STOCHASTIC FILTERING OF TWO-PHOTON IMAGING USING REWEIGHTED ℓ_1

Adam S. Charles[1]*, Member, IEEE , Alexander Song*[2] *, Sue Ann Koay*[1] *,*
David W. Tank[1,3,4] *and Jonathan W. Pillow*[1,5]
[1]Princeton Neuroscience Institute, [2]Department of Physics, [3]Bezos Center for Neural Circuit Dynamics,
[4]Department of Molecular Biology, [5]Department of Psychology
Princeton University, Princeton, NJ , 08540 USA
adamsc,as31,koay,dwtank,pillow@princeton.edu

ABSTRACT

Two-photon (TP) calcium imaging is an important imaging modality in neuroscience, allowing for large-scale recording of neural activity in awake, behaving animals at behavior-relevant timescales. Interpretation of TP data requires the accurate extraction of temporal neural activity traces, which can be accomplished via manual or automated methods. In this work we seek to improve the accuracy of both manual and automated TP microscopy demixing methods by introducing a denoising algorithm based on a statistical model of TP data which includes spatial contiguity, sparse activity and Poisson observations. Our method leverages recent developments in stochastic filtering of structured signals based on Laplacian-scale mixture models (LSMs) to model the neural activity in TP data as a set of spatially correlated sparse variables. We apply our method on TP images taken from the visual cortex of an awake, behaving mouse, and demonstrate improved neural activity demixing over current preprocessing techniques.

***Index Terms*—** Two-photon microscopy, Stochastic filtering, Reweighted ℓ_1

1. INTRODUCTION

Simultaneous recordings of neural populations at behavior-relevant time-scales are vital for studying the functionality and activity of neural systems. Two-photon (TP) microscopy offers the ability to record the activity of hundreds of neurons over relatively large field-of-views (FOVs) at neurally relevant time scales (e.g. 500 μm x 500 μm areas at 30 Hz frame-rates). TP microscopy accomplishes this task by raster-scanning a thin slice of tissue in which fluorescent proteins, which interact with calcium ions [1], have been introduced (virally or through transgenic animals [2]). When a cell fires, the increased calcium levels in the cell activates the proteins and the response to illumination is increased over the spatial region occupied by the cell (the cell's spatial profile). The resulting TP movie encodes temporal activity-related calcium traces over the entire FOV. To fully utilize TP data, the neural activity time-traces of all the cells in the FOV need to be extracted as accurately as possible, a process that can be complicated in a number of ways. First, the spatial profiles for neurons may overlap. Second, the illumination in TP microscopes can result in complex noise statistics. Finally the rise in intensity of the cell is spread over both space and time, resulting in a lower per-pixel signal-to-noise ratio (SNR).

A number of algorithms currently exist to extract the neural activity traces from TP movies. These methods include both temporal deconvolution methods that rely on hand-selected spatial profiles for each neuron in the FOV [3–7] or fully automated methods that factor the entire video sequence into spatial profiles and time traces [8–11]. While these algorithms have been broadly utilized, often pre-processing in the form of spatial or temporal averaging or down-sampling is used to raise the SNR for efficient demixing (manual or automated). Furthermore, newer TP techniques spread the illumination energy to capture entire neural volumes at the cost of reduced SNR [12]. To make the best use of such data and extract the most of the underlying neural signals, known structure in the signal statistics needs to be leveraged. Specifically, the spatial contiguity of neural profiles in the FOV, the spiking nature of neural time traces, and known statistics of the measurement process should all be used to extract the signal-specific fluctuations from TP movies.

One recent model that has the capability of incorporating all of these required statistics is the reweighted ℓ_1 spatial filtering (RWL1-SF) model. This model uses a Laplacian-scale mixture(LSM) to model the temporal fluorescence variable as a set of sparse and spatially interdependent variables. While this model was originally described for observations contaminated by Gaussian noise, we modify the RWL1-SF to accommodate Poisson observations, which can better approximate TP data in low-light regimes. We demonstrate in this work that the resulting modified RWL1-SF algorithm can extract signal-relevant activity from TP data in order to improve both manual and automated demixing.

2. BACKGROUND

2.1. Demixing Calcium Imaging

Standard methods of TP data analysis (e.g. [8,9]) model the observed fluorescence movie $\boldsymbol{Y}$ as a linear mixture of the individual neural profiles (the spatial region in the image where a given neuron's activity is observed) and temporal activities of the cells as a function of time,

$$\boldsymbol{Y}(t) = \sum_k \boldsymbol{X}_k s_k(t) + \boldsymbol{B}(t) + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{X}_k$ are the neural spatial profiles, $\boldsymbol{s}_k(t)$ are the corresponding temporal activity traces, $\boldsymbol{B}(t)$ is the potentially time-varying background, and $\boldsymbol{\epsilon}$ is often modeled as white-Gaussian measurement noise. The time traces $\boldsymbol{s}_k$ for each neuron are particularly important

This work was supported in part by NIH grants R01MH083868 and U01NS09054, the Simons Collaboration on the Global Brain. A.S.C. was supported by an NIH NRSA Training Grant in Quantitative Neuroscience (T32MH065214). J.W.P was supported by grants from the McKnight Foundation, Simons Collaboration on the Global Brain (SCGB AWD1004351) and NSF CAREER Award (IIS-1150186).

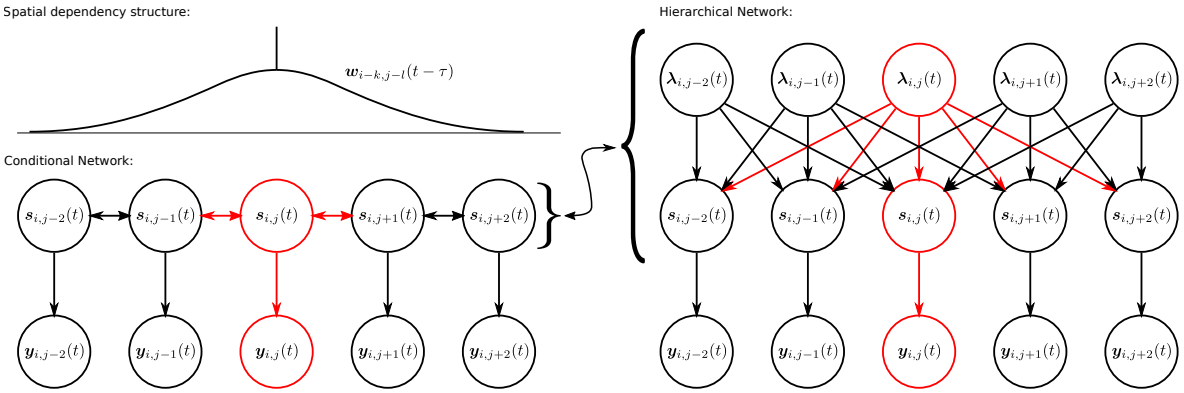

Fig. 1. Graphical model of the LSM for TP microscopy denoising. Left: Bayesian graphical model depicting latent activity parameters $\boldsymbol{s}_{i,j}(t)$ that generate the observed fluorescence variables $\boldsymbol{y}_{i,j}(t)$. Note that the model contains a Markov-filed like dependency structure (i.e. neighboring variables are dependent on each other). The weights $\boldsymbol{w}_{i,j}(t)$ shown above the graphical model depict the strength of the correlation as a function of distance. Right: An equivalent hierarchical model that decouples the activity parameters by introducing a set of hyper-parameters $\lambda_{i,j}(t)$. This hierarchical model can be used in an EM algorithm to perform inference over the activity parameters.

as they encode the activity-related calcium levels of the cell, a quantity closely related to the spiking activity of the cell. The time traces are often modeled as a convolution of the calcium concentration of the cell $\boldsymbol{c}_k$ with the impulse response of the fluorescing protein (i.e. the binding/unbinding dynamics) $\boldsymbol{h}$,

$$s_k(t) = (h * c_k)(t). \tag{2}$$

The per-pixel time traces $s_{i,j}$ can then be described as a linear combination of the per-neuron time traces s_k

$$s_{i,j}(t) = \sum_k X_{k,i,j} s_k(t), \tag{3}$$

where $X_{k,i,j}$ is the $i.j$ pixel of the k^{th} spatial profile.

To extract the time traces $c(t)$ of each neuron, a number of methods have been developed. These methods predominantly fall into two categories: methods that infer detailed information about the spiking times of a cell given a manual specification of the spatial profiles $\boldsymbol{X}$ [3–7] and methods that perform full demixing, uncovering both $\boldsymbol{X}$ and $\boldsymbol{s}$ simultaneously [8–11].

The methods that focus on temporal deconvolution rely on manually identified spatial profile to decide which pixels to average to obtain a higher-SNR time trace of that cell's activity. The resulting time traces are then deconvolved with an estimate of the protein impulse function $\boldsymbol{h}$ to obtain an estimate of the calcium trace $\boldsymbol{c}$. Many of these methods also take an additional step to infer spike times from the activity-dependent calcium trace. Methods along thse lines, for example, leverage Bayesian models [13] or Viterbi estimation under a biophysical model [5].

The methods that focus on full demixing take a more fully automated approach and infer the spatial profiles alongside the temporal activity. The most commonly used of these demixing methods, PCA/ICA [8], takes a two step approach to this problem. First, PCA is used to reduce the dimensionality of the data (essentially looking for the number of active profiles), and then uses ICA to allow non-orthogonal profiles and correlated time traces. More recent techniques take a similar dimensionality-reduction matrix factorization approach, using non-negative matrix factorization (NMF) [9], NMF incorporating additional sparsity and locality statistics [10], and dictionary learning techniques [11].

Methods in both these categories are often not applied to raw fluorescence data, but rather to pre-processed data that has been subject to spatial and/or temporal filtering. These steps are predominantly in place to improve the SNR of the data. Instead of pre-processing the data, we instead propose using a more detailed model which captures some of the key features of neural cell recordings: spatially contiguous firing, sparse temporal activity, and non-Gaussian observations.

2.2. Sparse Stochastic Filtering

Recent developments in sparse signal estimation have produced models that can characterize complex high-dimensional statistics using relatively simple inference algorithms. These stochastic filtering methods aim to infer sparse latent variables using very general local-consistency regularizers. Similar in spirit to the Gaussian Process literature, these models, based on Laplacian-scale mixture (LSM) [14, 15] impose spatial interdependencies via second-order statistics. One particular instantiation of this model, reweighted ℓ_1 spatial filtering (RWL1-SF) [15] solves for coupled sparse signals by alternating between solving a weighted LASSO optimization, and updating the LASSO weights based on the spatially neighboring signals. The update step allows neighboring signals to influence each other, instilling the desired interdependencies between the signals. These methods have shown improved performance for both dynamic filtering of sparse signals [16] and spatial hyperspectral imagery [15]. In this work we further extend this technique, adapting the underlying probabilistic model to the specifics of TP microscopy. Specifically, we introduce a Poisson measurement process to model photon incidence and a constant background term that conveys the static illumination level at a given location.

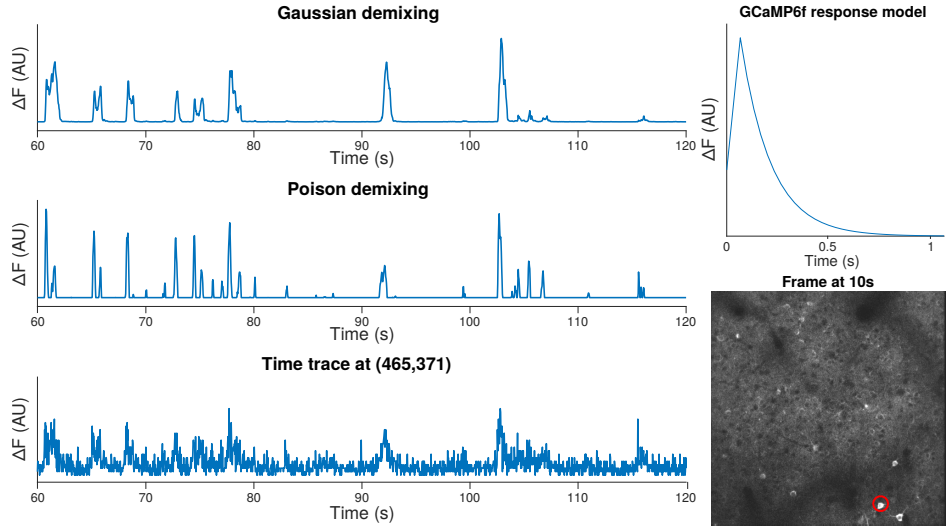

Fig. 2. Example comparison of Gaussian and Poisson deconvolution. Top Right: Model of the GCaMP6f protein impulse response $h(t)$. Bottom right: Single frame of TP microscopy from the mouse visual cortex. Left: Example demixing of a single pixel's time trace (pixel location circled in red in the bottom-right image) using Gaussian- and Poisson-based deconvolution. The overall activity patterns are similar, however the Poisson deconvolution achieves a sparser solution, indicating a better match to the data.

3. STOCHASTIC FILTERING FOR TWO-PHOTON IMAGING

Our method considers a spatio-temporal model of TP imaging based on Poisson modeling of the fluorescence values and the LSM spatial model used in RWL1-SF over the underlying activity. Mathematically, we formulate a forward model (shown graphically in Figure 1) of neuronal spike times influencing the activity at each pixel given the observed fluorescence $y_{i,j}(t)$,

$$\begin{aligned} y_{i,j}(t) &\sim \text{Poisson}((h * s_{i,j})(t) + b_{i,j}) \\ s_{i,j}(t) &\sim \text{Laplace}((w * \lambda)_{i,j}(t)) \\ b_{i,j} &\sim \text{Laplace}((\overline{w} * \gamma)_{i,j}), \end{aligned} \tag{4}$$

where $h(t)$ is the protein activation impulse response, $\lambda_{i,j}$ is the underlying activity-based photon rate for each pixel and time step, $\gamma_{i,j}$ is the underlying photon rate for each background pixel at each time step, and $w_{i,j}(t)$ and $\overline{w}_{i,j}$ are the dependency-inducing kernels indicating the how pixel activity is correlated across an image and across time[1]. These kernels essentially control the spatial extent of potential active regions. For example, if $w_{i,j}(t)$ spans a large spatial region, it is more likely that only small areas of the image will be active. This is because only a small number of hyper-parameters λ will be sufficient to suppress a large portion of the image. Consequently, $w_{i,j}(t)$ should not span a larger area than the largest expected active region (i.e. the radius of a neuron for TP data).

Nearby pixels will be encouraged to have similar activity histories (given the florescence data) via inter-linked Gamma hyper-priors over the rate parameters λ and γ,

$$\begin{aligned} \lambda_{i,j}(t) &\sim \text{Gamma}(\alpha, \theta) \\ \gamma_{i,j} &\sim \text{Gamma}(\overline{\alpha}, \overline{\theta}), \end{aligned} \tag{5}$$

where α, $\overline{\alpha}$, $\theta_{i,j}(t)$ and $\overline{\theta}_{i,j}$ are a-priori chosen parameters that control the moments of the Gamma distribution. To optimize this model and recover the 'clean' underlying traces, we optimize

$$\arg\min_{\boldsymbol{s},\boldsymbol{b}} -\log\left(p(\boldsymbol{s},\boldsymbol{b}|\boldsymbol{y})\right) = -\log\left(p(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{b})p(\boldsymbol{s})p(\boldsymbol{b})\right). \tag{6}$$

The marginal probability $p(\boldsymbol{s}) = \int p(\boldsymbol{s}|\boldsymbol{\lambda})p(\boldsymbol{\lambda})d\boldsymbol{\lambda}$ and $p(\boldsymbol{b}) = \int p(\boldsymbol{b}|\boldsymbol{\gamma})p(\boldsymbol{\gamma})d\boldsymbol{\gamma}$ is analytically computable in this case (resulting in a student-t style distribution [14]), however the resulting negative-log probabilities are not convex and difficult to optimize over in high dimensions. Instead, an expectation-maximization (EM) method can be utilized as in [14, 15, 17, 18] to iteratively calculate

$$\begin{aligned} \{\widehat{\boldsymbol{s}}, \widehat{\boldsymbol{b}}\} &= \arg\min_{\boldsymbol{s},\boldsymbol{b}} -\log(p(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{b})) - \log(p(\boldsymbol{s}|\widehat{\lambda})) - \log(p(\boldsymbol{b}|\boldsymbol{\gamma})) \\ \widehat{\boldsymbol{\lambda}} &= E_{\boldsymbol{\lambda}|\widehat{\boldsymbol{s}}}[\boldsymbol{\lambda}] \\ \widehat{\boldsymbol{\gamma}} &= E_{\boldsymbol{\gamma}|\widehat{\boldsymbol{b}}}[\boldsymbol{\gamma}]. \end{aligned} \tag{7}$$

Because the Gamma distribution is conjugate to the Laplacian distribution, we have closed form expressions for the expectation steps. The full form of the expectation is quite complex, however it is well approximated by the simple analytic expressions

$$\begin{aligned} \widehat{\lambda}_{i,j}(t) &= \frac{\xi}{|s_{i,j}(t)| + \sum_{\{l,m,\tau\}} w_{l,m}(\tau)|s_{l,m}(\tau)| + \beta} \\ \widehat{\gamma}_{i,j} &= \frac{\overline{\xi}}{|b_{i,j}| + \sum_{\{l,m\}} \overline{w}_{l,m}|b_{l,m}| + \beta}. \end{aligned} \tag{8}$$

where β and ξ are functions of α and θ, β is a fixed parameter describing the minimum variance of each λ and γ, and the weights $w_{i,j}(t) > 0$ and $\overline{w}_{i,j} > 0$ describe the correlation structure across the data cube (higher values of $w_{i,j}(t)$ indicate higher correlations at a distance of i, j, t). Since the neighborhood is spatially motivated

[1] We note that by the positivity constrains on s and b, the Laplacian distributions here are actually Exponential distributions, however we keep the general form to demonstrate that this model can be used in other datasets which may not have positivity constraints

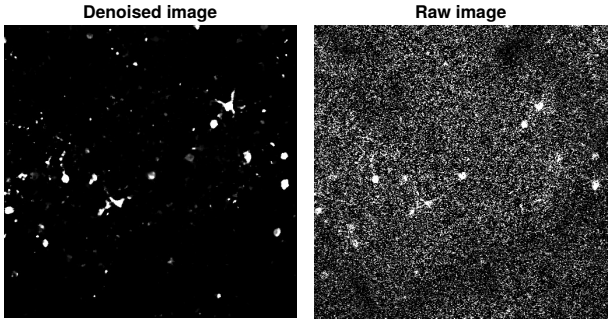

Fig. 3. Example frame of denoised data. Right: Original data frame obscures some of the activity. Left: Denoised frame clearly shows active cell bodies.

and identical at any shift, we can calculate the updates of $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ via 3D convolutions,

$$\begin{aligned} \widehat{\lambda}_{i,j}(t) &= \frac{\xi}{|[K * s]_{i,j}(t)| + \beta} \\ \widehat{\gamma}_{i,j} &= \frac{\overline{\xi}}{|(\overline{K} * b)_{i,j}| + \beta}, \end{aligned} \tag{9}$$

where K is a symmetric kernel with one at the origin and consisting of the weights w away from the origin.

The maximization step is more involved, due to the likelihood of $\boldsymbol{y}$ given $\boldsymbol{s}$ and $\boldsymbol{b}$. For a Gaussian likelihood, we would recover the RWL1-SF algorithm in [14, 15, 17, 18], however low-light TP data can be better modeled by a Poisson likelihood. Fortunately, methods, such as SPIRAL-TAP have can infer sparse causes for Poisson observations [19]. This algorithm is summarized in Algorithm 1.

Algorithm 1 Sparse spatial filtering for calcium image denoising.

Set α, $\boldsymbol{K}$
repeat
$\{\widehat{\boldsymbol{s}}, \widehat{\boldsymbol{b}}\} \leftarrow \arg\min_{\boldsymbol{s},\boldsymbol{b}} - \log(p(\boldsymbol{y}|\boldsymbol{s},\boldsymbol{b})) - \sum |s_{i,j}(t)|\widehat{\lambda}_{i,j}(t) - \sum |b_{i,j}|\widehat{\gamma}_{i,j}$
$\widehat{\lambda}_{i,j}(t) \leftarrow \frac{(\alpha+1)}{|[\boldsymbol{K}*\boldsymbol{S}]_{i,j}(t)|+\eta}$
$\widehat{\gamma}_{i,j} \leftarrow \frac{(\overline{\alpha}+1)}{|\overline{\boldsymbol{K}}*\boldsymbol{B}_{i,j}|+\eta}$
until Convergence
Output $\boldsymbol{b}$ and $\boldsymbol{s}$

4. RESULTS

We explore our denoising method by analyzing 2000 frames of TP microscopy data taken in the visual cortex of awake mice, stimulated in a virtual reality environment, expressing the GcAMP6f protein. As the microscope returns shifted and scaled noisy photon counts, we shifted the data by the smallest value in the data to make all values positive, scaled the data by the inverse of the Fano factor (the variance over the mean) to make the mean and variance of the data equal, and rounded the data to create integer data. The integer data was run through Algorithm 1 using parameters $\beta = 1$ and $\gamma = 80$. The third main parameter α was set as a function of position, $\alpha_{i,j}$, to account for variable illumination throughout the image. The weighting kernels $w_{i,j}(t), \overline{w}_{i,j}$ were set to Gaussian shapes with a spatial variance of 3 pixels, and the temporal kernel $h(t)$ was matched to empirically measured protein response functions [1].

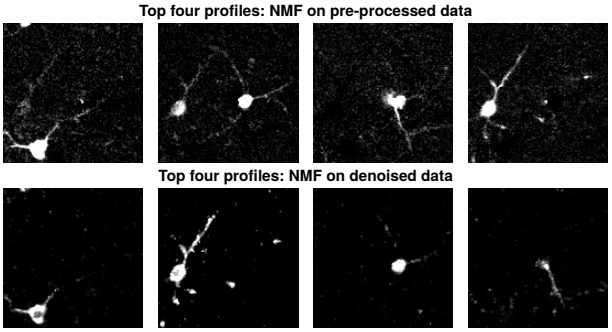

Fig. 4. Example neural spatial profiles found using NMF on the raw data and the denoised data. Top: Top four neural profiles as found by running NMF on data using standard pre-processing (in this case a running 5-frame temporal averaging). Note that some of the profiles include more than one neuron. Bottom: Top four spatial profiles found with the denoised data. The resulting profiles have less background noise and the second cells that appear in some of the spatial profiles in the top row are better demixed.

Figure 2 demonstrates the difference between Gaussian and Poisson likelihood functions. While both methods give similar results, the Poisson estimation recovered sparser estimates of the activity, indicating a better match to the data statistics. Figure 3 shows a single example frame of the original data and the output of the denoising algorithm. The reduction in noise is apparent, and the activity locations are more distinguishable by eye. To test the extent to which the denoising method helps isolate neural activity from different cells, we apply NMF both to pre-processed TP data from the upper-right hand portion of the image, as well as to the denoised data. Figure 4 shows the top four found profiles for each (pre-processed data on top and denoised data on the bottom). The results using the denoised data show far less contamination of neural profiles, demonstrating clear isolation of different cells.

5. CONCLUSIONS

We demonstrate in this work an alternative to pre-processing of TP microscopy data for neural recordings. Our method accounts for the expected temporal sparsity and spatial cohesion of neural activity, as well for providing a means to use Poisson measurement statistics in place of more common Gaussian statistics. The resulting reweighted algorithm denoises TP data and allows for improved demixing using either automated or manual methods.

A number of challenges remain in applying our method more broadly. First, the computational is significantly higher than standard pre-processing. Faster deconvolution algorithms would drastically reduce the run-time and should be explored. Additionally, parameter setting for our model was accomplished manually. As poor parameter selection can yield erroneous results, automated parameter setting would drastically improve the utility of our method. In particular, learning optimal dependency kernels (i.e. $w_{i,j}(t)$ in this work) directly from TP movies is especially important, as these kernels convey important information about the expected statistics of TP data. Additionally, while our algorithm is currently independent of demixing methods, our model could be adapted as a front-end to demixing algorithms, creating a single automated method that demixes given a more complete statistical model of the data. Finally, recent work in [20] can potentially be used to provide quantitative metrics for the success of our demixing steps.

6. REFERENCES

[1] T.-W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman *et al.*, "Ultrasensitive fluorescent proteins for imaging neuronal activity," *Nature*, vol. 499, no. 7458, pp. 295–300, 2013.

[2] L. Madisen, A. R. Garner, D. Shimaoka, A. S. Chuong, N. C. Klapoetke, L. Li, A. van der Bourg, Y. Niino, L. Egolf, C. Monetti, H. Gu, M. Mills, A. Cheng, B. Tasic, T. N. Nguyen, S. M. Sunkin, A. Benucci, A. Nagy, A. Miyawaki, F. Helmchen, R. M. Empson, T. Knopfel, E. S. Boyden, R. C. Reid, M. Carandini, and H. Zeng, "Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance," *Neuron*, vol. 85, no. 5, pp. 942–58, 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25741722

[3] E. Yaksi and R. W. Friedrich, "Reconstruction of firing rate changes across neuronal populations by temporally deconvolved ca2+ imaging," *Nature Methods*, vol. 3, no. 5, pp. 377–383, 2006.

[4] J. T. Vogelstein, B. O. Watson, A. M. Packer, R. Yuste, B. Jedynak, and L. Paninski, "Spike inference from calcium imaging using sequential monte carlo methods," *Biophysical journal*, vol. 97, no. 2, pp. 636–655, 2009.

[5] T. Deneux, A. Kaszas, G. Szalay, G. Katona, T. Lakner, A. Grinvald, B. Rózsa, and I. Vanzetta, "Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo," *Nature Communications*, vol. 7, 2016.

[6] O. Jon, S. R. Schultz, P. L. Dragotti *et al.*, "A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging," *Journal of neural engineering*, vol. 10, no. 4, p. 046017, 2013.

[7] B. F. Grewe, D. Langer, H. Kasper, B. M. Kampa, and F. Helmchen, "High-speed in vivo calcium imaging reveals neuronal network activity with near-millisecond precision," *Nature methods*, vol. 7, no. 5, pp. 399–405, 2010.

[8] E. A. Mukamel, A. Nimmerjahn, and M. J. Schnitzer, "Automated analysis of cellular signals from large-scale calcium imaging data," *Neuron*, vol. 63, no. 6, pp. 747–760, 2009.

[9] R. Maruyama, K. Maeda, H. Moroda, I. Kato, M. Inoue, H. Miyakawa, and T. Aonishi, "Detecting cells using non-negative matrix factorization on calcium imaging data," *Neural Networks*, vol. 55, pp. 11–19, 2014.

[10] E. Pnevmatikakis, Y. Gao, D. Soudry, D. Pfau, C. Lacefield, K. Poskanzer, R. Bruno, R. Yuste, and L. Paninski, "A structured matrix factorization framework for large scale calcium imaging data analysis," *arXiv preprint arXiv:1409.2903*, 2014.

[11] M. Pachitariu, A. M. Packer, N. Pettit, H. Dalgleish, M. Hausser, and M. Sahani, "Extracting regions of interest from biological images with convolutional sparse block coding," in *Advances in Neural Information Processing Systems*, 2013, pp. 1745–1753.

[12] A. Song, A. Charles, S. Koay, J. Gauthier, S. Thiberge, J. Pillow, and D. Tank, "Volumetric twp-photon imaging of neurons using stereoscopy (vtwins)," 2016, submitted (bioRxiv 073742).

[13] E. Pnevmatikakis and L. Paninski, "Sparse nonnegative deconvolution for compressive calcium imaging: algorithms and phase transitions," in *Advances in Neural Information Processing Systems*, 2013, pp. 1250–1258.

[14] P. Garrigues and B. Olshausen, "Group sparse coding with a laplacian scale mixture prior," *Advances in Neural Information Processing Systems*, vol. 24, 2010.

[15] A. Charles and C. Rozell, "Spectral super-resolution of hyperspectral imagery using reweighted ℓ_1 spatial filtering," *IEEE G. and Remote Sens. Lett.*, 2013, accepted.

[16] A. Charles, A. Balavoine, and C. Rozell, "Dynamic filtering of time-varying sparse signals via ℓ_1 minimization," *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5644–5656, 2016.

[17] A. S. Charles and C. J. Rozell, "Dynamic filtering of sparse signals using reweighted ℓ_1," *Proceedings of ICASSP*, 2013.

[18] A. S. Charles, "Dynamics and correlations in sparse signal acquisition," Ph.D. dissertation, Georgia Institute of Technology, 2015.

[19] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is spiral-tap: sparse poisson intensity reconstruction algorithms-theory and practice," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1084–1096, 2012.

[20] B. A. Wilt, J. E. Fitzgerald, and M. J. Schnitzer, "Photon shot noise limits on optical detection of neuronal spikes and estimation of spike timing," *Biophysical Journal*, vol. 104, no. 1, pp. 51–62, 2013.