica. The research was conceived and designed by NL, SB, TK, and DL; SB and DL collected *in vitro* data; NL and DL collected *in vivo* data; NL and DL analyzed the data and wrote the paper.

This manuscript entitled "Nonlinear transfer of signal and noise correlations in cortical networks" is currently being submitted.

Nonlinear transfer of signal and noise correlations in cortical networks

Dmitry R Lyamzin, Samuel J Barnes, Tara Keck, and Nicholas A Lesica

Summary

Correlated spiking is a prominent feature of cortical activity. During sensory processing, information transmission is controlled by the balance of signal and noise correlations, and their structure across a network is an important signature of its organization and function. Studies of the relationship between signal and noise correlations typically report a positive dependency, with variations due to changes in stimulus properties and behavioral context, but the factors that establish and modify correlations are not well understood. Here we show that the correlation structure in cortical networks is controlled by correlation transfer nonlinearities with complex dependencies on the properties of signal and noise inputs. We find that a change in the strength or correlation of either signal or noise inputs can affect both spike train signal and noise correlations. We first demonstrate these effects directly using whole-cell recordings of responses to injected currents and show that they are reproduced within an analytical framework based on a simple threshold nonlinearity. We then analyze the correlation structure of spike trains recorded in vivo and show that correlation transfer nonlinearities impose a variety of dependencies between signal and noise correlations across populations of neurons in auditory cortex, and that these effects underlie changes in correlation structure that result from changes in sound location. Our results demonstrate that correlation transfer nonlinearities can shape the correlation structure of cortical networks, even imposing a negative dependency between signal and noise correlations under certain conditions. It is likely that the correlation transfer nonlinearities that we observe operate across sensory systems and, thus, our results are a critical step toward a general understanding of how correlated spiking relates to the function of cortical networks.

Introduction

What are the factors that govern the structure of signal and noise correlations in cortical networks during sensory processing? Signal and noise correlations play an important role in cortical function by setting the balance between synergy and redundancy in the spike trains of neuronal populations (Averbeck & Lee, 2006; Oram, Földiák, Perrett, & Sengpiel, 1998; Panzeri, Schultz, Treves, & Rolls, 1999) and provide a window into network organization and dynamics over multiple timescales (Cohen & Kohn, 2011). Studies of the structure of signal and noise correlations across cortical areas typically report a positive dependency, i.e. noise correlation tends to increase with increasing signal correlation (Averbeck & Lee, 2003; Bair, Zohary, & Newsome, 2001; Kohn & Smith, 2005), but recent studies have shown that the this structure is dynamic and can be modulated by changes in stimulus properties (Kohn & Smith, 2005), task-dependent modulation (Cohen & Maunsell, 2009; Gutnisky & Dragoi, 2008; Mitchell, Sundberg, & Reynolds, 2009) and learning (Gu et al., 2011; Jeanne, Sharpee, & Gentner, 2013). It is generally assumed that the positive dependency between signal and noise correlations is a consequence of signal and noise being transmitted through the same shared inputs and recurrent connections, and that modulation of spike train signal and noise correlations reflects modulation of the correlations in these inputs. However, because of the thresholding associated with spike generation, the transfer of input correlations to spike train correlations is highly nonlinear (Binder & Powers, 2001; de la Rocha, Doiron, Shea-Brown, Josić, & Reyes, 2007; Dorn & Ringach, 2003; Shea-Brown, Josić, de la Rocha, & Doiron, 2008; Tchumatchenko, Malyshev, Geisel, Volgushev, & Wolf, 2010). For example, even when input correlations are constant, spike train correlations can vary with changes in

input strength, and it has been shown that this correlation transfer nonlinearity can impose a positive dependency between the mean spike rate and total correlation in cortical spike trains (de la Rocha et al., 2007). Understanding the effects of such correlation transfer nonlinearities is critical, as the relationship of spike train signal and noise correlations to network structure and function depends on whether they reflect the corresponding input correlations or arise as a consequence of a nonlinearity that is dependent on another input property. The effects of correlation transfer nonlinearities in the context of signal and noise correlations have not yet been studied, but the range of factors that modulate signal and noise correlations and the observation that spike train signal and noise correlations can covary with other spike train properties (for example, trial-to-trial variability (Cohen & Maunsell, 2009; Mitchell et al., 2009)) suggest that they may play an important role.

Results

To characterize how correlation transfer nonlinearities shape signal and noise correlations during sensory processing, we analyzed the structure of spike train signal and noise correlations in cortical networks. Figure 1a shows the spike trains for a typical population of cells in primary auditory cortex (A1) of an anesthetized gerbil recorded with a multi-tetrode array in response to repeated trials of a 2.5 second segment of speech presented from a speaker directed at the contralateral ear (figure 1A). We measured the correlation between the responses of pairs of neurons for spike counts in 50 ms bins and defined the signal correlation ρ_r^{sig} as the correlation that remained after shuffling the trial order, and the noise correlation as the difference between the total correlation and the signal correlation, $\rho_r^{noise} = \rho_r - \rho_r^{sig}$. Overall, there was a positive dependency between signal and noise correlations across the population ($r = 0.41$, $p < 0.001$; figure 1B), as is typically observed for studies of cortical spike trains (Averbeck & Lee, 2003; Bair et al., 2001; Kohn & Smith, 2005). One possible explanation for the positive dependency between spike train signal and noise correlations is that it reflects a corresponding dependency between input signal and noise correlations. However, there were also clear covariations of spike train signal and noise correlations with mean spike rate (figure 1C) and trial-to-trial variability (figure 1D; defined as signal-to-noise ratio (SNR), the ratio of the variance of the PSTH to the average variance of the deviation of each trial from the PSTH). These covariations suggest that the positive dependency between signal and noise correlations could also arise from the effects of correlation transfer nonlinearities dependent on the input strength or SNR, but the task of isolating the effects of correlation transfer nonlinearities on spike train correlations is complicated by the fact that multiple input properties can vary simultaneously during sensory processing. Our approach to characterizing the contribution of correlation transfer nonlinearities to the structure of signal and noise correlations was as follows: 1) we first made whole-cell recordings in vitro to investigate directly how systematic changes in individual input properties affect spike train correlations and to validate a method for inferring input properties from spike trains, 2) we next examined different changes in spike train signal and noise correlations during sensory processing in vivo and determined the extent to which each reflected changes in input correlations and/or the effects of correlation transfer nonlinearities.

To examine how changes in different input properties affect the transfer of signal and noise correlations from inputs to spike trains, we injected Gaussian white noise currents with specified variances and correlations into layer 2/3 pyramidal cells in mouse visual cortex and compared the resulting membrane potential correlations and spike train correlations between pairs (figure 2A). Figures 2B and 2C show two fundamental properties of correlation transfer that underlie the various dependencies between membrane potential and spike train signal and noise correlations that we will describe

below. The first of these fundamental properties, which we will call 'variance gain', is that the spike train correlation ρ_r increases with the membrane potential total variance σ_v^2, even when the membrane potential total correlation ρ_v is held constant. As shown in figure 2B, the spike train correlation ρ_r increases sublinearly with the membrane potential variance σ_v^2, i.e. each incremental increase in membrane potential variance results in a successively smaller increase in spike train correlation. The second fundamental property of correlation transfer, which we will call 'correlation gain', is that spike train total correlation ρ_r increases supralinearly with membrane potential total correlation ρ_v when all other membrane potential properties are held constant. Figure 2C shows the membrane potential and spike train correlations that result from injecting currents with varying correlation and constant variance. The spike train correlation ρ_r increases supralinearly with the membrane potential correlation ρ_v, i.e. each incremental increase in membrane potential correlation results in a successively larger increase in spike train correlation (for further explanation of the causes of these variance gain and correlation gain effects, see Supplementary Figure 1).

These correlation transfer nonlinearities can have complex consequences in the context of sensory processing, where correlations have both signal and noise components. In this context, variance gain and correlation gain impose a variety of interdependencies on the strength, SNR, and signal and noise correlations of membrane potential and spike trains. To examine these interdependencies, we injected currents with a signal component (same on each trial) and a noise component (different on each trial) that had specified variances and correlations. We calculated the membrane potential signal variance σ_s^2 and correlation ρ_s from the mean membrane potential across trials and the membrane potential noise variance σ_n^2 and correlation ρ_n from the deviations from the mean on each trial. First, to examine the effects of variance gain on signal and noise correlations, we injected currents with fixed signal and noise correlations and varied the signal and/or noise variance. Figure 2D shows the relationship between the membrane potential signal variance σ_s^2 and the spike train signal and noise correlations ρ_r^{sig} and ρ_r^{noise} that result from injecting currents with different signal variances and constant noise variance, signal correlation, and noise correlation. As the membrane potential signal variance σ_s^2 increases, the spike train signal correlation ρ_r^{sig} also increases, as might be expected from the variance gain effects described in figure 2B. More surprisingly, the spike train noise correlation ρ_r^{noise} decreases with increasing membrane potential signal variance σ_s^2, even though the membrane potential noise variance and noise correlation remain constant. This relationship between membrane potential signal variance and spike train noise correlation is also due to variance gain, though the cause is more subtle: because of the sublinearity of variance gain, an increase in the membrane potential signal variance σ_s^2 increases the membrane potential total variance $\sigma_v^2 = \sigma_s^2 + \sigma_n^2$ by the same amount, but, because the total variance is larger than the signal variance and the effects of variance gain are sublinear, the relative increase in the transfer of correlation due to the increase in variance is larger for the signal correlation ρ_r^{sig} than for the total correlation ρ_r, and, thus, the noise correlation $\rho_r^{noise} = \rho_r - \rho_r^{sig}$ decreases. The effects of membrane potential noise variance on the spike train correlations mirror those of membrane potential signal variance: an increase in the membrane potential noise variance σ_n^2 causes an increase in the spike train noise correlation ρ_r^{noise} and a decrease in the spike train signal correlation ρ_r^{sig} (see Supplementary Figure S2). These effects of changes in membrane potential signal and noise variance combine to produce the relationship between membrane potential SNR and spike train signal and noise correlations shown in figure 2E. As the membrane potential SNR σ_s^2/σ_n^2 is increased (with the total input variance $\sigma_v^2 = \sigma_s^2 + \sigma_n^2$ held constant), the spike train signal correlation ρ_r^{sig} is increased and the spike train noise correlation ρ_r^{noise} is decreased due to opposing variance gain effects.

Next, to examine the effects of correlation gain on signal and noise correlations, we injected currents with fixed signal and noise variances and varied either the signal or noise correlation. As shown in figure 2F, when the membrane potential variances are constant, an increase in the membrane potential signal correlation ρ_s causes an increase in both the spike train signal and noise correlations ρ_r^{sig} and ρ_r^{noise}, even when the membrane potential noise correlation ρ_n is held constant. This dependency of spike train noise correlation ρ_r^{noise} on the membrane potential signal correlation ρ_s arises from the supralinearity of correlation gain: an increase in the membrane potential signal correlation ρ_s increases the membrane potential total correlation $\rho_v = \rho_s + \rho_n$ by the same amount, but, because the total correlation is larger than the signal correlation and the effects of correlation gain are supralinear, the relative change in the transfer of correlation is larger for the total correlation ρ_r than for the signal correlation ρ_r^{sig}, and, thus, the noise correlation $\rho_r^{noise} = \rho_r - \rho_r^{sig}$ increases. This crossover of correlation gain between signal and noise correlations does not occur when only the input noise correlation ρ_n is changed; increasing the input noise correlation ρ_n results in an increase in the spike train noise correlation ρ_r^{noise}, but has no impact on the spike train signal correlation ρ_r^{sig} (because noise correlations are removed by shuffling the trial order to compute signal correlations; see Supplementary Figure 2). To summarize the effects of variance gain and correlation gain on signal and noise correlations observed in vitro, we replotted the results in figures 2E and 2F to visualize explicitly how changes in membrane potential SNR and signal correlation affect the relationship between spike train signal and noise correlations. As shown in figure 2G, variance gain imposes a negative dependency between spike train signal and noise correlations with changes in input SNR, and, as shown in figure 2H, correlation gain imposes a positive dependency between spike train signal and noise correlations with changes in membrane potential signal correlation.

During sensory processing in vivo, the different correlation transfer nonlinearities described above will act together with the potential to significantly transform the correlation structure in the network. However, from extracellular recordings, it is impossible to determine the extent to which the correlations between the spike trains of a neuronal population are dependent on correlation transfer nonlinearities without a method for isolating their effects. We have previously developed a method to determine the input variances and correlations required to generate spike trains with specified signal and noise correlations using a dichotomized Gaussian (DG) framework (Lyamzin, Garcia-Lazaro, & Lesica, 2012; Lyamzin, Macke, & Lesica, 2010; Macke, Berens, Ecker, Tolias, & Bethge, 2009). We assume that spike trains are generated by thresholding a Gaussian process with signal and noise components that have specified variances and correlations (figure 3A):

$$r(t) = \begin{cases} 1 & \text{if } s(t) + n(t) > 1 \\ 0 & \text{else} \end{cases}$$

with signal input $s \sim \mathcal{N}(0, \sigma_s^2)$ assumed to be the same on each trial, noise input $n \sim \mathcal{N}(0, \sigma_n^2)$ assumed to be different on each trial, and correlations ρ_s and ρ_n between the signal and noise inputs for each pair of cells. For every cell, there is a unique set of input parameters $\{\sigma_s^2, \sigma_n^2\}$ that correspond to the observed mean spike rate and spike train SNR, and for every pair of cells, there is a unique set of input parameters $\{\rho_s, \rho_n\}$ that correspond to the observed set of spike train signal and noise correlations. The DG input parameters can be estimated from spike trains and, while the results do not allow for inference of actual membrane potential values, they do provide an accurate reflection of its SNR, signal, and noise correlations. To test how well membrane potential statistics could be estimated from spike trains using the DG framework, we applied it to the whole-cell recordings described above. As shown in figure 3B, there was an excellent correspondence between the DG input parameters and the statistics of the actual membrane potentials, with correlation

coefficients of 0.97 for SNR, 0.91 for signal correlations, and 0.93 for noise correlations. There are two details regarding the inference of membrane potential properties using the DG framework that are worth noting: 1) to obtain accurate estimates of the absolute values of the signal and noise variances σ_s^2 and σ_n^2, knowledge of the actual value of the spike threshold relative to baseline is required; however, as the analyses below require only knowledge of the input SNR σ_s^2/σ_n^2, choosing an arbitrary value for the threshold is sufficient, and 2) the estimated noise inputs consistently have higher variance and lower correlation than the actual membrane potentials (due to variability in spike threshold that is not included in the DG framework, see Supplementary Figure 3), but these differences do not impact the analyses below for which the linear relationships between the estimated inputs and the actual membrane potential statistics implied by the high correlation coefficients are sufficient. The close correspondence between the DG input parameters and the actual membrane potential statistics across a wide range of input variances and correlations implies that a simple threshold nonlinearity is sufficient to capture the effects of correlation transfer nonlinearities in actual neurons. Indeed, as shown in figures 3C, 3D, and 3E, spike trains generated using the DG framework exhibited the same variance gain and correlation gain effects on spike train signal and noise correlations that we observed experimentally.

Having characterized the different correlation transfer nonlinearities that can operate during sensory processing and verified our ability to estimate membrane potential statistics from population spike trains using the DG framework, we returned to the analysis of spike train signal and noise correlations observed in vivo. We first continued our analysis of the A1 responses to speech shown in figure 1 for which there was a positive dependency between signal and noise correlations. However, as this sample includes pairs with a wide range of input variances, SNRs, and signal and noise correlations, the extent to which this dependency in the spike trains reflects a corresponding dependency between input signal and noise correlations and/or the effects of correlation transfer nonlinearities that are dependent on each of the different input properties is unclear. Estimation of membrane potential properties for these pairs using the DG framework revealed a positive dependency between input signal and noise correlations ($r = 0.24$, $p < 0.001$; figure 4A), but visualizing the gradients corresponding to changes in total input variance (figure 4B), input SNR (figure 4C), and input signal correlation (figure 4D), suggests that correlation transfer nonlinearities dependent on each of these properties may also play a role. As a first step toward isolating the effects of each input property, we restricted our analysis to pairs with similar total input variance (and, thus, mean spike rates), since the effects of changes in total input variance on spike train correlations have been studied in detail previously (de la Rocha et al., 2007). For a restricted subset of pairs with similar total input variance, the dependency between spike train signal and noise correlations was weak ($r = 0.01$, $p = 0.86$; figure 4E), as was the dependency between input signal and noise correlations ($r = 0.01$, $p = 0.84$; figure 4F), suggesting that a substantial fraction of the positive dependency observed for all pairs was due to variance gain effects dependent on total input variance, consistent with the results of previous studies (this result was robust across a wide range of total input variances, see Supplementary Figure 4). However, this weak dependency for pairs with similar total input variance does not necessarily imply that other input properties have no effect on spike train signal and noise correlations; the opposing gradients corresponding to changes in input SNR and input signal correlation shown above in figures 2G and 2H suggest that their effects may offset each other. Indeed, further restricting our analysis to pairs with similar input signal correlations revealed a negative dependency ($r = -0.4$, $p < 0.001$; figure 4G), while restricting to pairs with similar input SNRs revealed a positive dependency between signal and noise correlations ($r = 0.53$, $p < 0.001$; figure 4H). These dependencies were robust across a wide range of input signal correlations and SNRs (see Supplementary Figure 4).

For each of these restricted samples, we used the DG framework to determine the extent to which the observed dependency in the spike trains reflected a corresponding input dependency and/or the effects of a correlation transfer nonlinearity. First, we analysed the negative dependency between spike train signal and noise correlations for pairs with similar total input variance and input signal correlation shown in figure 4G. Because the pairs in this restricted subset all have similar total input variance and input signal correlation, the spread of spike train signal correlations and, thus, the dependency between spike train signal and noise correlations, must be due at least in part to variations in input SNR. As shown above in figure 2G, variance gain can impose a negative dependency between spike train signal and noise correlations with changes in input SNR. If the observed dependency between signal and noise correlations in the spike trains were due solely to these variance gain effects, then destroying any dependencies between input signal and noise correlations would not affect the dependency in the spike trains. However, after randomizing the pairing of signal and noise inputs across pairs and using the DG framework to generate new spike trains, the dependency between the spike train signal and noise correlations disappeared ($r = 0.02$, $p = 0.83$; figure 4I). This result suggests that negative dependency between spike train signal and noise correlations reflect a dependency in the input; however, as this subset is restricted to pairs with similar input signal correlation, the dependency cannot be between input signal and noise correlations, but rather must be between input SNR and noise correlation. Indeed, as shown in figure 4J, there was a negative dependency between input SNR and noise correlations for this sample of pairs ($r = -0.34$, $p < 0.001$). This analysis suggests that the negative dependency between the spike train signal and noise correlations for pairs with similar total input variance and input signal correlation arises from the combination of a dependency between SNR and noise correlations in the input and a dependency between SNR and signal correlations imposed by variance gain (the lack of variance gain effects on noise correlations is due to the sublinearity of variance gain, as discussed in detail below).

Next, we analysed the positive dependency between spike train signal and noise correlations for pairs with similar input total variance and SNR shown in figure 4H. This dependency must arise from either a corresponding dependency between input signal and noise correlations or the crossover of the correlation gain effects of changes in input signal correlation on spike train noise correlation shown above in figure 2F. As shown in figure 4K, there was a positive dependency between input signal and noise correlations for this sample of pairs ($r = 0.46$, $p < 0.001$). However, the existence (or, indeed, the strength) of this input dependency does not necessarily imply that it is the source of the observed dependency in the spike trains. To quantify the relative contributions of this input dependency and the crossover of correlation gain to the observed dependency in the spike trains, we must again determine how the spike train dependency changes when the input dependency is destroyed. Figure 4L shows the relationship between the spike train signal and noise correlations after removing the network structure, i.e. randomizing the pairing of signal and noise inputs across pairs, and using the DG framework to generate new spike trains for each pair of cells. Without any dependency between input signal and noise correlations, the dependency between the spike train signal and noise correlations persists, but its strength is substantially reduced ($r = 0.29$, $p < 0.001$). This analysis suggests that the positive dependency between the spike train signal and noise correlations for pairs with similar total input variance and SNR is due to a dependency between input signal and noise correlations that is strengthened by correlation gain.

Our analysis thus far suggests that correlation transfer nonlinearities play an important role in shaping the structure of spike train signal and noise correlations across a population for a given stimulus. It has also been shown that the spike train correlations for a given pair of cells can vary with changes in stimulus properties (Kohn & Smith, 2005). However, it is again

difficult to determine the extent to which any observed changes in spike train correlations reflect a corresponding change in input correlations due to the change in the stimulus and/or a change in the effects of correlation transfer nonlinearities. We compared A1 responses to speech presented from either side in virtual acoustic space (figure 5A). On average, as shown in figure 5B, changing the location of the speech from the side ipsilateral to the recording site to the side contralateral to the recording site did not change the mean spike rate (median change = -0.4%, $p = 0.82$), but caused a significant increase in the SNR of the spike trains (median change = 36.1%, $p < 0.001$). This increase in spike train SNR was accompanied by an increase in the spike train signal correlations and a decrease in the spike train noise correlations (median changes 48.8% and -13.1%, $p < 0.001$; figures 5D and 5E, black histograms). We first examined the extent to which the decrease in spike train signal correlation reflected a corresponding decrease in input signal correlation and/or the effects of correlation transfer nonlinearities dependent on input SNR and input signal correlation. As shown in figure 5C, there was a small decrease in input signal correlation when the sound was moved from the contralateral to the ipsilateral side (median change = 5.7%, $p < 0.001$). We used the DG framework to generate spike trains for each pair while changing the values for each input property from those estimated from responses to speech on the ipsilateral side to those estimated from responses to speech on the contralateral side. As shown in figure 5D, changing only input signal correlation had no impact on spike train signal correlations, while changing only input SNR resulted in changes in spike train signal correlations that were similar those observed in the experimental responses. This result suggests that the observed changes in spike train signal correlations with changing sound location are due primarily to variance gain effects dependent on input SNR.

We next performed a similar analysis for noise correlations. As shown in figure 5C, there was a small decrease in input noise correlation when the sound was moved from the contralateral to the ipsilateral side (median change = -9.2%, $p < 0.001$). We again used the DG framework to generate spike trains for each pair while changing the values for each input property and found that changing only input SNR or input signal correlation had no impact on the spike train noise correlations, while changing only input noise correlation resulted in changes in spike train noise correlations that were similar to those observed in the experimental responses. These results suggest that the observed changes in spike train noise correlations with changing sound location are due primarily to changes in input noise correlation.

Discussion

Taken together, our results demonstrate that the correlation structure in cortical networks can be shaped by correlation transfer nonlinearities in a number of complex ways. We have shown that correlation transfer nonlinearities dependent on the strength or correlation of either signal or noise inputs can potentially affect both spike train signal and noise correlations, imposing dependencies in correlations both across a population under fixed conditions and for a given pair of neurons under changing conditions. Our results also suggest that the degree to which changes in any particular input property will impact spike train correlations is predictable from the sublinearity of variance gain and the supralinearity of correlation gain, with large variance gain effects for inputs with small variances and large correlation gain effects for inputs with large correlations. For example, our analysis of population spike trains in A1 showed that changes in input SNR - either across a population in response to a sound at a fixed location, or for a given pair of neurons in response to sounds at different locations – impacted spike train signal correlations, but not spike train noise correlations. This is consistent with the sublinearity of variance gain; our data were recorded under anesthesia with the cortex in a synchronized state in which input signal variance is small and input noise variance is large (see estimated input SNRs in figure 4J), so a change in input

variance affects signal correlations, but not noise correlations. It is likely that in a desynchronized state in which input SNR is increased (for example, with attention (Cohen & Maunsell, 2009; Mitchell et al., 2009)) the effects of variance gain would shift from signal correlations to noise correlations (thus rendering the transfer of signal correlations robust to changes in variance when direct analysis of sensory inputs is a priority). The diversity of effects that we observed suggests that the precise role of correlation transfer nonlinearities in shaping the correlation structure of cortical networks may vary strongly across brain states and stimulus conditions.

Many of the covariations in spike train properties that we observed have also been observed in other studies. Most studies of the structure of signal and noise correlations across cortical areas report a positive dependency, i.e. noise correlation tends to increase with increasing signal correlation (Averbeck & Lee, 2003; Bair et al., 2001; Kohn & Smith, 2005) and, while it is generally assumed that this dependency reflects a corresponding dependency in input signal and noise correlations, our results demonstrate that this is not necessarily the case. We observed a positive dependency between spike train signal and noise correlations across our entire sample of pairs in auditory cortex, but this dependency disappeared when we considered only subsets of pairs with similar input variance (or mean spike rate), suggesting that the dependency was at least in part imposed by variance gain effects (de la Rocha et al., 2007). When we further restricted our analysis to pairs with similar total input variance and SNR, the positive dependency between spike train signal and noise correlations returned, but our analysis suggests that this was not due to a corresponding dependency in input correlations, but rather arose through the interplay between variance gain effects on spike train signal correlation that were dependent on input SNR and a dependency between input SNR and noise correlation. While some studies of dependencies between spike train signal and noise correlations have assessed the role of correlation transfer nonlinearities dependent on total input variance (by analyzing pairs with similar spike rates (Cohen & Maunsell, 2009; Gu et al., 2011; Jeanne et al., 2013)), few have considered the other effects that we have described here. Several studies have reported phenomena that are consistent with the effects that we observed: for example, increases in spike train SNR and decreases in spike train noise correlation have been observed with changes in attention in V4 (Cohen & Maunsell, 2009; Mitchell et al., 2009) and changes in contrast in V1 (Kohn & Smith, 2005). Without detailed analysis, it is impossible to determine the source of these changes in spike train noise correlations (though recent work suggests that noise correlations in inputs to V1 from LGN do not vary with contrast (Sadagopan & Ferster, 2012)), but our results (including that fact that our in vitro data are from V1) indicate that correlation transfer nonlinearities may play a significant role and that their ability to shape correlation structure may be a general principle of network function.

Methods

In vitro recordings

Deeply anaesthetized C57BL6 mice (P30 – 40) were transcardially perfused with 10 ml of ice cold (4 deg. C) dissection artificial cerebral spinal fluid (dASCF; in mM, 108 choline-Cl, 3 KCL, 26 NaHCO3 1.25 NaHPO4, 25 D-glucose, 3 Na pyruvate, 2 CaCl, and 1 MgCl2 bubbled with 95%O2/5%CO2). Coronal brain slices 300 µm thick were cut (Vibratome 3000 Leica, Wetzlar Germany) from V1. Slices were incubated for 30 minutes in a holding chamber and then recordings were made at room temperature (24 ºC) in recording ASCF (in mM, 120 NaCl, 3 KCL, 23 NaHCO3, 1.25 NaHPO4 , 10 D-glucose, 2 CaCl2 and 1 MgCl2 bubbled with 95%O2/5%CO2). Recordings were targeted to V1m based on coordinates from the Paxinos mouse brain atlas (2004). We made whole-cell recordings from L2/3 on a custom built microscope under differential interference contrast microscopy. L2/3 pyramidal neurons were recorded (based on spiking properties and pyramidal shaped soma) in current clamp mode (Multiclamp 700B, Molecular Devices, Sunnyvale, CA). The data was acquired using Ephus freeware (Janelia Farm, Svoboda Lab). All recordings were low-pass filtered at 3 kHz and digitized at 10 kHz. Patch pipettes (4 – 7 MΩ) contained the following in mM: 130 KMeSO4, 8 NaCl, 2 KH2PO4, 2 D-glucose and 10 HEPES. Recordings were discarded if the series resistance varied by ± 15% or the resting membrane potential or input resistance varied by ± 10 % across the recording period.

Current stimuli

For *in vitro* experiments, a constant current was injected into the soma to achieve a holding potential close to the action potential threshold (typically -55 mV) and zero mean signal and noise currents were injected to study the transfer of correlations. Signal and noise currents were Gaussian white noise initially created with a sampling rate of 200 Hz (to match the 5 ms time bins used for all initial analyses) and upsampled to 10 kHz with spline interpolation. Currents with a given set of variances and correlations were presented in blocks of 100 trials, with the signal current the same on each trial, and the noise current different on each trial. Currents were delivered in 1 s segments with a 1 s pause between each segment. All pairs of cells included in the analysis were recorded non-simultaneously to ensure the only correlations were those in the injected currents.

In vivo recordings

Adult male gerbils (70-90 g, P60-120) were anesthetized for surgery with an initial injection of a mix of ketamine, xylazine, and saline, and the same solution was infused continuously during recording. A small metal rod was mounted on the skull and used to secure the head of the animal in a stereotaxic device in a sound-attenuated chamber. A craniotomy was made over the primary auditory cortex between 2 mm and 4.5 mm anterior to lambda spanning the dorsoventral extent of the squamosal plate. An incision was made in the dura mater and a multi-tetrode array (figure 1A; Neuronexus) was inserted into the brain at a site approximately 0.8 mm medial from most lateral point of the exposed cortical surface and approximately 3 mm anterior to lambda, with variations due to surface vasculature. The electrode was advanced along the dorsoventral axis with the shanks oriented rostrocaudally and lowered into layer V/VI of A1 (between 1 mm and 1.5 mm). A1 was distinguished from other fields based on the direction of the tonotopic gradient (Thomas, Tillein, Heil, & Scheich, 1993). Recording sites were located in the low frequency area of A1 with preferred frequencies for most sites between 1 and 2 kHz.

Spike sorting

For *in vivo* experiments, the procedure for the isolation of single-unit spikes consisted of 1) bandpass filtering each channel between 500 and 5000 Hz, 2) whitening each tetrode, i.e. projecting the signals from the 4 channels into a space in which they are uncorrelated, 3) identifying potential spikes as snippets with energy (Choi, Jung, & Kim, 2006) that exceeded a threshold (with a minimum of 0.7 ms between potential spikes), 4) projecting each of the snippets into the space defined by the first three principal components for each channel, 5) identifying clusters of snippets within this space using KlustaKwik (http://klustakwik.sourceforge.net) and Klusters (Hazan, Zugaro, & Buzsáki, 2006), and 6) quantifying the likelihood that each cluster represented a single unit using isolation distance (Schmitzer-Torbert, Jackson, Henze, Harris, & Redish, 2005). Isolation distance assumes that each cluster forms a multi-dimensional Gaussian cloud in feature space and measures, in terms of the standard deviation of the original cluster, the increase in the size of the cluster required to double the number of snippets within it. The number of snippets in the 'noise' cluster (non-isolated multi-unit activity) for each tetrode was always at least as large as the number of spikes in any single-unit cluster. Only single-unit clusters with an isolation distance greater than 20 were analyzed.

Auditory stimuli

For *in vivo* experiments, sounds were generated with a 48 kHz sampling rate, attenuated, and delivered to speakers. In some experiments, a free-field speaker (TDT MF1) was positioned 10 cm from the ear contralateral to the recording site. In these experiments, sounds were filtered such that the effective frequency response of the speaker measured at the location of the ear was flat ($\pm$ 5 dB SPL) between 0.5 and 20 kHz. In other experiments, speakers (Etymotic ER2) coupled to tubes were inserted into both ear canals for diotic sound presentation along with microphones for calibration. The frequency response of these speakers measured at the entrance of the ear canal was flat ($\pm$ 5 dB SPL) between 0.2 and 5 kHz. At each recording site, a sequence of tones with different frequencies and intensities with 5 ms cosine on and off ramps were presented to characterize basic response properties. In free-field experiments, a 2.5 second segment of female speech from the UCL SCRIBE database (http://www.phon.ucl.ac.uk/resource/scribe) was presented 512 times. In closed-field experiments, a 3 second segment of speech from the same database was presented 120 times with an interaural time difference of +160 μs or -160 μs, corresponding to azimuthal angles of approximately +90 deg. or -90 deg. in virtual acoustic space (interaural level differences for gerbils are negligible at frequencies below 5 kHz(Maki & Furukawa, 2005)).

Calculation of spike train SNR and correlations

The analyses of spike trains shown in all figures were performed as follows: Spike trains were initially considered as binary vectors r with $r_i^p(t)$ = 1 if cell p fired an action potential in time bin t on trial i (and $r_i^p(t)$ = 0 otherwise) with 5 ms time bins. The spike train signal-to-noise ratio (SNR) was computed as

$$\mathrm{SNR} = \frac{var(\bar{r})}{\langle var(r_i - \bar{r}) \rangle_i}$$

where $\bar{r}$ is the average of the spike trains across trials (the PSTH). Spike train correlations for 5 ms bins were computed directly from the binary spike trains (figures 3 and 5). Spike train correlations for 50 ms bins were computed after converting binary spike trains to spike count vectors by summing the spikes from non-overlapping groups of 10 successive bins (figures 1, 2, and 4). For both the binary spike trains and the spike count vectors, the total spike train correlation

between the spike trains of two cells *p* and *q* was calculated as the correlation coefficient between spike trains on the same trials:

$$\rho_r = \frac{\langle cov(r_i^p, r_i^q)\rangle_i}{\sqrt{\langle var(r_i^p)\rangle_i \langle var(r_i^q)\rangle_i}}$$

The spike train signal correlation between spike trains of two cells *p* and *q* was calculated as the correlation between spike trains on different trials:

$$\rho_r^{sig} = \frac{\langle cov(r_i^p, r_j^q)\rangle_{i\neq j}}{\sqrt{\langle var(r_i^p)\rangle_i \langle var(r_i^q)\rangle_i}}$$

The spike train noise correlation was calculated as the difference between total and signal correlations:

$$\rho_r^{noise} = \rho_r - \rho_r^{sig} = \frac{\langle cov(r_i^p, r_i^q)\rangle_i - \langle cov(r_i^p, r_j^q)\rangle_{i\neq j}}{\sqrt{\langle var(r_i^p)\rangle_i \langle var(r_i^q)\rangle_i}}$$

Calculation of membrane potential variances and correlations

The analyses of membrane potentials recorded in vitro shown in figure 2 were performed as follows: The membrane potential for cell p on trial i in each 5 ms time bin k, $v_i^p(t)$, was obtained by first removing action potentials and then downsampling from 10 kHz by computing the mean value in each time bin. The total membrane potential variance for each cell *p* was computed as the average variance of these signals across trials, $\langle var(v_i^p)\rangle_i$, and the total membrane potential correlation between two cells *p* and *q* was calculated as the correlation coefficient between membrane potentials on the same trials:

$$\rho_v = \frac{\langle cov(v_i^p, v_i^q)\rangle_i}{\sqrt{\langle var(v_i^p)\rangle_i \langle var(v_i^q)\rangle_i}}$$

The signal component of the membrane potential was estimated from the mean membrane potential averaged across all trials. The membrane potential signal variance for each cell *p* was computed as the variance of the signal component $var(\langle v_i^p\rangle_i)$ and the membrane potential signal correlation between two cells *p* and *q* was calculated as the correlation coefficient between signal components:

$$\rho_s = \frac{cov(\langle v_i^p\rangle_i, \langle v_i^q\rangle_i)}{\sqrt{var(\langle v_i^p\rangle_i) var(\langle v_i^q\rangle_i)}}$$

The noise component of the membrane potential was estimated from the deviation of the membrane potential on each trial from the mean averaged across all trials. The membrane potential noise variance for each cell *p* was computed as the average variance of these noise components across trials $\langle var(v_i^p - \langle v_i^p\rangle_i)\rangle_i$ and membrane potential noise correlation between two cells *p* and *q* was calculated as the correlation coefficient between these noise components:

$$\rho_n = \frac{\langle cov\left(v_i^p - \langle v_i^p \rangle_i, v_i^q - \langle v_i^q \rangle_i\right) \rangle_i}{\sqrt{\langle var\left(v_i^p - \langle v_i^p \rangle_i\right) \rangle_i \langle var\left(v_i^q - \langle v_i^q \rangle_i\right) \rangle_i}}$$

A Dichotomized Gaussian framework for spike trains

We have previously developed a Dichotomized Gaussian framework to generate spike train with specified second-order statistics (Lyamzin et al., 2012, 2010). Within this framework, there are a unique set of input parameters that correspond to the mean spike rates, spike train SNRs, and spike train signal and noise correlations for any pair of cells, and these input parameters can be estimated from the spike trains. For a pair of cells p and q, we assume that binary spike trains are generated for each cell on each trial i by thresholding the sum of signal and noise inputs in each time bin t:

$$r_i^p(t) = \begin{cases} 1 & \text{if } s^p(t) + n_i^p(t) > 1 \\ 0 & \text{else} \end{cases} \quad \text{and} \quad r_i^q(t) = \begin{cases} 1 & \text{if } s^q(t) + n_i^q(t) > 1 \\ 0 & \text{else} \end{cases}$$

The signal input s was assumed to be the same on each trial for each cell and the noise input n was assumed to be different on each trial for each cell.

DG version 1: For analyses of correlations in binary spike trains with 5 ms time bins that required separate estimates of input signal variance and input signal correlation, the signal was assumed to be Gaussian $s \sim \mathcal{N}(0, \Sigma_s)$ with covariance $\Sigma_s = \begin{bmatrix} \sigma_{s^p}^2 & \rho_s \sigma_{s^p} \sigma_{s^q} \\ \rho_s \sigma_{s^p} \sigma_{s^q} & \sigma_{s^q}^2 \end{bmatrix}$ and the noise was assumed to be Gaussian $n \sim \mathcal{N}(0, \Sigma_n)$ with covariance $\Sigma_n = \begin{bmatrix} \sigma_{n^p}^2 & \rho_n \sigma_{n^p} \sigma_{n^q} \\ \rho_n \sigma_{n^p} \sigma_{n^q} & \sigma_{n^q}^2 \end{bmatrix}$.

DG version 2: For analyses of correlations in spike count vectors with 50 ms time bins that required matching correlations across multiple 5 ms time bins in the binary spike trains, the signal input was assumed to be an arbitrary deterministic function and the noise was assumed to be Gaussian $n \sim \mathcal{N}(0, \Sigma_n)$ with covariance $\Sigma_n = \begin{bmatrix} 1 & \rho_n \\ \rho_n & 1 \end{bmatrix}$, with the correlation ρ_n specified for lags up to 6 time bins (30 ms), which was sufficient to match the spike train correlations with 50 ms bins. When generating spike trains with DG version 2, the noise process was generated by a multivariate autoregressive process(Gutnisky & Josić, 2010).

Our methods for inferring the input signal and noise parameters from spike trains are described below. For further details, see (Lyamzin et al., 2012, 2010).

Estimation of input SNR from spike trains

The estimates of input SNR from spike trains used in the analyses in figures 3, 4, and 5 were obtained as follows: we assumed that binary spike trains with 5 ms time bins were generated by a dichotomized Gaussian process with Gaussian signal and noise inputs as described above (DG version 1). Within this framework, the mean spike rate r_0 can be written in terms of the input signal and noise variances as

$$r_0 = \Phi(-1, \sigma_s^2 + \sigma_n^2)$$

where $\Phi(x, \sigma^2)$ is the cumulative distribution function (CDF) for a Gaussian with zero mean and variance σ^2 evaluated at x. The spike train SNR can also be written in terms of the input signal and noise variances. The numerator of the equation spike train SNR given above can be rewritten as

$$var(\bar{r}) = var\left(\frac{1}{I}\sum_{i=1}^{I} r_i\right) = \frac{1}{I^2}\left(I\langle var(r_i)\rangle_i + 2I(I-1)\langle cov(r_i, r_j)\rangle_{i\neq j}\right)$$

where

$$\langle var(r_i)\rangle_i = \Phi(-1, \sigma_s^2 + \sigma_n^2)(1 - \Phi(-1, \sigma_s^2 + \sigma_n^2))$$

and

$$\langle cov(r_i, r_j)\rangle_{i\neq j} = \Phi_2\left(\begin{bmatrix}-1\\-1\end{bmatrix}, \begin{bmatrix}\sigma_s^2 + \sigma_n^2 & \sigma_s^2\\ \sigma_s^2 & \sigma_s^2 + \sigma_n^2\end{bmatrix}\right) - \Phi(-1, \sigma_s^2 + \sigma_n^2)^2$$

where $\Phi_2(x, \Sigma)$ is the CDF for a 2-D Gaussian with zero mean and covariance Σ evaluated at x. The denominator of the equation spike train SNR given above can be rewritten as

$$\langle var(r_i - \bar{r})\rangle_i = var(\bar{r}) + \langle var(r_i)\rangle_i - 2\langle cov(\bar{r}, r_i)\rangle_i$$

where

$$\langle cov(\bar{r}, r_i)\rangle_i = \langle \bar{r} \cdot r_i\rangle_{t,i} - r_0^2 = \langle \bar{r}^2\rangle_t - r_0^2 = \langle \frac{1}{I}\left(\sum_{i=1}^{I} r_i\right)^2\rangle_t - r_0^2 = \frac{1}{I^2}\left(I\langle r_i^2\rangle_{t,i} + I(I-1)\langle r_i \cdot r_j\rangle_{t,i\neq j}\right) - r_0^2$$

$$= \frac{1}{I^2}\left(I\Phi(-1, \sigma_s^2 + \sigma_n^2) + I(I-1)\Phi_2\left(\begin{bmatrix}-1\\-1\end{bmatrix}, \begin{bmatrix}\sigma_s^2 + 1 & \sigma_s^2\\ \sigma_s^2 & \sigma_s^2 + 1\end{bmatrix}\right)\right) - \Phi(-1, \sigma_s^2 + \sigma_n^2)^2$$

The values of σ_s^2 and σ_n^2 that satisfied these equations for a given mean spike rate r_0 and spike train SNR were determined numerically. The estimated input SNR was σ_s^2/σ_n^2.

Estimation of input signal and noise correlations from spike trains

The estimates of input signal and noise correlations were computed in two different ways depending on whether spike train correlations were measured from binary vectors with 5 ms time bins or spike count vectors with 50 ms time bins. Estimates of input signal and noise correlations corresponding to spike train correlations in 5 ms time bins (figures 3 and 5) were obtained as follows: we assumed that binary spike trains with 5 ms time bins were generated by a dichotomized Gaussian process with Gaussian signal and noise inputs as described above (DG version 1). Within this framework, the spike train total and signal correlations can be written in terms of the input signal and noise variances and correlations.

$$\rho_r^{sig} = \frac{\langle cov(r_i^p, r_j^q)\rangle_{i\neq j}}{\sqrt{\langle var(r_i^p)\rangle_i \langle var(r_i^q)\rangle_i}} \quad \text{and} \quad \rho_r^{noise} = \frac{\langle cov(r_i^p, r_i^q)\rangle_i - \langle cov(r_i^p, r_j^q)\rangle_{i\neq j}}{\sqrt{\langle var(r_i^p)\rangle_i \langle var(r_i^q)\rangle_i}}$$

The covariance terms in the numerators of the equations for the spike train correlations can be written as

$$\langle cov\left(r_i^p, r_j^q\right)\rangle_i = \Phi_2\left(\begin{bmatrix}-1\\-1\end{bmatrix}, \begin{bmatrix}\sigma_{s^p}^2 + \sigma_{n^p}^2 & \rho_s\sigma_{s^p}\sigma_{s^q} + \rho_n\sigma_{n^p}\sigma_{n^q}\\ \rho_s\sigma_{s^p}\sigma_{s^q} + \rho_n\sigma_{n^p}\sigma_{n^q} & \sigma_{s^q}^2 + \sigma_{n^q}^2\end{bmatrix}\right) - \Phi\left(-1, \sigma_{s^p}^2 + \sigma_{n^p}^2\right)\Phi\left(-1, \sigma_{s^q}^2 + \sigma_{n^q}^2\right)$$

and

$$\langle cov\left(r_i^p, r_j^q\right)\rangle_{i\neq j} = \Phi_2\left(\begin{bmatrix}-1\\-1\end{bmatrix}, \begin{bmatrix}\sigma_{s^p}^2 + \sigma_{n^p}^2 & \rho_s\sigma_{s^p}\sigma_{s^q}\\ \rho_s\sigma_{s^p}\sigma_{s^q} & \sigma_{s^q}^2 + \sigma_{n^q}^2\end{bmatrix}\right) - \Phi\left(-1, \sigma_{s^p}^2 + \sigma_{n^p}^2\right)\Phi\left(-1, \sigma_{s^q}^2 + \sigma_{n^q}^2\right)$$

where Φ and Φ_2 are Gaussian CDFs as described above. The variance terms in the denominators of the equations for the spike train correlations can be written as

$$var\left(r_i^p\right) = \Phi\left(-1, \sigma_{s^p}^2 + \sigma_{n^p}^2\right)\left(1 - \Phi\left(-1, \sigma_{s^p}^2 + \sigma_{n^p}^2\right)\right).$$

The values of ρ_s and ρ_n that satisfied these equations were determined numerically.

Estimates of input signal and noise correlations corresponding to spike train correlations in 50 ms time bins (figures 3 and 5) were obtained as follows: we assumed that binary spike trains with 5 ms time bins were generated by a dichotomized Gaussian process with deterministic signal input and Gaussian noise input as described above (DG version 2). Within this framework, the value of the deterministic signal input for each cell in each time bin can be estimated from the PSTH:

$$\bar{r}(t) = \langle r_i(t)\rangle_i = \Phi(s(t),1)$$

The values of s that satisfied this equation in each time bin were determined numerically. The input signal correlation for each pair of cells was estimated as the correlation coefficient between signal inputs after downsampling by averaging values from non-overlapping groups of 10 successive bins. The noise correlations can be written in terms of the input signal and noise variances and correlations as

$$\rho_r^{noise} = \frac{\langle cov\left(r_i^p, r_i^q\right)\rangle_i - \langle cov\left(r_i^p, r_j^q\right)\rangle_{i\neq j}}{\sqrt{\langle var\left(r_i^p\right)\rangle_i \langle var\left(r_i^q\right)\rangle_i}}$$

The covariance terms in the numerator of the equations for the spike train correlations can be written as

$$\langle cov\left(r_i^p, r_j^q\right)\rangle_i = \langle \Phi_2\left(\begin{bmatrix}s^p\\s^q\end{bmatrix}, \begin{bmatrix}1 & \rho_n\\ \rho_n & 1\end{bmatrix}\right)\rangle_t - \langle \Phi(s^p(t),1)\rangle_t \langle \Phi(s^q(t),1)\rangle_t$$

and

$$\langle cov\left(r_i^p, r_j^q\right)\rangle_{i\neq j} = \langle \Phi(s^p(t),1)\Phi(s^q(t),1)\rangle_t - \langle \Phi(s^p(t),1)\rangle_t \langle \Phi(s^q(t),1)\rangle_t$$

where Φ and Φ_2 are Gaussian CDFs as described above. The variance terms in the denominators of the equation for the spike train correlations can be written as

$$var\left(r_i^p\right) = \langle \Phi(s^p(t),1)\rangle_t (1 - \langle \Phi(s^p(t),1)\rangle_t).$$

The values of ρ_n that satisfied these equations were determined numerically for lags up to 6 time bins (30 ms), which was sufficient to match the spike train correlations with 50 ms bins. The input signal correlation for each pair of cells was estimated by summing the values of ρ_n across time bins.

Acknowledgements

This work was supported by the German Research Foundation (DFG) and the Wellcome Trust. The authors thank X for comments on the manuscript.

References

Averbeck, B. B., & Lee, D. (2003). Neural noise and movement-related codes in the macaque supplementary motor area. *The Journal of neuroscience*, *23*(20), 7630–41.

Averbeck, B. B., & Lee, D. (2006). Effects of noise correlations on information encoding and decoding. *Journal of neurophysiology*, *95*(6), 3633–44. doi:10.1152/jn.00919.2005

Bair, W., Zohary, E., & Newsome, W. T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The Journal of neuroscience*, *21*(5), 1676–97.

Binder, M. D., & Powers, R. K. (2001). Relationship Between Simulated Common Synaptic Input and Discharge Synchrony in Cat Spinal Motoneurons Relationship Between Simulated Common Synaptic Input and Discharge Synchrony in Cat Spinal Motoneurons, *Journal of neurophysiology, 86,* 2266–2275.

Choi, J. H., Jung, H. K., & Kim, T. (2006). A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios. *IEEE transactions on bio-medical engineering*, *53*(4), 738–46. doi:10.1109/TBME.2006.870239

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature neuroscience*, *14*(7), 811–9. doi:10.1038/nn.2842

Cohen, M. R., & Maunsell, J. H. R. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, *12*(12), 1594–600. doi:10.1038/nn.2439

De la Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., & Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature*, *448*(7155), 802–6. doi:10.1038/nature06028

Dorn, J. D., & Ringach, D. L. (2003). Estimating membrane voltage correlations from extracellular spike trains. *Journal of neurophysiology*, *89*(4), 2271–8. doi:10.1152/jn.000889.2002

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), 750–61. 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), 220–4. doi:10.1038/nature06563

Gutnisky, D. a, & Josić, K. (2010). Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process. *Journal of neurophysiology*, *103*(5), 2912–30. doi:10.1152/jn.00518.2009

Hazan, L., Zugaro, M., & Buzsáki, G. (2006). Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. *Journal of neuroscience methods*, *155*(2), 207–16. doi:10.1016/j.jneumeth.2006.01.017

Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron*, *78*(2), 352–63. doi:10.1016/j.neuron.2013.02.023

Kohn, A., & Smith, M. a. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *The Journal of neuroscience*, *25*(14), 3661–73. doi:10.1523/JNEUROSCI.5106-04.2005

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 (Bristol, England)*, *23*(1-2), 76–103. doi:10.3109/0954898X.2012.679334

Lyamzin, D. R., Macke, J. H., & Lesica, N. a. (2010). Modeling Population Spike Trains with Specified Time-Varying Spike Rates, Trial-to-Trial Variability, and Pairwise Signal and Noise Correlations. *Frontiers in computational neuroscience*, *4*(November), 144. doi:10.3389/fncom.2010.00144

Macke, J. H., Berens, P., Ecker, A. S., Tolias, A. S., & Bethge, M. (2009). Generating spike trains with specified correlation coefficients. *Neural computation*, *21*(2), 397–423. doi:10.1162/neco.2008.02-08-713

Maki, K., & Furukawa, S. (2005). Acoustical cues for sound localization by the Mongolian gerbil, Meriones unguiculatus. *The Journal of the Acoustical Society of America*, *118*(2), 872. doi:10.1121/1.1944647

Mitchell, J. F., Sundberg, K. a, & Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron*, *63*(6), 879–88. doi:10.1016/j.neuron.2009.09.013

Oram, M. W., Földiák, P., Perrett, D. I., & Sengpiel, F. (1998). The “Ideal Homunculus”: decoding neural population signals. *Trends in neurosciences*, *21*(6), 259–65.

Panzeri, S., Schultz, S. R., Treves, a, & Rolls, E. T. (1999). Correlations and the encoding of information in the nervous system. *Proceedings. Biological sciences / The Royal Society*, *266*(1423), 1001–12. doi:10.1098/rspb.1999.0736

Sadagopan, S., & Ferster, D. (2012). Feedforward origins of response variability underlying contrast invariant orientation tuning in cat visual cortex. *Neuron*, *74*(5), 911–23. doi:10.1016/j.neuron.2012.05.007

Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., & Redish, a D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience*, *131*(1), 1–11. doi:10.1016/j.neuroscience.2004.09.066

Shea-Brown, E., Josić, K., de la Rocha, J., & Doiron, B. (2008). Correlation and Synchrony Transfer in Integrate-and-Fire Neurons: Basic Properties and Consequences for Coding. *Physical Review Letters*, *100*(10), 108102. doi:10.1103/PhysRevLett.100.108102

Tchumatchenko, T., Malyshev, A., Geisel, T., Volgushev, M., & Wolf, F. (2010). Correlations and Synchrony in Threshold Neuron Models. *Physical Review Letters*, *104*(5), 058102. doi:10.1103/PhysRevLett.104.058102

Thomas, H., Tillein, J., Heil, P., & Scheich, H. (1993). Functional organization of auditory cortex in the mongolian gerbil (Meriones unguiculatus). I. Electrophysiological mapping of frequency representation and distinction of fields. *The European journal of neuroscience*, *5*(7), 882–97.

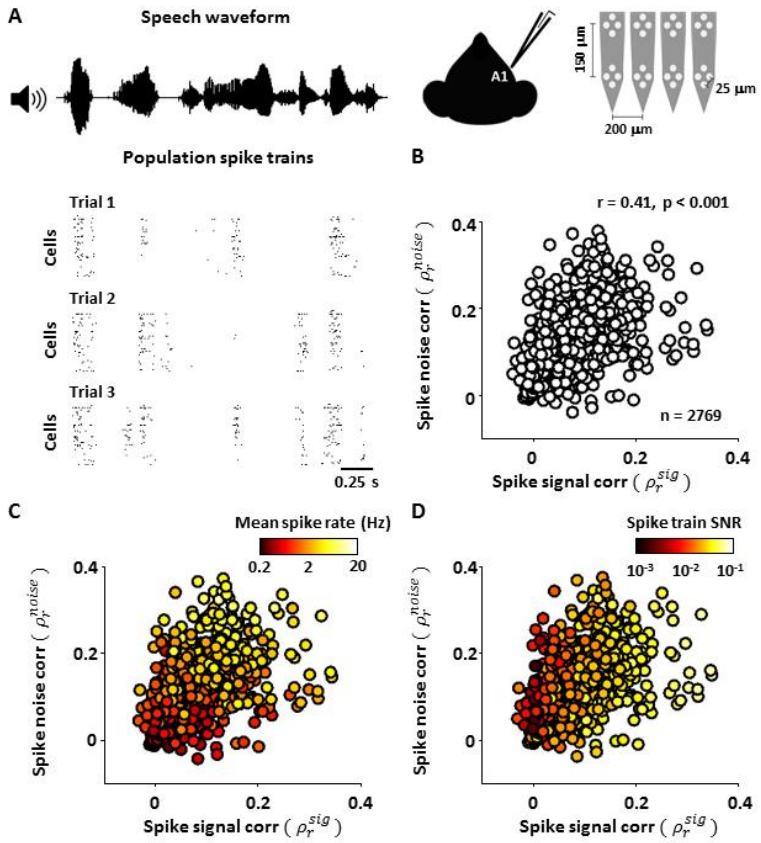


Figure 1: Signal and noise correlations in A1 responses to speech

A) Spike trains were recorded from primary auditory cortex (A1) in anesthetized gerbils with a multi-tetrode array in response to repeated trials of a segment of speech presented from a speaker directed at the contralateral ear. The raster plots show the spike trains of the population on three successive trials. B) The spike train signal and noise correlations ρ_r^{sig} and ρ_r^{noise} in A1 responses to speech. Correlations were computed for spike counts in 50 ms time bins from responses to 512 repeated trials of a 2.5 s segment of speech. C,D) The same correlations as in B, with the dots colored according to the geometric mean of the mean spike rates or spike train SNRs for each pair.

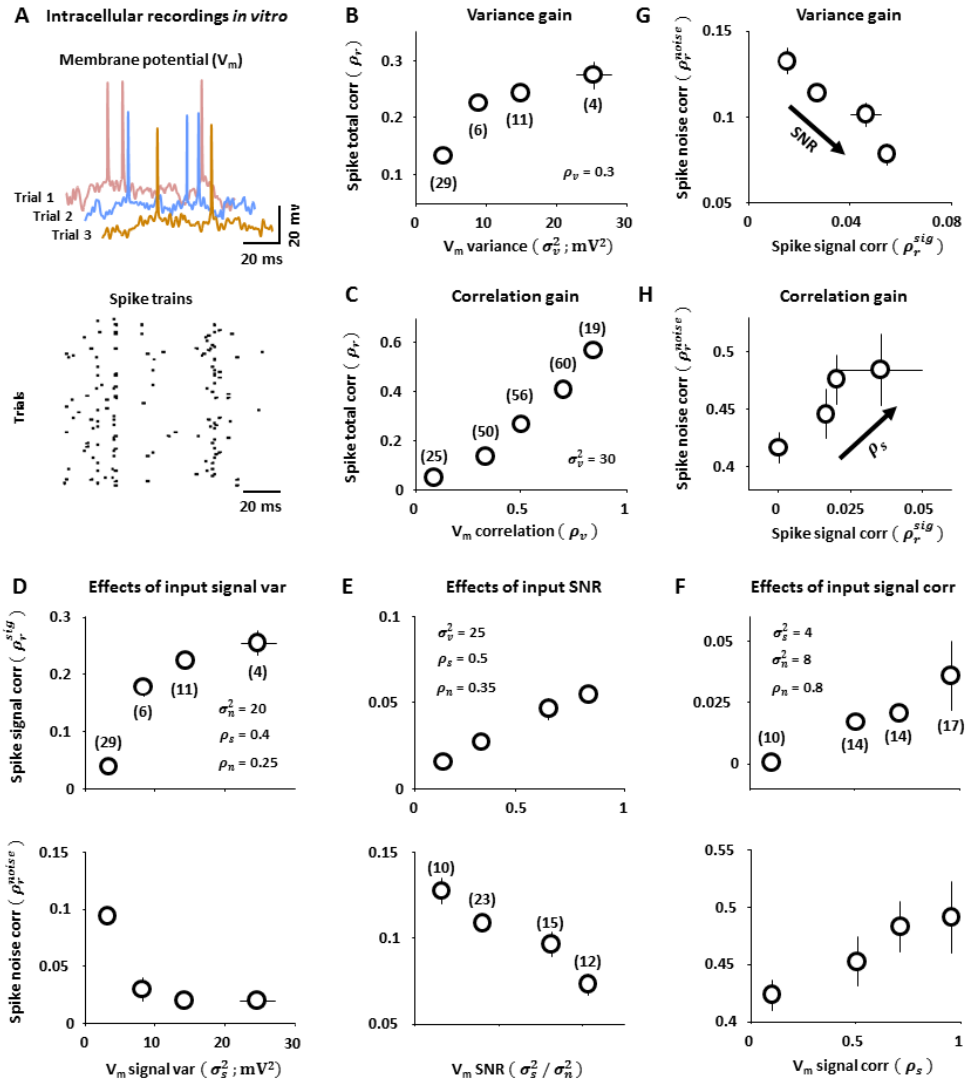


Figure 2: Nonlinear transfer of signal and noise correlations *in vitro*

A) Gaussian white noise currents with specified signal and noise variances and correlations were injected into layer 2/3 pyramidal cells in mouse V1 *in vitro* and the resulting membrane potential correlations and spike train correlations between pairs were compared. Spike train correlations were computed for spike counts in 50 ms time bins. Membrane potential statistics were computed for 5 ms time bins. Membrane potential variances for a pair of cells were computed as the geometric mean of the variances for each cell. B-F) The spike train correlations for pairs of cells with the specified membrane potential variances and correlations. For each plot, pairs were grouped by the value of the membrane potential statistic that was varied as indicated on the horizontal axes. The number of pairs in each group is indicated. Vertical and horizontal error bars indicate SEMs. The mean values of the membrane potential statistics that were not varied are shown. G, H) The data from E and F replotted to show the relationship between spike train signal and noise correlations. Dots are colored by membrane potential SNR or signal correlation with low values in red and high values in yellow.

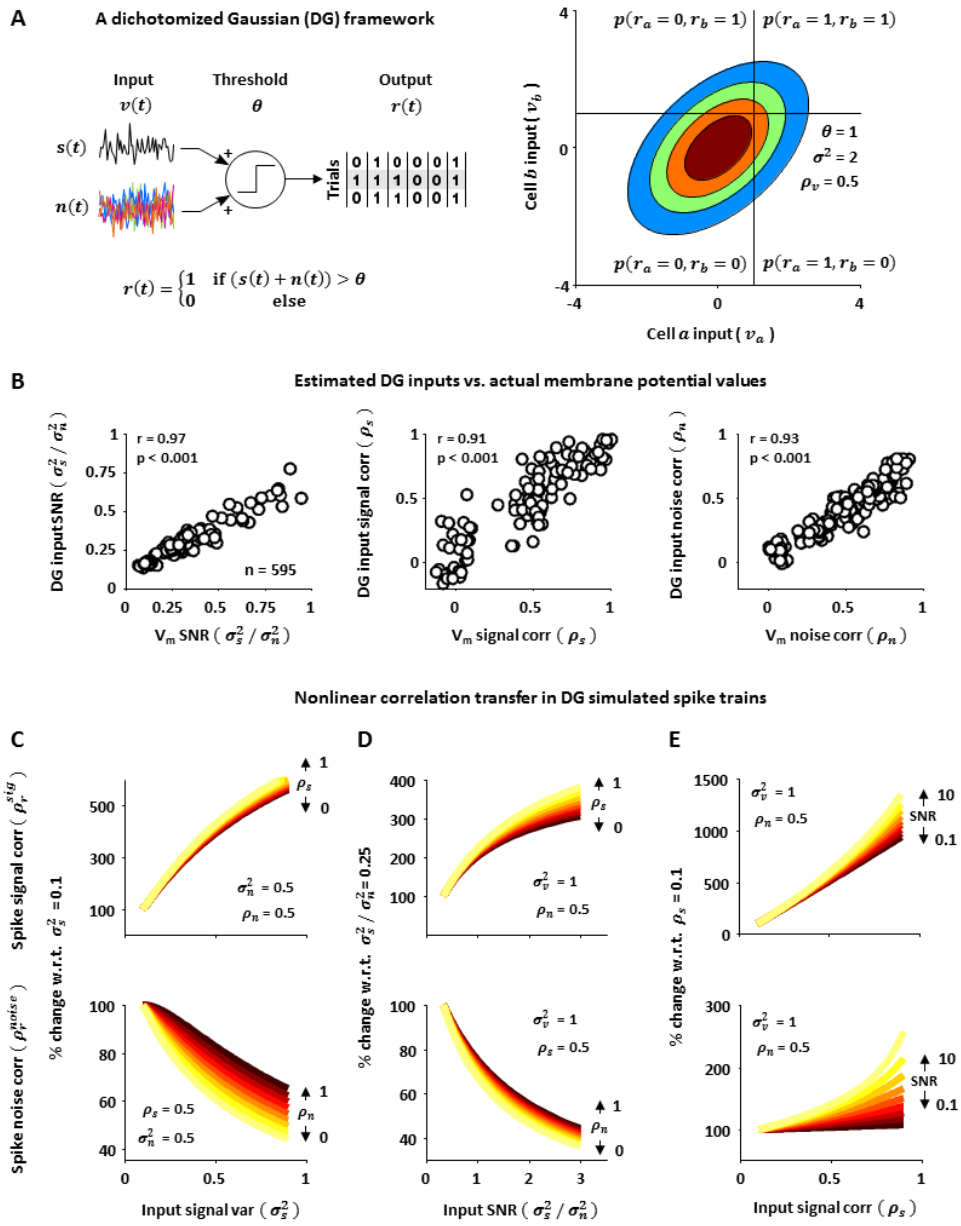


Figure 3: A dichotomized Gaussian (DG) framework for analysis of correlation transfer

A) Schematic diagrams illustrating the DG framework; see Methods for a detailed description. B) The values of input SNR, signal correlation, and noise correlation estimated from spike trains as a function of the actual values for pairs recorded in vitro as described in figure 2. The correlation coefficients between the estimated and actual values are shown. C-E) The spike train signal correlations (top panels) and noise correlations (bottom panels) for pairs of spike trains generated by the DG framework with the specified input variances and correlations.

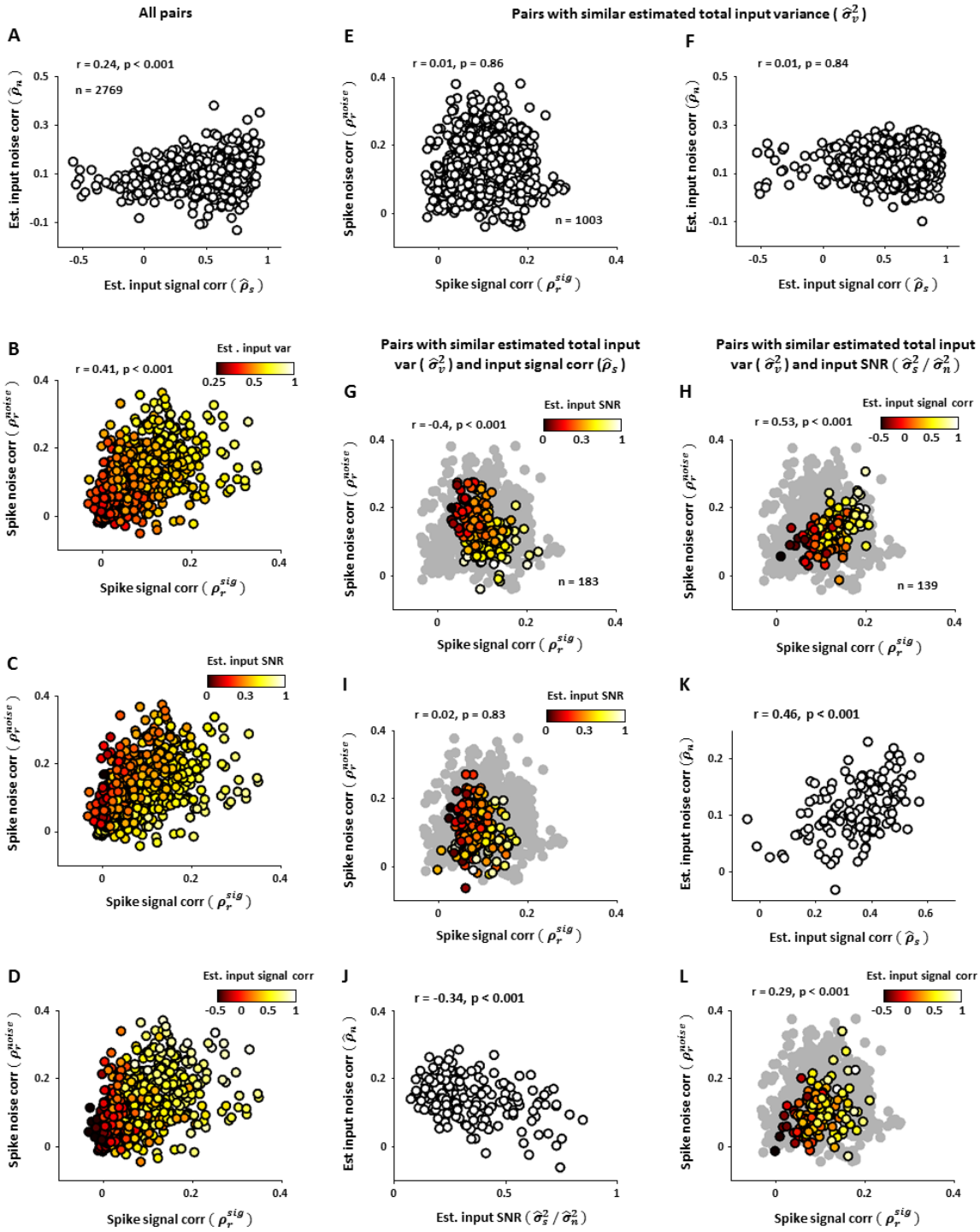


Figure 4: Analysis of spike train signal and noise correlations in A1 response to speech

A) The estimated input signal and noise correlations corresponding to the spike train signal and noise correlations in the A1 responses to speech in figure 1B. B-D) The same spike train signal and noise correlations as in figure 1B, with the dots colored according to the geometric mean of the estimated total input variances, the geometric mean of the estimated input SNRs, or the estimated input signal correlations for each pair. E) The spike train signal and noise correlations in A1 responses to speech for a subset of pairs with estimated total input variance between 0.5 and 0.6. G) The spike train signal and noise correlations in

A1 responses to speech for a subset of pairs with estimated total input variance between 0.5 and 0.6 and estimated input signal correlation between 0.6 and 0.7. The dots are colored according to the geometric mean of the estimated input SNRs for each pair. H) The spike train signal and noise correlations in A1 responses to speech for a subset of pairs with estimated total input variance between 0.5 and 0.6 and estimated input SNR between 0.35 and 0.45. The dots are colored according to the estimated input signal correlation for each pair. I) The signal and noise correlations in spike trains generated by the DG framework after randomizing the pairing of signal and noise inputs for the subset of pairs in G. J) The estimated input SNR and noise correlations for the subset of pairs in G. K) The estimated input signal and noise correlations for the subset of pairs in H. L) The signal and noise correlations in spike trains generated by the DG framework after randomizing the pairing of signal and noise inputs for the subset of pairs in H.

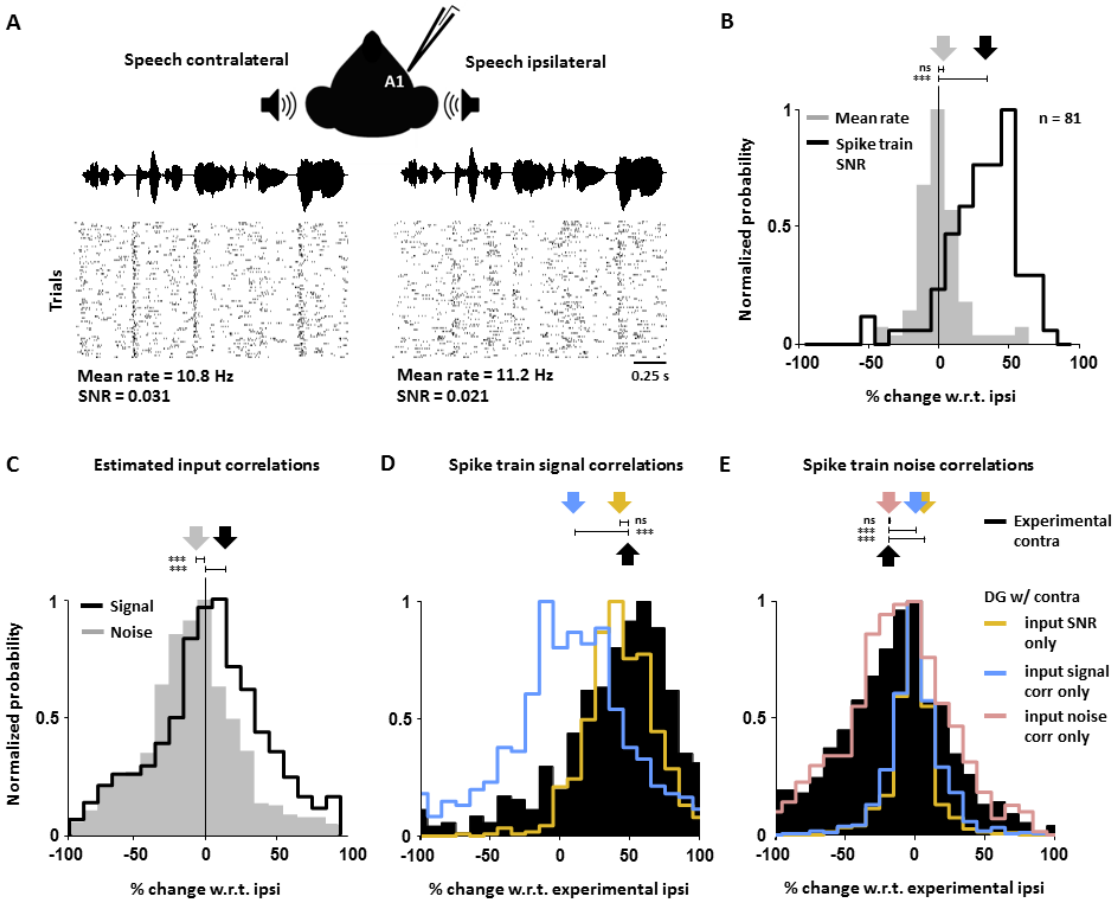
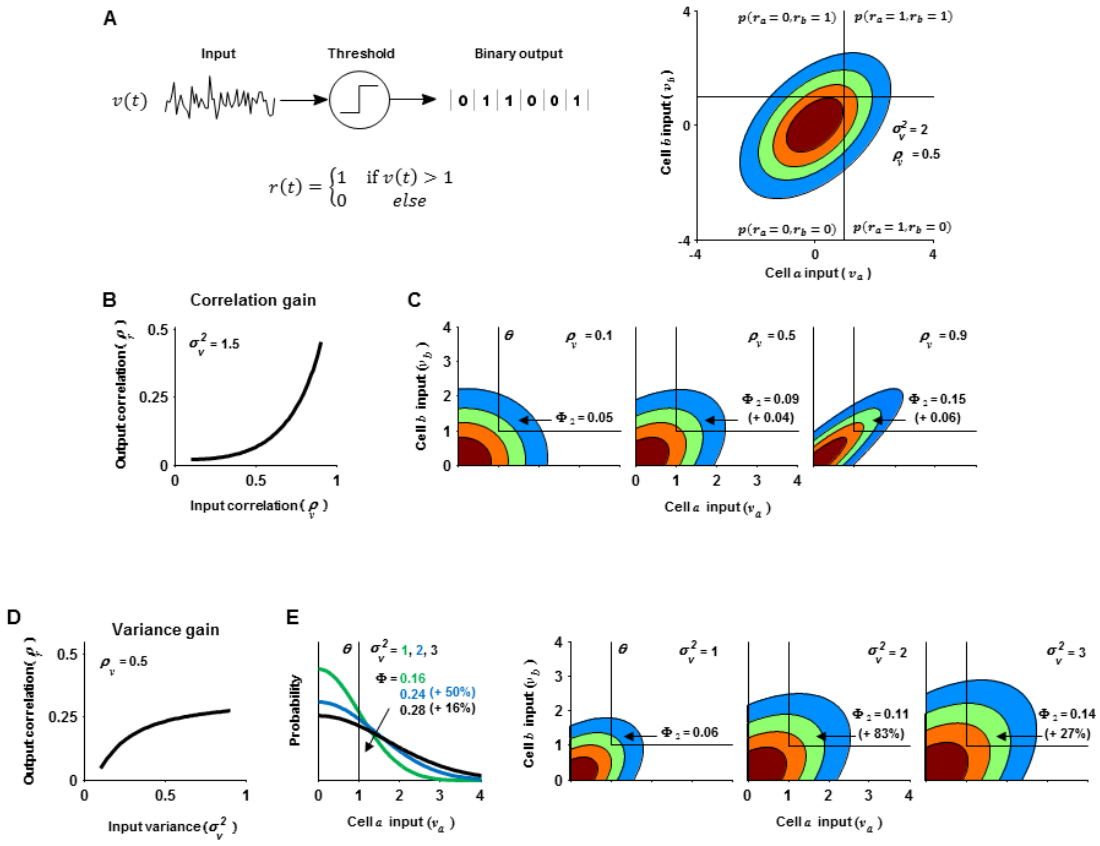


Figure 5: Analysis of changes in spike train signal and noise correlations in A1 responses to speech at different locations

A) Spike trains were recorded in response to repeated trials of a segment of speech presented from a location either contralateral or ipsilateral to the recording site in virtual acoustic space. The raster plots show the spike trains of typical cell across repeated trials. B) Histograms of the % change in the mean spike rate and spike train SNR for individual cells in responses to speech presented contralaterally with respect to speech presented ipsilaterally. C) Histograms of the % change in the input signal and noise correlations for pairs of cells estimated from responses to speech presented contralaterally with respect to speech presented ipsilaterally. D,E) The spike train signal and noise correlations for pairs of cells for actual responses to speech presented contralaterally (black) and simulated responses with input SNRs, signal correlations, and noise correlations estimated from responses to speech presented either ipsilaterally or contralaterally (colored lines). Histograms show % change with respect to the correlations in the actual responses to speech presented ipsilaterally. Correlations were computed for binary spike trains in 5 ms time bins from responses to 100 repeated trials of a 3 s segment of speech.



Supplementary Figure 1: Understanding correlation gain and variance gain within a dichotomized Gaussian framework

This figure is intended to provide intuition into the nonlinear variance gain and correlation gain effects on the transfer of correlations from membrane potential to spiking.

We assume binary output spike trains for a pair of neurons are generated within a dichotomized Gaussian (DG) framework:

$$r(t) = \begin{cases} 1 & \text{if } v(t) > 1 \\ 0 & \text{else} \end{cases}$$

with Gaussian input $v \sim \mathcal{N}(0, \sigma_v^2)$ for each cell, correlation ρ_v between the inputs for the pair, and threshold θ = 1 (see schematic diagrams in figure S1A).

As shown in figure S1B, assuming positive correlations, the correlation ρ_r in the spike trains increases supralinearly as input correlation ρ_v is increased, even when input strength σ_v^2 is constant.

We refer to this effect as 'correlation gain' and we can gain insight into its origin by analyzing the factors that determine the transfer of correlation from input to output within the DG framework.

If we consider two cells a and b with identical input variance σ_v^2 and positive input correlation ρ_v, the output correlation ρ_r can be specified in terms of Gaussian cumulative distribution functions (CDFs)

$$\rho_r = \frac{\mathrm{cov}(r_a, r_b)}{\sqrt{\mathrm{var}(r_a)\mathrm{var}(r_b)}} = \frac{\mathrm{cov}(r_a, r_b)}{\mathrm{var}(r_a)} = \frac{\mathrm{E}[r_a r_b] - \mathrm{E}[r_a]E[r_b]}{\mathrm{E}[r_a^2] - \mathrm{E}[r_a]^2} = \frac{\Phi_2(-\theta, \Sigma) - \Phi(-\theta, \sigma^2)^2}{\Phi(-\theta, \sigma^2) - \Phi(-\theta, \sigma^2)^2}$$

with input covariance matrix $\Sigma = \begin{bmatrix} \sigma_v^2 & \sigma_v^2\rho_v \\ \sigma_v^2\rho_v & \sigma_v^2 \end{bmatrix}$.

Φ and Φ_2 are the values of univariate and bivariate cumulative distribution functions (CDFs) for Gaussians with zero mean and the specified (co)variances evaluated at $-\theta$.

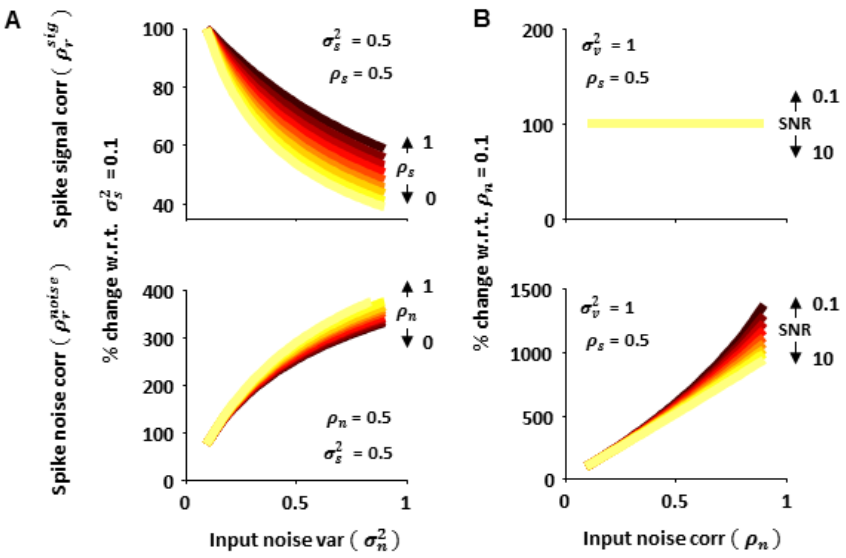
The supralinear correlation gain results from the fact that each increase in ρ_v results in successively more central parts of the bivariate Gaussian (and thus successively larger fractions of probability density) crossing threshold in the Φ_2 term of equation 2, as illustrated in figure S1C (note that the fraction of density above threshold in the Φ_2 term reflects the probability of the cells firing together, but does not reflect the mean spike rate of each cell, which is controlled by the input variance σ_v^2 and is independent of the input correlation ρ_v).

A second important correlation transfer nonlinearity, which we will call 'variance gain', is shown in figure S1D: the spike train correlation ρ_r increases as the input variance σ_v^2 is increased, even when the input correlation ρ_v is constant.

Variance gain can also be understood by comparing the Φ_2 and Φ terms in equation 2.

As illustrated in figure S1E, variance gain results from the fact that a change in σ_v^2 has a larger impact on the Φ_2 term than on the Φ terms (a positive correlation shifts probability density into the quadrant containing the threshold for the bivariate process and, thus, the same increase in variance results in more probability density crossing threshold for the bivariate process than for the univariate one).

Note that variance gain is sublinear: each successive increase in input variance causes a smaller increase in output correlation.



Supplementary Figure 2: Effects of input noise variance and input noise correlation on spike train signal and noise correlations

This figure shows the effects of correlation transfer nonlinearities dependent on the two input properties that are not shown in the main text: input noise variance and input noise correlation.

We generated spike trains using the dichotomized Gaussian (DG) framework described in the main text:

$$r(t) = \begin{cases} 1 & \text{if } s(t) + n(t) > 1 \\ 0 & \text{else} \end{cases}$$

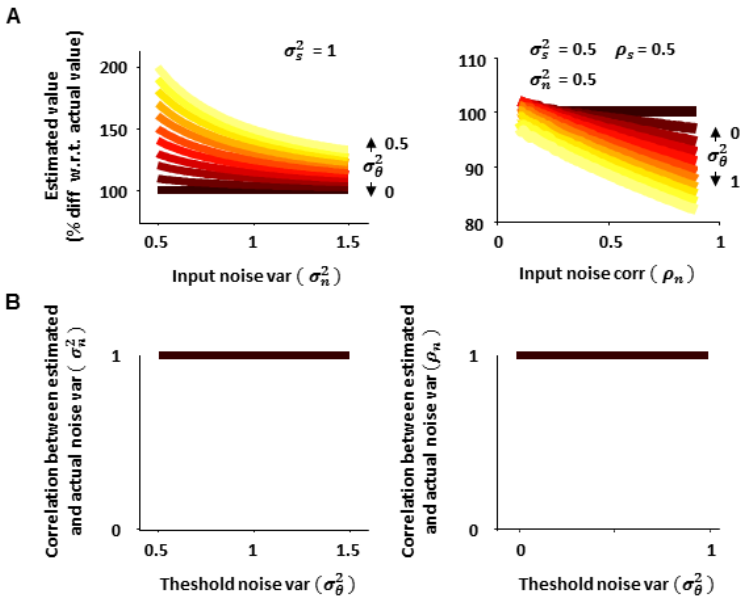
with signal input $s \sim \mathcal{N}(0, \sigma_s^2)$ assumed to be the same on each trial, noise input $n \sim \mathcal{N}(0, \sigma_n^2)$ assumed to be different on each trial, and correlations ρ_s and ρ_n between the signal and noise inputs for each pair of cells.

Figure S2A shows the effects of increasing input noise variance on spike train signal and noise correlations: an increase in the input noise variance σ_n^2 causes a decrease in the spike train signal correlations ρ_r^{sig} and an increase in the spike train noise correlations ρ_r^{noise}, even when the input signal variance and all input correlations are constant.

These effects can be understood as follows: 1) The total correlation in the spike trains ρ_r increases with increasing input noise variance σ_n^2 due to variance gain effects. 2) The spike train signal correlation ρ_r^{sig} decreases with increasing input noise variance σ_n^2 because the increase in σ_n^2 increases the variance of the trial-shuffled spike trains used to calculate ρ_r^{sig}, but does not increase their covariance, so the correlation between them, which is ratio of the covariance to the variance, decreases. 3) The spike train noise correlation ρ_r^{noise} increases with increasing input noise variance σ_n^2 because ρ_r increases and ρ_r^{sig} decreases, so the difference between them $\rho_r^{noise} = \rho_r - \rho_r^{sig}$ increases.

Figure S2B shows the effects of increasing input noise correlation on spike train signal and noise correlations: an increase in the input noise correlation ρ_n causes an increase in the spike train signal correlations ρ_r^{noise}, but has no effect on the spike train signal correlations ρ_r^{sig}.

The reason for the independence of the spike train signal correlation ρ_r^{sig} and the input noise correlation ρ_n is straightforward: a change in ρ_n does not affect either the variance or covariance of the trial-shuffled spike trains used to calculate ρ_r^{sig}.



Supplementary Figure 3: Effects of variable spike threshold on estimates of input noise variance and input noise correlation

This figure shows the effects of a variable spike threshold on estimates of input noise variance and input noise correlation in the dichotomized Gaussian (DG) framework.

We generated spike trains using the dichotomized Gaussian (DG) framework as described in the main text, but with a variable threshold:

$$r(t) = \begin{cases} 1 & \text{if } s(t) + n(t) > 1 + \theta(t) \\ 0 & \text{else} \end{cases}$$

with signal input s and noise input n defined as in the main text, and threshold noise $\theta(t) \sim \mathcal{N}(0, \sigma_\theta^2)$ assumed to be different on each trial.

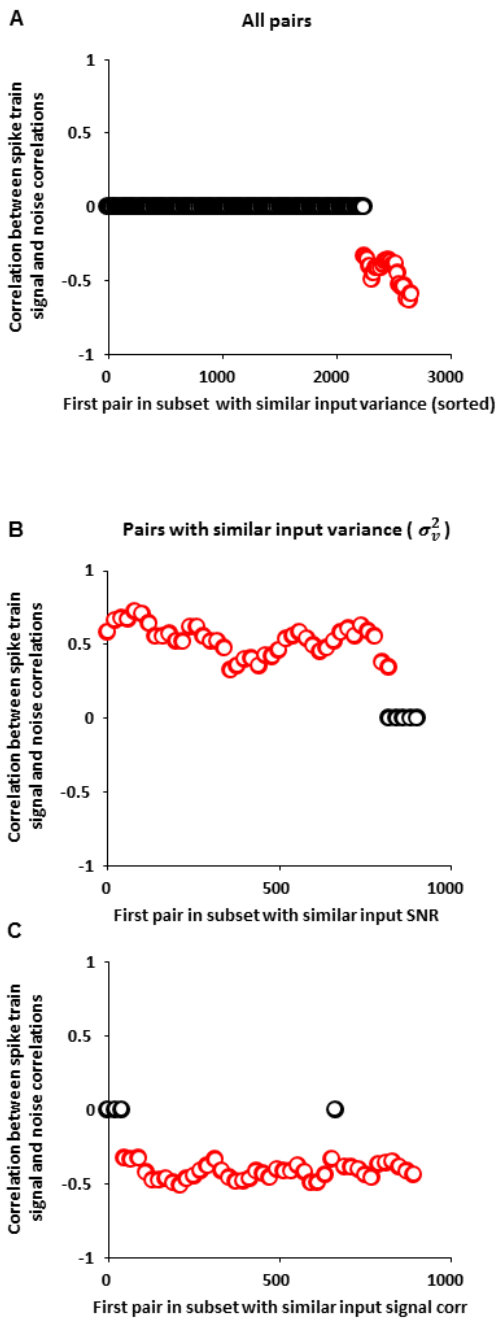
We then estimated the input parameters from the spike trains assuming the DG framework described in the main text with a fixed threshold:

$$r(t) = \begin{cases} 1 & \text{if } s(t) + n(t) > 1 \\ 0 & \text{else} \end{cases}$$

The variable spike threshold is an additional source of noise that is uncorrelated between cells.

Because the standard DG framework does not include a variable threshold, it attributes this additional noise to the noise input n, thus overestimating the input noise variance σ_n^2 and underestimating the input noise correlation ρ_n, as shown in figures S3A.

However, the effect of the variable threshold on the estimated input properties is only a linear transformation and does not affect the correlation between the actual and estimated values, as shown in figure S3B.



Supplementary Figure 4: Dependencies between spike train signal and noise correlations in vivo persist across a wide range of subpopulations

This figure shows that the dependencies between spike train signal and noise correlations in A1 persist across a wide range of subpopulations.

Figure S4A shows the correlation coefficient between the spike train signal and noise correlations for different subsets of pairs of A1 cells with similar total input variance.

The subsets were formed by sorting all pairs in order of increasing total input variance and taking groups of 100 consecutive pairs starting with every tenth pair.

Only the significant correlations ($p < 0.001$) marked in red are shown with their actual values; insignificant correlations are marked in black with a value of zero.

Spike train signal and noise correlations were uncorrelated for subsets including pairs with the lowest 81% of total input variances, and were negatively correlated for subsets including pairs with highest 19% of total input variances.

Figure S4B shows the correlation coefficient between the spike train signal and noise correlations for different subsets of pairs of A1 cells with similar total input variance and SNR.

The subsets were formed by selecting only those pairs with total input variance between 0.5 and 0.6, sorting them in order of increasing input SNR, and taking groups of 100 consecutive pairs starting with every tenth pair.

Spike train signal and noise correlations were positively correlated for subsets including pairs with the lowest 83% of total input variances and uncorrelated for subsets including pairs with highest 17% of total input variances.

Figure S4C shows the correlation coefficient between the spike train signal and noise correlations for different subsets of pairs of A1 cells with similar total input variance and input signal correlation.

The subsets were formed by selecting only those pairs with total input variance between 0.5 and 0.6, sorting them in order of increasing input signal correlation, and taking groups of 100 consecutive pairs starting with every tenth pair.

Spike train signal and noise correlations were negatively correlated for nearly all subsets of pairs.

VI Discussion

In this thesis we have developed a framework based on dichotomized Gaussian that operates on spike trains of populations of neurons obtained in response to repeated presentations of sensory stimuli. The framework is designed to operate on binary matrices, although certain quantities (e.g. signal and noise correlations) can be calculated for larger time bins and therefore arbitrary sets of values, as shown in Chapter V.

Dichotomized Gaussian (DG) process has been proposed over a century ago (Pearson, 1909), and its properties have been studied in detail (Cox & Wermuth, 2002; Emrich & Piedmonte, 2013), including its applications to the modeling of the activity of neuronal populations (Gutnisky & Josić, 2010; Macke et al., 2009) in which the binary matrices generated as the output of DG correspond to spike trains of individual cells in response to repeated presentations of a stimulus.

As follows from the name DG is based on a discrete-time Gaussian process that is thresholded at every point in time. The value of the output of DG at a given time bin equals to one if the value of the Gaussian in that time bin is larger than the threshold value, and equals to zero otherwise. The desired rate of ones (or spike rate in case of neural modeling) is achieved by manipulating mean of the Gaussian, its variance, and threshold. The desired pairwise correlations between binary output matrices are fitted by setting correlation coefficients between the components of the Gaussian process to the appropriate value. As shown in (Macke et al., 2009), DG can be successfully fitted to the spike train data, accurately replicating firing rates and correlations between pairs of neurons.

In the first paper we extended the existing DG framework by representing the latent Gaussian process as a sum of two components: signal and noise. We set the signal component to be the part of the latent Gaussian that was the same on every trial and therefore resulted in different threshold crossing probabilities at different time bins.

The noise component of the Gaussian was different on every trial, although the mean, variance, and correlation coefficients between its components were the same for a given stimulus and a given pair of cells.

Having augmented the existing DG model this way, we found expressions for signal and noise correlations of the resulting binary spike trains.

The first paper illustrates applications of the model: fitting the responses of the experimentally observed pairs of neurons, arbitrarily changing their response properties, and generating population responses with arbitrary statistics.

An important feature of the model that was not stressed in Chapter III, but has been exploited throughout this thesis and in Chapter V in particular, was that all of the single cell parameters and pairwise correlations of the spike trains were linked with each other through internal model parameters, which means that given a change in one of parameters of spike trains, the resulting changes in the others could be calculated numerically, without sampling from the model.

In the second paper we used the ability of the model to reproduce temporal correlations and found expressions for Fano factors, coefficients of variation of the ISI distributions and count correlations of pairs of neurons, depending on the auto- and cross-correlation functions of the latent Gaussians.

We showed among other things that it is possible to have different configurations of ISI distributions for a given variability of counts across trials, even though both these values are defined by the autocorrelation functions of the Gaussian process corresponding to the neuron.

In the third, unpublished work, we focused on the relationships between all parameters of the model that arise due to the nonlinear mechanism of spike generation - thresholding.

We showed that fitting the parameters of spike trains using DG imposes nonlinear relationships between the spike train correlations and input (Gaussian process) correlations.

In particular, increasing input variance led to the increase in spike train correlations, even when the input correlations were kept constant, as observed previously by (de la Rocha et al., 2007). We called this effect *variance gain* and found it to be sublinear: for equal increments of variance the respective changes in spike train correlations were decreasing.

The other fundamental effect that we found was by analogy called *correlation gain* (Binder & Powers, 2001), and it manifested itself in ever increasing spike train correlations when input correlations were incremented by a fixed amount. Thus, correlation gain was found to be supralinear.

According to the existing literature (de la Rocha et al., 2007) the variance gain effect that DG produces is also observed in whole-cell recordings in slice preparations, and therefore represents an actual effect in neurons.

We too found that 1) correlation gain was manifested in the activity of neurons in slice preparations 2) both effects take place for signal and noise components of the input currents 3) these nonlinearities are partially responsible for the dependencies between spike train signal and noise correlations that we observe in the in vivo A1 data.

The nonlinear relationships between spike train correlations and input correlations were found to qualitatively replicate the predictions of DG. In the cases of input SNR and input signal and noise correlations that were of particular interest to us, the predictions of DG model were highly correlated with the results we obtained from whole-cell recordings in slices.

We used the latter finding to decouple the effects of the spike generating nonlinearities from the relationships between correlations and variances already existing in the inputs. To stress this point in particular, the third manuscript shows that some dependencies observed

experimentally e.g. the positive relationship between signal and noise correlations are not exclusively a reflection of the dependencies existing in the network, but also are the result of thresholding.
We show estimate to what extent the observed dependencies are due to inputs or due to the correlation transfer nonlinearities.

In Chapter V we were particularly interested in the dependencies between spike train signal and noise correlations when input SNR or input signal correlations were changing.
We found that in our data input noise correlations are inversely related to input SNR for groups of cells with similar firing rates and input signal correlations. In other words, we found that if two cells get more noise in their inputs, this noise is likely to be correlated. Similarly, we found that for pairs of cells with fixed input firing rates and fixed input SNRs, input signal and noise correlations are positively related. In both of these cases however nonlinearities strengthen the relationships present in the inputs.

Finally we used DG framework for the analysis of population responses to stimuli with varying ITD in the cortex.
We found that 1) with the change in location of the sound from the side contralateral to ipsilateral firing rate does not change while SNR decreases 2) signal correlations decrease 3) noise correlations increase 4) the decrease in signal correlations is predominantly explained by the change in SNR 5) the increase in noise correlations is predominantly explained by the change in input noise correlations.
Summing up these results, with the change of the location of the sound source from contralateral to ipsilateral side, the responses of a population of neurons becomes less temporally precise, however the spiking remains correlated across the population, which is manifested in the increase in input noise correlations.

Variance gain curve in synchronized state

Possibly no less important than understanding how particular correlation structures arise in responses to particular classes of stimuli, the DG framework can give an intuition of how the correlation transfer nonlinearities might ultimately affect behavior through shaping population activity.

As mentioned before, the extent to which the spiking response of a population of sensory neurons carries *signal,* depends not only on how tuning properties of this population are related to the stimulus features, but also on the brain state, and in some cases may be completely defined by the ongoing cortical activity (Marguet & Harris, 2011). Such ongoing cortical activity constitutes in our terminology *noise*, which at rest or during sleep bears most of the power.

In one of such states due to the sublinearity of variance gain a change in input noise variance will not change the extent to which noise is correlated across the local population significantly. On the other hand a similar change of input signal variance will result in a significant increase of signal correlations across the population, which during sleep or rest of an animal will mean an instantaneous forced engagement of the brain in sensory processing.

Such a rapid shift in the correlation structure of the neural activity may lead to desynchronization of local population activity, and could be a mechanism of a bottom-up analogue of attention, saliency. Indeed, salient stimuli are known to involuntarily attract attention, and in case of an idle animal, a sudden appearance of a sensory input can be behaviorally relevant and require immediate reaction. It is also known that *punctuate* stimuli (Kenneth D Harris & Thiele, 2011) produce larger instantaneous responses in a synchronized state, in contrast to sustained stimuli the power of response to which is normally higher in desynchronized state.

This consideration is of course speculative, as it is not exactly clear how an animal would benefit from the correlation transfer nonlinearities in desynchronized state.
In case of desynchronized state the input variance of noise is close to zero, input variance of signal is high and its changes won't affect output signal correlations significantly.
One can speculate that in such state a slight increase in noise variance may be beneficial to noise correlations that arise due to the spiking that is jittered in time, and therefore carries information about suboptimal stimuli (low contrast or ipsilateral ITD) therefore enhancing their coding across the population. Or variance gain can amplify the neural oscillations emerging when an animal is going to sleep or rest, i.e. when a broad correlated engagement of the cortex is necessary. However this argumentation is of course very speculative.

Both correlation and variance gain curves, as shown in Chapter V, depend on the firing properties of a particular neuron and can be steeper or shallower, saturate or reach zero for different values of input correlations. Therefore, the variance gain and correlation gain may have quite different effects from case to case.

The situation is complicated by the fact that DG can obtain good estimates of input correlations of signal and noise, but not input variances. The estimation of variances will require finding an appropriate model for input currents and for the transformation of these currents by cell membrane.
This task goes beyond the scope of current thesis, although it might be a promising direction for future research.

The analogy between visual contrast and ITD processing

Another curious finding that DG helped uncover is a seeming analogy of the ITD processing in A1 and processing of contrast in V1.

In Chapter V we found that when the sound source was moved from the side contralateral to the hemisphere we recorded from to the ipsilateral side, the precision of spikes from individual cells reduced significantly, which resulted in smaller spike train SNR. We also found that noise correlations increased and signal correlations decreased at the same time. At least partially the change in the relative magnitude of signal and noise correlations can be explained by the *redistribution* of spikes between signal and noise: once the position of the sound source became suboptimal, the spiking events across the population lost their *absolute* temporal precision and became *noise* spikes according to our definition, but didn't lose timing precision relative to the rest of the population and remained correlated.

Similar effect was observed in LGN (Desbordes et al., 2008) and V1 (Lampl, Reichova, & Ferster, 1999; Sadagopan & Ferster, 2012), in the latter case cross-correlation functions being steeper for high contrast stimuli not only for spike trains, but also for membrane potential fluctuations.

This mechanism could also be at play in (Kohn & Smith, 2005) where spike train cross-correlations were shown to be shallower for low contrast stimuli and the synchrony was admittedly lower. Although the observed large timescale correlations that resulted in high count correlations were probably due to the synchronous ongoing cortical activity characteristic of the ketamine anesthesia.

Interestingly, suppression of the ongoing cortical activity by a high contrast (salient) stimulus happened even in the case where the neurons were tuned to the orthogonal orientation. In that case the neurons didn't fire any spikes and cross-correlations were flat.

The degraded temporal precision of population response can therefore carry information on the ITD and contrast itself that could be extracted had the cortical neurons had the access to the *absolute time* of their spikes.

Of course, the behavior is based on a single presentation of a stimulus, and given that the degraded temporal precision is correlated across the population it can't therefore be averaged out across the population either.

However it is possible that the information about absolute timing could be extracted when comparing the spike timing to the neurons with stereotyped responses (Brasselet, Panzeri, Logothetis, & Kayser, 2012), or to the spike timing of the other hemisphere.

Measures adopted in this Thesis and in other studies

Our definition of noise correlations is in a slight contrast with works that define noise correlations as fluctuations of mean firing rate or spike count across trials (Bair et al., 2001; Jung et al., 2000; Kohn & Smith, 2005; Lee, Port, Kruse, & Georgopoulos, 1998; Liu et al., 2013). We call these *count correlations* and show the link between count correlations and noise correlations in Chapter II.

Similarly, studies that characterize neuronal responses with firing rates, normally use Fano factor as a measure of trial-to-trial variability (Cohen & Maunsell, 2009; Kara, Reinagel, & Reid, 2000). Our measure of trial to trial variability, SNR (Borst & Theunissen, 1999) is more appropriate for the peripheral sensory neurons and neurons of primary sensory cortices.

Signal correlations in these cases are often characterized as correlations of the average firing rates of two cells across stimulus conditions, or analogously, correlations between their tuning curves (Bair et al., 2001; Kohn & Smith, 2005). Unlike this approach, we get one value of signal correlation per stimulus condition per pair of cells.

The principal difference between measures adopted here and other common measures may seem insignificant. Indeed, any complex naturalistic stimulus can be represented as a series of time bins with certain firing probabilities and certain variability, and can effectively be regarded as a set of stimuli.

However in our definition noise spikes for example can occur due to jitter from the PSTH peak, as discussed before, and thus also have the property of time, and differ for the spike trains with equal numbers of spikes on every trial, but different temporal precision.
This quality is particularly useful if a feature is encoded in timing rather than firing rate, which, as we will see further, is indeed the case for certain classes of stimuli in vision and audition. As an outcome of such temporally-resolved approach, we get that both signal and noise correlations take on a certain value for one stimulus, which makes it possible to compare changes in both signal and noise correlations with changes in stimulus.

Importantly, the measures of signal and noise correlations used in this work represent the comparison of signal and noise correlations both across pairs and across stimuli as equivalent, which might make sense from a certain perspective: for the activity of a neuronal population in response to a sequence of N different stimuli, one can find N subpopulations of neurons that would produce the same set of N responses to one stimulus, provided that the population of neurons is sufficiently heterogeneous.

Mutual information and the specifics of auditory data in local cortical populations

The development of DG framework was partially motivated by its potential use in information theoretic analysis. We did preliminary calculations of the effects that signal and noise correlations have on mutual information in the population of neurons in primary auditory cortex.
In responses of local populations of A1 neurons to diotically presented sounds we found that mutual information measured in one time bin words did not change significantly with a change in noise correlations.

It is quite likely therefore that most of the information at the level of A1 is carried in signal. However, given the inverse relationship between signal and noise correlations found for stimuli varying in their ITD in Chapter V, it could be possible that mutual information will behave differently for that class of stimuli.

Another possible avenue for the information theoretic analysis is taking into account cells with OFF tuning curves found in layer 5 in A1 (Zhou et al., 2010), that way signal correlation between a pair of ON- and OFF-cells will be negative which may ultimately be beneficial for coding (Averbeck & Lee, 2006; Romo, Hernández, Zainos, & Salinas, 2003).

These results have not been published or presented in any form.

Looking at whether changes in correlations affect mutual information is meant to clarify if these changes play a role in sensory processing. However answering this is not always possible primarily because we don't always know the decoding principles of the brain (Latham & Nirenberg, 2005; Sheila Nirenberg & Latham, 2003; Schneidman, Bialek, & Berry, 2003). The approach of measuring mutual information while being numerically accurate and free of any assumptions about the identity of the neural code is not always effective in understanding the coding principles, and should be used in conjunction with other methods when the criteria of optimality are not clear.

In sensory periphery where the primary goal of the neural circuit is to convey information about stimulus features with highest possible precision, the amount of mutual information contained in the population response is a good measure of the optimality of neural code.

As the sensory information travels more centrally, the brain makes use of it for more complex computations, e.g. to calculate higher-order features, to integrate it with the information from other sensory modalities, or to guide the behavior. In these conditions it is sometimes more efficient to adjust population spiking to these computations, even, if necessary, at the expense of losing mutual information.

Ironically, noise correlations in sensory periphery are close to negligible (S Nirenberg, Carcieri, Jacobs, & Latham, 2001) and their contribution to mutual information becomes significant only in the cortex, where the results obtained with information theoretic approach are harder to interpret.

In this respect, the intermediate stages of sensory processing may be of particular interest: while preserving most of the information about the stimulus, neurons in primary sensory areas are the first to adapt the incoming sensory inputs to the ongoing processing of information in the brain.

Bibliography

Abbott, L., & Dayan, P. (2001). Theoretical neuroscience: Computational and mathematical modeling of neural systems.

Averbeck, B. B., & Lee, D. (2003). Neural noise and movement-related codes in the macaque supplementary motor area. *The Journal of neurosciencee*, *23*(20), 7630–41.

Averbeck, B. B., & Lee, D. (2006). Effects of noise correlations on information encoding and decoding. *Journal of neurophysiology*, *95*(6), 3633–44. doi:10.1152/jn.00919.2005

Bair, W., Zohary, E., & Newsome, W. T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The Journal of neuroscience*, *21*(5), 1676–97.

Binder, M. D., & Powers, R. K. (2001). Relationship Between Simulated Common Synaptic Input and Discharge Synchrony in Cat Spinal Motoneurons Relationship Between Simulated Common Synaptic Input and Discharge Synchrony in Cat Spinal Motoneurons. *Journal of neurophysiology*, 86(5): 2266–2275.

Borst, a, & Theunissen, F. E. (1999). Information theory and neural coding. *Nature neuroscience*, *2*(11), 947–57. doi:10.1038/14731

Brasselet, R., Panzeri, S., Logothetis, N. K., & Kayser, C. (2012). Neurons with stereotyped and rapid responses provide a reference frame for relative temporal coding in primate auditory cortex. *The Journal of neuroscience*, *32*(9), 2998–3008. doi:10.1523/JNEUROSCI.5435-11.2012

Chornoboy, E. S., Schramm, L. P., & Karr, a F. (1988). Maximum likelihood identification of neural point process systems. *Biological cybernetics*, *59*(4-5), 265–75.

Cohen, M. R., & Maunsell, J. H. R. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, *12*(12), 1594–600. doi:10.1038/nn.2439

Cox, D., & Wermuth, N. (2002). On some models for multivariate binary variables parallel in complexity with the multivariate Gaussian distribution. *Biometrika*, 462–469.

De la Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., & Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature*, *448*(7155), 802–6. doi:10.1038/nature06028

Desbordes, G., Jin, J., Weng, C., Lesica, N. a, Stanley, G. B., & Alonso, J.-M. (2008). Timing precision in population coding of natural scenes in the early visual system. *PLoS biology*, *6*(12), e324. doi:10.1371/journal.pbio.0060324

Devroye, L. (1986). *Non-Uniform Random Variate Generation*. doi:10.1.1.85.8760

Emrich, L. J., & Piedmonte, M. R. (2013). A Method for Generating High-Dimensional Multivariate Binary Variates, *45*(4), 302–304.

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), 750–61. 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), 220–4. doi:10.1038/nature06563

Gutnisky, D. a, & Josić, K. (2010). Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process. *Journal of neurophysiology*, *103*(5), 2912–30. doi:10.1152/jn.00518.2009

Harris, K D, Henze, D. a, Csicsvari, J., Hirase, H., & Buzsáki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *Journal of neurophysiology*, *84*(1), 401–14.

Harris, Kenneth D, & Thiele, A. (2011). Cortical state and attention. *Nature reviews. Neuroscience*, *12*(9), 509–23. doi:10.1038/nrn3084

Headrick, T. C. (2002). JMASM3: A Method for Simulating Systems of Correlated Binary Data, *I*(I).

Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron*, *78*(2), 352–63. doi:10.1016/j.neuron.2013.02.023

Jung, M. W., Qin, Y., Lee, D., & Mook-Jung, I. (2000). Relationship among discharges of neighboring neurons in the rat prefrontal cortex during spatial working memory tasks. *The Journal of neuroscience*, *20*(16), 6166–72.

Kara, P., Reinagel, P., & Reid, R. C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, *27*(3), 635–46.

Kohn, A., & Smith, M. a. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *The Journal of neuroscience*, *25*(14), 3661–73. doi:10.1523/JNEUROSCI.5106-04.2005

Lampl, I., Reichova, I., & Ferster, D. (1999). in Neurons of the Cat Visual Cortex, *22*, 361–374.

Latham, P. E., & Nirenberg, S. (2005). Synergy, redundancy, and independence in population codes, revisited. *The Journal of neuroscience*, *25*(21), 5195–206. doi:10.1523/JNEUROSCI.5319-04.2005

Lee, D., Port, N. L., Kruse, W., & Georgopoulos, a P. (1998). Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *The Journal of neuroscience*, *18*(3), 1161–70.

Liu, S., Gu, Y., DeAngelis, G. C., & Angelaki, D. E. (2013). Choice-related activity and correlated noise in subcortical vestibular neurons. *Nature neuroscience*, *16*(1), 89–97. doi:10.1038/nn.3267

MacKay, D. (2003). *Information theory, inference and learning algorithms*.

Macke, J. H., Berens, P., Ecker, A. S., Tolias, A. S., & Bethge, M. (2009). Generating spike trains with specified correlation coefficients. *Neural computation*, *21*(2), 397–423. doi:10.1162/neco.2008.02-08-713

Marguet, S. L., & Harris, K. D. (2011). State-dependent representation of amplitude-modulated noise stimuli in rat auditory cortex. *The Journal of neuroscience*, *31*(17), 6414–20. doi:10.1523/JNEUROSCI.5773-10.2011

Martin, K. a C., & Schröder, S. (2013). Functional heterogeneity in neighboring neurons of cat primary visual cortex in response to both artificial and natural stimuli. *The Journal of neuroscience*, *33*(17), 7325–44. doi:10.1523/JNEUROSCI.4071-12.2013

Mitchell, J. F., Sundberg, K. a, & Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron*, *63*(6), 879–88. doi:10.1016/j.neuron.2009.09.013

Moore, G. P., Segundo, J. P., Perkel, D. H., & Levitan, H. (1970). Statistical signs of synaptic interaction in neurons. *Biophysical journal*, *10*(9), 876–900. doi:10.1016/S0006-3495(70)86341-X

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784. doi:10.1126/science.4023713

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

Nirenberg, Sheila, & Latham, P. E. (2003). Decoding neuronal spike trains: how important are correlations? *Proceedings of the National Academy of Sciences of the United States of America*, *100*(12), 7348–53. doi:10.1073/pnas.1131895100

O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*(3), 317–30. doi:10.1002/hipo.450030307

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), 617–21. doi:10.1038/nature09178

Pearson, K. (1909). On a New Method of Determining Correlation Between a Measured Character A, and a Character B, of which Only the Percentage of Cases Wherein B Exceeds (or Falls Short of) a Given Intensity is Recorded for Each Grade of A. *Biometrika*, *7*(1/2), 96. doi:10.2307/2345365

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., & Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, *454*(7207), 995–9. doi:10.1038/nature07140

Pola, G., Thiele, a, Hoffmann, K. P., & Panzeri, S. (2003). An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network (Bristol, England)*, *14*(1), 35–60.

Poulet, J. F. a, & Petersen, C. C. H. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature*, *454*(7206), 881–5. doi:10.1038/nature07150

Qaqish, B. (2003). A family of multivariate binary distributions for simulating correlated binary variables with specified marginal means and correlations. *Biometrika*, 455–463.

Romo, R., Hernández, A., Zainos, A., & Salinas, E. (2003). Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron*, *38*(4), 649–57.

Sadagopan, S., & Ferster, D. (2012). Feedforward origins of response variability underlying contrast invariant orientation tuning in cat visual cortex. *Neuron*, *74*(5), 911–23. doi:10.1016/j.neuron.2012.05.007

Schneidman, E., Bialek, W., & Berry, M. J. (2003). Synergy, redundancy, and independence in population codes. *The Journal of neuroscience*, *23*(37), 11539–53. doi:10.1523/JNEUROSCI.5319-04.2005

Wilson, M. a, & McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science (New York, N.Y.)*, *261*(5124), 1055–8.

Zagha, E., Casale, A. E., Sachdev, R. N. S., McGinley, M. J., & McCormick, D. a. (2013). Motor cortex feedback influences sensory processing by modulating network state. *Neuron*, *79*(3), 567–78. doi:10.1016/j.neuron.2013.06.008

Zhou, Y., Liu, B., Wu, G. K., Kim, Y.-J., Xiao, Z., Tao, H. W., & Zhang, L. I. (2010). Preceding inhibition silences layer 6 neurons in auditory cortex. *Neuron*, *65*(5), 706–17. doi:10.1016/j.neuron.2010.02.021