ELASTIC SOURCE SELECTION FOR IN VIVO IMAGING OF NEURONAL ENSEMBLES

Logan Grosenick[1], *Todd Anderson*[2], *and Stephen J. Smith*[1]

[1]Neurosciences Program, Stanford University,
Stanford, CA, USA

[2]Biology, Stanford University,
Stanford, CA, USA

ABSTRACT

Advances in microscopy and biochemistry now allow investigators to image the calcium dynamics of hundreds to thousands of neurons in awake behaving animals. However, as speed and resolution of such techniques rapidly increase, so do the dimension and complexity of the data collected. ICA has been widely employed to reveal independent non-Gaussian sources underlying large data sets consisting of mixed sources. We apply a recently developed sparse regression method, the Elastic Net (ENET), to the columns of the mixing matrix of a Independent Component Analysis (ICA) procedure. This regression method automatically selects only those columns of the mixing matrix relevant to a dependent variable of interest. Further, because ICA is a linear operator, we can easily project the "relevance filtered" data back into the native data space for interpretation. We demonstrate the utility of this method on 3D calcium imaging data collected from the optic tectum of an awake behaving larval zebrafish watching a prey-like stimulus.

Keywords: Neuroimaging, Calcium Imaging, ICA, Sparse Regression, In Vivo

I. INTRODUCTION

Recently, the ability to image calcium signals associated with neural activity has provided a means of interrogating populations of neurons in living animals. Various techniques for "bulk-loading" neural cells using AM ester calcium indicators have provided opportunities for imaging activity patterns *in vivo* across networks of cells in intact neural tissues [6]. Most of these studies to date have employed two-dimensional confocal or two-photon laser scanning methods to maximize depth discrimination and to minimize interfering effects of tissue light scattering. However, these methods entail unwelcome limits to volume acquisition rates.

In the current study we employ a new technology: Light Field Microscopy (LFM) [5] to rapidly acquire dynamic 3D images from the brain of a transparent larval zebrafish. While this method is currently limited to transparent animals, developing techniques in 3D microscopy promise to improve volume imaging in turbid tissues as well. Such technologies are already revolutionizing the way neuroscientists collect data, allowing the acquisition of megavoxels in milliseconds.

With the ability to quickly acquire these large data sets comes the need for signal processing approaches able to extract meaningful structure from such correlated, high-dimensional data. In previous calcium imaging studies, analyses have been largely limited to inspecting averaged time courses [7] or choosing peak responses of a particular neuron to one stimulus versus another [8]. Here we combine modern unsupervised (ICA [4]) and supervised (ENET [10]) machine learning methods to first automatically extract single-neuron sources and their time courses, and then to automatically select sources of interest. In the latter step the elastic net is used to select only those sources that predict the value of a target variable, in our example the location of a prey-like visual stimulus presented to the larval zebrafish.

Because we are able to perform this variable selection on the columns of the ICA mixing matrix rather than on the sources themselves, our approach also results in a massive reduction in problem dimensionality (and thus in computational cost). In addition to reduced computational burden, this reduction in input space will allow future work to use non-linear models that project into higher dimensional subspaces or work on expanded basis sets. As both methods we employ are linear, we are able to project the relevance-cleaned data back into image space to view a volume-rendered movie of only that part of the data related by our model to the external variable of interest.

II. EXPERIMENTAL PROCEDURE

II-A. Light Field Microscopy

The type of microscope used this study is unique in that it captures a 3D volume at every exposure. By inserting an array of thousands of microlenslets at the intermediate image plane of the microscope, and positioning a CCD sensor at the image plane formed by this array, we capture the four-dimensional "light-field" (radiance along rays as a function of position and direction in space) [5]. Using synthetic

aperture photography, we convert this light-field into a stack of focal slices, in this case with 25 slices 3 μm apart, and 72 μm range along the optical axis (Z). This volume represents the intensity of the specimen at a single instant in time.

II-B. Calcium Imaging

Following anesthetization in 0.02% MS-222 and embedding in 1.5% low melting point agarose (LMA), the tectal neuropil of a larval zebrafish (7 days post-fertilization) was injected with calcium indicator Oregon Green BAPTA-1 AM ester dye dissolved in DMSO / 20% pluronic solution. Over the next hour the dye was absorbed by tectal neurons as the fish swam freely in an incubator. Next the fish was again immobilized in LMA (unanesthetized) and positioned in the stimulation chamber such that the eye opposite the injected tectum directly faced the OLED screen. Light-field images were acquired at 4 frames per second through a 63x 0.9 NA water-dipping objective using a Retiga 4000R camera attached to a custom-built microlens array and mount.

II-C. Visual Stimulation

Visual stimuli were delivered using a 0.96" diagonal organic light-emitting diode screen (uOLED96-G1) mounted on the side of the chamber and driven over the serial port by custom software. Stimulation consisted of a small dot (subtending $\sim$ 2 degrees on the fish retina) moving along a circular path at 90 angular degrees per second. Stimulus position was triggered by the camera shutter.

III. METHODS

III-A. Spatial ICA

Given an $n \times p$ data matrix X consisting of n rows containing n observations in time on p variables in space (e.g., voxels in the case of a volume), ICA represents the data as a linear combination of maximally statistically independent sources S and a "mixing matrix" A. so

$$X = AS \tag{1}$$

Here, the rows of S are static spatial components, and the columns of A the time courses associated with each of these static sources. While the spatial sources have been chosen to be nearly statistically independent, the rows of the mixing matrix might be very correlated (i.e. sources may be spatially independent but temporally correlated).

III-B. Automatic variable selection, L1-regression (LASSO), and the Elastic Net

In principle, automatic variable selection should improve both model parsimony and interpretability. In practice, investigators must take care to ensure the stability of coefficients across samples when fitting correlated data. Traditionally, penalization (or "regularization") has provided an effective means of stabilizing correlated coefficients. More recently, related models such as the Least Absolute Shrinkage and Selection Operator (LASSO) have extended penalized regression to include automatic variable selection, which sets irrelevant coefficients to exactly zero (allowing researchers to remove the corresponsing variables from their model).

In general, the penalized linear regression models we consider have coefficient estimates given by:

$$\hat{\beta} = \arg\min_{\beta} ||y - Z\beta||_2^2 + \lambda J(\beta) \tag{2}$$

where y is a real-valued vector of outputs (dependent variables), Z is our input variable matrix (with observations in the rows), β is the vector of coefficients with jth entry $\beta_j; j \in \{1, ..., p\}$ and $\hat{\beta}$ the corresponsing estimates, the function $J(\beta)$ is some penalty function in terms of β, λ is a penalty parameter, and $||\beta||_2$ is the L_2 norm of β.

The LASSO [9] uses the penalty function $J(\beta) = ||\beta||_1$ in equation (2) above, where $||\beta||_1 = \sum_j |\beta_j|$ is the L_1 norm of β. When the number of non-zero coefficients in the model is expected to be sparse (e.g., N for $p >> N$), the LASSO provides an attractive alternative to other regression methods, as it performs simultaneous variable subset selection and prediction, and its entire penalty path is easily computed using the LARS algorithm [1]. The LASSO has also been shown to perform well at prediction in many problems, competing favorably with ridge regression (where $J(B) = ||\beta||_2^2$) and the more general "bridge regression" (where $J(\beta) = \sum_j |\beta_j|^\gamma$, for $0 \leq \gamma \leq 2$) [2].

Although the LASSO performs well in variable selection and prediction, it also has limitations, particularly given multicollinear input variables [10]. Given a group of highly correlated input variables, the LASSO is likely to arbitrarily select just one variable from the group, generating potentially unstable coefficients over different samples of the same variables and failing to include correlated groups of relevant variables [10]. Model performance also suffers given correlated inputs. For instance, ridge regression empirically dominates the LASSO in typical $N > p$ regression settings when the input variables are correlated [9].

Because of these problems, the LASSO is likely not best suited for application to the columns of our ICA mixing matrix, which are potentially very correlated. However, a generalization of the LASSO called the elastic net (ENET) addresses correlations between inputs by implementing a hybrid penalty with both ridge (L_2) and LASSO (L_1) properties [10]. ENET coefficient estimates are given by

$$\hat{\beta}^{ENET} = (1 + \lambda_2) \arg\min_{\beta} ||y - ZB||_2^2 + \lambda_1 ||\beta_1|| + \lambda_2 ||\beta||_2^2 \tag{3}$$

and thus implement a hybrid penalty involving two penalty parameters λ_1 and λ_2 , with the former essentially modulating automatic variable selection while the latter allows

relevant but correlated variables to remain together in the model. The estimates $\hat{\beta}^{ENET}$ can be rewritten as

$$\hat{\beta}^{ENET} = \arg\min_{\beta} \beta^T \left(\frac{Z^T Z + \lambda_2 I}{\sqrt{1+\lambda_2}} \right) \beta - 2y^T Z\beta + \lambda_1 ||\beta||_1 \quad (4)$$

where standard LASSO estimates obtain when $\lambda_2 = 0$ so

$$\hat{\beta}^{LASSO} = \arg\min_{\beta} \beta^T Z^T Z \beta - 2y^T Z\beta + \lambda_1 ||\beta||_1 \quad (5)$$

From (4) and (5) we see that ENET represents a stabilized version of the LASSO, in which the estimate covariance matrix $\hat{\Sigma} = Z^T Z$ is shrunk towards the $p \times p$ identity matrix I as λ_2 increases, and the stabilized sample covariance matrix in equation (4) can be written

$$\frac{Z^T Z + \lambda_2 I}{\sqrt{1+\lambda_2}} = (1-\gamma)\hat{\Sigma} + \gamma I \quad (6)$$

(with $\gamma = \lambda_2/(1+\lambda_2)$) [10].

III-C. Elastic Source Selection (ESS) for neuroimaging

A priori knowledge or inspection of the singular values of the input data may be used to choose a reasonable number of independent components N_{IC} (we chose N_{IC} = 300 using the latter approach). Applying ICA to raw input data matrix X yields the mixing and source matrices A and S as described in equation (1), where A is $n \times N_{IC}$ and S is $N_{IC} \times p$. It is important to note that in most imaging applications $n >> N_{IC}$ (e.g., in our application $n = 375,150$ while $N_{IC} = 300$). As the spatial distributions of the neural sources in S are static, the dynamics of the process are captured by the columns of A. Thus if we wish to relate our sources to a time-varying variable y, we can do so by regressing y on A, which results in a massive reduction in input dimension as compared to regressing y on X.

We used the fastICA algorithm [4] to estimate A and S, and the LARS-EN algorithm [10] to fit the elastic net solutions with A as the input matrix (so letting $Z = A$) in (3)). As is standard for such models [3], inputs were centered and standardized prior to entry into the model. The LARS-EN yields the entire λ_1 path solution in approximately the time of an ordinary least-squares fit. To choose parameters (λ_1, λ_2) we therefore found λ_1 path solutions for different λ_2 values (ranging from 10^{-5} to 10^{10} and including 0 for the LASSO solution) and chose the most parsimonious model (smallest λ_1) within 10% of the minimum Bayesian information criterion (BIC) over this grid of possible values. While there are a variety of commonly used tuning parameter selection methods (e.g. cross-validation, generalized cross-validation, AIC and BIC) it has been shown that for variable selection efficiency BIC is a better tuning parameter selector than cross-validation, GCV and AIC, which tend to overfit [3]. To estimate BIC, we used the trace of the hat matrix of the active set (variables with non-zero coefficients) as an estimate of the degrees of freedom for each (λ_1, λ_2) [3].

To estimate our extra-sample error, we used 10-fold cross validation, fitting the ENET parameters as described above using 9/10 of the data and holding out the remining 1/10 to test the fit of the chosen model. This yielded 10 models and fits, one for each validation step. As a final model, we took the median over the 10 fits of elastic source selection (ESS) coefficient estimates $\hat{\beta}^{ESS}$. Since regressions using an L_1 penalty on the coefficients are equivalent to maximum a posteriori (MAP) estimators of the mode with Laplacian priors, the median (rather than the mean) provided a good estimation of central tendency for cross-validated coefficients.

Finally, once we had chosen our final model coefficients $\hat{\beta}^{ESS}$, we set to zero vectors the columns of the mixing matrix corresponding to zero coefficients to yield a new matrix A^* such that the jth column a_j^* of A^* satisfies

$$a_j^* = \begin{cases} a_j & \text{if } \hat{\beta}_j^{ESS} \neq 0 \\ 0 & \text{if } \hat{\beta}_j^{ESS} = 0 \end{cases}$$

Then our relevance cleaned data was reconstructed as

$$X^* = A^* S \quad (7)$$

IV. RESULTS

Applying ICA as in (1) to our $3000 \times 375,150$ data matrix X yielded a 3000 × 300 mixing matrix A and a $300 \times 375,150$ matrix S. Some of the resulting sources (rows of S) and their corresponding time courses (columns of A) are shown in Figure 1. Note that ICA extracts spatial components that are single neuronal cell bodies from the volume (the mean volume is shown above for comparison). We found ICA performed well at extracting such single-neuron sources and their time courses. Following ICA, we applied the elastic net (3) using A as the input matrix to predict two stimulus dimensions (y in (3)): the x- and y-axis positions of the moving stimulus presented to the larval zebrafish. The final model resulting from taking the median (absolute) coefficient values across the 10-fold cross-validation chose 23 neuronal sources, two of which are shown in Figure 1.

In Figure 2, the black lines show the ENET prediction of the y-coordinate of stimulus location overlayed on the true y-coordinate value (blue lines) (x-coordinate results are similar; we show the first and last 50 seconds of the 4.5 minute trial). These are predicted only from neural activity in the zebrafish tectum. We see that the ENET does a reasonably good job of predicting the oscillations in the stimulus coordinates on held-out data–although as it does not take into account the smoothness of the target variables, the solutions are somewhat rough. Still, the quality of the fit suggests that the underlying model is good, and thus that the sources selected are appropriate. In future work we will add a temporal smoothness penalty to the model.

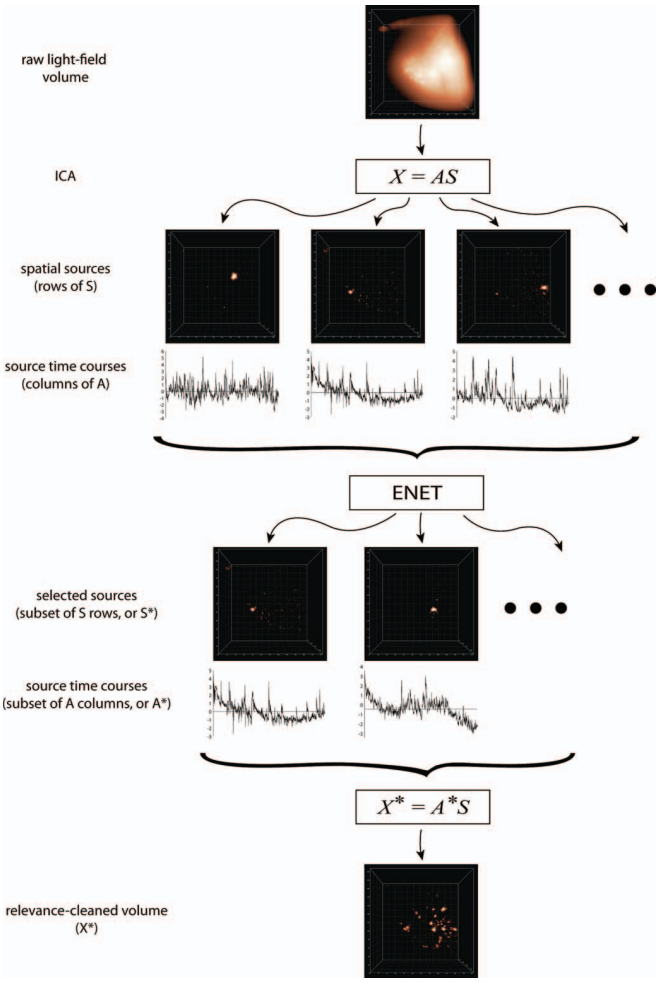

Figure 1: The raw LF volume is reshaped into matrix X, in which each row is a volume. ICA factors X into a "mixing" matrix A and a "source" matrix S. Volume renderings of individual rows of S are shown as "example spatial sources", with the corresponding time series (column of A) below. ENET automatically selects a relevant subset of the A columns given a target variable (stimulus position). We can project "relevance cleaned" data back into the native data space by zeroing out the columns of A not selected by our model (yielding A^*) and using this matrix to transform relevant sources back into image space (X^*).

V. DISCUSSION AND CONCLUSION

In this paper we develop a method for extracting neuronal sources and their time courses from volumes of *in vivo* calcium imaging data, and then use these neural signals to predict the position of a visually presented stimulus. Overall, we found that applying our method to data from the zebrafish optic tectum: (1) identified individual neurons as sources, (2) successfully chose a subset of these neurons that predicted the location of a prey-like stimulus shown to the zebrafish, and (3) allowed relevance filtering of the data, that is, projection of a subset of the data related to a variable of interest back into the native data space for interpretation

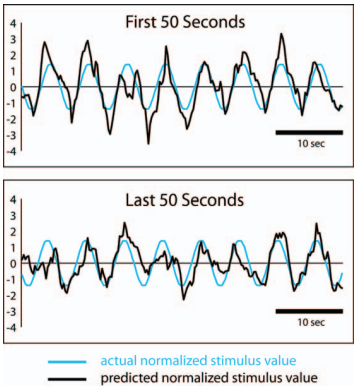

Figure 2: Plots showing the actual stimulus position (blue line) and the position predicted by the ESS model on held-out calcium imaging data from the first and last 50 seconds of a trial (black line).

and further analysis. This short paper shows our method to be a promising approach to extracting single-neuron signals from large volumetric imaging data, although further studies on additional data will be needed to support this result.

VI. REFERENCES

[1] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshitani. Least angle regression. *Annals of Statistics*, 32(2):407–499, 2004.

[2] Ildiko E. Frank and Jerome H. Friedman. A statistical view of some chemometrics regression tools. *Chemometrics*, 35(2):109–135, 1993.

[3] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, New York, Second edition, 2009.

[4] Aapo Hyvarinen, Juha Karhunen, and Erkki Oja. Independent component analysis. May, 2001.

[5] Marc Levoy, Ren Ng, Andrew Adams, Matthew Footer, and Mark Horowitz. Light field microscopy. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pages 924–934, New York, NY, USA, 2006. ACM.

[6] Cristopher Niell and Stephen J. Smith. Functional imaging reveals rapid development of visual response properties in the zebrafish tectum. *Neuron*, 45(6):941–51, 2005.

[7] Pavan Ramdya and Florian Engert. Emergence of binocular functional properties in a monocular neural circuit. *Nat Neurosci*, 11(9):1083–90, 2008.

[8] German Sumbre, Akira Muto, Herwig Baier, and Mu-Ming Poo. Entrained rythmic activities of neuronal ensembles as perceptual memory of time interval. *Nature*, 456:102–106, 2008.

[9] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B*, 58(1):267–288, 1996.

[10] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B*, 67(2):301–320, 2005.