Spatially-filtered temporal dictionary learning for calcium imaging analysis

Gal Mishne  
Yale University  
New Haven, CT 06520  
Email: gmishne@yale.edu

Nathan Cermak, Jackie Schiller  
Technion - Israel Institute of Technology  
Haifa, Israel, 31096  
Email: jackie@technion.ac.il

Adam Charles  
Princeton University  
Princeton, NJ, 08544  
Email: adamsc@princeton.edu

Abstract—Optical calcium imaging is a versatile imaging modality that permits the recording of neural activity, including deep neural populations as well as single dendrites and spines using two-photon microscopy (TPM). Calcium imaging analysis relies on extracting the temporal fluorescence fluctuations of each component (e.g., spines, cell bodies or dendrites) from the full video. Current methods rely on strong a-priori spatial and temporal regularization to isolate each component, which can bias time-trace estimation and restrict applicability across imaging scales. We reverse the modeling and instead aim to minimize the spatial inference, while focusing on finding the set of temporal traces present in the data. We reframe the problem in a dictionary learning setting, where the dictionary contains the time-traces and the sparse coefficient are spatial maps. Our method is based on spatially-filtered Laplacian-scale mixture models, and we introduce additional constraints to implicitly infer the number of components and suppress their redundancy. We demonstrate our approach on calcium imaging at two different scales: population and dendritic imaging. First we compare our method to a current state-of-the-art algorithm, Suite2p, on the publicly available Neurofinder dataset. The flexible spatial constraints allow our model to isolate both disconnected portions of the same neuron and small components otherwise over-shadowed by stronger components. This latter case is important as such configurations can cause scientifically-misleading false transients. On dendritic data we demonstrate how our method isolates spines and dendritic firings.

I. INTRODUCTION

Calcium imaging (CI), or the optical recording of calcium concentrations in neural tissue, is an important neural imaging technique due to its ability to simultaneously record large neural populations at single cell resolution in awake behaving animals [1]–[3]. In particular, calcium imaging via two-photon microscopy enables minimally-invasive imaging hundreds of microns beneath the surface.

II. MODEL

Existing factorization methods for calcium imaging data [4], [5] weight both spatial and temporal components, and validation often assesses only the accuracy of spatial profiles. The main quantities of interest, however, for scientific inquiry, are the time-traces, as they are the variables linked to behavior, stimuli, etc. Thus, we reverse the current modeling philosophy of looking for spatial components to find time traces, and rather directly model temporal traces. Given T motion-corrected frames Yt ∈ R N × Ny , we model the per-pixel activity as yt,j = ∑K k=1 φk T aij,k + εt,j, where φk ∈ R T for k ∈ [1, ... , K] are the K time-traces, aij,k is the strength of each component’s fluorescence at each pixel, and εt,j ∈ R T is the sensor noise. This modeling shift is similar to the difference, for example, in modeling the spatial statistics of hyperspectral imaging data versus the predominant spectral end-member analysis. We define a cost function of the time-traces φk and the spatial presence coefficients aij,k, that captures both the data model and the a-priori information that few neurons overlap at any given pixel (aij,k = [aij,k1, ... , aij,kK]T ∈ R K) is sparse) and that both Φ and A are non-negative.

We adopt a dictionary learning (DL) approach, where the dictionary Φ = [φ1, ... , φK] ∈ R T × K is composed of the time traces and the sparse coefficients A are the spatial profiles. The basic DL optimization minimizes the cost function J(Φ, A) = ∑t,i,j ∥yt,j − Φaij,k∥2 + λ ∥aij,k∥1, where λ is a parameter that trades off data fidelity and sparsity in ai,j,k and is minimized via alternating updates for Φ and ai,j,k. We expand on this model by both introducing new penalties to regularize the dictionary, and present a new graph-based sparse coding for the spatial profiles. For the penalties over Φ, a Frobenius norm serves to remove unused dictionary elements, and a penalty over intra-dictionary correlations penalizes time-traces with trivial differences. To ensure stable convergence, we also further penalize the change in Φ between updates. The update for Φ at algorithmic iteration t is given ai,j,k = arg min ∥ Y − ΦA ∥2 + λ1∥Φ∥2 + λ2∥Φ − Φt−1∥2 + λ3∥Φ − diag(Φ T Φ)∥F 2

Re-weighted ℓ1 spatial filtering (RWL1-SF): To remedy the lack of spatial cohesion in traditional DL, we adapt RWL1-SF [6], which is an expansion of re-weighted ℓ1 [7] and the Laplacian-scale mixture model [8]. Inferring ai,j,k (via approximate expectation-maximization) involves iteratively solving for all {i, j}

\[ a_{i,j,k} = \arg \min_{\alpha \geq 0} \frac{1}{2\sigma^2} \left|\|\mathbf{y}_{i,j} - \Phi \mathbf{a}\|_2^2 + \sum_k \lambda_{i,j,k} |a_{i,j,k}| \right|, \]

\[ \lambda_{i,j,k} = \xi \left( \beta + |a_{i,j,k}| + \left| W \cdot \mathbf{A}_k \right|_{i,j} \right), \]

where \( W \) contains the presence coefficients for a single component across the FOV and \( W \cdot \mathbf{A}_k \) indicates a 2D convolution. Here \( \xi \) and \( \beta \) depend only on model parameters \( \alpha, \theta, \sigma^2 \) [6].

The weights \( \lambda \) in RWL1-SF incorporate spatial information into per-pixel solutions by sharing second-order statistics. Coefficients “activated” in the initial optimization lower the weights for neighboring coefficients, encouraging them to activate in subsequent iterations, whereas non-active coefficients penalize activation with higher weights. The kernel \( W \) specifies the coefficient influence radius, dictating the neighborhood where interactions are strongest.

III. RESULTS & CONCLUSIONS

We apply our method to simulated, population-level and dendritic calcium imaging data. Simulations revealed how the algorithm converges from a random initialization to capture the true underlying time-traces (Fig. 1B). Comparisons to Suite2p [4] on the standard Neurofinder dataset revealed that our method better located nested apical and discontinuous components (Fig. 1C). Dendritic data was decomposed into dendritic and spine components (Fig. 1D).
Fig. 1. A: Our DL method uses a per-pixel generative model with spatially correlated coefficients. B: Our algorithm converges from random time-courses to a minimum. C: Temporal DL finds subtle features in the Neurofinder dataset. (top) An apical dendrite (blue) significantly overlapping with a soma (green) was isolated. Manually labeled soma (yellow) and Suite2p (red) do not account for the apical, resulting in contaminated time traces. (bottom) Spatially disjoint pieces of the same neuron were correctly identified as one time trace, while Suite2p splits the component in two. D: Applications to dendritic data extracts both dendrite and spine activity (bottom), as seen by the spatial maps where each component is colored differently (top).

REFERENCES


ACKNOWLEDGEMENT

G.M. is supported by the NIBIB and the NINDS of the NIH, Award R01EB026936.