Neural data science: accelerating the experiment-analysis-theory cycle*

Liam Paninski

Departments of Statistics and Neuroscience
Center for Theoretical Neuroscience
Zuckerman Mind Brain Behavior Institute
Grossman Center for the Statistics of Mind
Columbia University
http://www.stat.columbia.edu/~liam
liam@stat.columbia.edu

Support: NSF, NIH, DARPA, IARPA, ARO, ONR, Google, Simons, CZI, Wellcome.

*Review article: Paninski and Cunningham, Curr. Opin. Neurobio. 2018.

A golden age of statistical neuroscience

- fast, cheap computation
- powerful new statistical machine learning tools
- optical / optogenetic revolution
- large-scale, high-density multi-electrode arrays
- plenty of exciting big data to analyze
- more important: opportunities to develop new scientific directions — many fundamental neuroscience questions are statistics problems in disguise

$\Longrightarrow$ **bottleneck: need more neural data scientists!**

Some statistical neuroscience questions

- what information is encoded in the activity of neural populations? can we decode this information?
- can we predict circuit function / dynamics from structure, or vice versa?
- what does one region of the brain tell another?
- can we optimally control neural activity with light or electrical inputs?
- what is a cell type?
- what is the best experiment I should run next?
- ...

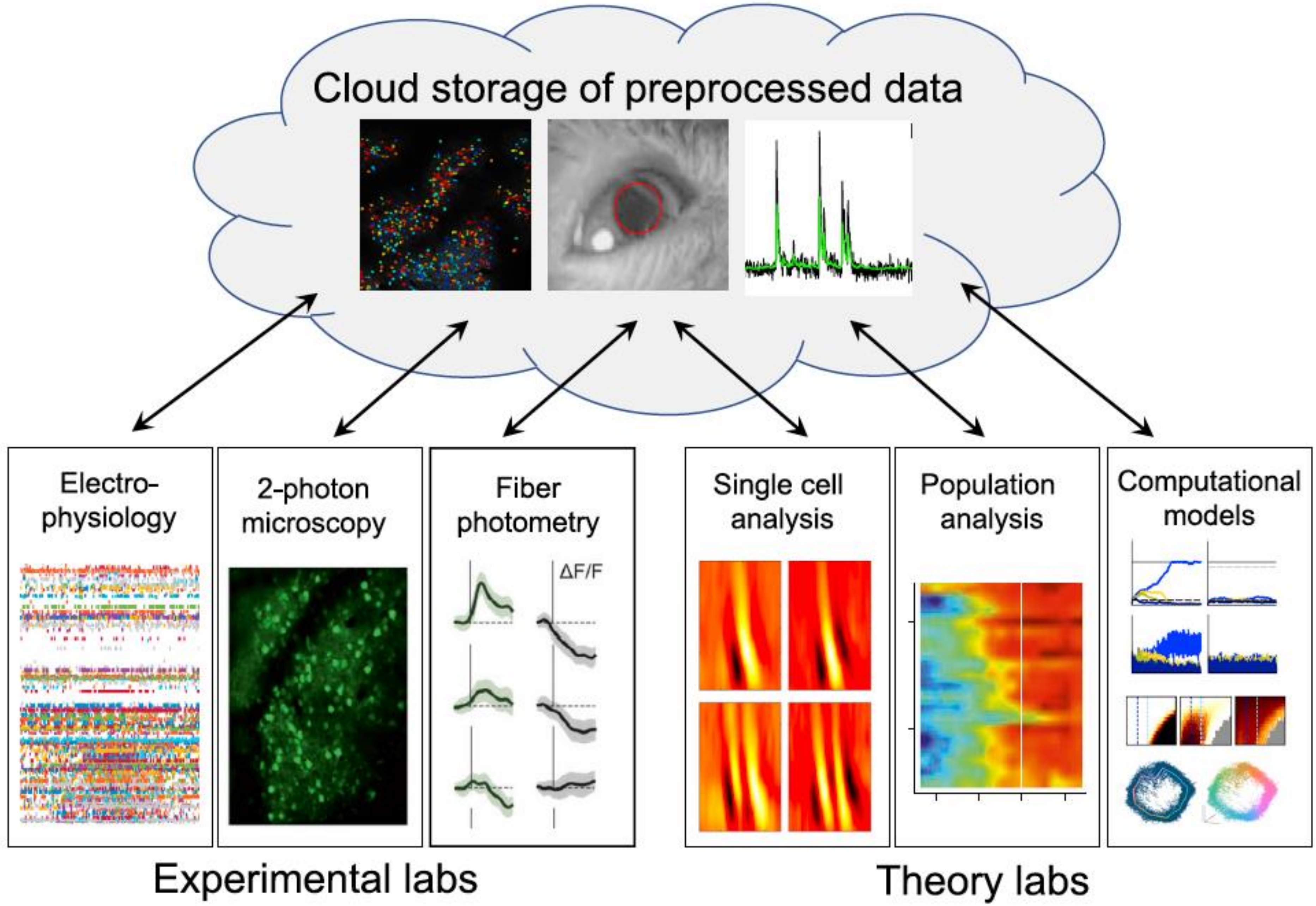
Neuron
NeuroView

An International Laboratory for Systems and Computational Neuroscience

The International Brain Laboratory*
*Correspondence: churchland@cshl.edu
https://doi.org/10.1016/j.neuron.2017.12.013

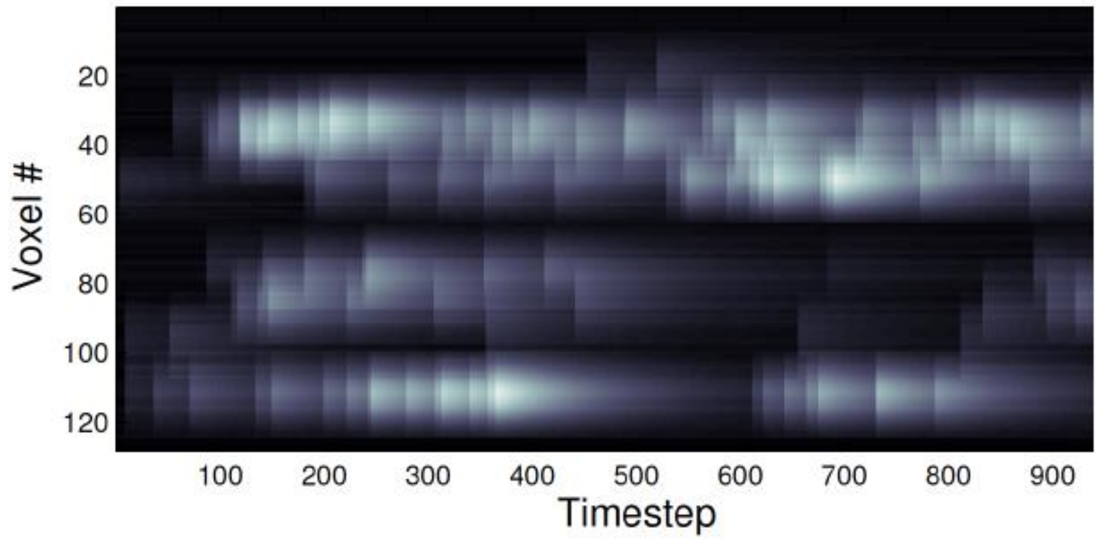
The neural basis of decision-making has been elusive and involves the coordinated activity of multiple brain structures. This NeuroView, by the International Brain Laboratory (IBL), discusses their efforts to develop a standardized mouse decision-making behavior, to make coordinated measurements of neural activity across the mouse brain, and to use theory and analyses to uncover the neural computations that support decision-making.

~20 labs; brain-wide recordings during a single standardized behavioral task



Cloud storage of preprocessed data
Electro-physiology
2-photon microscopy
Fiber photometry
ΔF/F
Single cell analysis
Population analysis
Computational models
Experimental labs
Theory labs

Constrained non-negative matrix factorization



$$
\begin{aligned}
Y &= \hat{Y} + background + noise \\
\hat{Y}(x,t) &= \sum_{i=1}^{r} a_i(x) c_i(t) \\
c_i(t) &= K_i * n_i(t) \\
n_i(t), a_i(x) &\geq 0
\end{aligned}
$$

Goal: infer low-rank matrix C from noisy Y. Rank r = number of visible neurons

Additional structure: $s_i(t)$ sparse, local in space. $n_i(t)$ sparse

CNMF (Pnevmatikakis et al, Neuron 2016) / CaImAn; OnACID (Giovanucci et al, NIPS 2017); see also Suite2p, SCALPEL, others

Model-based estimation of spike rates

Forward model:

$$\begin{aligned} y_t &= f(C_t) + \epsilon_t \\ C_t &= K_\tau * n_t \end{aligned}$$

n_t: spikes; C_t: calcium; y_t: observed fluorescence; K_τ: convolution with a single- or double-exponential

Inference:

- Hidden Markov model (Vogelstein et al 2009, Deneux et al 2016)
- Optimization methods (Vogelstein et al 2010; Grewe et al 2010; Pnevmatikakis et al 2016, Friedrich et al 2017, Jewell et al 2018)
- Markov chain Monte Carlo methods (Pnevmatikakis et al 2014)
- Neural net like methods (Theis et al 2016, Aitchison et al 2017, Berens et al 2018)
- Many other approaches (e.g., Ganmor et al 2016).

Fast maximum a posteriori (MAP) estimation

Recipe: biophysical model, then likelihood, then computation.

Start by writing out the posterior:

$$
\begin{aligned}
\log p(C|Y) &= \log p(C) + \log p(Y|C) + const. \\
\log p(Y|C) = \log p(Y|n) &= -\frac{1}{2\sigma^2}||f(K_\tau * n_t) - y_t||^2 + const.
\end{aligned}
$$

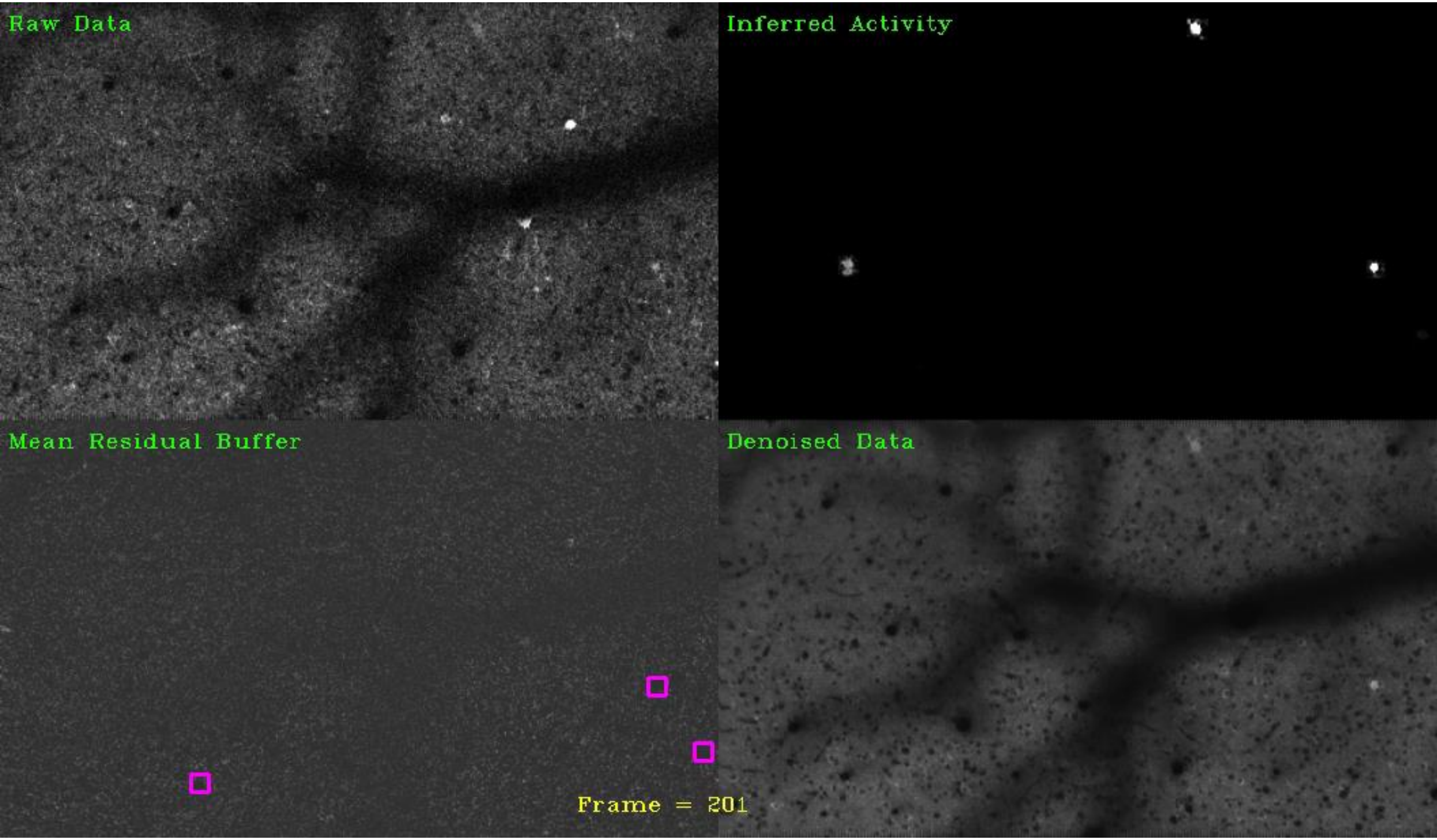
Next impose constraints:

- n_t is nonnegative and sparse
- $||f(K_\tau * n_t) - y_t||^2$ should be close to $\sigma^2 T$.

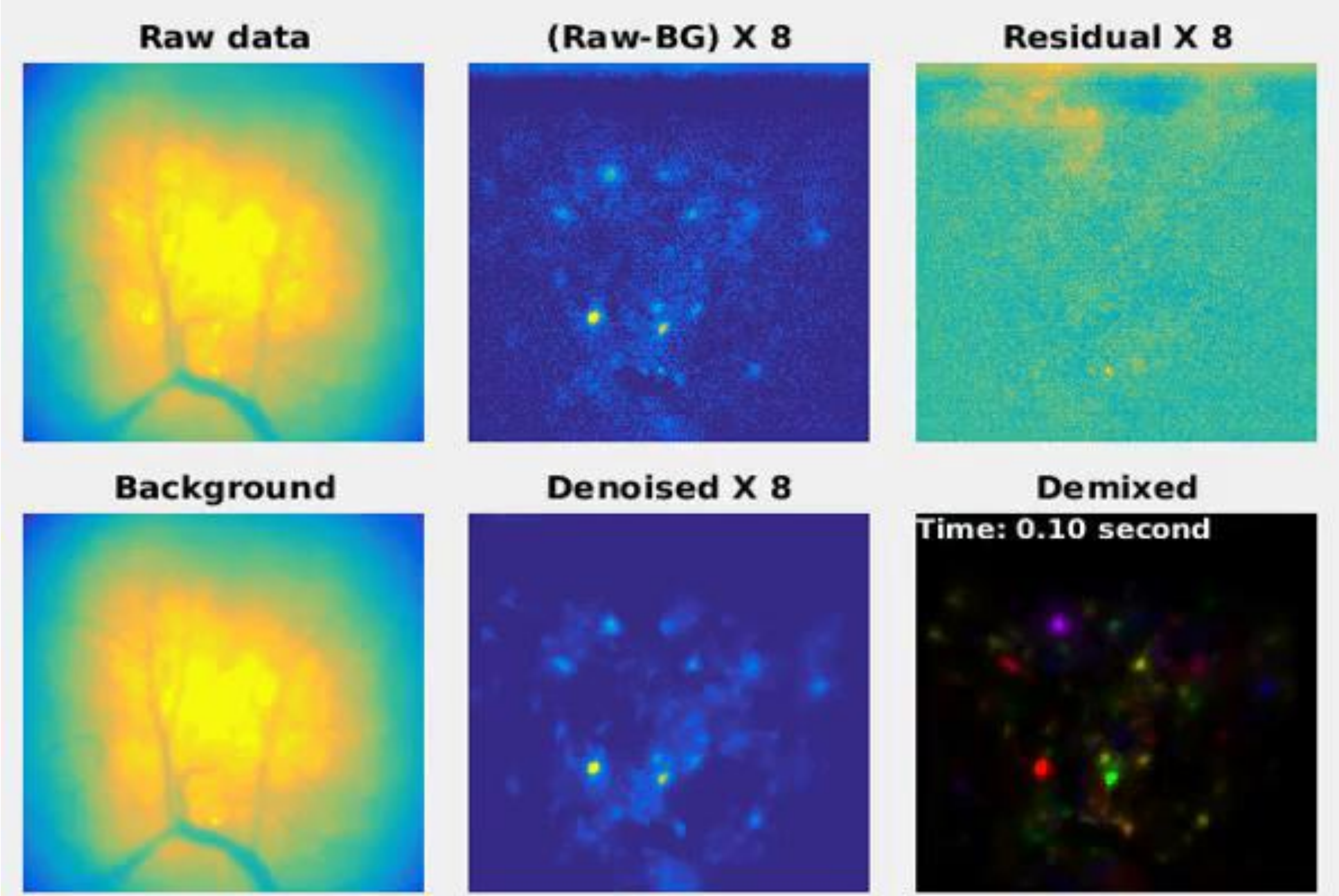
Then solve a constrained convex optimization problem:

$$
\min \sum_t n_t \;\; s.t. \;\; n_t \geq 0, \; ||f(K_\tau * n_t) - y_t||^2 \leq \sigma^2 T
$$

Special structure of $K_\tau \implies$ fast optimization (Vogelstein et al 2010; Pnevmatikakis et al 2016; Friedrich et al 2017). Enables closed loop experiments.



OnACID; Giovanucci et al, NIPS 2017. Mesoscope data from A. Tolias lab



CNMF-E; Zhou et al, eLife 2018. Very different background model required for 1p data

Compress and denoise before CNMF?

CNMF demixes, deconvolves, denoises, and compresses simultaneously

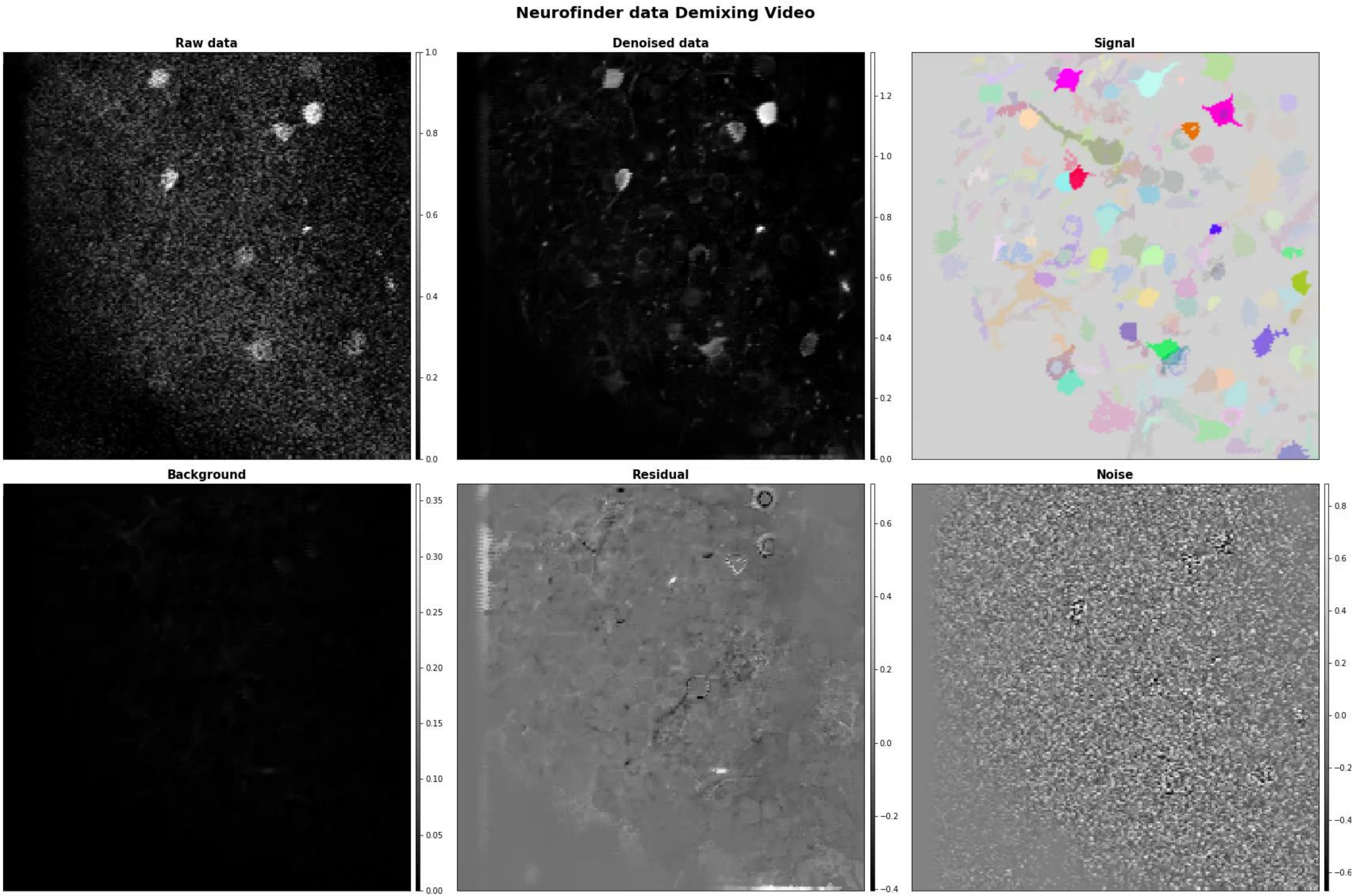
But compression and denoising should be easier than demixing – PCA is mathematically easier than NMF

Idea: compress and denoise first, then demix.

How? Local penalized matrix decomposition. Souped-up PCA on local spatial patches in parallel (with penalties to discard spatial and temporal noise), then glue the patches back together. Embarrassingly parallel: perfect for AWS

Automated selection of PCA rank and spatial / temporal smoothness to avoid discarding signal; Buchanan, Kinsella, Zhou et al 2018

Now big datasets fit on GPU for fast demixing



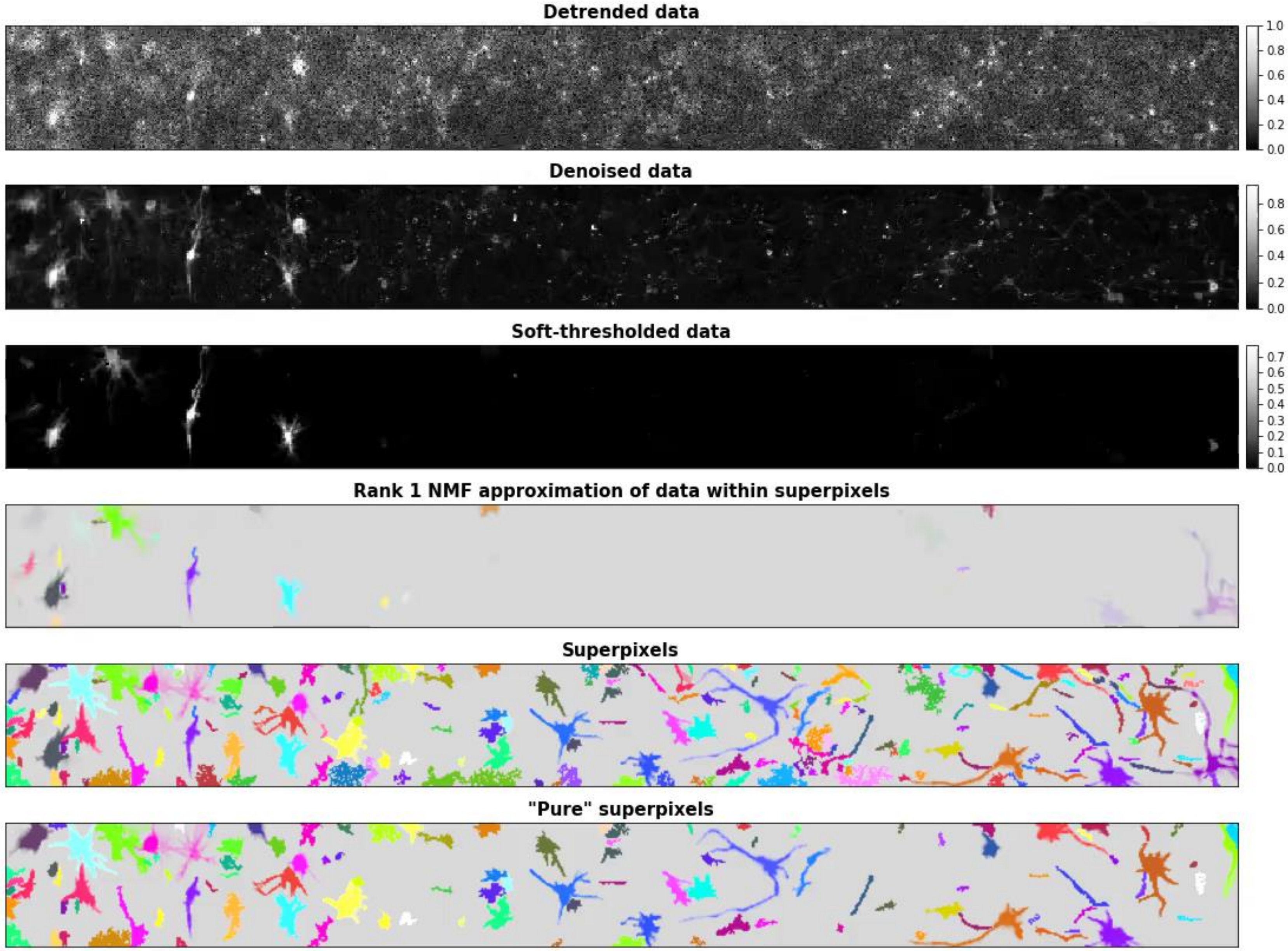
~100x compression. Buchanan, Kinsella, Zhou et al, 2018

Compression / denoising works well across a wide variety of data types

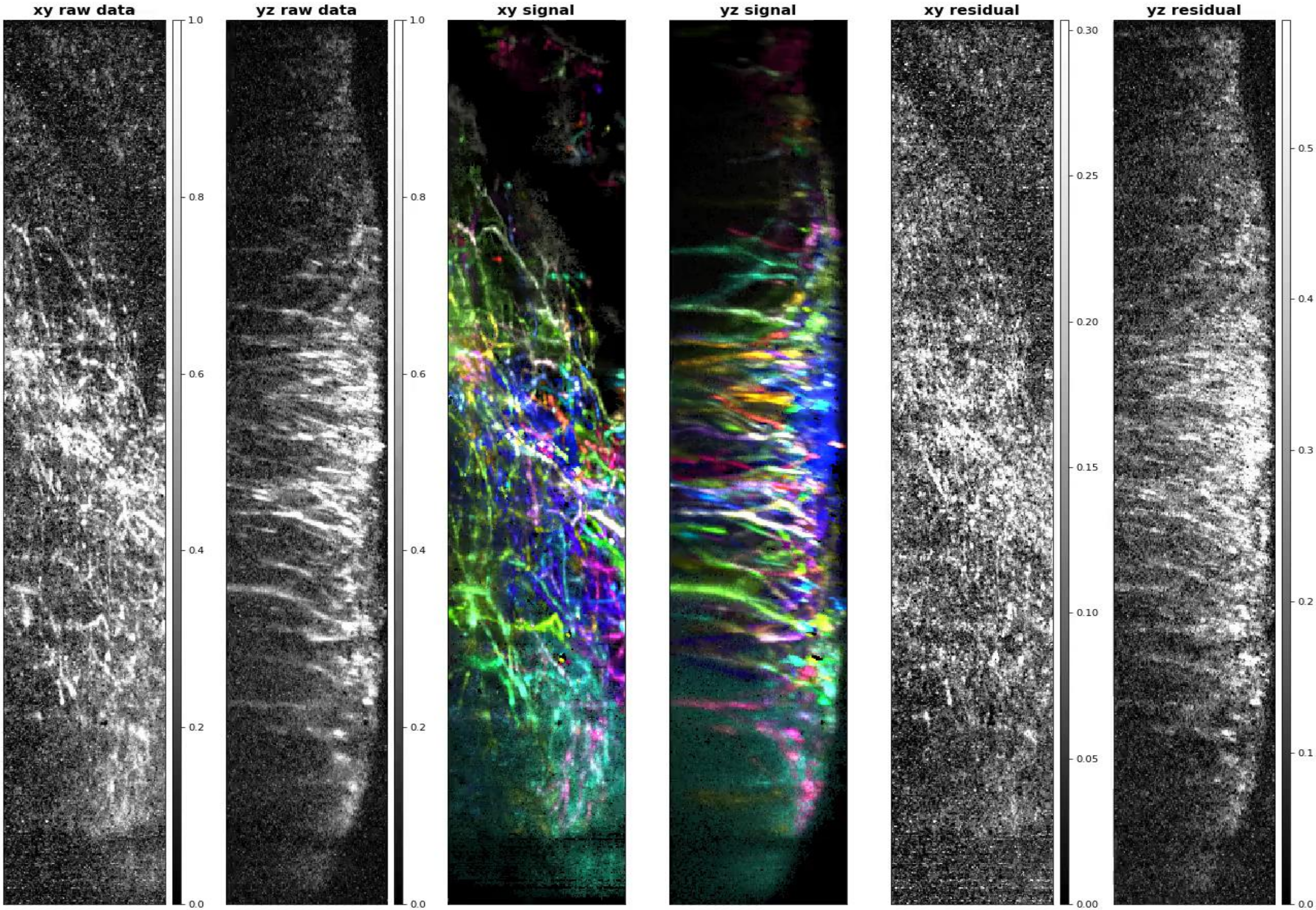
Dataset	Dimensions			Method	Compression	Total	SNR
	Frames	FOV	Patch		ratio	runtime (s)	metric
Endoscopic	6000	256x256	16x16	Patch-wise PMD	23	220.4	2.3
			NA	Standard PCA	2	595.5	1.3
Dendritic	1000	192x192	16x16	Patch-wise PMD	52	3.2	3.7
			NA	Standard PCA	2	18.3	1.1
Three-photon	3650	160x240	20x20	Patch-wise PMD	94	12.4	1.8
			NA	Standard PCA	2	187.2	1.0
Widefield	1872	512x512	32x32	Patch-wise PMD	298	12.5	3.5
			NA	Standard PCA	10	80.1	1.6
Voltage	6834	80x800	40x40	Patch-wise PMD	180	30.5	2.8
			NA	Standard PCA	8	185.1	1.0

(All on a 6-core machine; significantly faster on AWS.)

Superpixelization Video

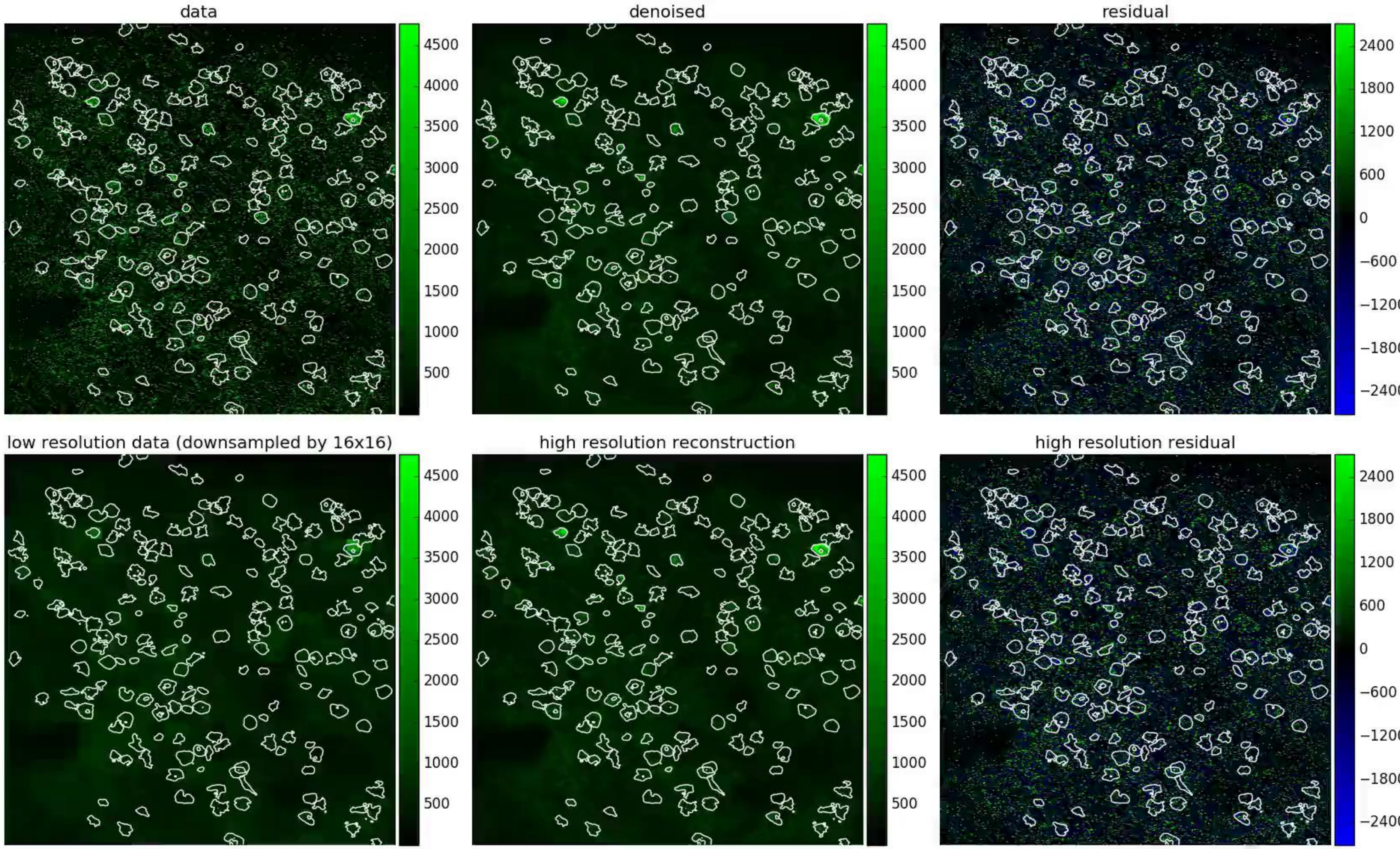


See also Adam et al (Nature 2019) for in vivo data

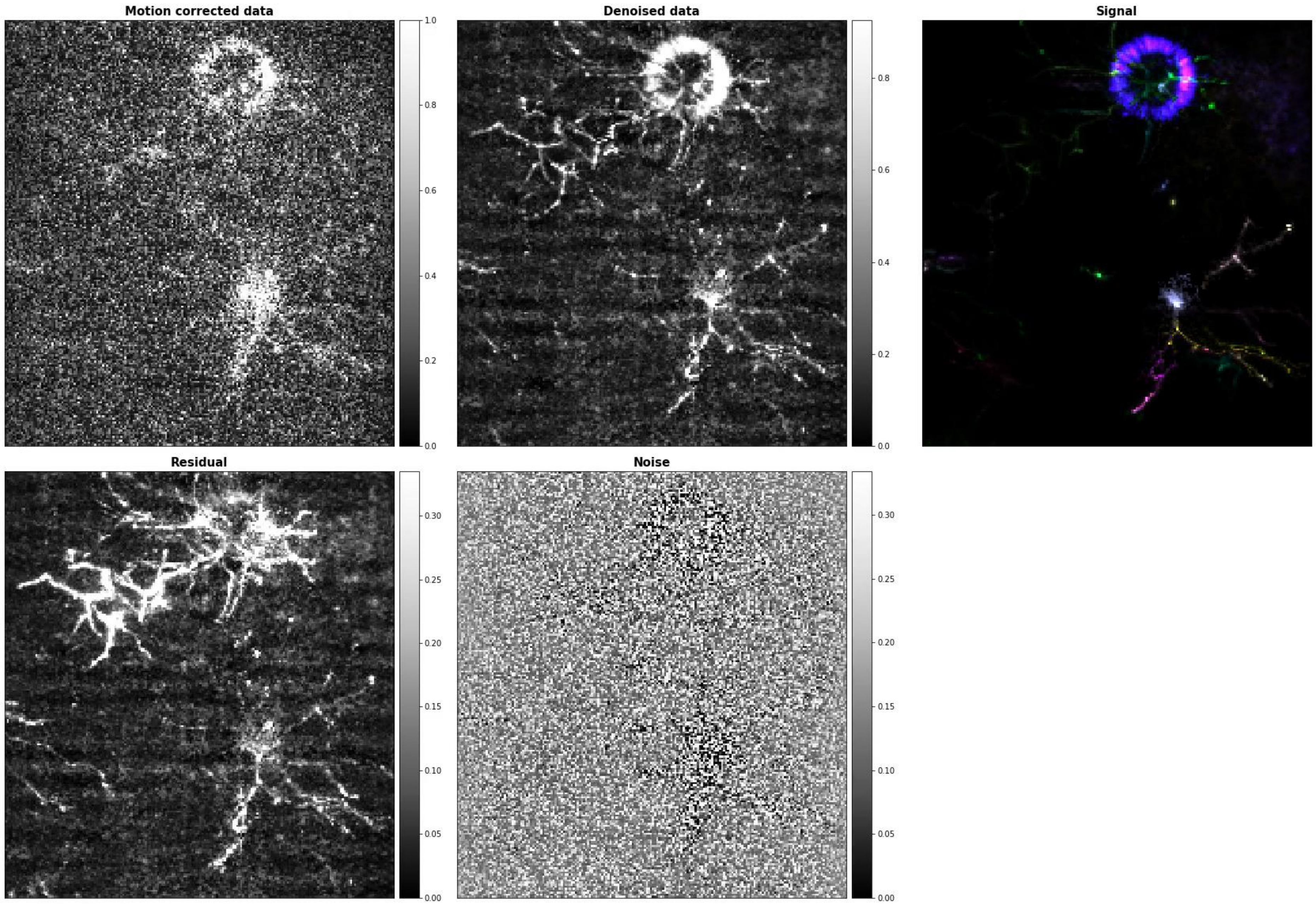


Ding Zhou; 3d SCAPE data from S. Benezra, R. Bruno + E. Hillman labs

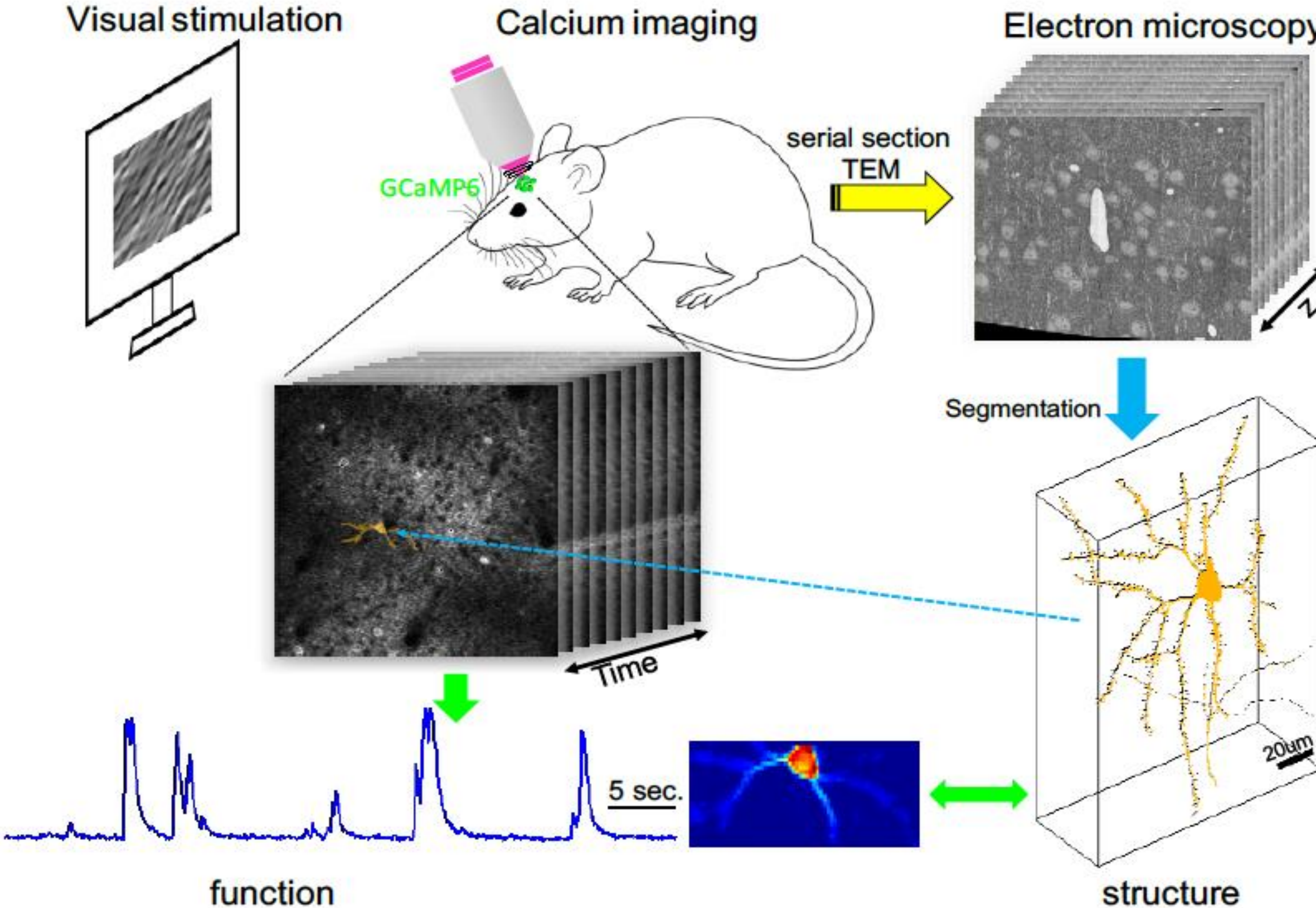
Compression then demixing enables faster acquisition



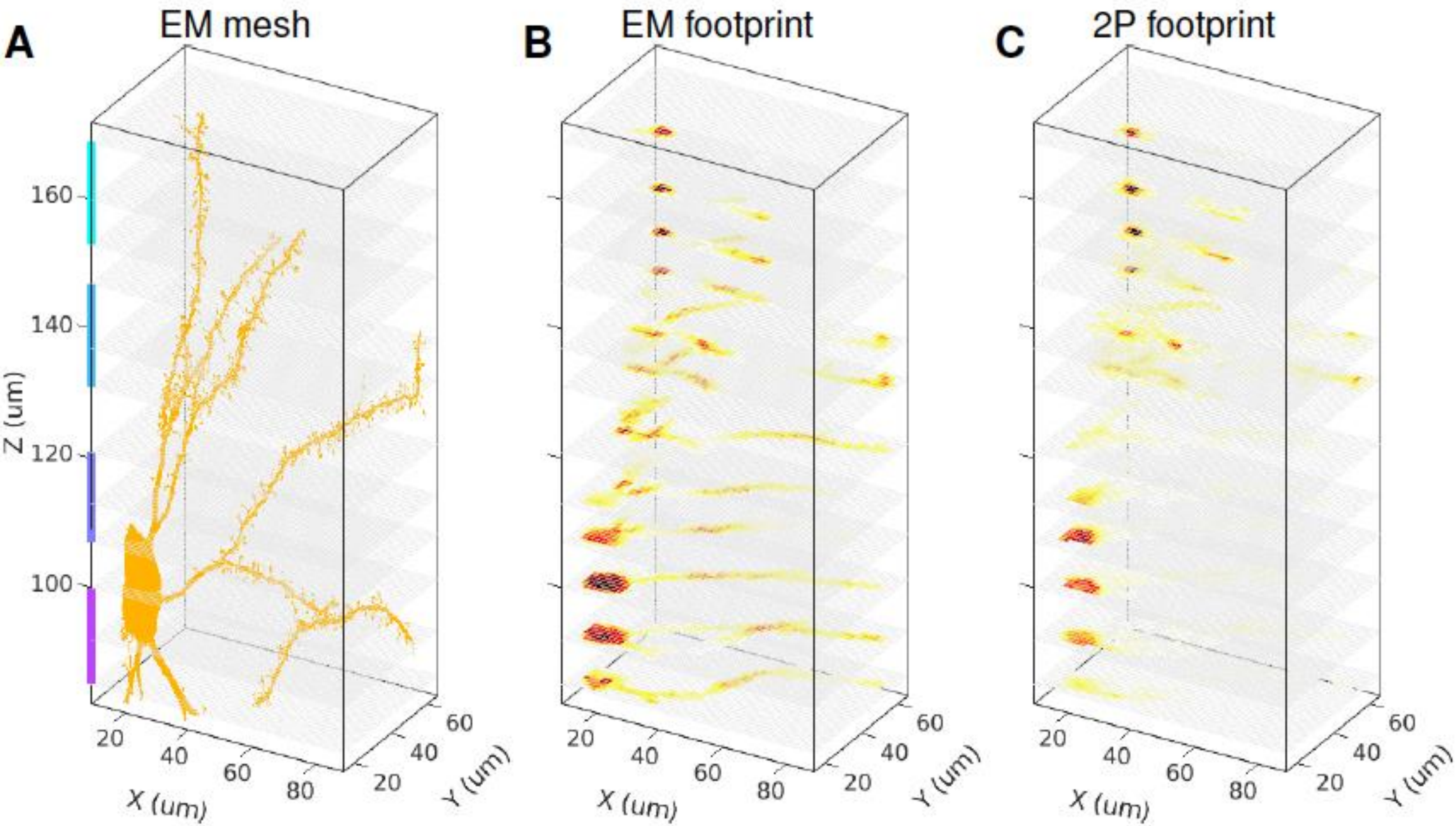
Friedrich + Paninski, PLoS Comp. Bio. 2017



Buchanan, Kinsella, Zhou et al, 2018. Bessel beam data from Na Ji lab

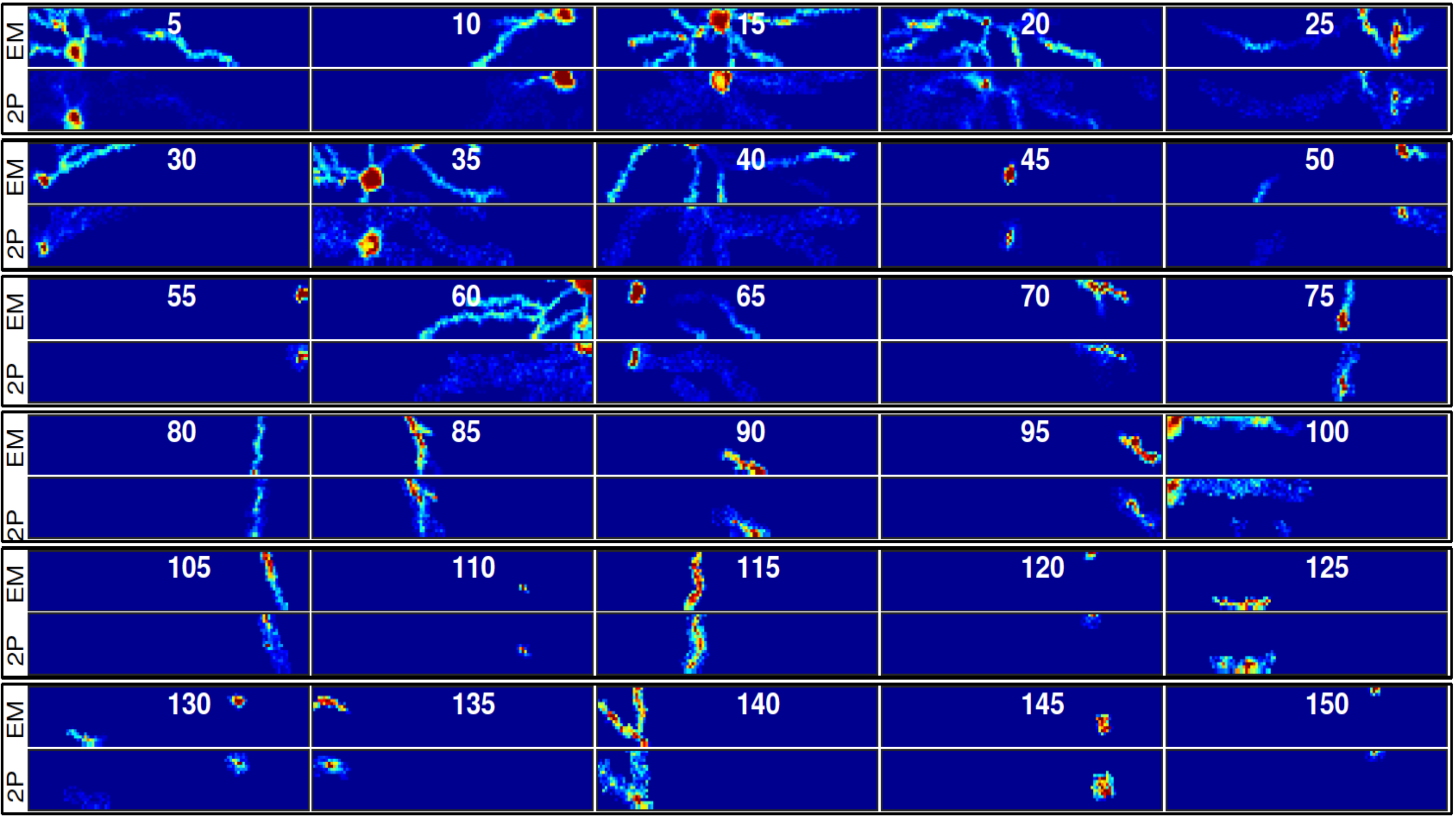


MICRONS collaboration; Tolias and Seung labs, Allen Institute; P.C. Zhou

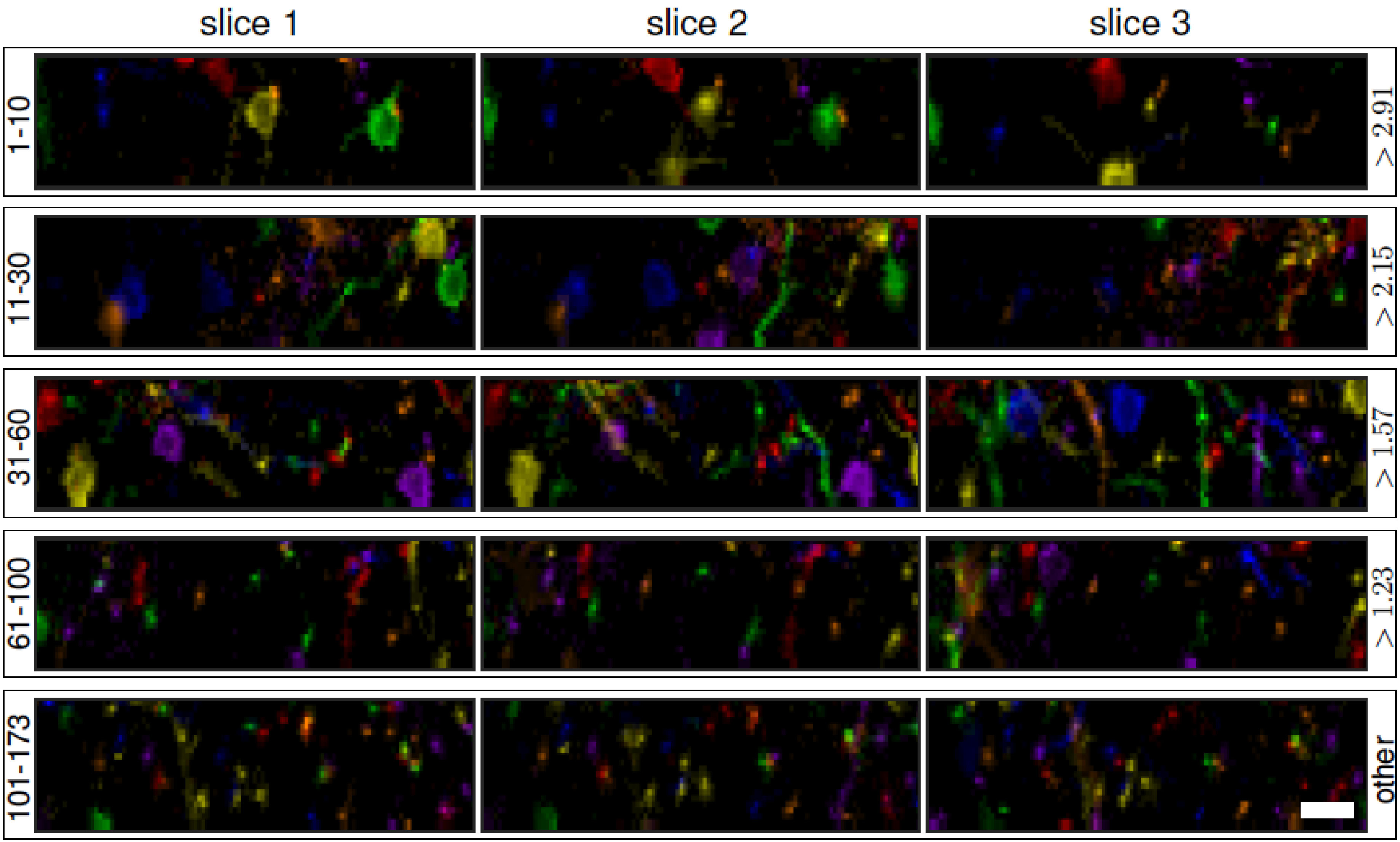


MICRONS collaboration; Tolias and Seung labs, Allen Institute; P.C. Zhou

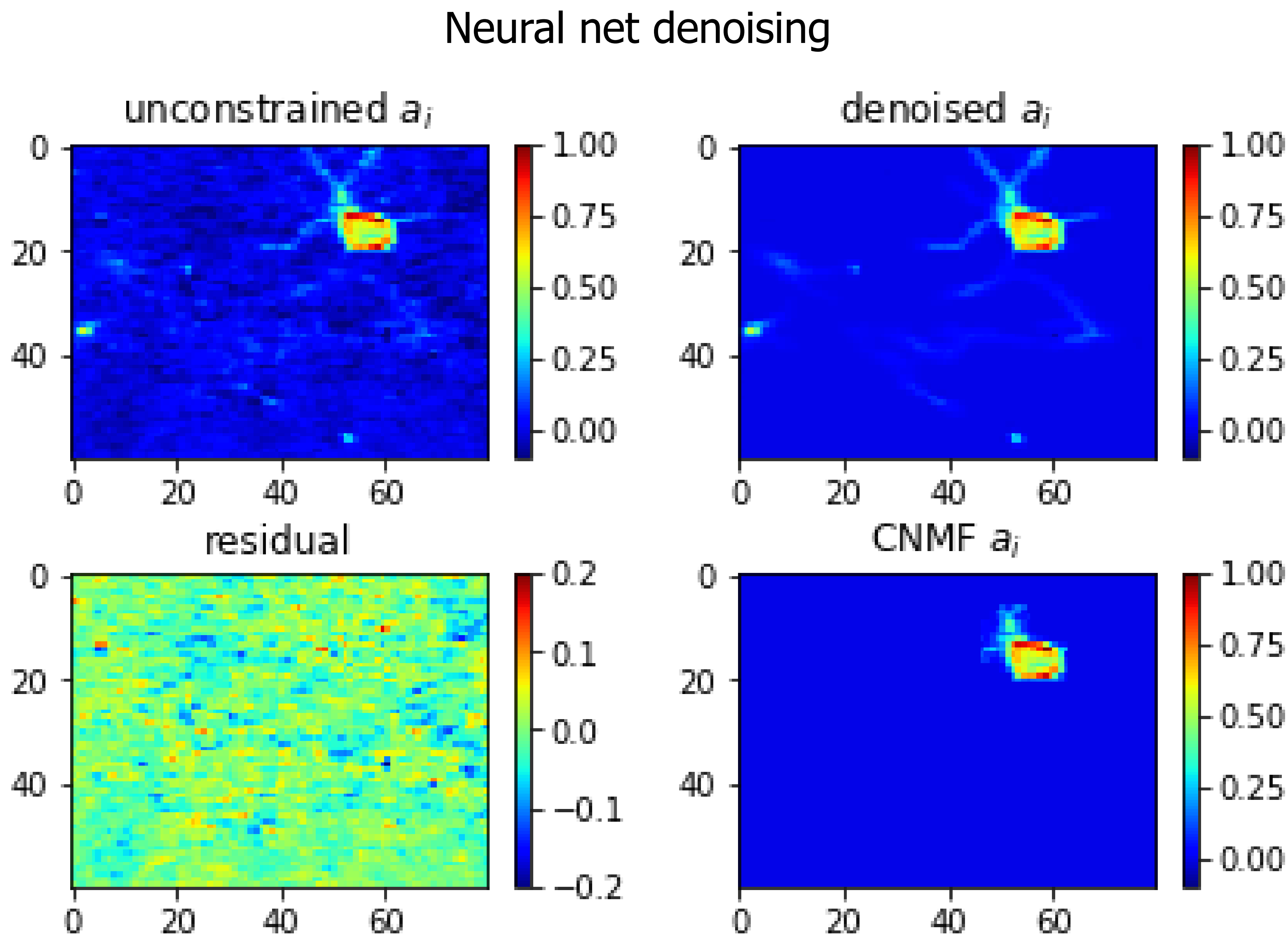
EM-constrained gold-standard demixing output



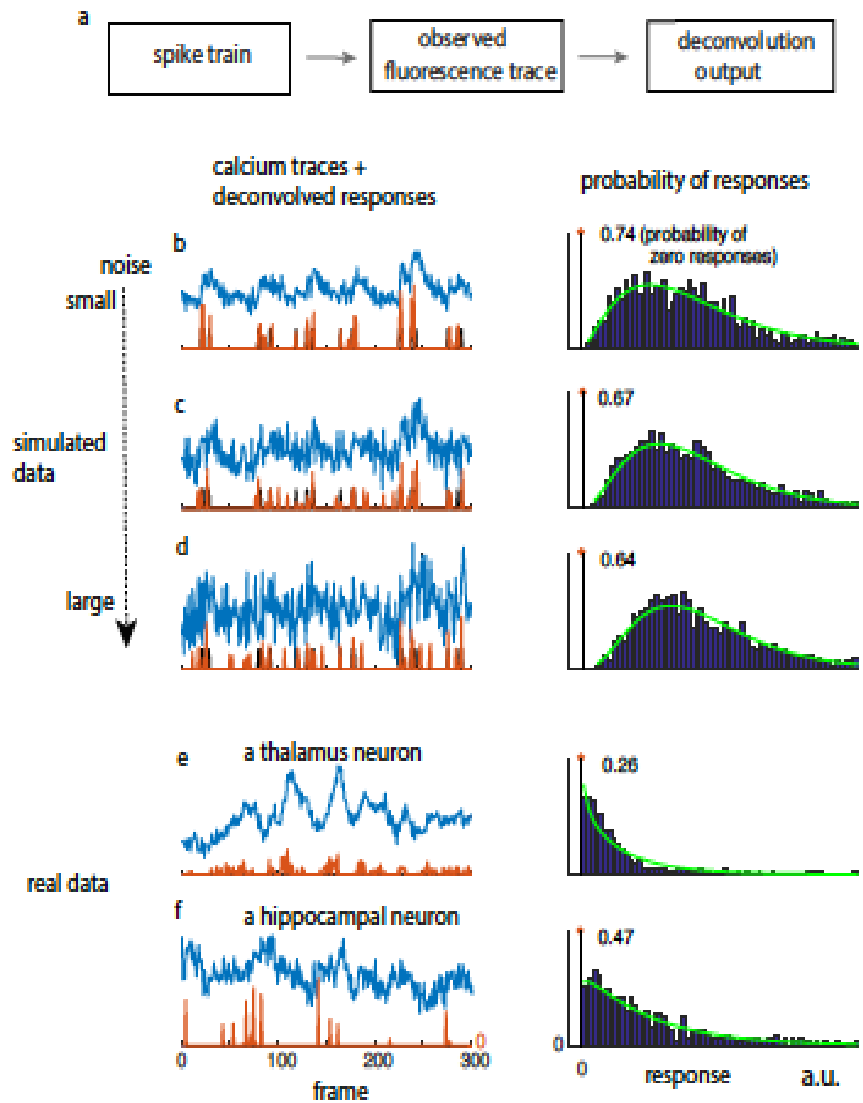
MICRONS collaboration; Tolias and Seung labs, Allen Institute; P.C. Zhou



(Most components are non-somatic...)

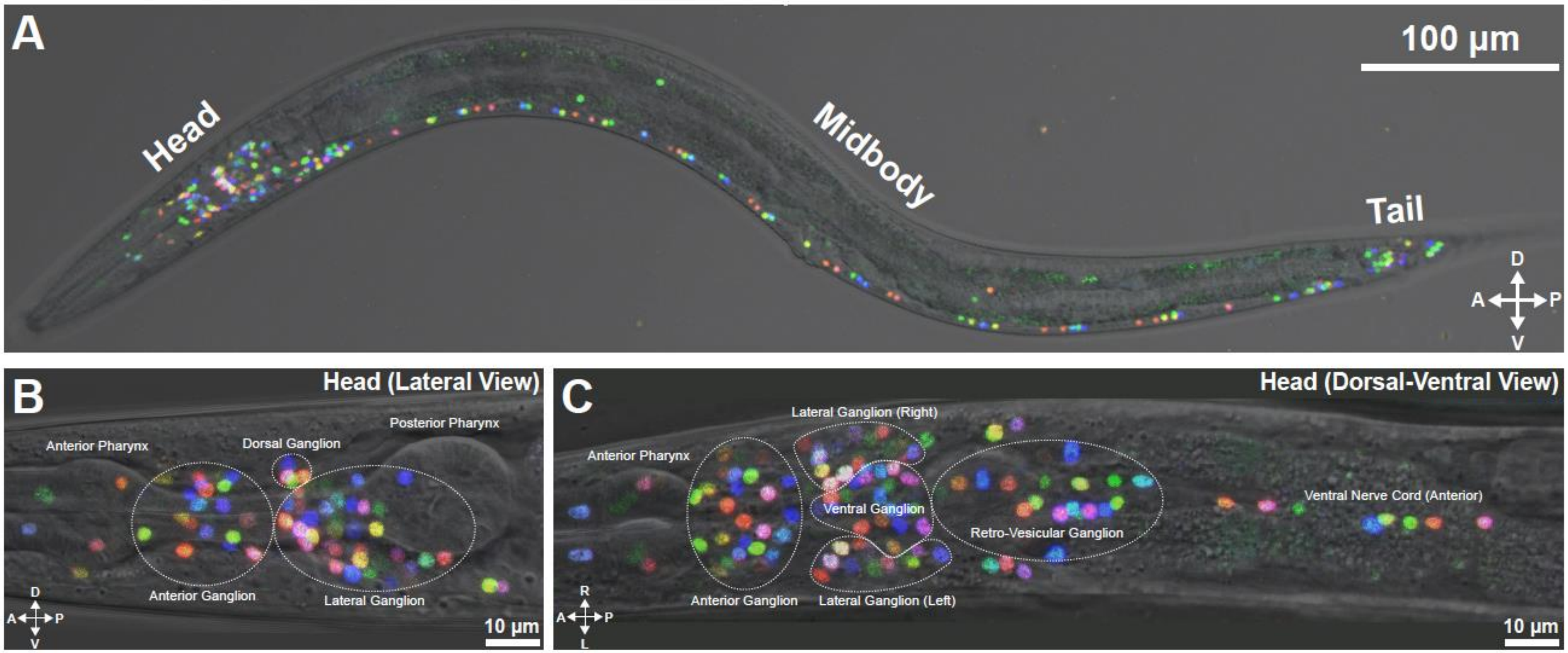

Neural net denoising
unconstrained a_i
denoised a_i
residual
CNMF a_i
0
20
40
60
1.00
0.75
0.50
0.25
0.00
0.2
0.1
0.0
−0.1
−0.2

How to model post-deconvolved output?



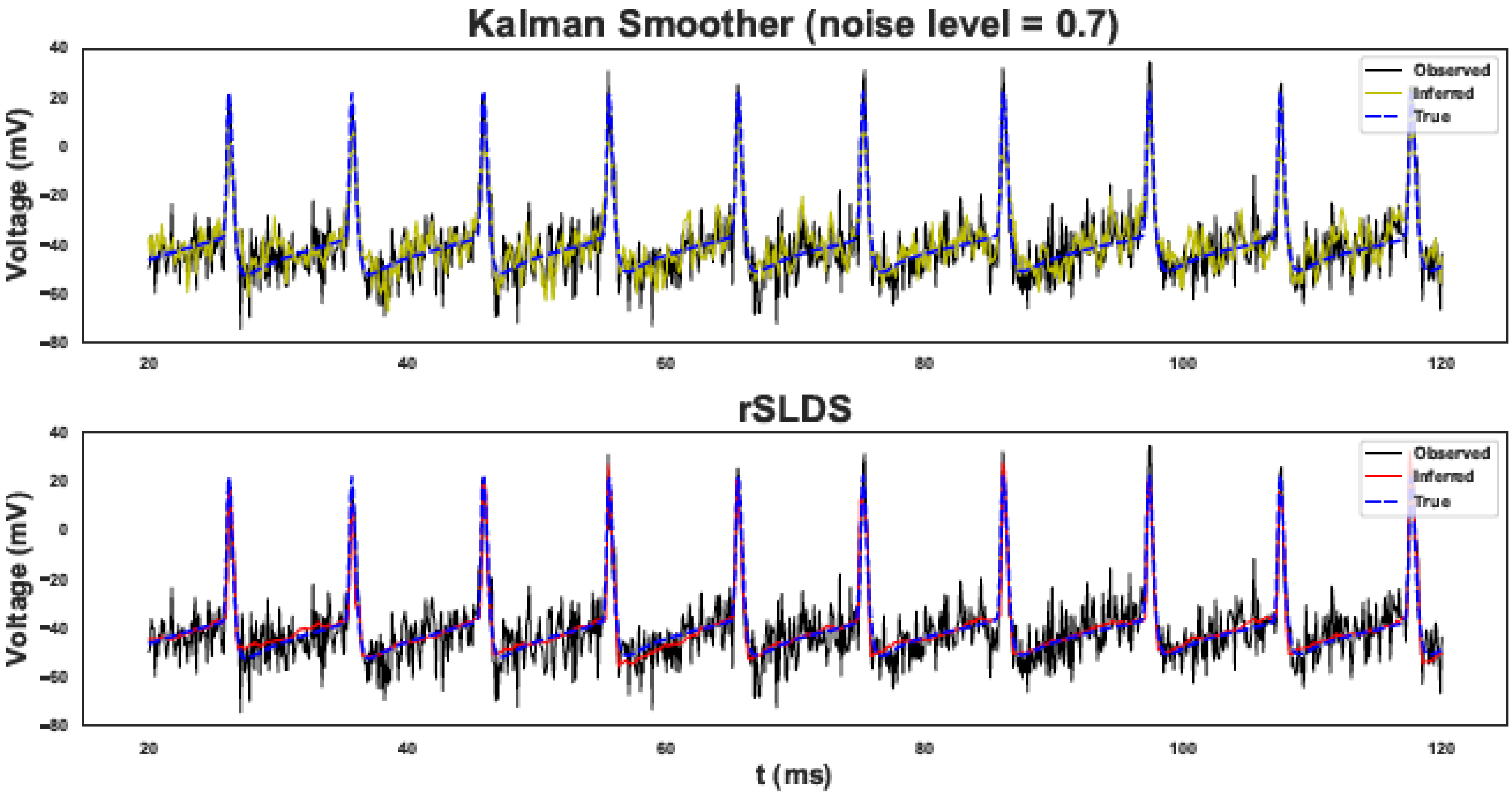
Wei et al, Biorxiv 2019

How to handle really bendy animals?

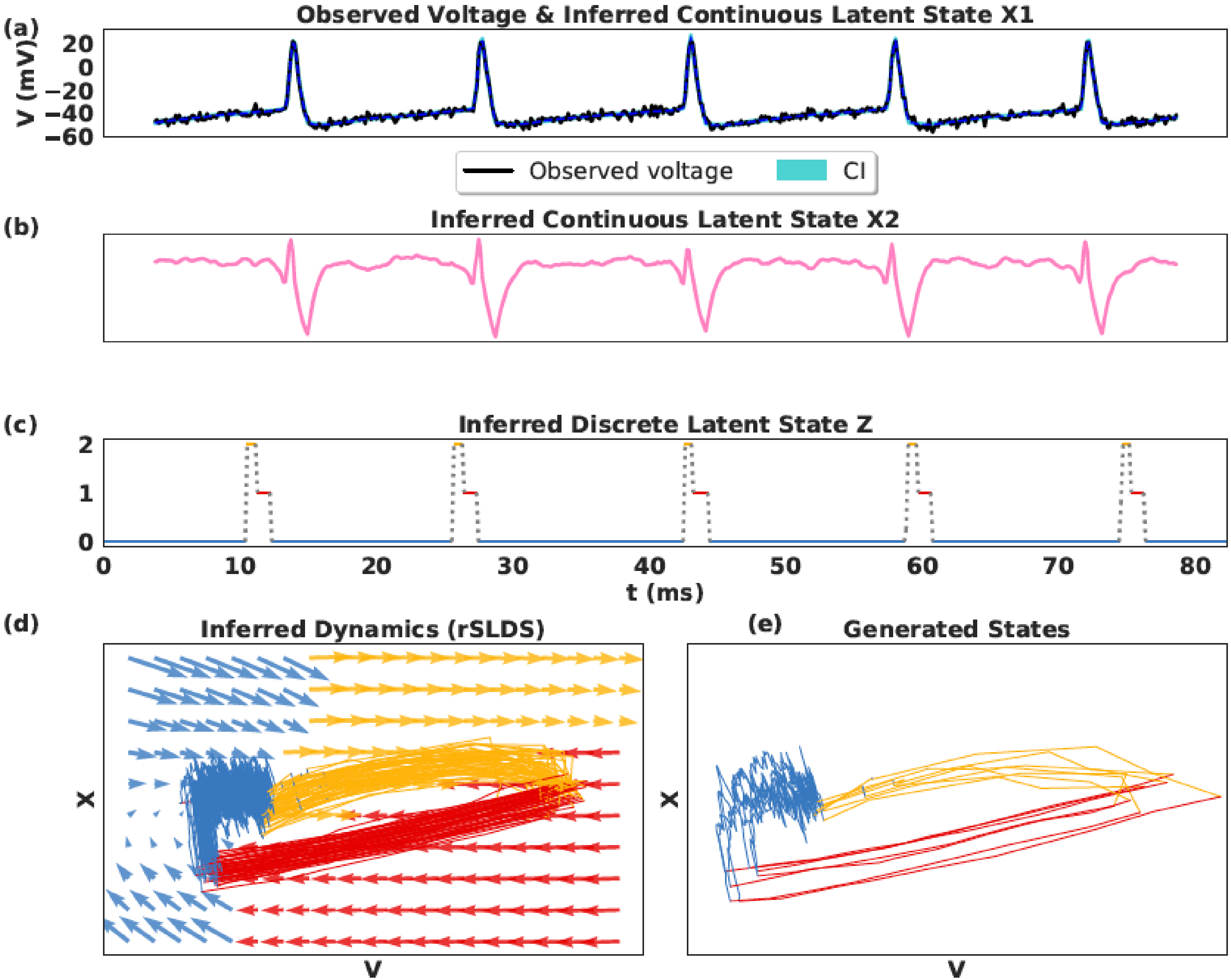


NeuroPAL: Yemini et al Biorxiv 2019

Denoising subthreshold activity on dendritic trees?

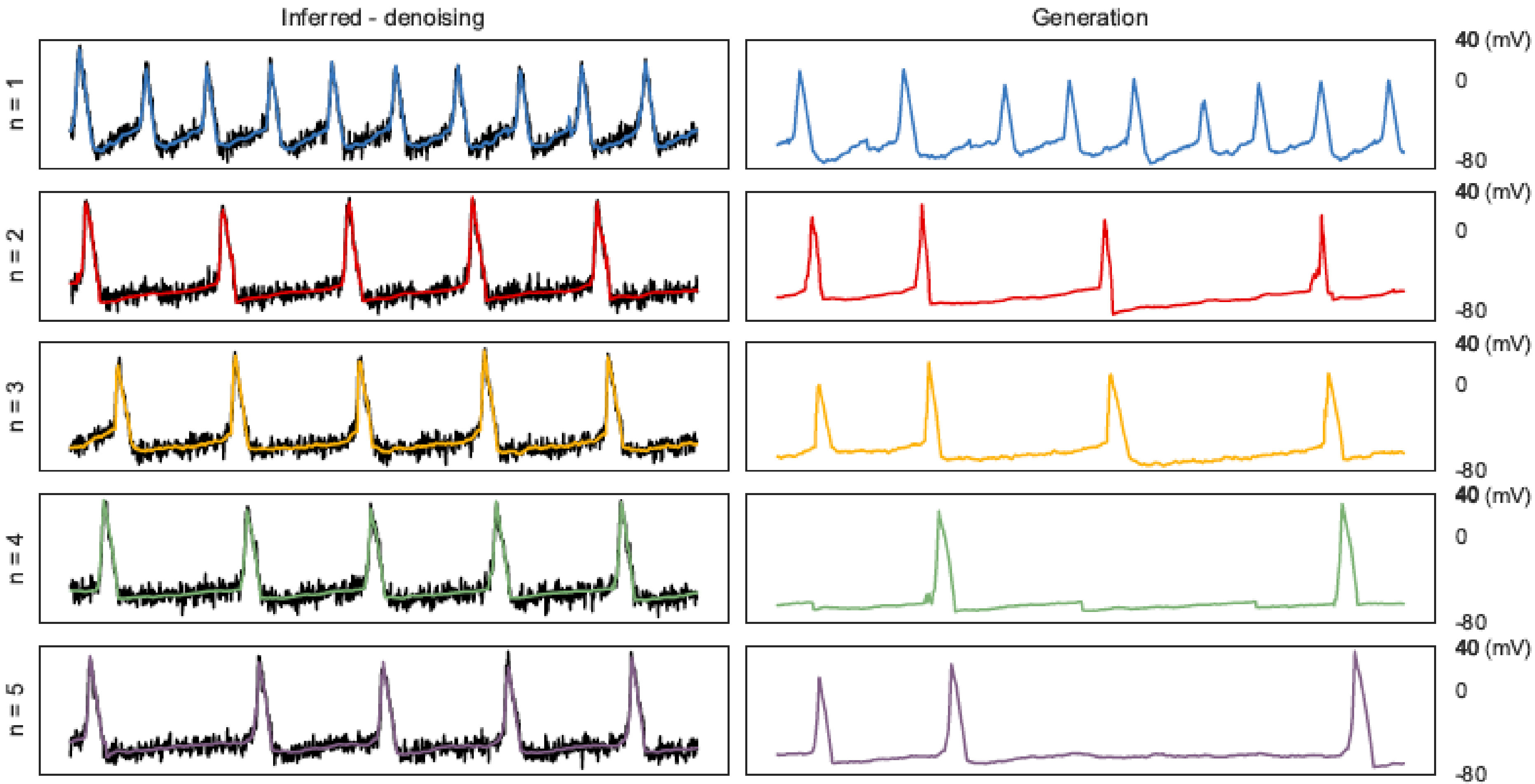


Sun et al 2019

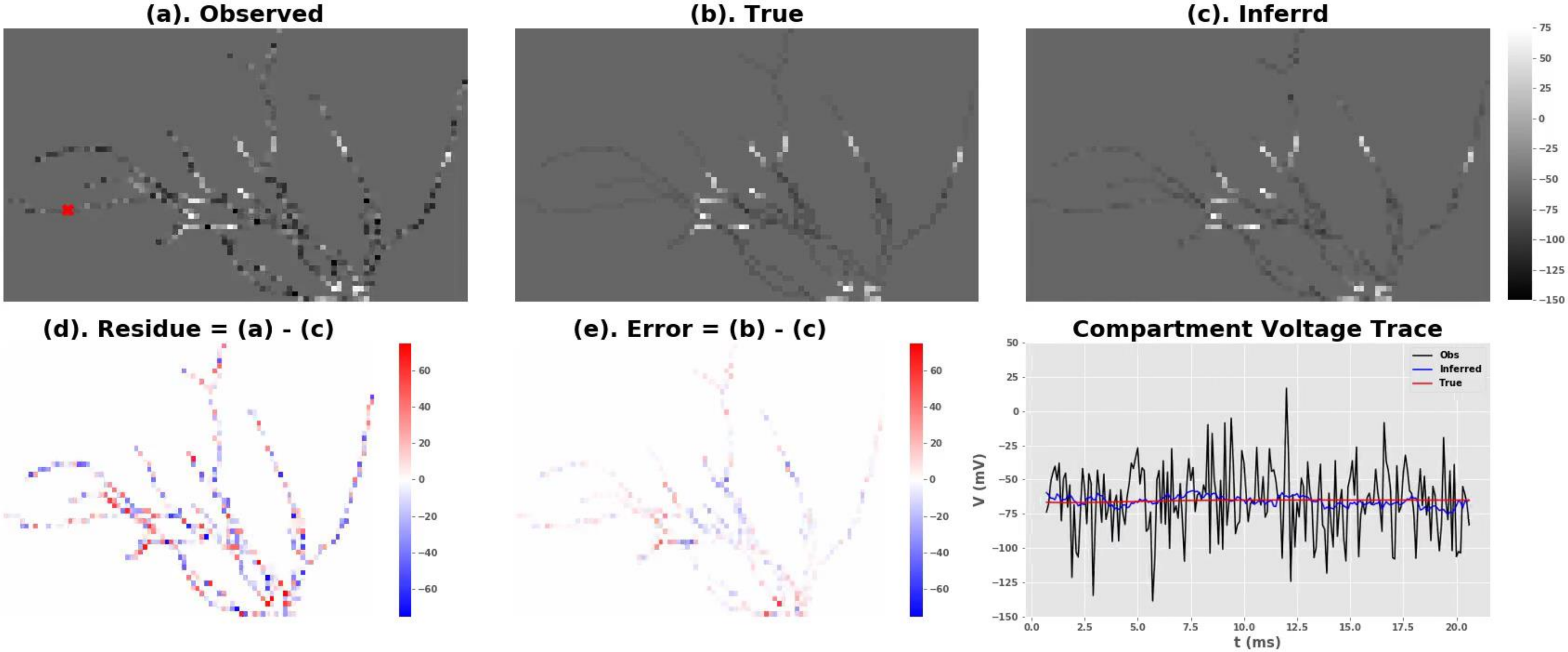


(a)
Observed Voltage & Inferred Continuous Latent State X1
V (mV)
20
0
−20
−40
−60
Observed voltage
CI
(b)
Inferred Continuous Latent State X2
(c)
Inferred Discrete Latent State Z
2
1
0
0
10
20
30
40
50
60
70
80
t (ms)
(d)
Inferred Dynamics (rSLDS)
X
V
(e)
Generated States
X
V

Multicompartment case



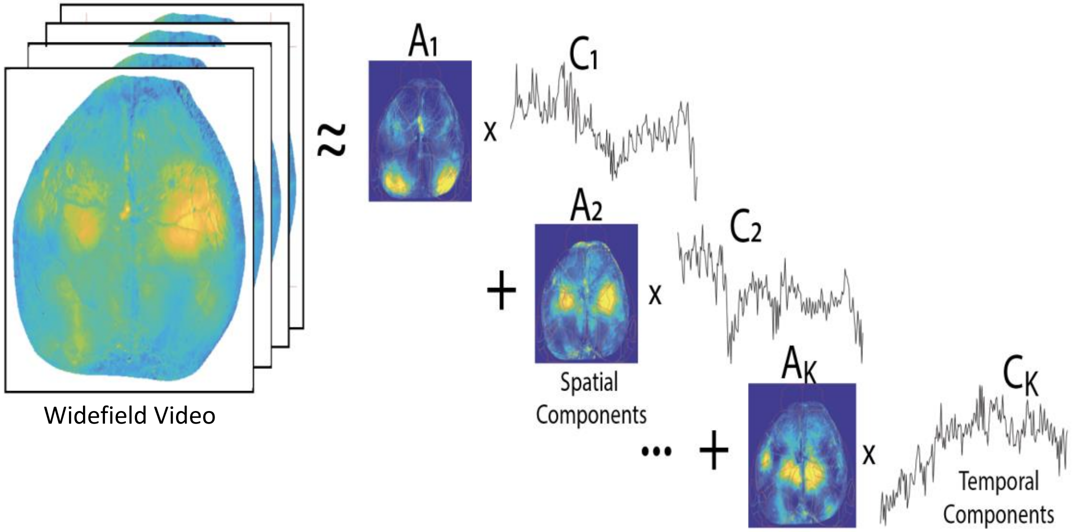
Sun et al 2019

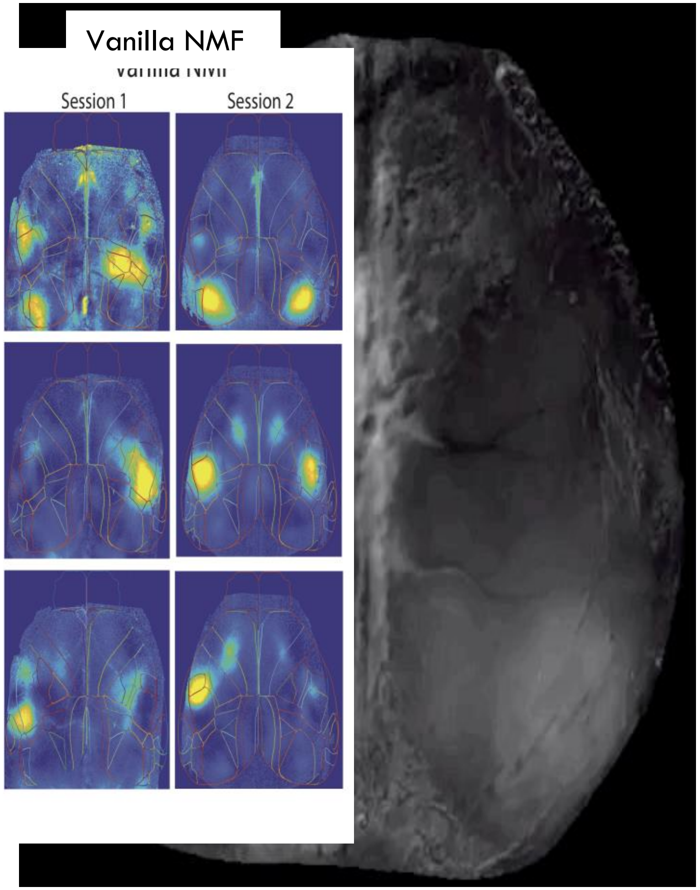


Sun et al 2019

What about widefield data?

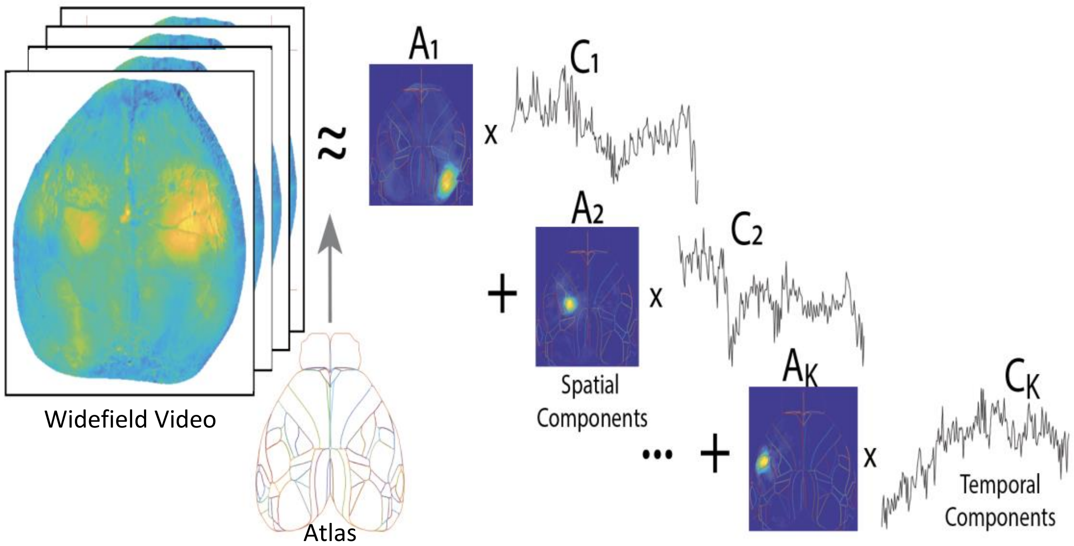
Vanilla Non-negative Matrix Factorization (NMF)

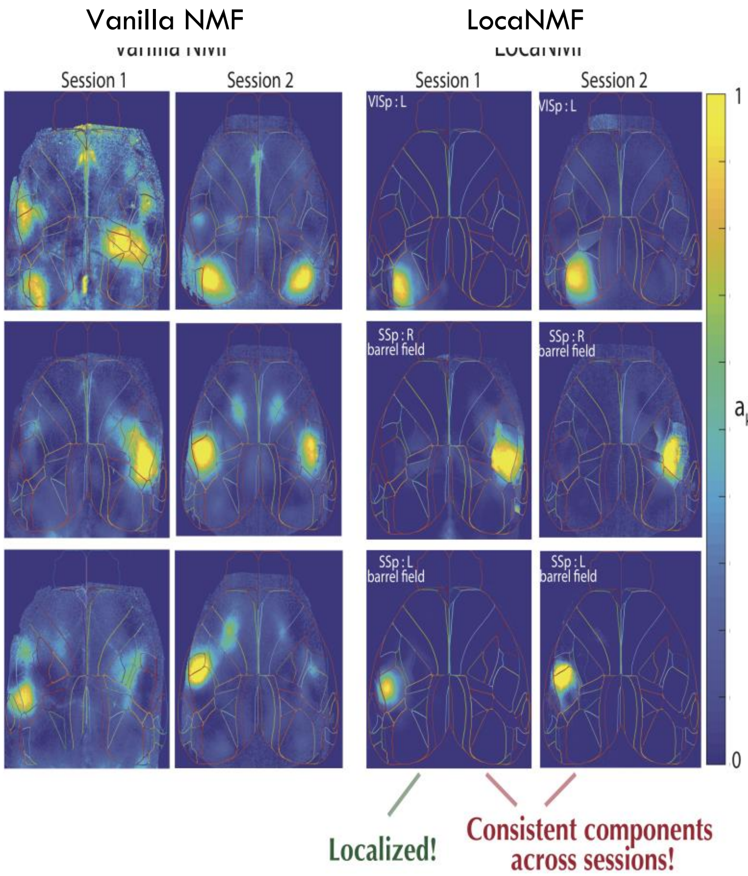




Reproducible and Interpretable Signal Extraction

LocaNMF : Localized Non-negative Matrix Factorization



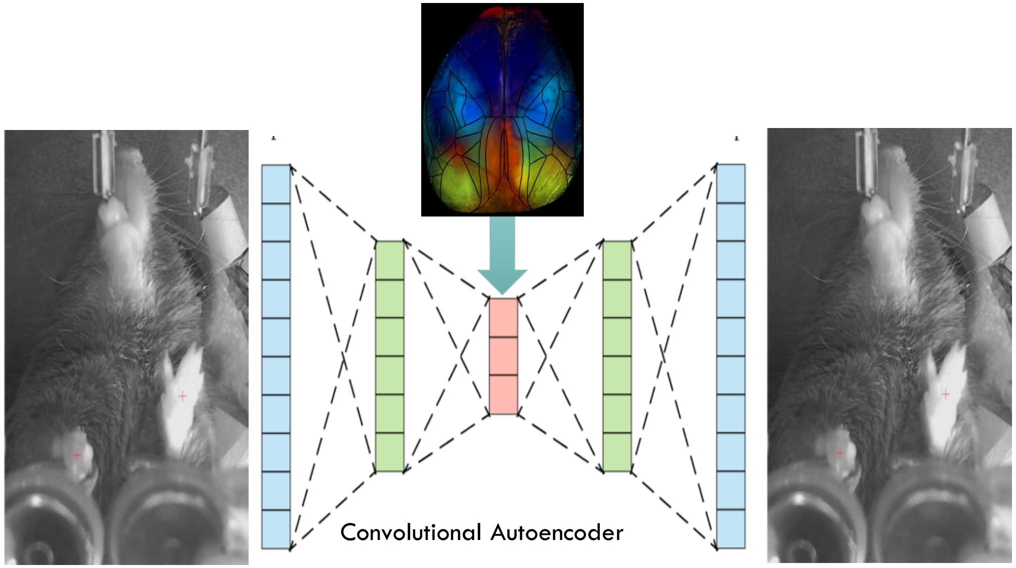


- Allows comparisons across sessions, mice, experimental conditions

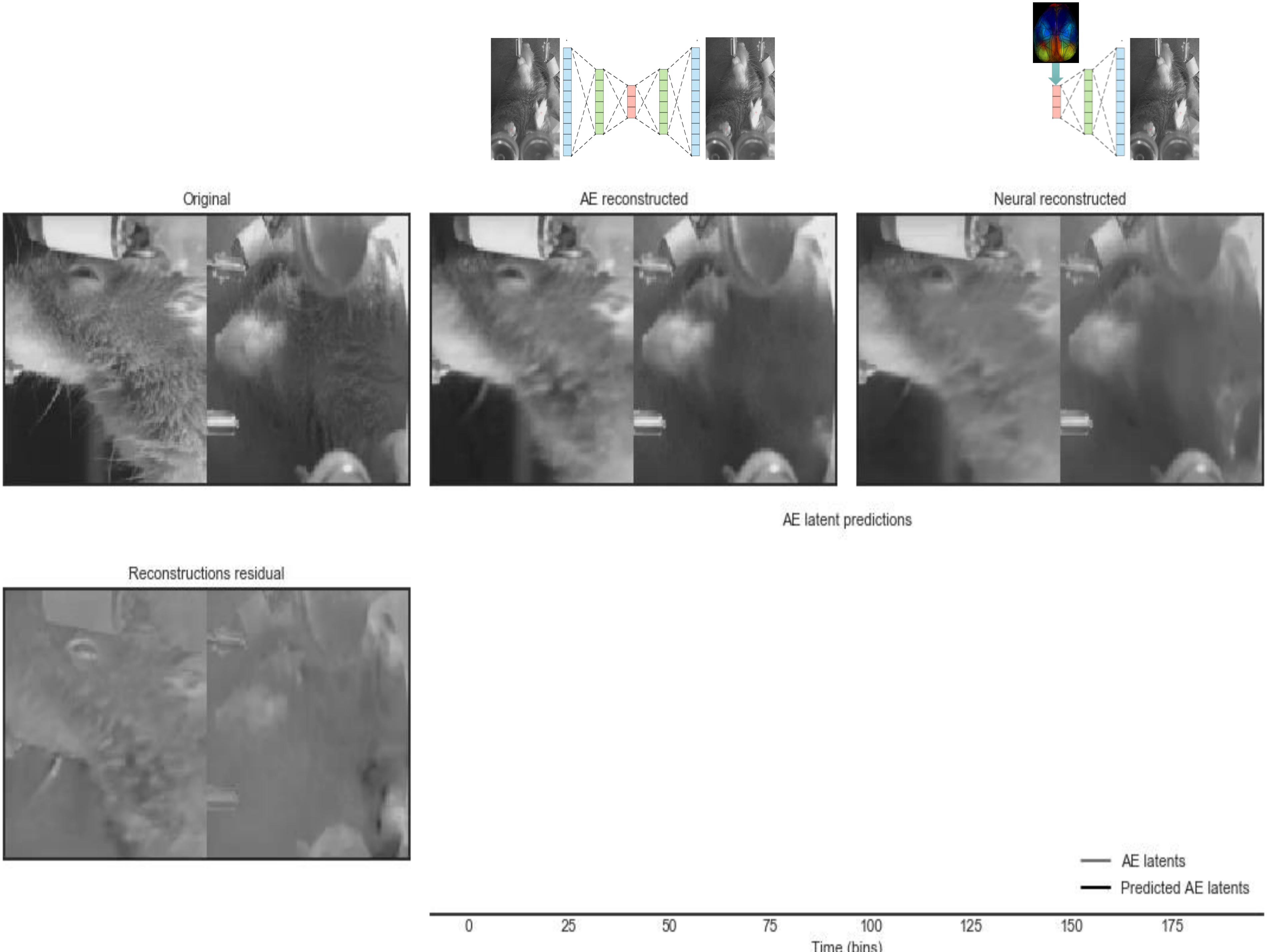
Saxena S., Kinsella I., Musall S., Kim S., Meszaros J., Thibodeaux D., Kim C., Cunningham J., **Hillman E.**, Churchland A., **Paninski L.**, *bioRxiv*, 2019

Compression and decoding of high dimensional behavior

- ❑ Dimensionality reduction on behavioral video
- ❑ Decode behavioral latents using LocaNMF components

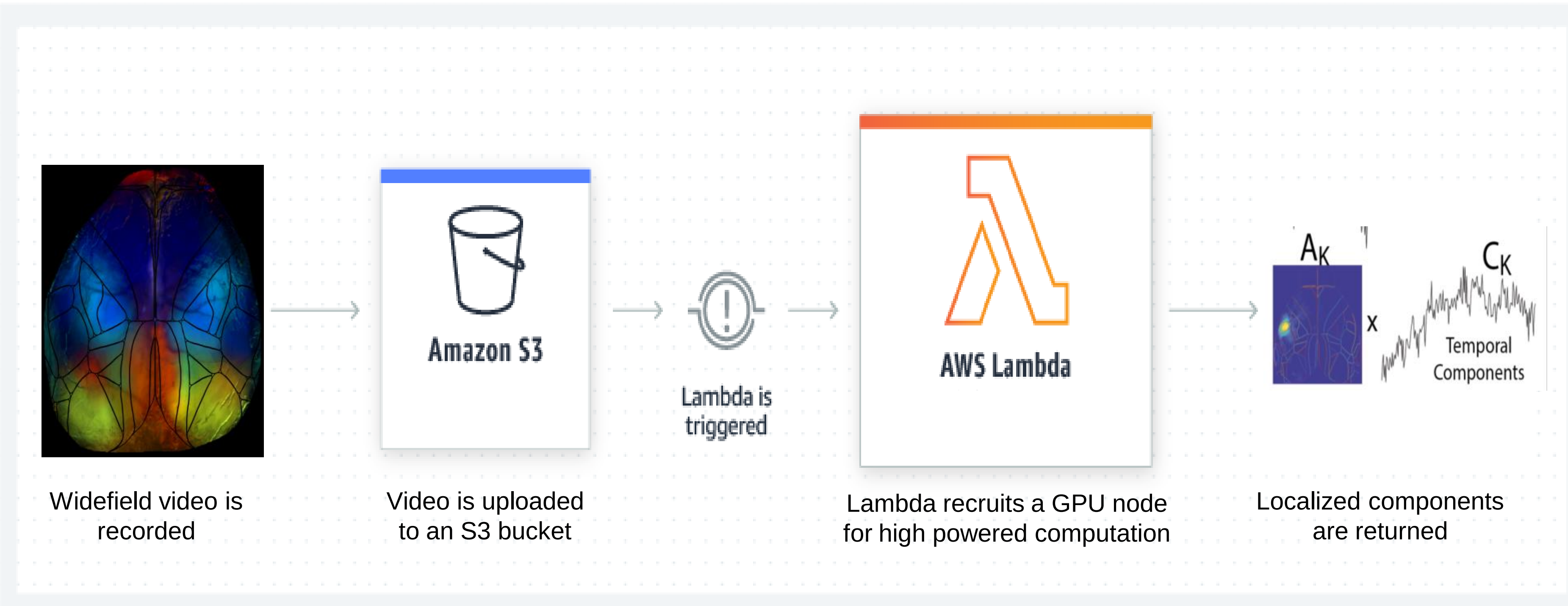


Matt Whiteway, Ella Batty, Shreya Saxena, Scott Linderman


Original
AE reconstructed
Neural reconstructed
AE latent predictions
Reconstructions residual
AE latents
Predicted AE latents
0
25
50
75
100
125
150
175
Time (bins)

AWS platform

- Drag-and-drop computing – easier for user (no installation headaches…)
- Scalable, no need to buy big hardware
- Easier for developers to manage (no installation headaches…)
- Cheaper for most applications



With: Taiga Abe, Ian Kinsella, John Cunningham

Concluding thoughts: some remaining bottlenecks

Not enough neural data scientists.

Data sharing remains depressingly rare and primitive. Imagine how much faster we could move if anyone in the world could easily access data from any lab… Need better infrastructure/support (make it easier to share) and stronger enforcement (make it harder not to share).

Software engineering – how to go from grad student prototype code to fully engineered, tested, easy-to-use, open-source, scalable, robust, well-supported, sustainable tools? Good steps from Allen, CZI, Simons/Flatiron, but we lack a good model for this in academic setting.

Thanks!

Eftychios Pnevmatikakis – CNMF
Andrea Giovanucci – OnACID, CaImAn
P.C. Zhou – CNMF-E, EM+2p
Kelly Buchanan, Ian Kinsella, Ding Zhou – compression, denoising
Johannes Friedrich – OASIS deconvolution, multiscale imaging
Shreya Saxena - LocaNMF
Matt Whiteway, Dan Biderman – behavioral video decoding
Daniel Soudry – shotgun imaging
Xuexin Wei – spike+gamma model for post-deconvolved activity
Ruoxi Sun, Scott Linderman – spatiotemporal voltage denoising

Cat Mitelut, Peter Lee, Hooshmand Razaghi, Nishchal Dette, Eduardo Blancas, Denis Turcu, Shenghao Wu, David Carlson – spike sorting
Ella Batty, Nikhil Parthasarathy - decoding

Collaborators: Ji, Chichilnisky, Churchland, Datta, Sabatini, Cohen, Ahrens, Hillman, Losonczy, Tolias, Seung, Yuste, Peterka, Cunningham, Hobert labs, plus IBL, Q-State Biosciences and Allen Institute