Improved estimation of neural correlations suggests detailed interactions in visual cortex

Dimitri Yatsenko,[1,*], Krešimir Josić[2], Alexander S. Ecker[1,3,4], Emmanouil Froudarakis[1], R. James Cotton[1], Andreas S. Tolias[1,5]

1 Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

2 Department of Mathematics and Department of Biology and Biochemistry, University of Houston, Houston, TX, USA

3 Werner Reichardt Center for Integrative Neuroscience and Institute for Theoretical Physics, University of Tübingen, Germany

4 Bernstein Center for Computational Neuroscience, Tübingen, Germany

5 Department of Computational and Applied Mathematics, Rice University, Houston, TX, USA

∗ **E-mail: yatsenko@cns.bcm.edu**

Abstract

Ambitious projects currently under way aim to record the activity of ever larger and denser subsets of neurons *in vivo*. Correlations measured in such recordings are anticipated to uncover important aspects of the functional organization of neural circuits. However, estimation and interpretation of large correlation matrices from finite recordings are challenging. Estimation can be improved by regularization: biasing of the estimate towards a low-dimensional, less variable approximation. The amount of improvement depends on how parsimoniously the reduced approximation captures the dominant dependencies in the data. Therefore, the selection of the most efficient estimator is an empirical question that informs about the types of dominant dependencies present in the system.

In this study, we sought the most statistically efficient estimator of neural correlations in recordings from large, dense groups of cortical neurons. Using fast 3D random-access laser scanning microscopy of calcium signals, we recorded the activity of nearly every neuron in volumes of about 100 μm in diameter (150–350 cells) in mouse visual cortex. We hypothesized that, in these dense recordings, the correlation matrix would be most efficiently represented by a sparse network of linear interactions between pairs of observed neurons combined with common latent factors representing unobserved inputs and global network fluctuations. Indeed, in cross-validation tests, the covariance matrix estimator based on this correlation structure outperformed estimators regularized toward other plausible low-dimensional approximations. Furthermore, we found consistent relationships between the organization of sparse interactions inferred by the estimator and the physical distances and preferred orientation differences of pairs of cells. Consistent with previous synaptic connectivity studies, the density of positive interactions decreased rapidly with distance and with differences in preferred orientation, whereas negative interactions were less selective. To further corroborate physiological interpretation of the inferred functional structure, future experiments will augment this analysis with measured synaptic connectivity, cell types, and brain states.

Author Summary

Correlations of the activity of neurons have proven useful as descriptors of the functional organization of neural circuits with implications for stimulus coding and circuit architecture. Estimation of correlation matrices can be improved by imposing some kind of structure on the estimate with greatest improvement attained when the imposed structure closely matches the real dependencies in the data. Using fast 3D two-photon imaging of calcium signals, we recorded the activity of large and dense groups of cells in mouse visual cortex and evaluated the cross-validated performance of correlation matrix estimators that imposed different kinds of structure. The correlation structure of the estimator that proved most efficient

comprised a sparse network of partial correlations between pairs of neurons combined with several common latent factors, with both components required for efficient estimation. Because of its superior benefit for estimation, we proposed that this inferred structure could prove more relevant for the description of functional connectivity than the usual correlation matrix in densely sampled neural recordings. As only a first application of this approach, we analyzed how the inferred connectivity related to distances between cells and differences in their preferred orientations and found basic agreement with previous studies of synaptic connectivity.

Introduction

Pearson correlations between the spiking activity of pairs of neurons, or simply *neural correlations*, are the most familiar descriptive statistics of neural population activity [1–5]. For example, *noise correlations*, *i.e.* the correlations of stimulus response variability between pairs of neurons, have been shown to have profound theoretical implications for stimulus coding [1, 2, 6–8]. In addition, neural correlations are hypothesized to reflect aspects of functional connectivity in neural circuits. Such interpretation is supported by a series of discoveries of nontrivial relationships between neural correlations and other aspects of circuit organization such as the physical distance separating the neurons [9, 10], their synaptic connectivity [11], stimulus tuning similarity [3, 11], cortical layer specificity [12, 13], progressive changes in development and in learning [14], changes due to sensory stimulation and global brain states [5, 15–17], and others.

However, neural correlations do not submit to ready or unambiguous mechanistic interpretation. Theoretical studies and simulations have shown that neural correlations on various temporal scales may arise from combinations of multiple mechanisms including direct synaptic interactions, common inputs or correlated inputs, chains of multiple synaptic connections, oscillations, top-down modulation, and background network fluctuations [18–22].

A correlation matrix provides more information than the equivalent number of pairwise correlations assessed in isolation. Yet early studies of neural correlations were based on measurements from isolated pairs of neurons; the effects of correlations on neural coding were then extrapolated in simulations and theoretical studies [2, 19]. Modern multineuronal recordings from increasingly large and dense subsets of neurons in intact circuits allow estimating the entire correlation matrix of large populations of highly interconnected neurons. Such estimates can be transformed into various representations that express different aspects of the correlation structure and may suggest different mechanistic interpretations. For example, factor analysis or principle component analysis may reveal common activity fluctuations across the entire population [23]. The correlation matrix also allows computing the *partial correlations* between specific pairs of neurons, *i.e.* correlations that remain after accounting for the linear effects of all other recorded cells. Under several assumptions (predominance of linear effects (sufficiently complete sampling of interacting neurons, and sufficient sample size), partial correlations may better reflect the direct statistical associations or interactions between pairs of neurons. In multivariate normal distributions, zero partial correlations between pairs of variables indicate conditional independence or lack of interaction between the pair and the network of non-zero interactions constitute Gaussian Graphical Models or Gauss-Markov Random Fields [24]. With other distributions, the correspondence between conditional dependence and partial correlations diminishes or breaks down [25]. Estimation of networks of partial correlations (sometimes called *association networks*) have been applied to map gene interactions [26, 27] or to establish functional connectivity between brain regions from fMRI data [28, 29].

Multineuronal recordings also present major challenges for the estimation and interpretation of correlations. As the population size increases, the amount of recorded data grows only linearly whereas the number of estimated coefficients grows quadratically. This mismatch leads to increased opportunities for spurious correlations, overestimation of shared activity (*i.e.* overestimation of large eigenvalues) [30], and poorly conditioned estimates of the partial correlations [26].

In this study, we pursued two related aims: (a) improved estimation of neural correlation matrices and (b) discovery of low-dimensional structure of correlations in recordings of multineuronal activity to facilitate their interpretation.

Estimation of the correlation matrix relies on estimating the means and variances of the activity of individual cells. To estimate the mean activity, we resort to the usual *sample mean* without attempting to improve it. Estimation of a *covariance matrix* is equivalent to estimating both variances and correlations. Therefore, we rephrase the problem as one of optimal estimation of the covariance matrix given an estimate of the mean firing rates.

Estimation can be improved through *regularization*: the deliberate biasing of the estimate toward one of several possible low-dimensional approximations referred to here as *target estimates* [26, 31]. The usual estimate, the *sample covariance matrix* has the advantage of being unbiased but, on average, it falls far from the true covariance matrix due to its sensitivity to sampling noise in the data. Low-dimensional estimates of various forms are typically less susceptible to sampling noise but are also liable to introduce their respective biases away from the true covariance matrix. Regularization works by striking a favorable balance between bias and variability. Regularization can produce *some* improvement even with an arbitrary target estimate that has no relation to the true covariance structure in the data. Yet, when the target estimate is well suited for capturing the important features of the true covariance matrix with few terms, it will introduce minimal bias and outperform other estimators.

To illustrate key concepts, Figure 1 illustrates a regularization scheme based on *covariance selection* [32]. In covariance selection, the estimate is produced by fitting only an optimal subset of the coefficients of the *precision matrix* (the inverse of the covariance matrix) while setting the remaining coefficients to zero. Scaling the precision matrix so that its diagonal is composed of all -1's produces the matrix of partial correlations. Thus zeros in the precision matrix produce zero partial correlations between the corresponding pairs of neurons. Panel 1 A depicts the sample correlation matrix of a period of somatic calcium signals of a group of 298 neurons in mouse visual cortex. Due to the low-pass filtration effect of calcium dye kinetics, correlations in unprocessed calcium signals are much higher than typical firing rate correlations on shorter temporal scales. Panel 1 B magnifies the fragment of the sample correlation matrix outlined in 1 A whereas 1 C shows the corresponding fragment of the regularized estimate. The regularized estimate is produced from the same data by zeroing the optimal set of 31,501 (71.2%) off-diagonal coefficients of the precision matrix and fitting the remaining coefficients. Panels 1 D and E show the same fragments of the respective estimates of the partial correlation matrix. The apparent similarity of the two estimates of the correlation matrix belies the utterly different partial correlation structure. The regularized estimate is less sensitive to sample noise in the data, is likely to be closer to truth in most cases than the unbiased estimate, which may be ascertained by cross-validation studies.

It may be tempting to interpret the regularized sparse estimate of partial correlations as a more accurate representation of the functional connectivity than the original correlation coefficients: after all, the regularized solution consistently produces better estimates and has fewer parameters which are easier to interpret. However, an alternative regularization scheme using 43 latent factors interacting linearly with all neurons while having about the same number of parameters, performed even better than covariance selection on this dataset.

It is therefore important to compare multiple families of regularized estimators on the system under investigation. The family that consistently outperforms the others may be proposed as candidate low-dimensional representation of the correlation structure. The greater efficiency of one estimator over another is not a mathematical necessity but an empirical finding depending on the prevalent types of correlation structure in the investigated system.

In this study, we compared four families of regularized combining partial interactions and latent units: 'shrinkage toward diagonal', 'shrinkage toward a multifactor model', 'sparse partial correlations', and 'sparse+latent partial correlations'. First, in simulation with known low-dimensional structure of correlations, we demonstrate that regularized estimators from matching families of low-dimensional targets

generally outperform all other estimators. We then repeat this study on neural data from dense groups of neurons in mouse visual cortex under visual stimulation. We find that most efficient estimation was achieved by modeling correlations as a sparse network of partial correlations combined with several latent units interacting with the entire population. Finally, we analyze the relationship between the elements of such estimates and other properties of the circuit such as physical distances between neurons and their preferred orientations.

Results

Covariance estimation

We aim to estimate the true covariance matrix

$$\Sigma = \mathbb{E}\left[(x-\mu)(x-\mu)^{\mathsf{T}}\right] \tag{1}$$

where $\mathbb{E}[\cdot]$ denotes expectation; x is the $p \times 1$ vector of real-valued instantaneous firing rates in bins of duration Δt. The vector of mean firing rates is $\mu = \mathbb{E}[x]$.

The usual estimator of the covariance matrix is the *sample covariance matrix* C_0 computed from the empirical sample of observations $x(1), \ldots, x(n)$ as

$$C_0 = \frac{1}{\nu}\sum_{t=1}^{n}(x(t)-\mu)(x(t)-\mu)^{\mathsf{T}} \tag{2}$$

where ν is the number of degrees of freedom per neuron in the sample ($\nu = n-1$ if observations are independent). The sample covariance matrix is unbiased, *i.e.* $\mathbb{E}[C_0] = \Sigma$. For finite sample sizes, however, C_0 is not as close to Σ as a number of biased estimators.

We considered four regularized estimators based on distinct families of low-dimensional target estimates: 'independent', 'latent factors', 'sparse interactions', and 'sparse+latent' (Fig. 2 row 1).

In the first regularized estimator C_{diag}, the target estimate is the diagonal matrix D containing on its diagonal estimates of the variances. The regularized estimate is obtained by linear *shrinkage* of the unbiased estimate C_0 toward D controlled by the scalar *shrinkage intensity* parameter $\lambda \in [0,1]$:

$$C_{\mathsf{diag}} = (1-\lambda)C_0 + \lambda D \tag{3}$$

The diagonal target estimate expresses the idea of a lack of dependence (or of linear association) between the activity of observed neurons (Fig. 2 A). If this assumption aptly describes recorded data, then strong shrinkage toward D will add little bias while strongly reducing the variability of the estimate. Shrinkage allows for partial commitment to the low-dimensional representation.

In the second regularized estimator C_{factor}, the target estimate is the factor model $F = LL^{\mathsf{T}} + \Psi$ with d factors so that L is the $p \times d$ matrix of *factor loadings* and the diagonal matrix Ψ contains the independent variances of each neuron. Then the estimate is

$$C_{\mathsf{factor}} = (1-\lambda)C_0 + \lambda F \tag{4}$$

This estimator has two hyperparameters: the number of factors d and shrinkage intensity λ. The target estimate F expresses the assumption that correlated fluctuations in the population activity are driven by a small number of latent factors that affect many cells while direct interactions between cells are insignificant (Fig. 2 B).

The third estimator C_{sparse} is based on the assumption that all correlations are the result of direct linear interactions between pairs of observed cells and that such interactions occur only between a fraction of such pairs (Fig. 2 C). This assumption is enforced by reducing to zero the majority of pairwise

partial correlations in the recorded population. While usual correlations are calculated from the marginal distribution without conditioning on all the other neurons, the partial correlation expresses the pairwise correlation conditioned on the activity of all the other cells. If cells only exert linear effects on each other and all neurons are recorded, partial correlations express the direct interactions between neurons. Therefore, this estimator is biased toward the assumption that correlations arise due to interactions between a subset of pairs of recorded neurons (Fig. 3C). When the partial correlation between a pair of neurons is zero, then the corresponding element of the inverse of the covariance matrix (often referred to as the *precision matrix* or *concentration matrix*) must be zero as well. Then the estimator is

$$C_{\mathsf{sparse}} = S^{-1} \tag{5}$$

where S is a sparse matrix with a large fraction of zeros in its off-diagonal elements. The estimate has one hyperparameter to regulate the sparsity (fraction of off-diagonal zeros) in S.

Finally, we consider the fourth estimator $C_{\mathsf{sparse+lowrank}}$, which provides for both common latent factors interacting with all recorded neurons and sparse interactions between the recorded neurons (Fig. 2 D). This estimator has the form [33, 34]

$$C_{\mathsf{sparse+latent}} = (S - LL^{\mathsf{T}})^{-1} \tag{6}$$

where, as above, S is a sparse matrix and L is a $d \times p$ matrix of factor loadings. The estimator has two hyperparameters: the number of latent units d and the sparsity of S.

Simulation

To verify our approach and to illustrate the performance of the four regularized estimators, we constructed four model populations of size $p = 50$ neurons each with correlation structures of the same type as the target estimates of the first (Fig. 2 Row 2). Panel A2 contains a diagonal correlation matrix matching the target of C_{diag}. Panel B2 is a factor model with 3 factors conforming to the family of target estimates of C_{factor}. Panel B3 is a correlation matrix with sparse partial correlations (76% off diagonal zeros in the precision matrix). Finally, Panel B4 is a correlation matrix whose inverse is composed the sum of a sparse matrix (82% sparse) and a low-rank component (rank=1). Row 3 contains sample correlation matrices calculated from samples of size $n = 1000$ from a multivariate normal distributions with the respective correlation matrices from Row 2.

To evaluate the performance of a covariance matrix C, we define a real-valued *loss function* $\mathcal{L}(C, \Sigma)$ such that it attains its minimum when $C = \Sigma$. The loss function quantifies the error of the estimate, *i.e.* the deviation of the estimate C from truth Σ.

In this study, we adopted *negative normal log-likelihood loss*[1] defined as

$$\mathcal{L}(C, \Sigma) = \frac{1}{p}\left[\ln\det C + \mathrm{tr}(C^{-1}\Sigma)\right] \tag{7}$$

We normalized the loss function by $\frac{1}{p}$ to make its value comparable across different population sizes.

This choice is motivated by mathematical convenience. Other popular choices for the loss function are the Frobenius norm of the difference $\|C - \Sigma\|_F$ [26, 30], Stein's entropy loss, and quadratic loss [35, 36]. We expect that the main conclusions of our study will not change qualitatively under other well behaved loss functions.

[1] The loss function $\mathcal{L}(C, \Sigma)$ is related to the multivariate log-likelihood function $L(\Sigma|C_0)$:

$$L(\Sigma|C_0) = -\frac{p}{2}\ln(2\pi) - \frac{p}{2}\mathcal{L}(\Sigma, C_0)$$

To make the loss function to assume zero at its minimum when $C = \Sigma$, we define *excess loss* as

$$\mathcal{L}_0(C, \Sigma) = \mathcal{L}(C, \Sigma) - \mathcal{L}(\Sigma, \Sigma) \tag{8}$$

In Figure 2, row 4 shows the mean excess losses and their standard errors calculated from 30 samples of sizes n=250, 500, 1000, 2000, and 4000 for each estimator and each kind of ground truth. For each estimator, its hyperparameters were estimated by nested cross-validation (See Methods for more details.)

As expected, estimators with matching low-dimensional structures typically outperformed the other estimators. There were two exceptions to this observation. For small sample sizes, before the data were sufficient to reveal the true correlation structure and to allow the correct model to dominate, estimates with simpler targets often outperformed the matching estimator.

When ground truth Σ is not accessible, loss can be estimated solely from the data through *validation*. In validation, an additional independent *testing sample* is used to compute an independent sample covariance estimate C_0'. Then *validation loss* $\mathcal{L}(C, C_0')$ can be used as an unbiased estimate[2] of loss $\mathcal{L}(C, \Sigma)$. Thus, estimators resulting in consistently lower validation loss can be inferred to produce estimates that are closer to truth than estimators with higher validation loss.

With empirical data, acquiring a separate testing sample is not practical. Instead, K-fold cross-validation is used. In cross-validation, the sample is divided into K subsets of approximately equal size. In all computations in this paper $K = 10$ was used. Then each subset is used as the validation sample with the other $K - 1$ serving as the training dataset. The validation losses from each of such 'folds' are averaged to produce *cross-validation loss* or *CV-loss* for short. Let $C_0^{\{k\}}$ denote the sample covariance matrix computed from the k^{th} subset and $C^{\{\setminus k\}}$ denote the results of estimator C trained on the remaining $K - 1$ subsets. Then cross-validation loss for estimator C is

$$\ell_C = \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}\left(C^{\{\setminus k\}}, C_0^{\{k\}}\right) \tag{9}$$

In the present formulation, ℓ_C is known as the cross-validated Gaussian log-likelihood (up to a constant offset and multiplier).

Since we only need to compare estimators against each other, we are only interested in *relative* CV-loss of estimator C with respect to reference estimator C_{ref}:

$$\ell_{C, C_{\mathsf{ref}}} = \frac{1}{K} \sum_{k=1}^{K} \left[\mathcal{L}\left(C^{\{\setminus k\}}, C_0^{\{k\}}\right) - \mathcal{L}\left(C_{\mathsf{ref}}^{\{\setminus k\}}, C_0^{\{k\}}\right)\right] \tag{10}$$

In simulation, CV-loss accurately reproduced the differences between the estimators' excess losses, although with greater variability (Figure 2, Row 5). For each kind of ground truth (Row 2), relative CV-losses were computed with respect to the estimator whose regularization target matched the structure of the respective ground truth: $\ell_{C, C_{\mathsf{diag}}}$, $\ell_{C, C_{\mathsf{factor}}}$, $\ell_{C, C_{\mathsf{sparse}}}$, and $\ell_{C, C_{\mathsf{sparse+lowrank}}}$. Just as with excess loss in Row 4, the means and their standard errors were computed from 30 samples taken for each ground truth and for each sample size.

These simulation results demonstrated that, with sufficiently large sample sizes, the selection of the most efficient of several regularized estimators could be used to infer the type of low-dimensional structure present in the data, at least when one of the regularized estimators could capture such structure.

[2] Validation loss $\mathcal{L}\left(C, C_0'\right)$ is an unbiased estimate of loss $\mathcal{L}(C, \Sigma)$ when $\mathcal{L}(\cdot, \cdot)$ is *additive* in its second argument so that

$$\mathcal{L}(C, z_1) + \mathcal{L}(C, z_2) \equiv \mathcal{L}(C, z_1 + z_2)$$

Then the absence of bias can be shown:

$$\mathbb{E}\left[\mathcal{L}\left(C, C_0'\right)\right] = \mathcal{L}\left(C, \mathbb{E}\left[C_0'\right]\right) = \mathcal{L}(C, \Sigma)$$

This property does not hold for some other popular losses such as Stein's entropy loss, for example, which prevents their substitution with corresponding validation losses. However, various other losses do conform to this constraints and could be used in this study.

Covariance estimation in neural data

We recorded the calcium activity of dense populations of neurons in the supragranular layers in primary visual cortex of anesthetized mice using fast random-access 3D scanning two-photon microscopy [37–39]. We presented 300 repetitions of full-field drifting gratings with two directions of motion or 150 repetitions with five directions (Fig. 3 A) on one side of the visual field of anesthetized mice. Groups of cells loaded with a calcium-sensitive fluorescent dye were imaged and localized in 3D space (Fig. 3 B and E) in the visual cortex on the contralateral side to the stimulus. With acousto-optic modulator (AOD) 3D steering of the laser, the located cells were imaged at high sampling rates with concurrent motion detection. This technique allowed fast sampling (100–150 Hz) from large numbers (150–350) of cells in a small volume of cortical tissue ($200 \times 200 \times 100$ μm^3) in layers 2/3 and 4 (Fig. 3 C). After downsampling the signal to 20 Hz, relative firing rates were inferred using sparse nonnegative deconvolution [40] (Fig. 3 C and D). Only cells that produced detectable calcium activity were included in the analysis. The average stimulus response was subtracted from each trial; the remaining signals were further downsampled into 150 ms bins to compute the noise correlation matrix (Fig. 3 F and G). The durations of recordings dedicated to estimating the noise correlations ranged between 15–20 minutes. Besides the high-repetition stimulus protocol for noise correlations, an orientation tuning stimulus protocol was used to map the orientation tuning of each cells (Fig. 3 E). Overall, 31 imaged sites from 24 animals were included in the study.

In these highly localized populations both direct interactions between cells and common diffuse inputs are likely to contribute to the overall population variability, leading us to hypothesize that regularized estimates capable of accommodating partial correlations between specific pairs of cells (*e.g.* C_{sparse} and $C_{\mathsf{sparse+lowrank}}$) would be more efficient than others. At the same time, average correlations were relatively small (Fig. 3 G), to the advantage of estimates biased toward independence (*e.g.* C_{diag}). Finally, there is the possibility that the correlations structure is best explained by common fluctuations across the entire population, to the advantage of estimates biased toward low-rank correlation structure (*e.g.* C_{factor}, $C_{\mathsf{factor+lowrank}}$).

We found that estimator $C_{\mathsf{sparse+lowrank}}$ was more efficient than the other estimators (Fig. 4). The median relative CV-validation loss (Eq. 10) of each of the estimates C_0, C_{diag}, C_{factor}, and C_{sparse} with respect to $C_{\mathsf{sparse+lowrank}}$ was significantly above zero ($p < 0.001$, Wilcoxon signed rank test, for each).

Relationship of correlation structure to circuit architecture

Having demonstrated that $C_{\mathsf{sparse+lowrank}}$ dominated the other estimators, we examined the elements of the low-dimensional structure revealed by this estimator for individual imaged sites. By its design, the sparse+lowrank estimator finds the optimal balance between a sparse network of partial correlations and shared common latent units. If this estimate dominates all other estimators evaluated thus far, it seems reasonable to hypothesize that these interaction may reflect underlying physiological interactions more closely than the usual correlations. In particular, the sparse component of the partial correlation matrix suggests direct interactions between pairs of neurons, whereas the low-rank component suggest common fluctuations such as those caused by common inputs or other collective synchronized activity.

We examined the structure of the sparse+lowrank estimate in individual sites such as the example site first depicted in Fig. 3. In this site, as in others, the regularized estimate of the correlation matrix appeared very similar to the unregularized sample correlation matrix (Fig. 5 A and D). However, the corresponding partial correlation matrices differed dramatically (Fig. 5 B and E). The partial correlation was decomposed into the sparse and low-rank components (Fig. 5 C). Although correlations were mostly positive, the sparse partial correlations (or 'interactions'), although mostly positive, had a much larger proportion of negative values. The sparse component had 82.1% sparsity (or 17.9% connectivity), which corresponded to the average node degree (interactions per cell) of 52.5 (Fig. 5 G). The low-rank component was of rank 17.

In previous publications that looked at the structure of correlations, the sparse network was obtained

by thresholding the correlations coefficients above some level deemed significant and examining the network of correlations above that threshold [14, 41]. There was fairly little overlap between the network of interactions revealed by this thresholding method and those revealed by the sparse+lowrank estimator (Fig. 5 F). When thresholded to the same sparsity (82.1%), only 42% of cell pairs connected in one network were connected in the other while the average magnitude of such correlations was much lower in the case of the regularized estimator (Fig. 5 F, H, and I). In particular, many low sample correlations translated into negative interactions in the regularized estimate. Indeed, the absence of a correlation between pairs of cells that both correlate similarly to several of their neighbors should be considered as significant as a high correlation coefficient. Regularized partial correlations reveal such phenomena whereas regular sample correlations cannot.

The average partial correlations revealed by the regularized estimator in all 31 sites were about 5 times lower than the sample correlations and less variable across sites (Fig. 6 A). The average node degree of the sparse component of the partial correlations and the number of inferred latent units varied widely between sites but generally increased with the recorded population size (Fig. 6 B and C). However, there was an inverse relationship between the number of latent units and the average node degree (Fig. 6 D). Several sites, even with relatively large population sizes, had fairly few pairwise interactions and were dominated by latent units. These differences have multiple explanations, including differences in brain states and recording quality (neuropil contamination, motion).

We also examined the relationship of differences in orientation preference and physical distances (lateral and by depth) between pairs of cells to the average sample correlations, average regularized partial correlations, and inferred connectivity between them (Fig. 7). The average partial correlations fell more rapidly with with difference in preferred orientation (Fig. 7 A and D) and lateral displacements at equal depths ($\pm 25\mu$m) (Fig. 7 B and E), and differences in depth at small ($\pm 25\mu$m) lateral displacements (Fig. 7 B and E).

The connectivity of partial correlations in the regularized estimate had different organization for positive and negative interactions. Positive interactions fell rapidly as a function of difference in preferred orientation (Fig. 7 G and J), lateral displacements (Fig. 7 H and K), and displacement in depth (Fig. 7 I and L). Negative interactions were much less selective (Fig. 7 G–L).

Discussion

Big neural data

Meaning of interactions

Fitting probabilistic models (*e.g.* GLMs) vs. covariance estimation.

Future directions

Methods

Data Acquisition

Fast 3D two-photon imaging of calcium signals *in vivo* was performed as described in [39]

Covariance Estimation

TODO

Cross-validation

TODO

Simulation

TODO

Acknowledgments

Genevera Allen
Eftychios Pnevmatikakis

References

1. Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7: 358–366.

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

3. Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J Neurosci 25: 3661-73.

4. Bair W, Zohary E, Newsome WT (2001) Correlated firing in macaque visual area mt: time scales and relationship to behavior. The journal of Neuroscience 21: 1676–1697.

5. Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, et al. (2010) The asynchronous state in cortical circuits. Science 327: 587-90.

6. Abbott L, Dayan P (1999) The effect of correlated variability on the accuracy of a population code. Neural computation 11: 91–101.

7. Berens P, Ecker AS, Gerwinn S, Tolias AS, Bethge M (2011) Reassessing optimal neural population codes with neurometric functions. Proc Natl Acad Sci U S A 108: 4423-8.

8. Ecker AS, Berens P, Tolias AS, Bethge M (2011) The effect of noise correlations in populations of diversely tuned neurons. The Journal of Neuroscience 31: 14272–14283.

9. Smith MA, Kohn A (2008) Spatial and temporal scales of neuronal correlation in primary visual cortex. J Neurosci 28: 12591–12603.

10. Denman DJ, Contreras D (2013) The structure of pairwise correlation in mouse primary visual cortex reveals functional organization in the absence of an orientation map. Cereb Cortex .

11. Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, et al. (2011) Functional specificity of local synaptic connections in neocortical networks. Nature 473: 87–91.

12. Hansen BJ, Chelaru MI, Dragoi V (2012) Correlated variability in laminar cortical circuits. Neuron 76: 590–602.

13. Smith MA, Jia X, Zandvakili A, Kohn A (2013) Laminar dependence of neuronal correlations in visual cortex. Journal of neurophysiology 109: 940–947.

14. Golshani P, Gonçalves JT, Khoshkhoo S, Mostany R, Smirnakis S, et al. (2009) Internally mediated developmental desynchronization of neocortical network activity. J Neurosci 29: 10890-9.

15. Goard M, Dan Y (2009) Basal forebrain activation enhances cortical coding of natural scenes. Nat Neurosci 12: 1444-9.

16. Kohn A, Zandvakili A, Smith MA (2009) Correlations and brain states: from electrophysiology to functional imaging. Curr Opin Neurobiol 19: 434-8.

17. Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, et al. (2010) Decorrelated neuronal firing in cortical microcircuits. Science 327: 584-587.

18. Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes: II. simultaneous spike trains. Biophysical journal 7: 419–440.

19. Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18: 3870-96.

20. Salinas E, Sejnowski TJ (2001) Correlated neuronal activity and the flow of neural information. Nature Reviews Neuroscience 2: 539–550.

21. Ostojic S, Brunel N, Hakim V (2009) How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. The Journal of Neuroscience 29: 10234–10253.

22. Rosenbaum R, Josić K (2011) Mechanisms that modulate the transfer of spiking correlations. Neural computation 23: 1261–1305.

23. Yu BM, Cunningham JP, Santhanam G, Ryu SI, Shenoy KV, et al. (2009) Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. Journal of neurophysiology 102: 614–635.

24. Koller D, Friedman N (2009) Probabilistic graphical models: principles and techniques. MIT press.

25. Loh PL, Wainwright MJ (2012) Structure estimation for discrete graphical models: Generalized covariance matrices and their inverses. arXiv preprint arXiv:12120478 .

26. Schäfer J, Strimmer K, et al. (2005) A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Statistical applications in genetics and molecular biology 4: 32.

27. Peng J, Wang P, Zhou N, Zhu J (2009) Partial correlation estimation by joint sparse regression models. Journal of the American Statistical Association 104.

28. Varoquaux G, Gramfort A, Poline JB, Thirion B (2012) Markov models for fmri correlation structure: is brain functional connectivity small world, or decomposable into networks? Journal of Physiology-Paris 106: 212–221.

29. Ryali S, Chen T, Supekar K, Menon V (2012) Estimation of functional connectivity in fmri data using stability selection-based sparse partial correlation with elastic net penalty. Neuroimage 59: 3852–3861.

30. Ledoit O, Wolf M (2004) A well-conditioned estimator for large-dimensional covariance matrices. Journal of multivariate analysis 88: 365–411.

31. Bickel PJ, Li B, Tsybakov AB, van de Geer SA, Yu B, et al. (2006) Regularization in statistics. Test 15: 271–344.

32. Dempster A (1972) Covariance selection. Biometrics : 157–175.

33. Chandrasekaran V, Parrilo PA, Willsky AS (2010) Latent variable graphical model selection via convex optimization. In: Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on. IEEE, pp. 1610–1613.

34. Ma S, Xue L, Zou H (2013) Alternating direction methods for latent variable gaussian graphical model selection. Neural computation : 1–27.

35. James W, Stein C (1961) Estimation with quadratic loss. In: Proceedings of the fourth Berkeley symposium on mathematical statistics and probability. volume 1, pp. 361–379.

36. Fan J, Fan Y, Lv J (2008) High dimensional covariance matrix estimation using a factor model. Journal of Econometrics 147: 186–197.

37. Reddy GD, Saggau P (2005) Fast three-dimensional laser scanning scheme using acousto-optic deflectors. J Biomed Opt 10: 064038.

38. Katona G, Szalay G, Maák P, Kaszás A, Veress M, et al. (2012) Fast two-photon in vivo imaging with three-dimensional random-access scanning in large tissue volumes. Nat Methods .

39. Cotton RJ, Froudarakis E, Storer P, Saggau P, Tolias AS (2013) Three-dimensional mapping of microcircuit correlation structure. Frontiers in Neural Circuits 7: 151.

40. Vogelstein JT, Packer AM, Machado TA, Sippy T, Babadi B, et al. (2010) Fast nonnegative deconvolution for spike train inference from population calcium imaging. Journal of neurophysiology 104: 3691–3704.

41. Malmersjö S, Rebellato P, Smedler E, Planert H, Kanatani S, et al. (2013) Neural progenitors organize in small-world networks to promote cell proliferation. Proceedings of the National Academy of Sciences 110: E1524–E1532.

Figure Legends

Figure 1. Illustration of regularized estimation of partial correlations. A. The sample correlation matrix of unprocessed somatic calcium signals from a population of cells in mouse visual cortex. The outlined square fragment is magnified in **B**. **C**. The same fragment of another estimate of the correlation matrix regularized to yield sparse partial correlations. Corresponding fragments of partial correlations matrices of the unregularized and regularized estimated are shown in **D** and **E**, respectively.

Figure 2. Estimators whose low-dimensional regularization targets can represent the structure of the true covariance matrix outperform other estimators. Row 1. Graphical representations of four types of low-dimensional structures of interactions between observed neurons (green spheres) and latent units (light-shaded spheres). In the 'independent model' (**A**), observed neurons exert no linear effects on one another neither directly nor through interactions with common latent units. In 'latent factors' (**B**), the correlated activity of observed cells is driven by several latent units. In 'sparse interactions' (**C**), the correlation matrix is defined by a set of linear interactions between observed neurons. In 'sparse+latent' (**D**), correlations arise through direct linear interactions between some pairs of observed neurons and through interactions with common latent units. **Row 2.** Examples of 50×50 correlation matrices corresponding to each type of low-dimensional structure. The factor model (**B**) has three latent units. The partial correlation matrix of the sparse model (**C**) is 73% sparse. The 'sparse+latent' matrix has one latent unit and its direct interactions are 78% sparse. **Row 3.** Examples of sample correlation matrices calculated from samples of 1000 observations taken from simulated random processes with corresponding correlation matrices from row 2. **Row 4.** Mean *excess loss* (Eq. 8) attained by each of the five estimators as a function of sample size. The error bars indicate the standard error of the mean based on 30 samples. **Row 5.** Mean *cross-validation loss* (Eq. 10) of covariance estimators with respect to the matching estimator . The values are relative to the validation loss of the estimator that matches the low-dimensional structure of the true covariance matrix. The error bars indicate the standard error of the mean based on 30 samples.

Figure 3. Acquisition of neural signals for the estimation of noise correlations. Visual stimuli comprising full-field drifting gratings interleaved with blank screens (**A**) were presented to anesthetized mice while two-photon recordings of somatic calcium signals were collected using fast 3D random-access microscopy (**B**). The visual stimulus included an initial period with 16 directions of motion for orientation tuning followed by a longer (15–20 min) period of stimulation with only 2 or 5 directions of motion for the computation of the noise correlation matrix. **C.** Representative calcium signals from eight cells out of 298 cells downsampled to 20 Hz. The inferred firing rate binned in 150 ms intervals are indicated by red ticks below each trace. **D.** The raster plot of the inferred firing rates, binned in 150 ms intervals, from the entire population from the first (left) and last (right) minute of the entire recording. The traces from **C** are highlighted in red. **E.** The spatial arrangement and orientation tuning of the 298 cells from the imaged site. **F.** The noise correlation matrix of the activity of the neural population. **G.** The histogram of the noise correlation coefficients with the mean indicated by the red line.

Figure 4. The sparse+lowrank estimator $C_{\mathsf{sparse+lowrank}}$ outperforms the other estimators on neural data. A–D. Histograms of average cross-validation loss differences of the respective estimators C_0, C_{diag}, C_{factor}, and C_{sparse} from $C_{\mathsf{sparse+lowrank}}$. The histograms are based on 31 imaged sites in 24 mice. All medians (red dashed lines) were significantly greater than zero, indicating the dominance of $C_{\mathsf{sparse+lowrank}}$ over the other estimators. The green arrows indicate the results for the site shown in Fig. 3 and Fig. 5

Figure 5. Example of low-dimensional correlation structure revealed by the sparse+low-rank estimator. A. The regularized estimate of the correlation matrix (top-right) closely approximates the sample correlation matrix (bottom left). This close approximation is also demonstrated by the scatter plot of the correlation coefficients produced by the two estimates (**D**). However, the partial correlation matrices from the two estimate show more pronounced differences (**B** and **E**). **C.** Furthermore, the partial correlation matrix of the regularized estimate is decomposed into a sparse component with 82.2% off-diagonal zeros (bottom-left) and low-rank component of rank 15 (top-right). **F.** The sparse component of the regularized partial correlation matrix had little resemblance to the sample correlations. The gray interval indicates the range of correlations containing 82.2% of cells pairs, equal to the fraction of zeros in the sparse partial correlation matrix. This interval contained 58.9% of the partial correlations. **G.** The graphical depiction of the positive (green) and negative (magenta) partial correlations as edges between observed neurons. The line density is proportional to the magnitude of the correlation. **H.** A subset of neurons from the center of the cluster shown in **G** showing the regularized partial correlations. **I.** The same subset with sample correlations thresholded to match the sparsity of the regularized interactions.

Figure 6. Properties of sparse+low-rank regularized estimates from all imaged sites A. The average sample correlations vs. average partial correlations for each imaged site. In each plot, the red asterisk indicates the site shown in figures 3 and 5. **B.** The average node degree for sparse partial correlations vs. population size in each imaged site. **C.** The number of inferred latent units vs. population size in each imaged site. **D.** The number latent units vs. average node degree for sparse partial correlations for each site.

Figure 7. Dependence of correlations and partial correlations on orientation tuning differences and physical distance between cell pairs. A–C. Average sample correlations (black) and regularized partial correlations (red) between pairs of cells in the example site shown in previous figures. The correlations were normalized by the respective average correlations (6 A). **A.** Average correlations between pairs of neurons tuned to orientation with differences in preferred orientation in the intervals of 0–15°, 15–45° and 45–90°. **B.** Average correlations between pairs of neurons located at the same depth ($\pm 25\mu$m) separated by lateral distances in the intervals of 0–25 μm, 25–75 μm, 75–150 μm, and 150+ μm. **C.** Average correlations between pairs of neurons displaced laterally by less than 25 μm separated in depth by distances in the intervals of 0-25 μm, 25–60 μm, and 60+ μm. **D–F.** Same measurements as **A**–**D** averaged across multiple sites. Only sites that had at least 20 qualifying cell pairs in each of the intervals were included in the averages. The error bars indicate the standard error of the mean (often too small to be seen). **G–I.** Normalized connectivity of positive (green) and negative (dark red) interactions from the sparse component of the regularized partial correlations in the example site show in previous figures. Normalized connectivity was computed as the fraction of pairs connected by interactions of corresponding signs in each interval divided by fraction of non-zero interactions across the entire site. The intervals are identical to those in **A**–**C**. **J–L.** Same measurements as in **G**–**I** averaged across multiple sites. The error bars indicate the error of the mean. Only sites that had at least 20 qualifying pairs in each interval were included in the averages.