Unsupervised discovery of neural sequences in large-scale recordings



Emily Mackevicius



Andrew Bahle



Alex Williams



Shijie Gu



Natasha Denissenko



Mark Goldman

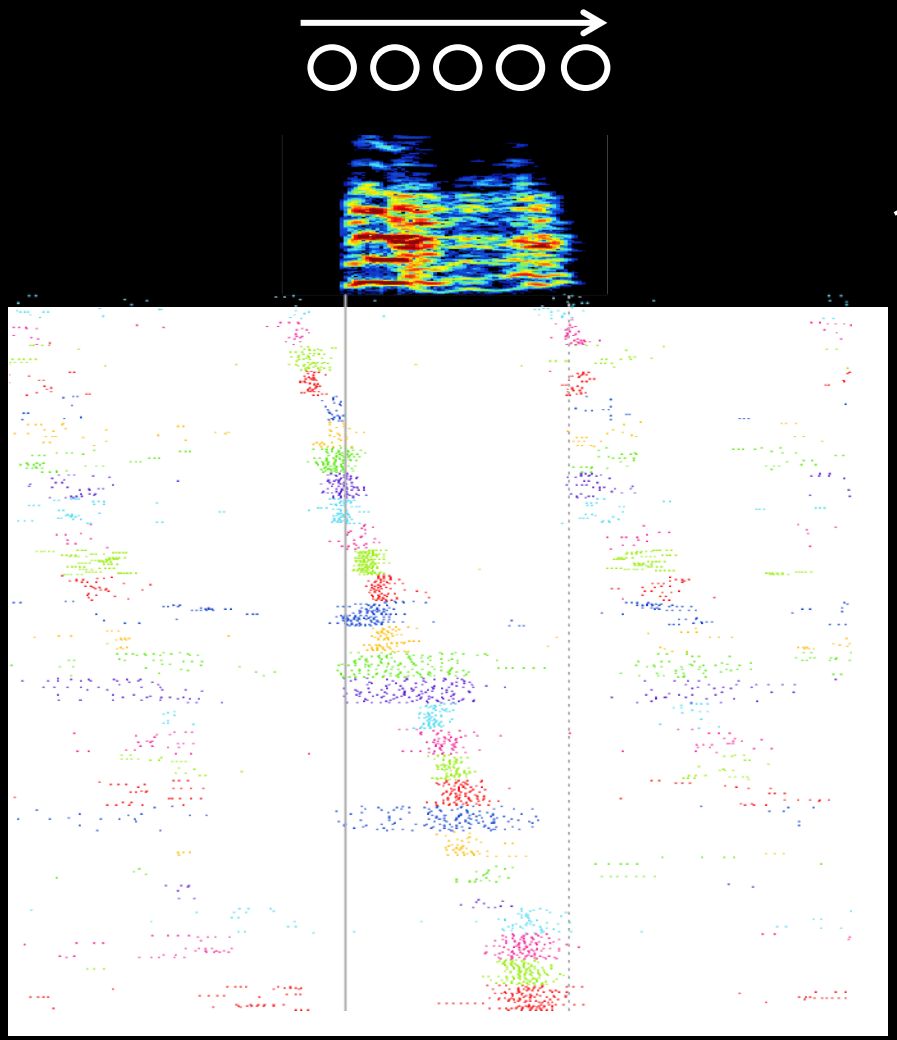


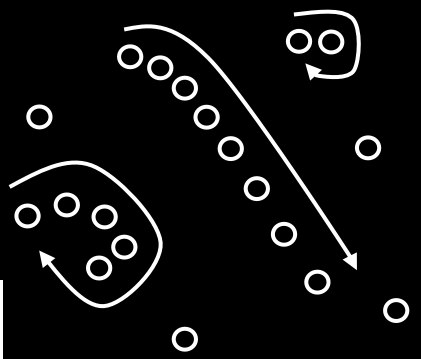
Michale Fee

github.com/Feelab/seqNMF

BioRxiv preprint

Neurons form sequences

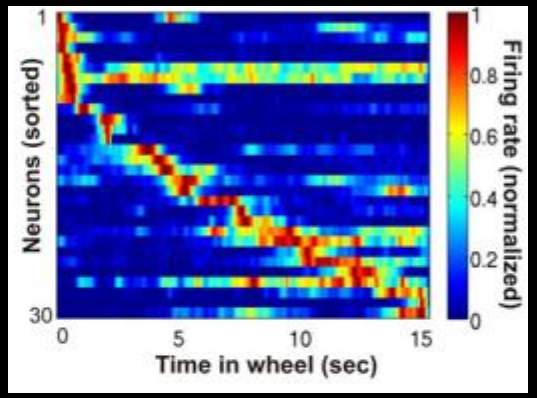




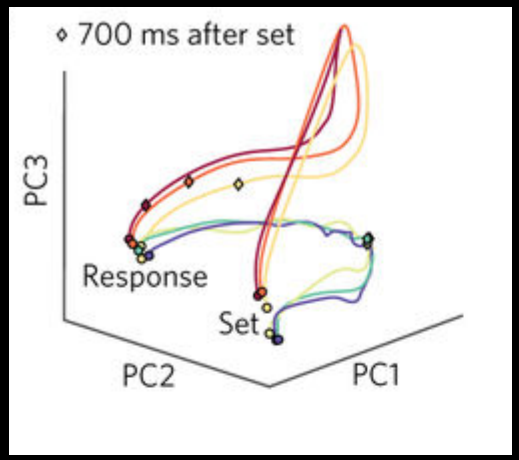
Songbird HVC (premotor cortex),
Okubo et al. 2015

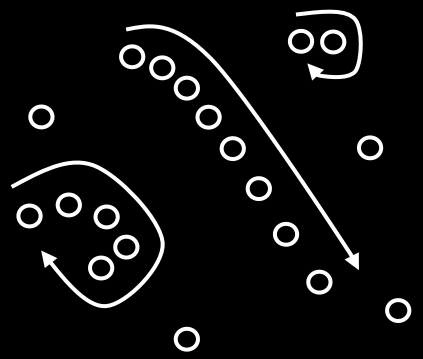
Neurons form sequences

Rat hippocampus,
Pastalkova et al. 2008

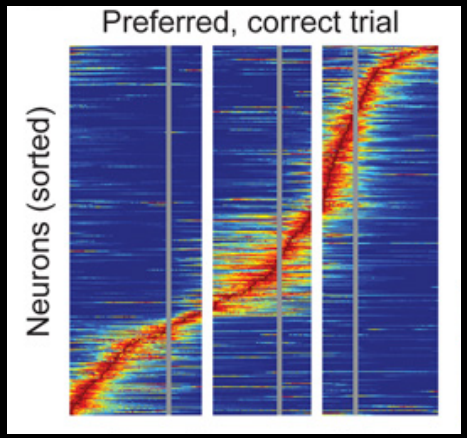


Macaque medial frontal cortex,
Wang et al. 2018

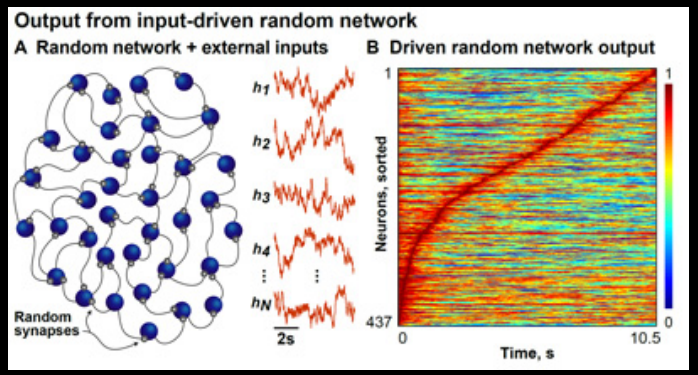




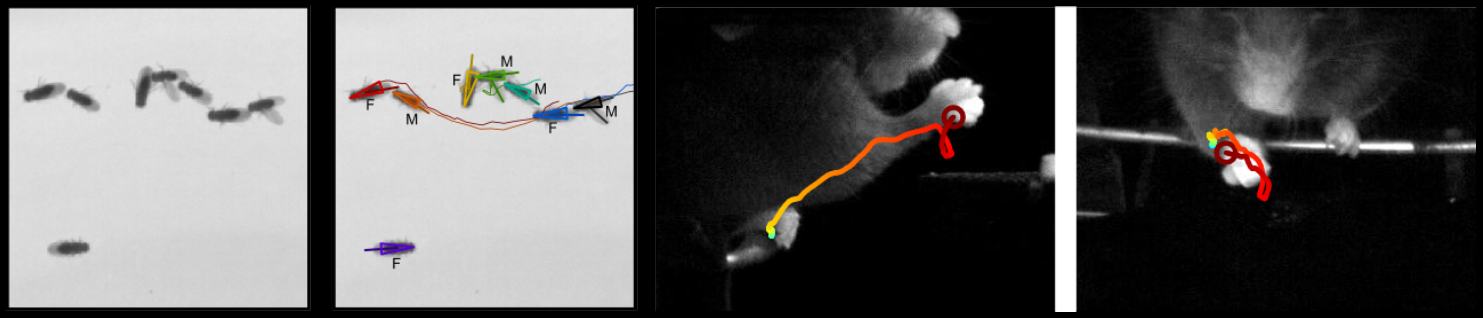
Rat posterior parietal cortex,
Harvey et al. 2012



Simulated recurrent network,
Rajan et al. 2017

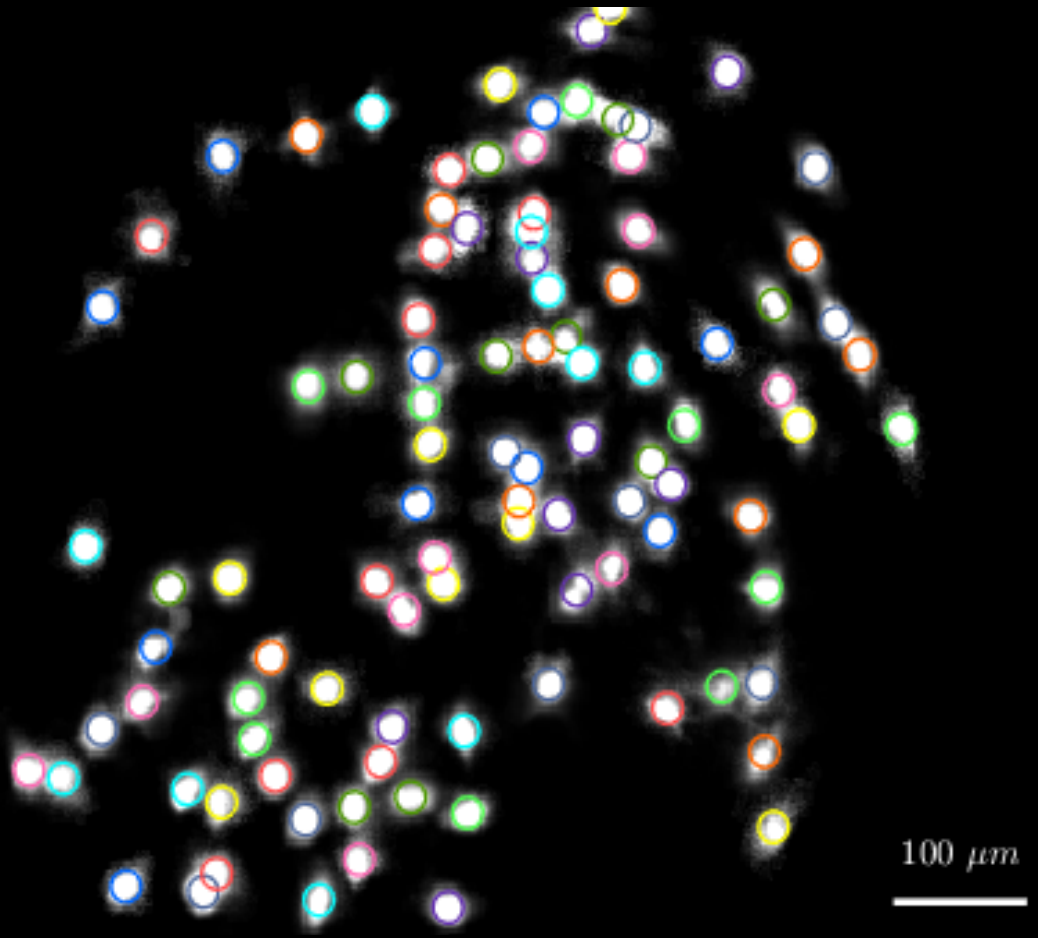


Behavioral sequences are also relevant to neuroscience



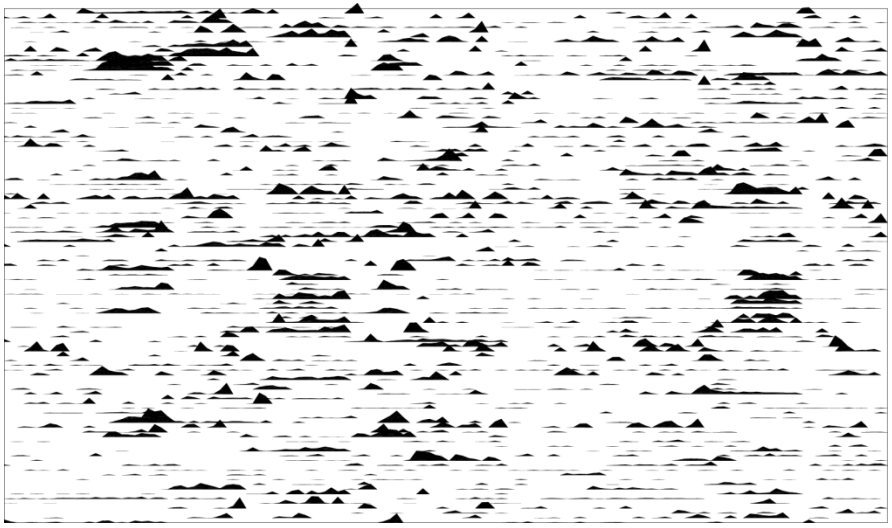
Kristin Branson lab,
see JAABA behavioral analysis tool


100 μm



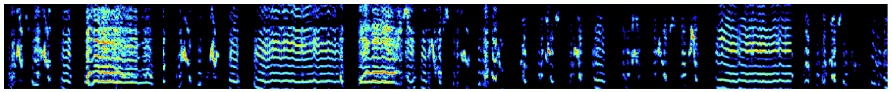
CNMF_E cell extraction (PC Zhou, Paninski lab, with Shijie Gu)

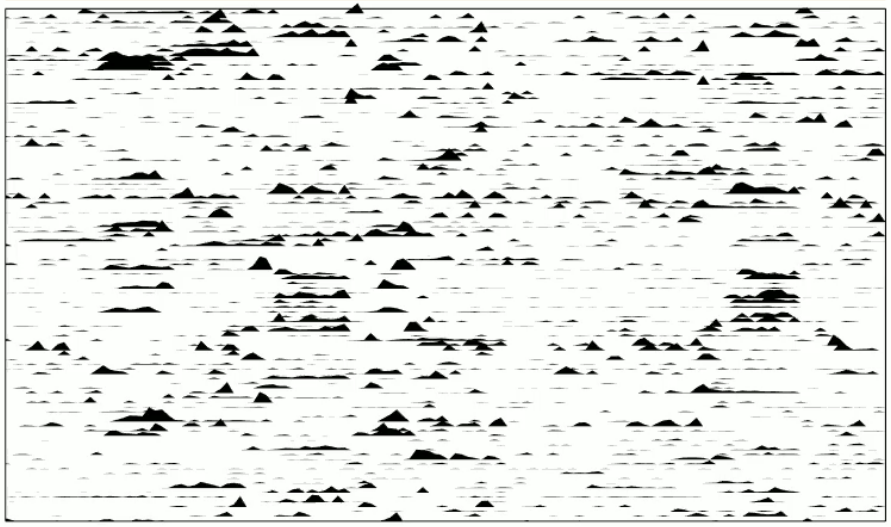
Neurons

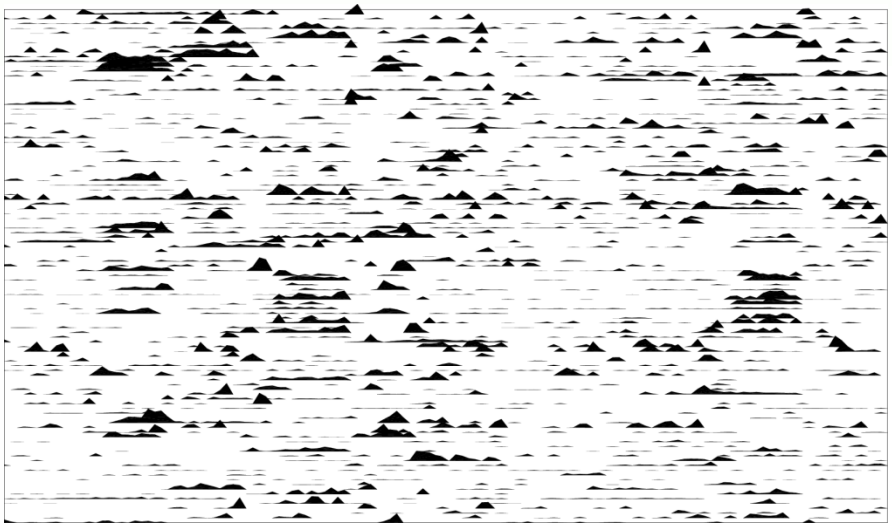


Time

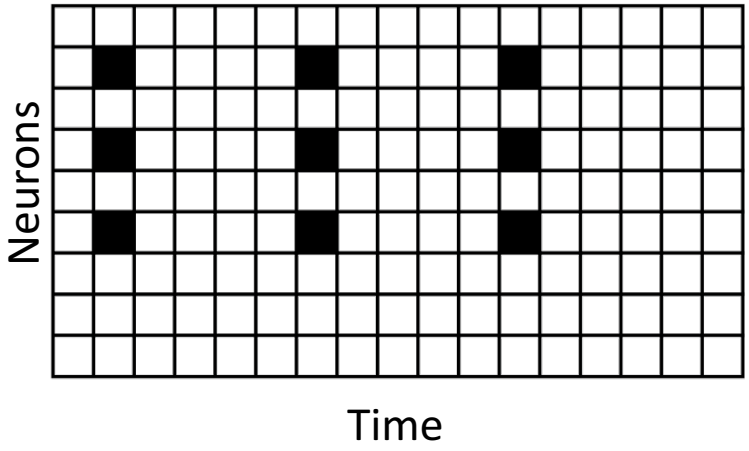
Neurons
(re-sorted)







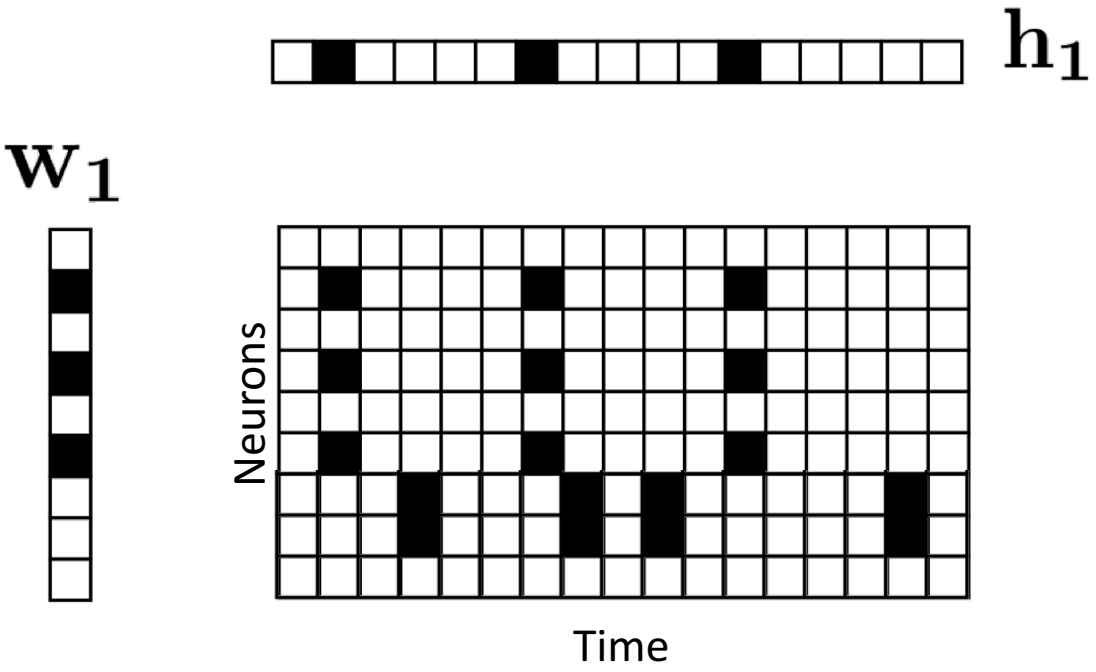
Non-negative matrix factorization (NMF)



Non-negative matrix factorization (NMF)

(like PCA/SVD, but factors must be positive)

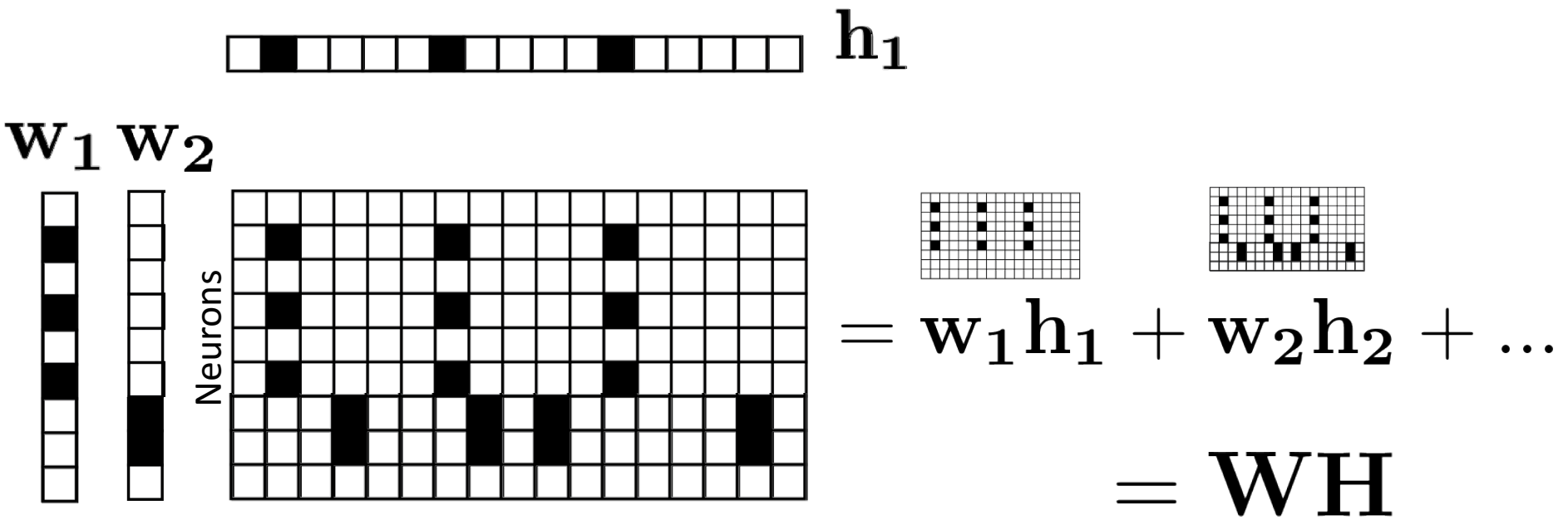
Non-negative matrix factorization (NMF)



Non-negative matrix factorization (NMF)

(like PCA/SVD, but factors must be positive)

Non-negative matrix factorization (NMF)



Non-negative matrix factorization (NMF)

(like PCA/SVD, but factors must be positive)

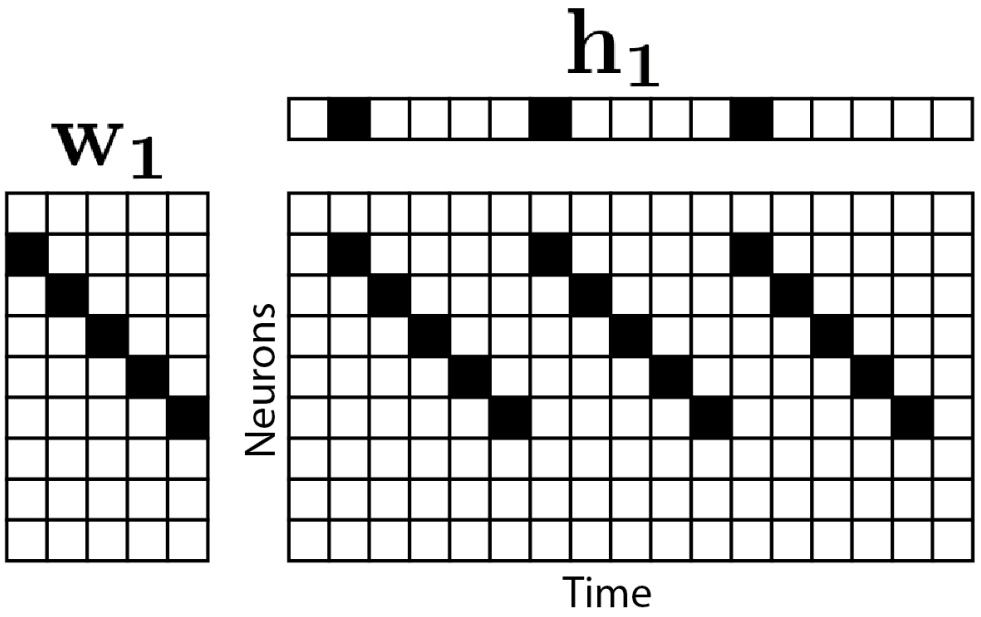
Non-negative matrix factorization (NMF)

$$\mathbf{W} \leftarrow \mathbf{W} \times \frac{\mathbf{X}\mathbf{H}^{\top}}{\mathbf{W}\mathbf{H}\mathbf{H}^{\top}}$$

$$\mathbf{H} \leftarrow \mathbf{H} \times \frac{\mathbf{W}^{\top}\mathbf{X}}{\mathbf{W}^{\top}\mathbf{W}\mathbf{H}}$$

Lee and Seung 1999

Convolutional NMF



$$+ \mathbf{w_2} \circledast \mathbf{h_2} + \cdots$$

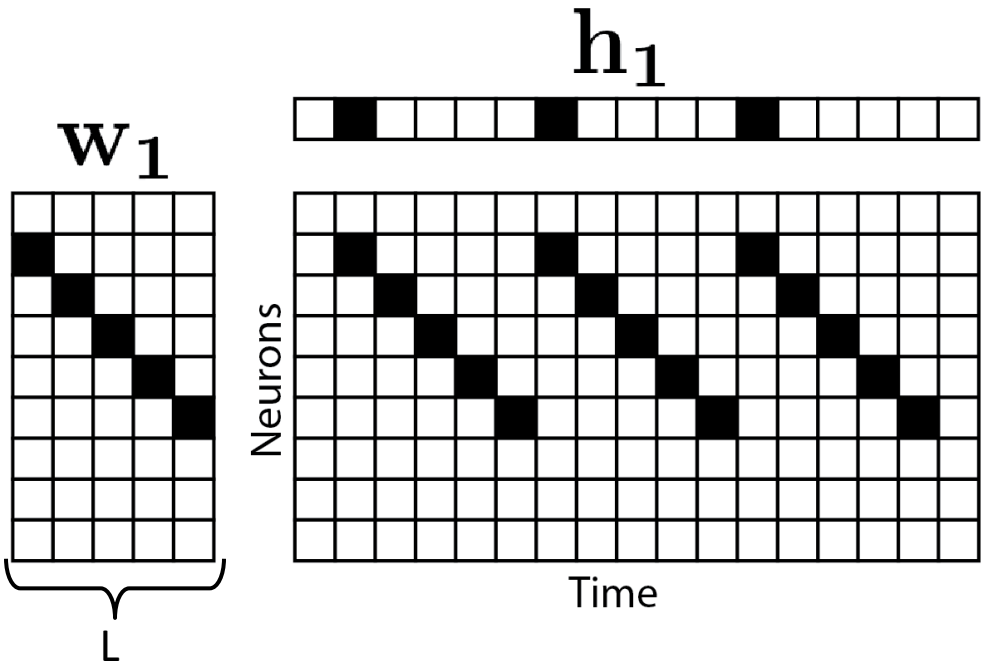
$$= \mathbf{W} \circledast \mathbf{H}$$

Easy to fit

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 \right)$$

Smaragdis, 2004, 2007

Convolutional NMF

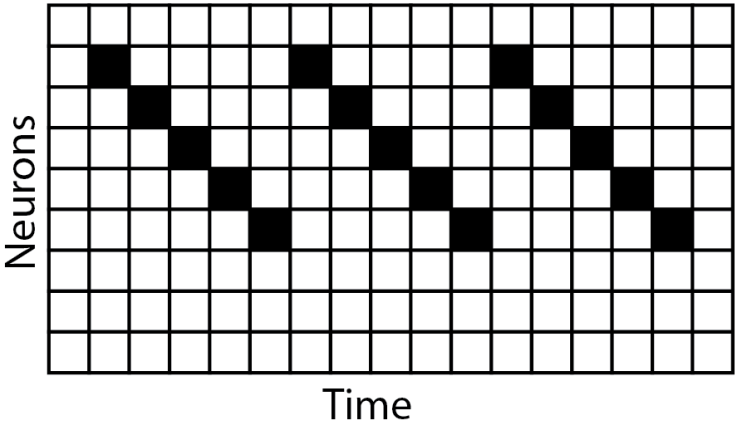


$$\mathbf{H} \leftarrow \mathbf{H} \times \frac{\sum_\ell \mathbf{W}_{..\ell}^\top \overleftarrow{\mathbf{X}}^\ell}{\sum_\ell \mathbf{W}_{..\ell}^\top \overleftarrow{\tilde{\mathbf{X}}}^\ell}$$

$$\mathbf{W}_{..\ell} \leftarrow \mathbf{W}_{..\ell} \times \frac{\mathbf{X} \overrightarrow{\mathbf{H}}^{\ell\top}}{\tilde{\mathbf{X}} \overrightarrow{\mathbf{H}}^{\ell\top}}$$

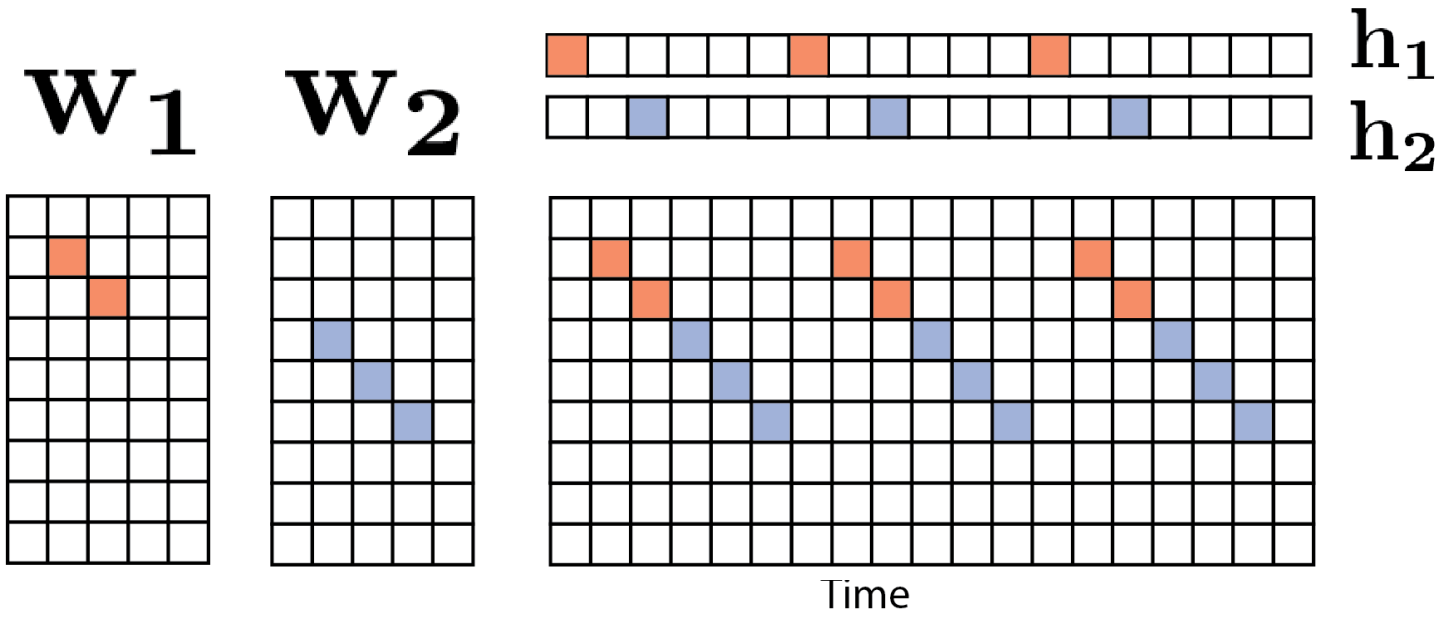
Smaragdis, 2004, 2007

In practice … redundant factors



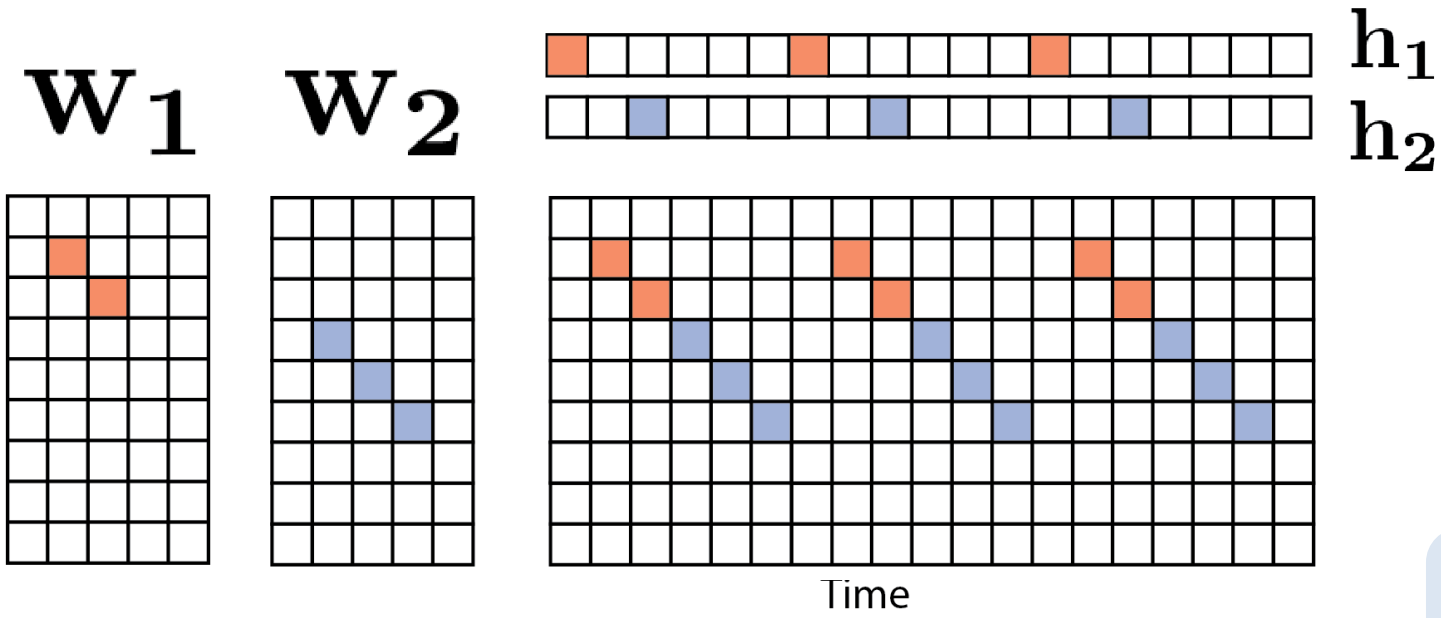
$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 \right)$$

In practice … redundant factors



$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 \right)$$

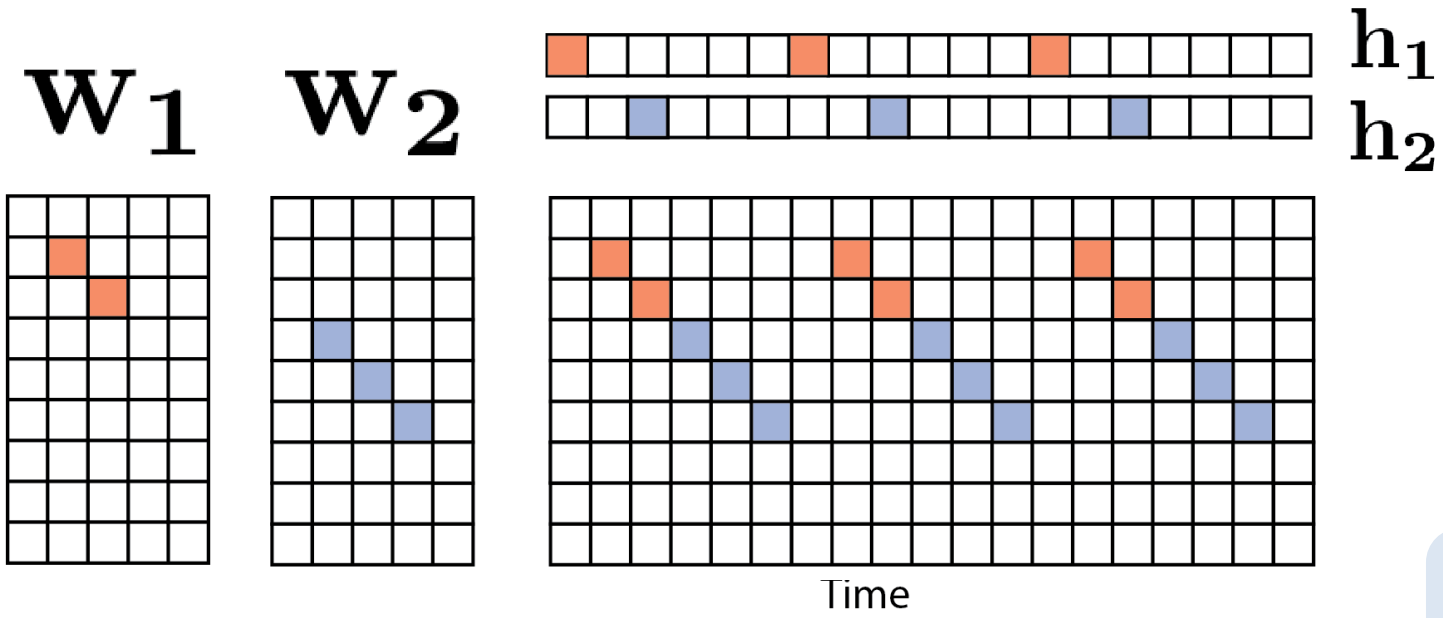
In practice … redundant factors



Penalty for correlations

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \overbrace{\lambda \mathcal{R}} \right)$$

In practice … redundant factors

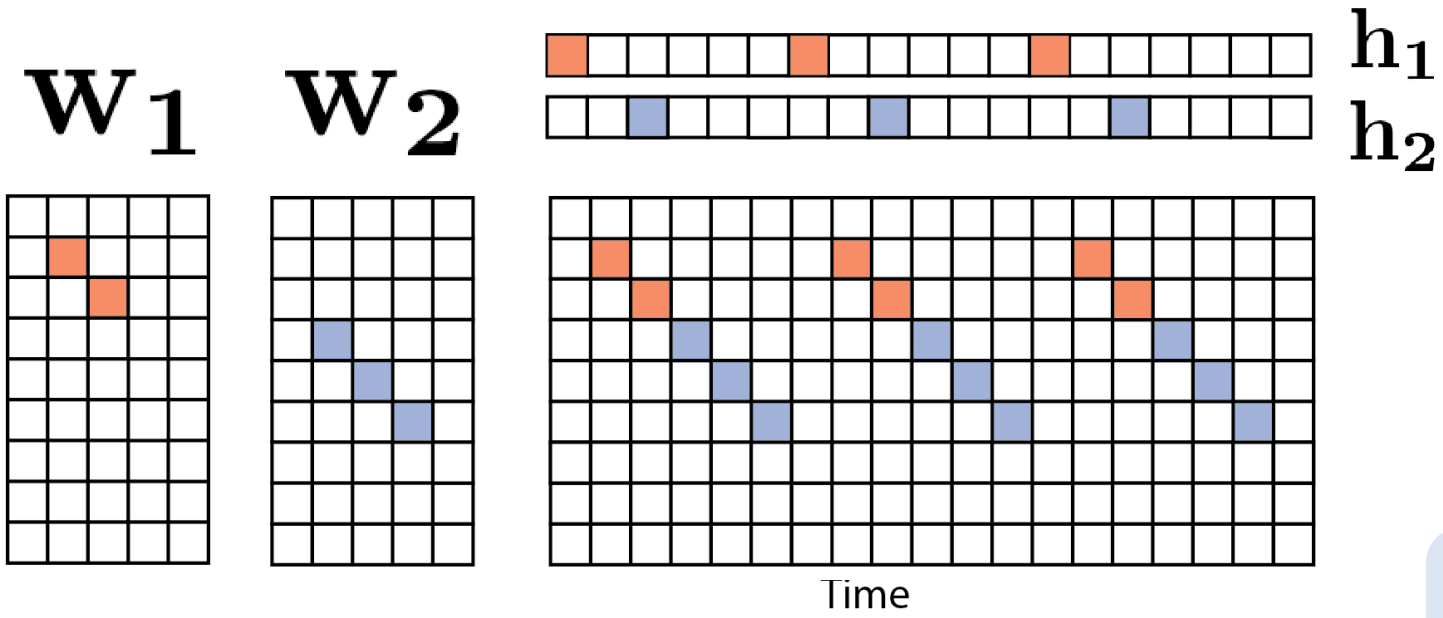


Penalty for correlations

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \overbrace{\lambda \mathcal{R}} \right)$$

$$\mathcal{R} = ||\mathbf{H}\mathbf{H}^\top||$$

In practice … redundant factors

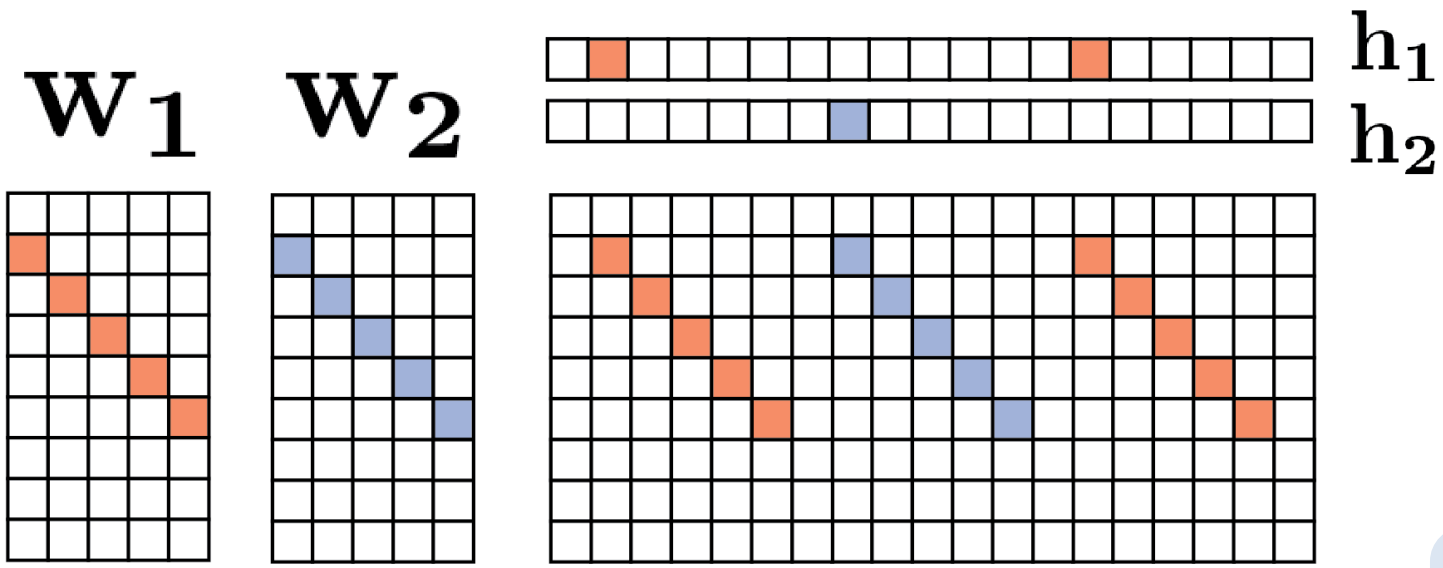


Penalty for correlations

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \overbrace{\lambda \mathcal{R}} \right)$$

$$\mathcal{R} = ||\mathbf{H}\mathbf{S}\mathbf{H}^\top||$$

In practice … redundant factors

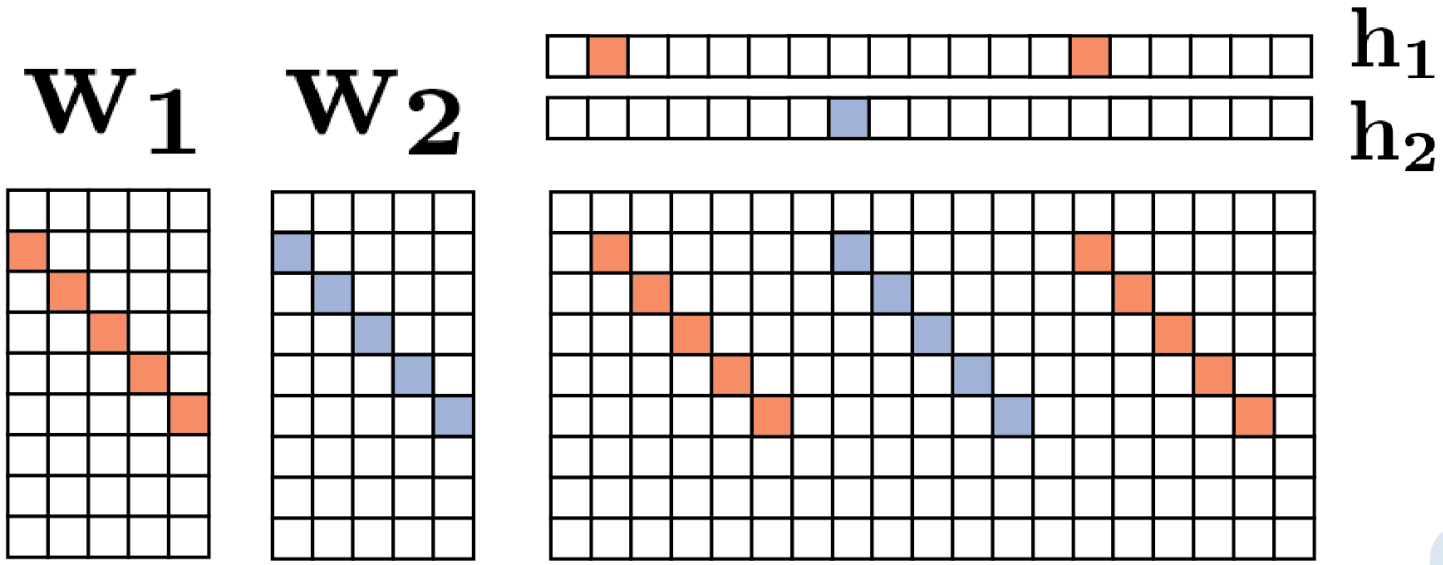


Penalty for correlations

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \overbrace{\lambda \mathcal{R}} \right)$$

$$\mathcal{R} = ||\mathbf{H}\mathbf{S}\mathbf{H}^\top||$$

In practice … redundant factors



Penalty for correlations

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \overbrace{\lambda \mathcal{R}} \right)$$

$$\mathcal{R} = ||\mathbf{W} \overset{\top}{\circledast} \mathbf{X}\mathbf{S}\mathbf{H}^\top||$$

Simple multiplicative update rules

$$\mathbf{W}_{..\ell} \leftarrow \mathbf{W}_{..\ell} \times \frac{\mathbf{X}\left(\overset{\ell\rightarrow}{\mathbf{H}}\right)^{\top}}{\widetilde{\mathbf{X}}\left(\overset{\ell\rightarrow}{\mathbf{H}}\right)^{\top} + \lambda \overset{\leftarrow\ell}{\mathbf{X}}\mathbf{S}\mathbf{H}^{\top}(1-\mathbf{I})}$$

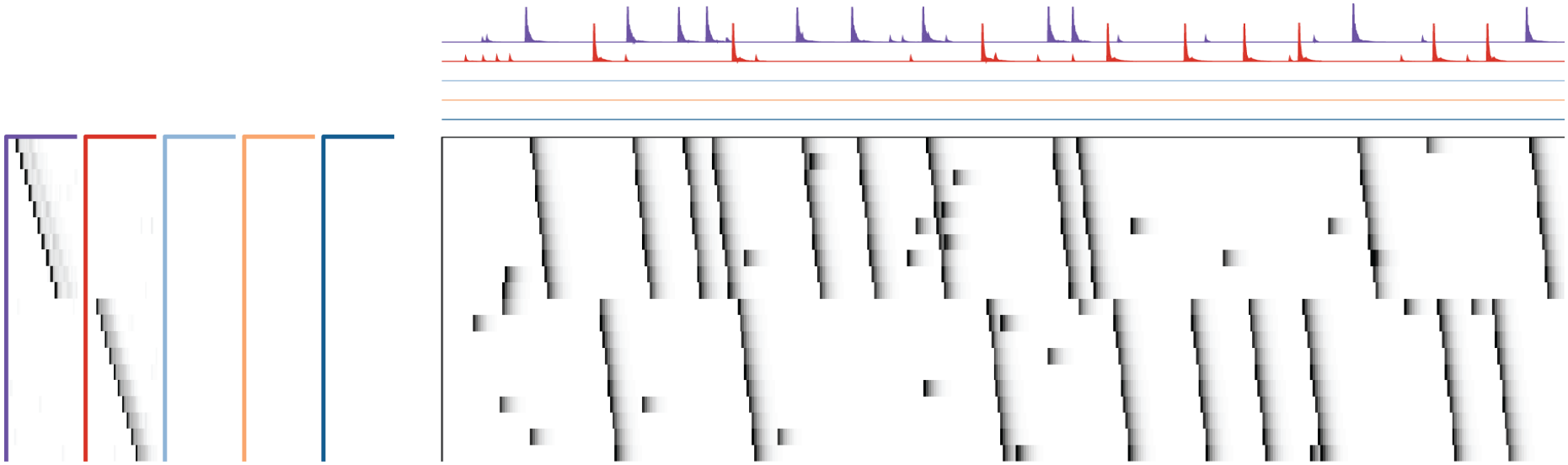
$$\mathbf{H} \leftarrow \mathbf{H} \times \frac{\overset{\top}{\mathbf{W}} \circledast \mathbf{X}}{\overset{\top}{\mathbf{W}} \circledast \widetilde{\mathbf{X}} + \lambda(1-\mathbf{I})(\overset{\top}{\mathbf{W}} \circledast \mathbf{X}\mathbf{S})}$$

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \lambda \mathcal{R} \right)$$

$$\mathcal{R} = ||\mathbf{W} \overset{\top}{\circledast} \mathbf{X}\mathbf{S}\mathbf{H}^{\top}||$$

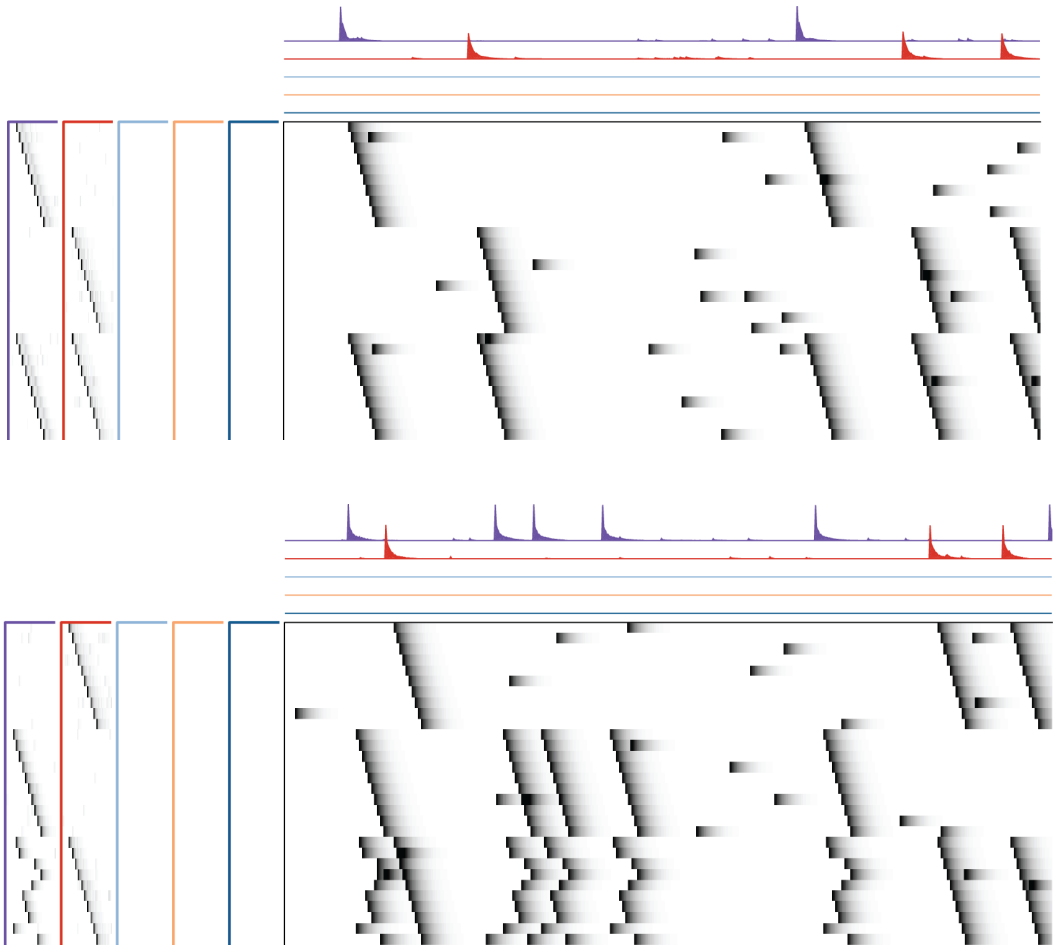
Mackevicius and Bahle, et al., 2018

Testing seqNMF on simulated sequences



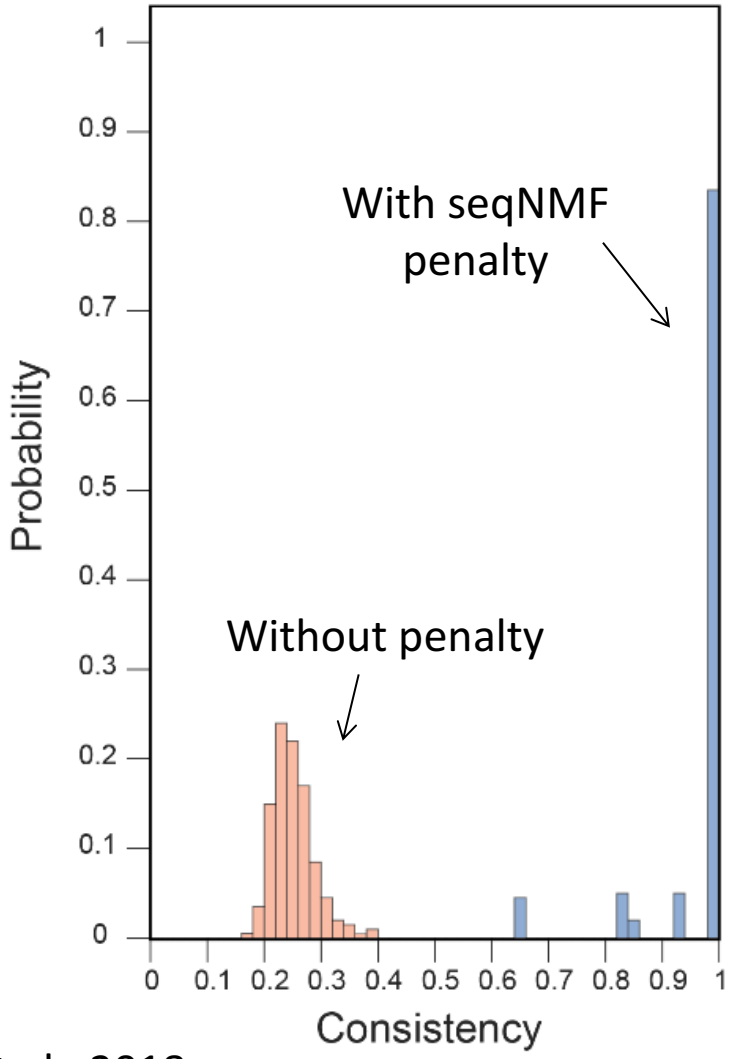
Mackevicius and Bahle, et al., 2018

Testing seqNMF on simulated sequences



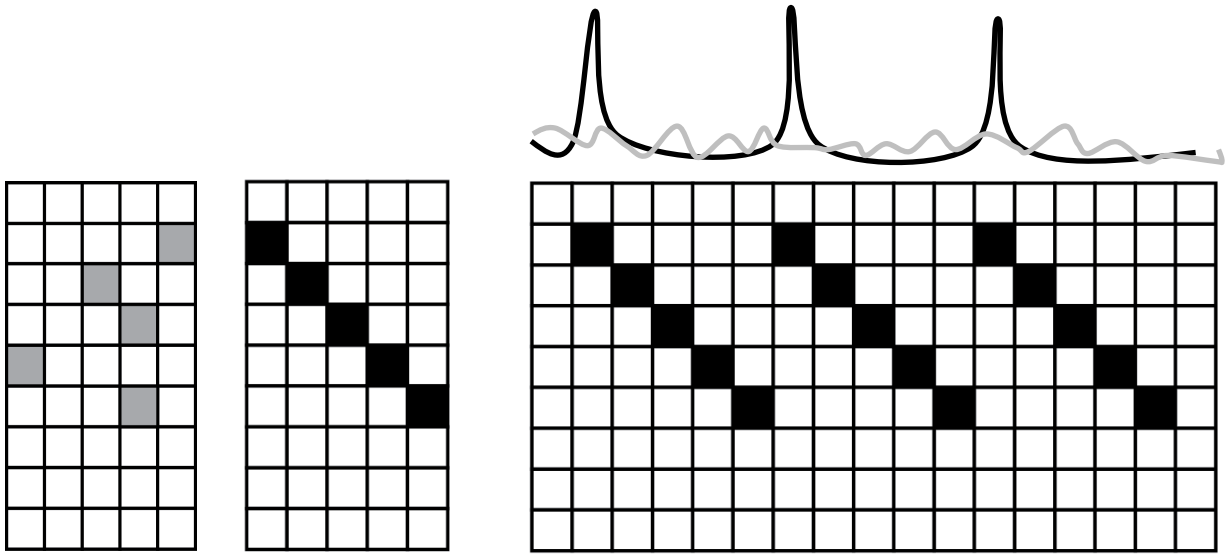
Mackevicius and Bahle, et al., 2018

SeqNMF factorizations are highly consistent



Mackevicius and Bahle, et al., 2018

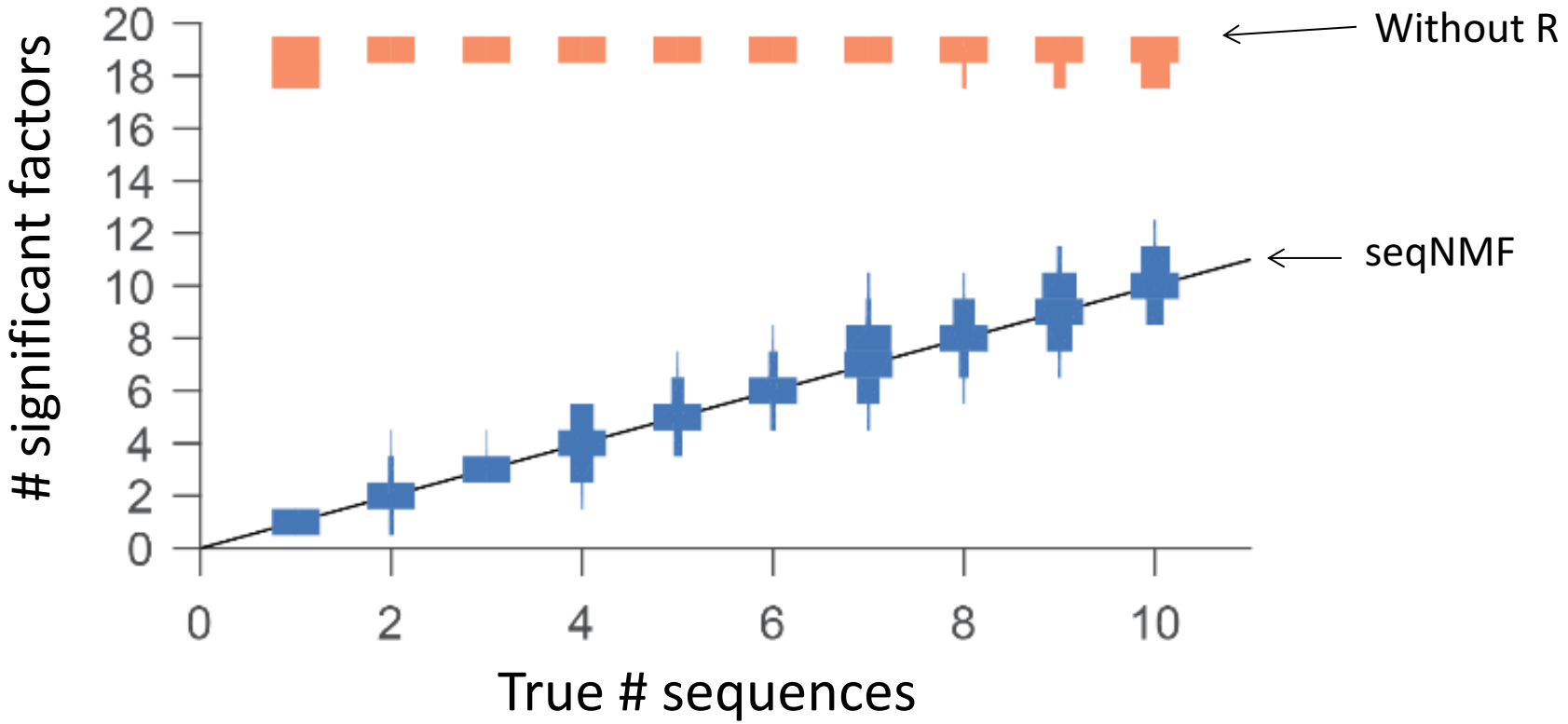
Testing significance of each factor on held-out data



Many moments of high overlap, compared to null

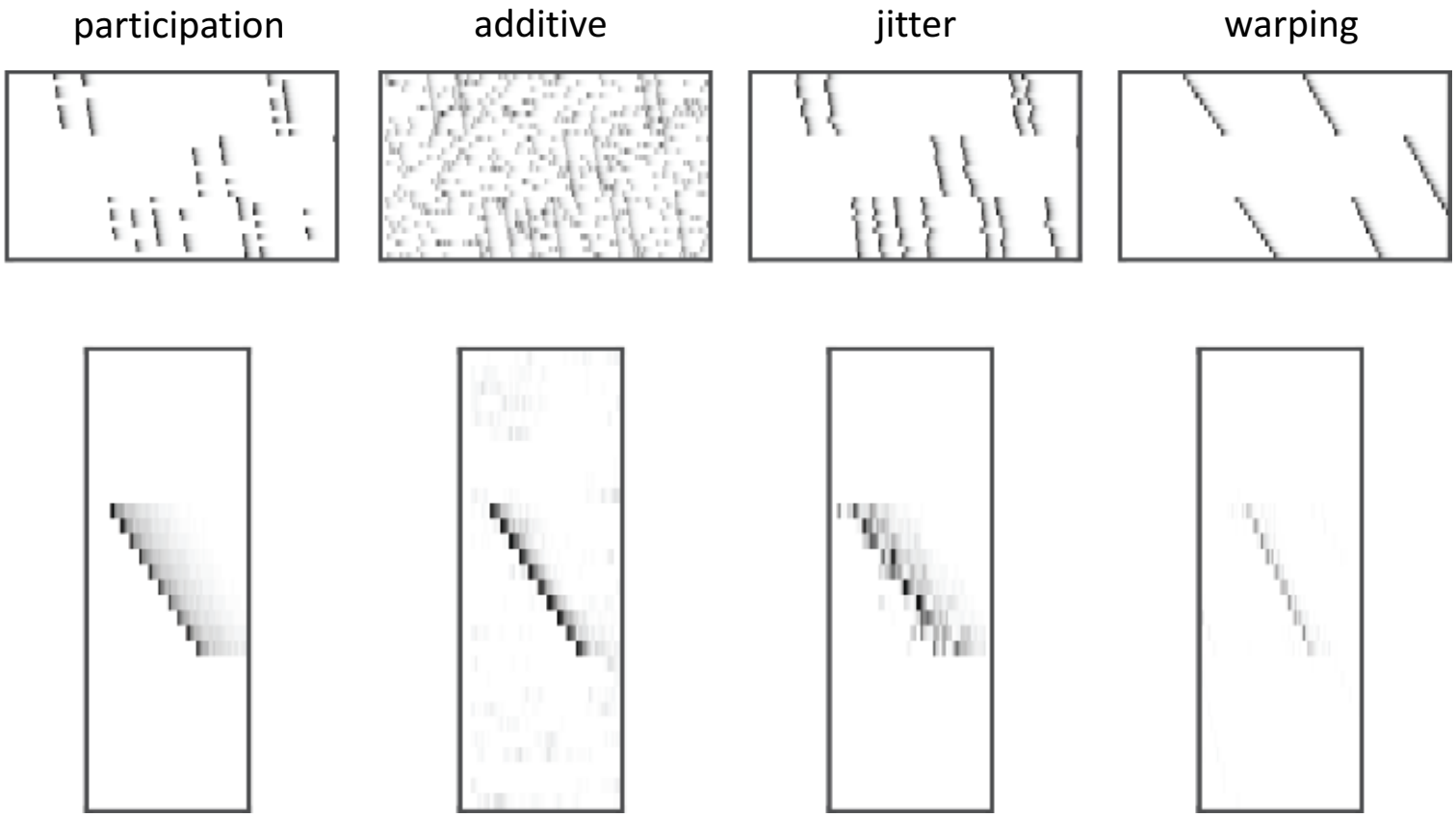
Mackevicius and Bahle, et al., 2018

SeqNMF discovers the correct number of sequences

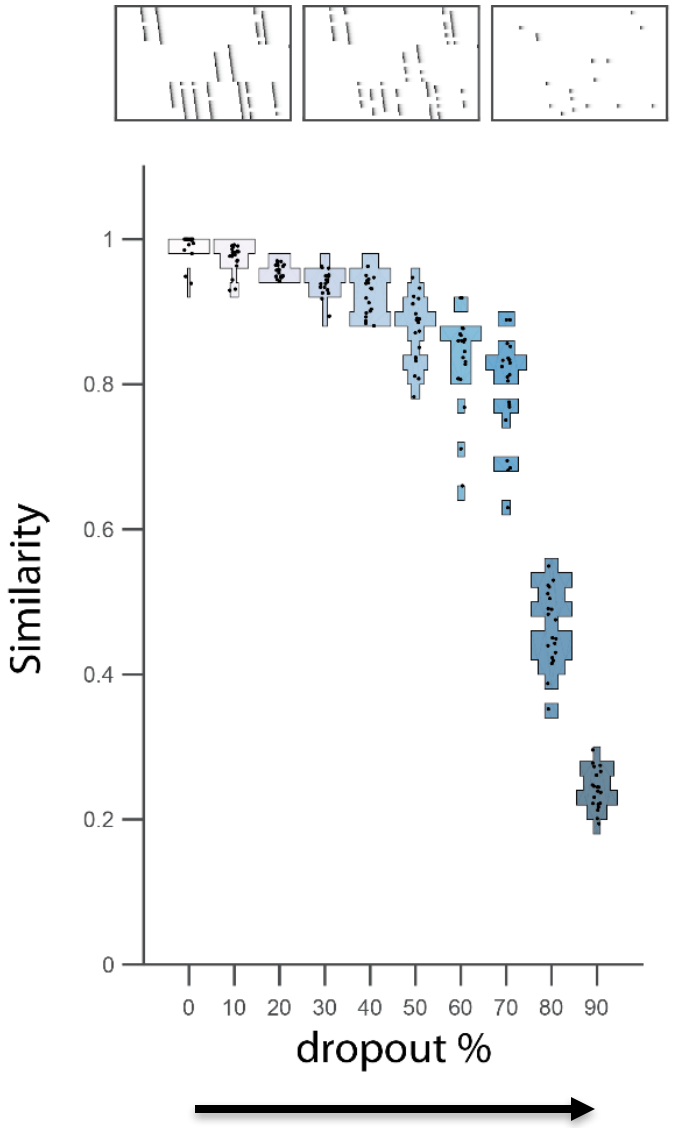


Mackevicius and Bahle, et al., 2018

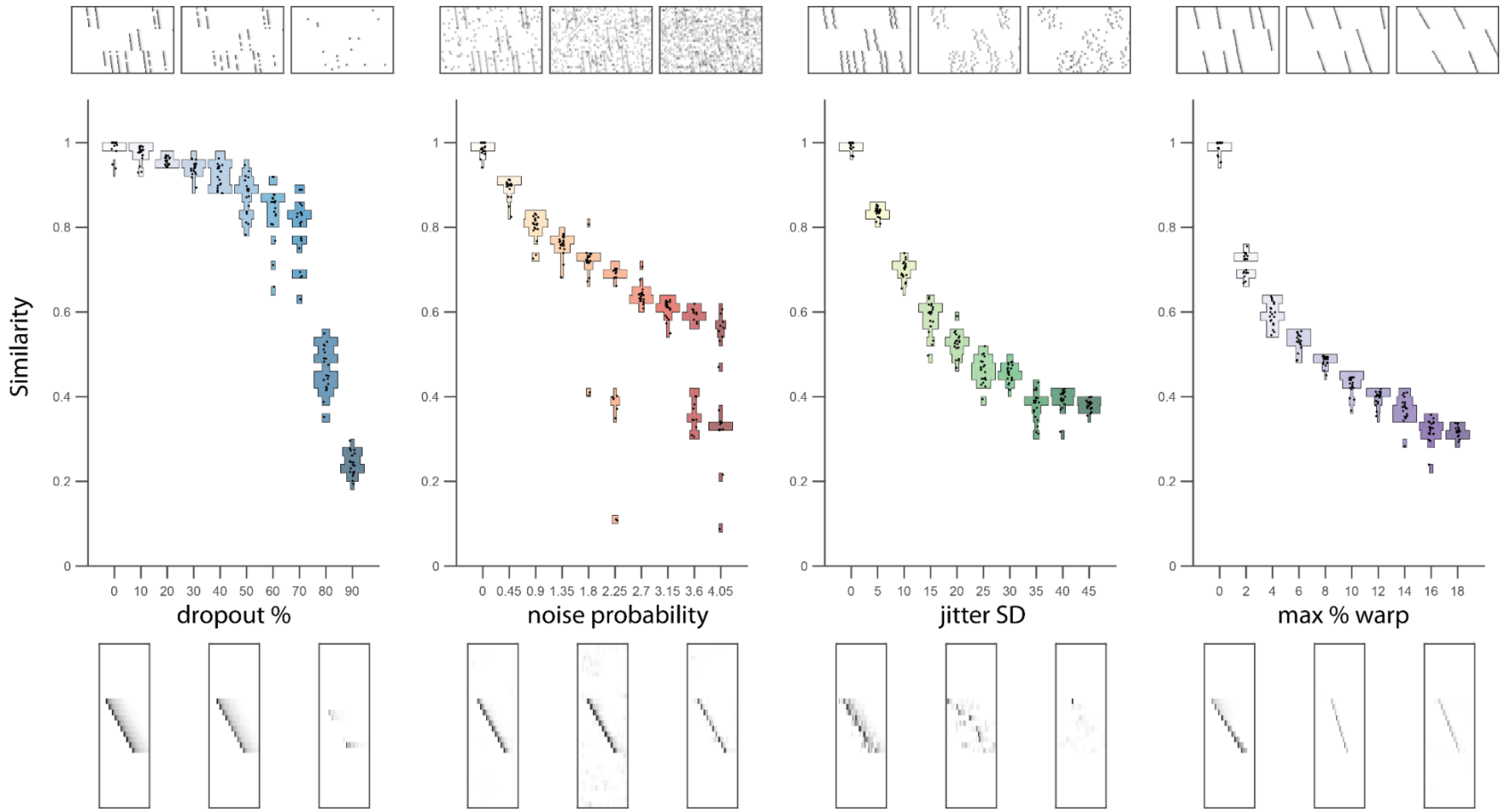
SeqNMF is robust to noise



Mackevicius and Bahle, et al., 2018

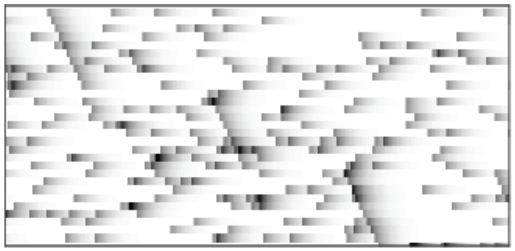


Mackevicius and Bahle, et al., 2018



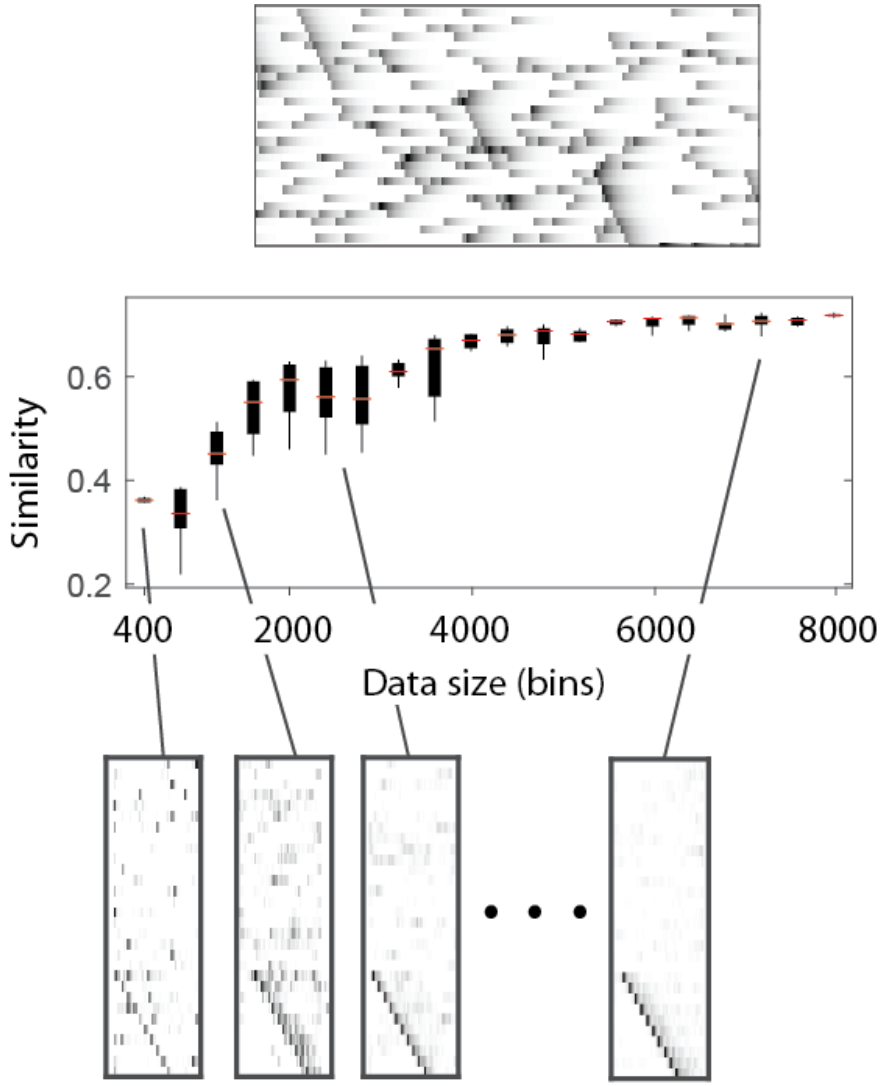
Mackevicius and Bahle, et al., 2018

Can SeqNMF discover sequences in small datasets?



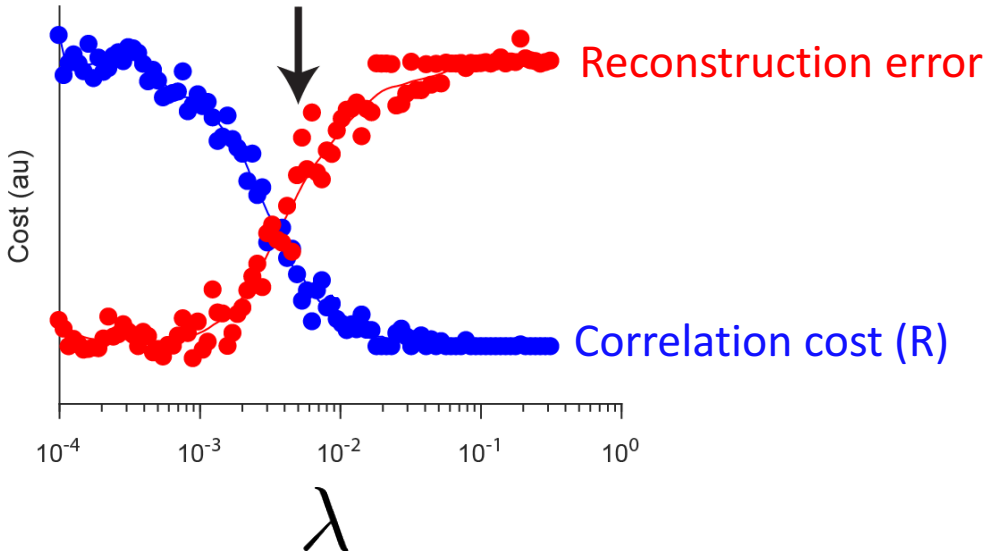
Mackevicius and Bahle, et al., in preparation

Can SeqNMF discover sequences in small datasets?



Mackevicius and Bahle, et al., in preparation

Method to choose lambda

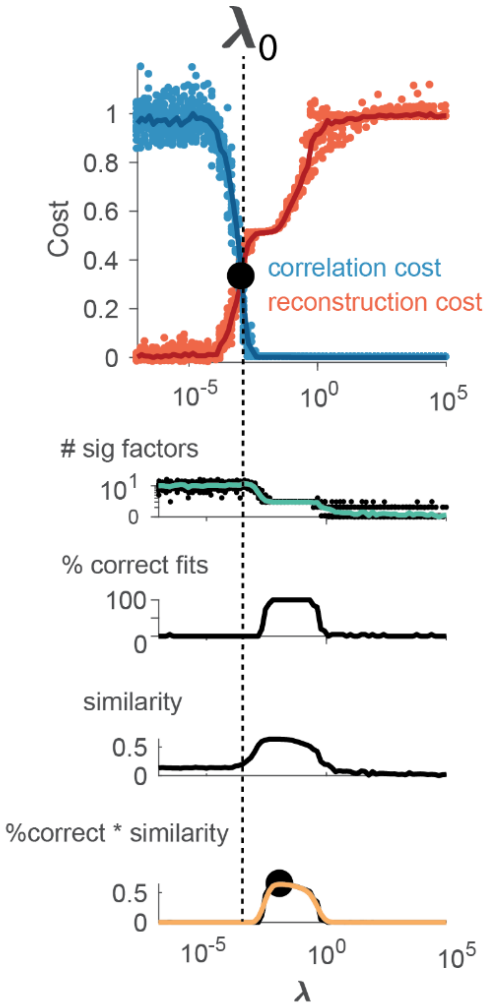


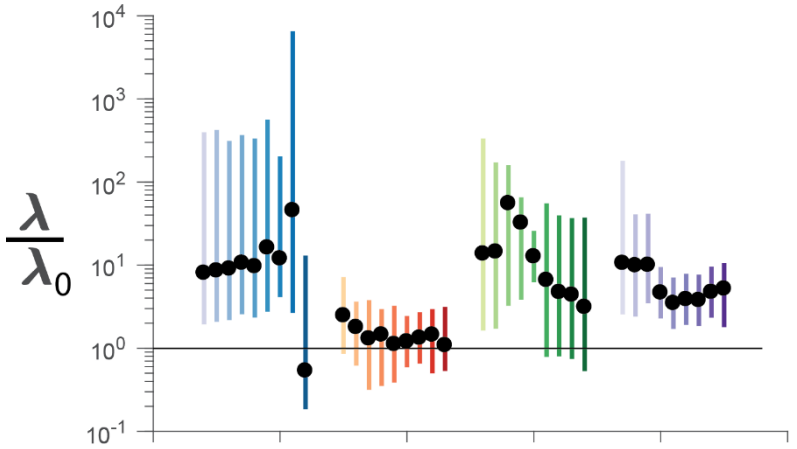
Penalty for correlations

$$(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}}) = \arg\min_{\mathbf{W},\mathbf{H}} \left(||\mathbf{X} - \mathbf{W} \circledast \mathbf{H}||_F^2 + \lambda \mathcal{R} \right)$$

Mackevicius and Bahle, et al., 2018

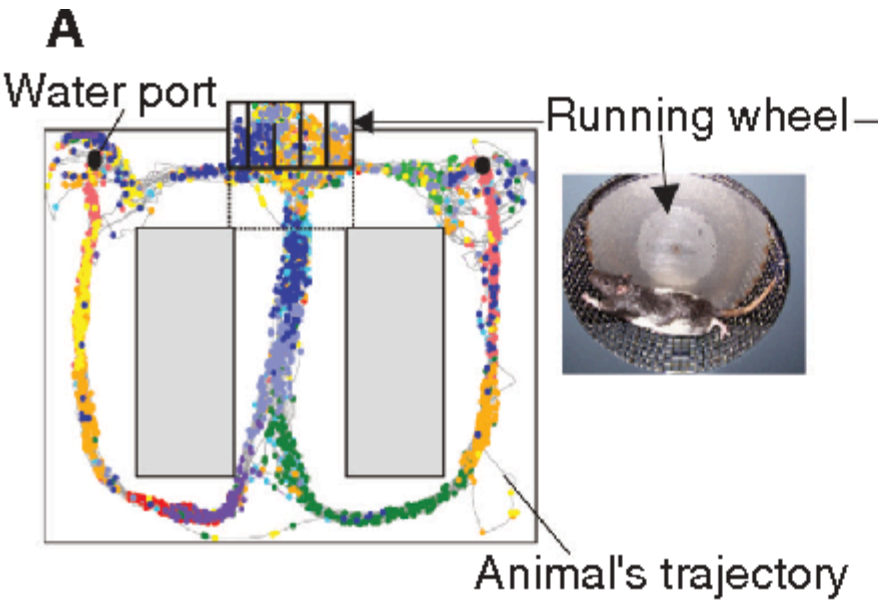
Method to choose lambda





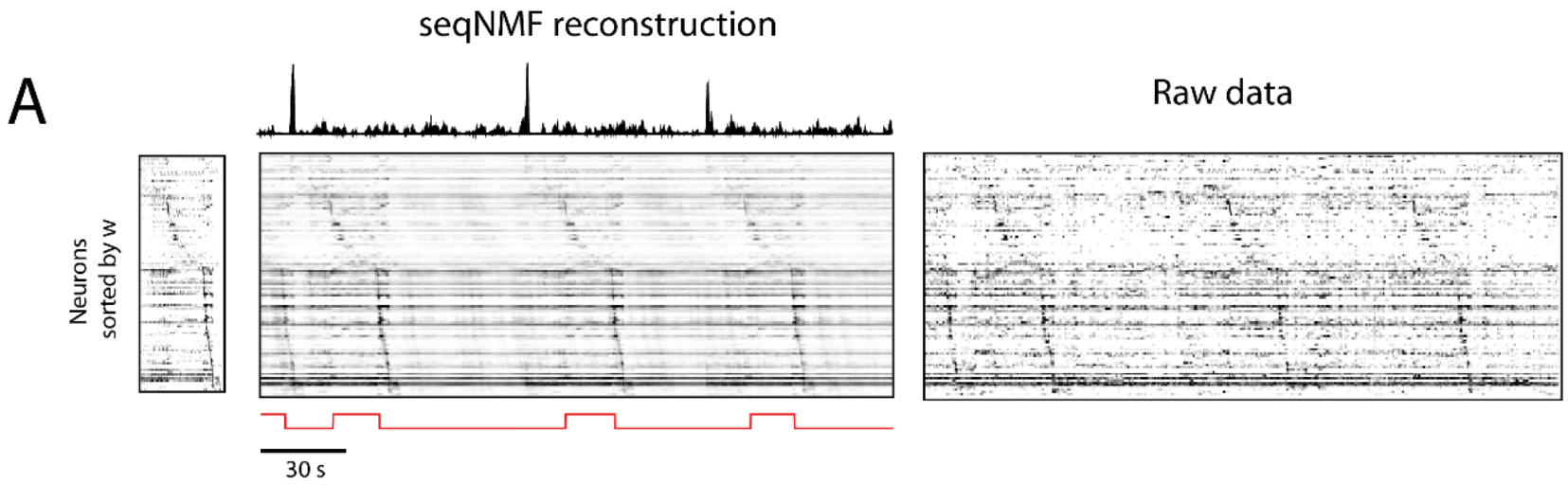
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on hippocampal sequences



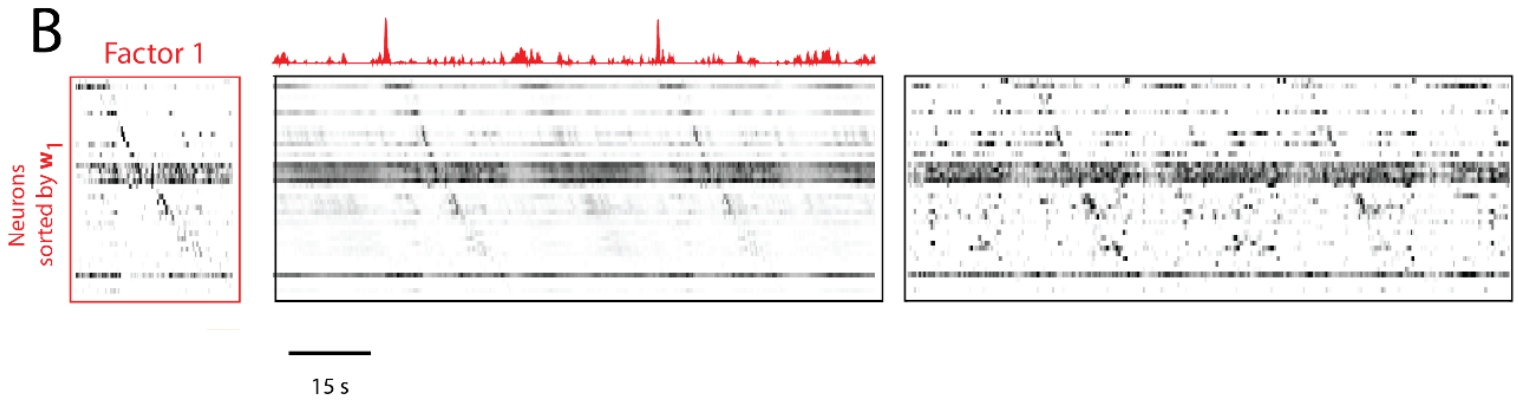
Pastalkova et al., 2008

Testing SeqNMF on hippocampal sequences



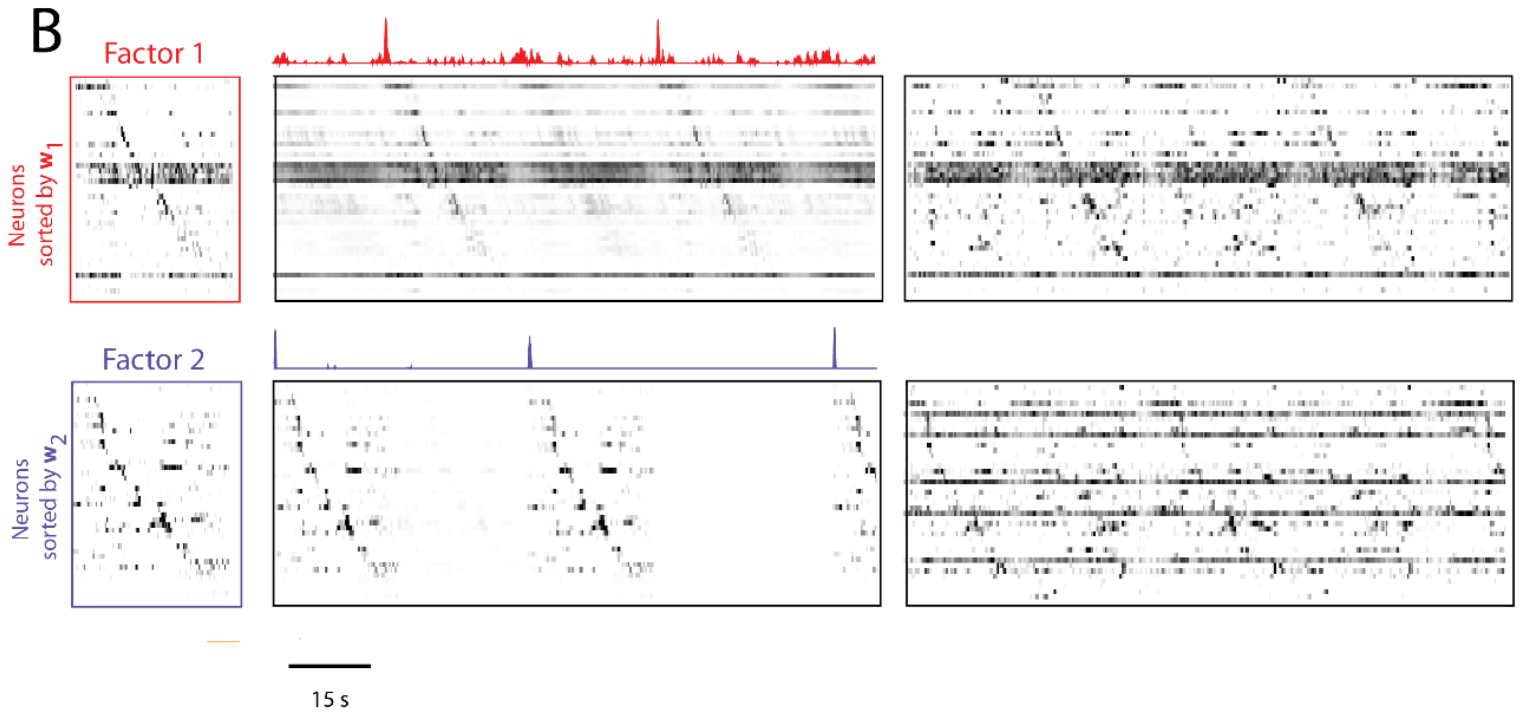
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on hippocampal sequences



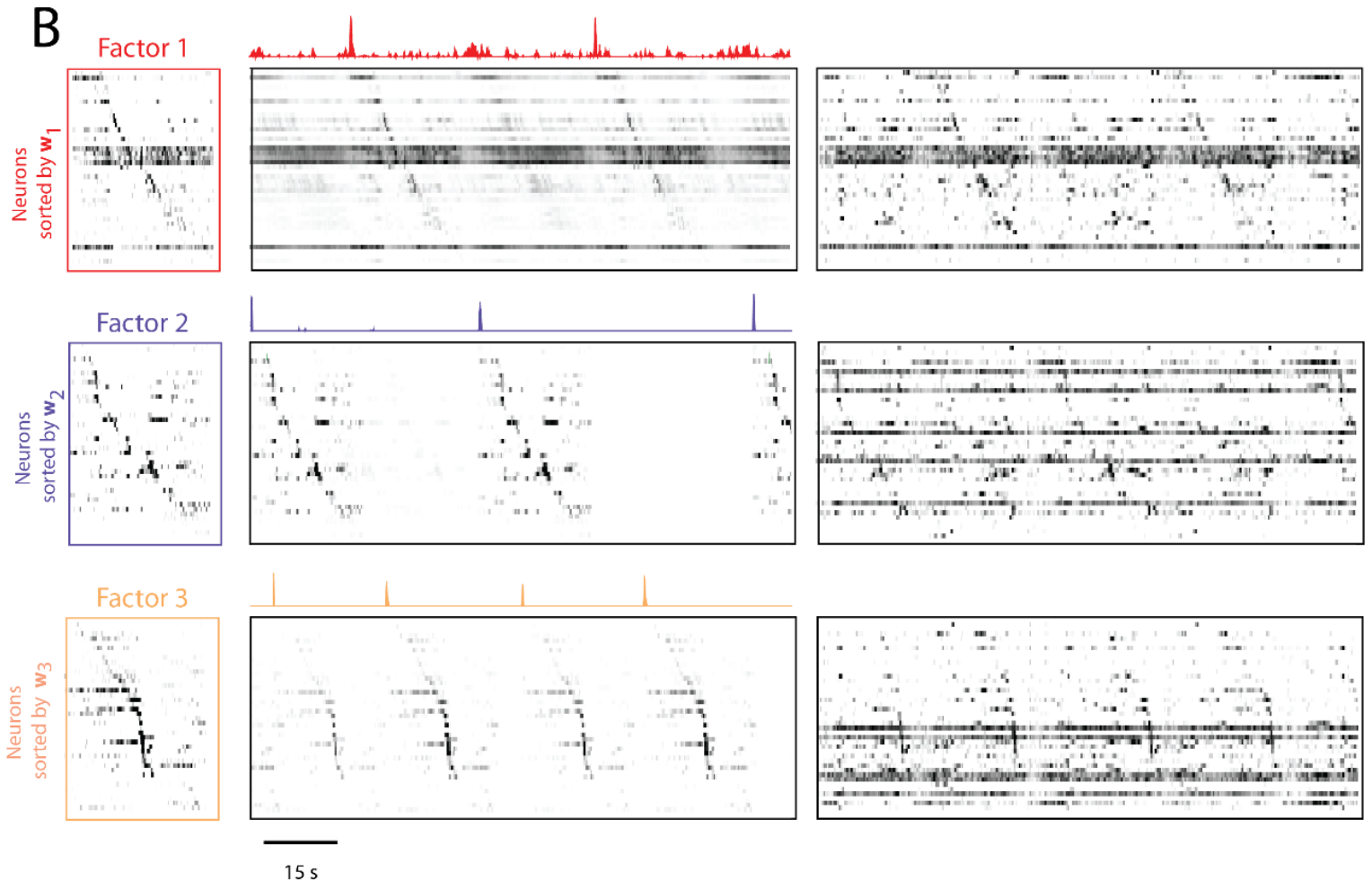
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on hippocampal sequences



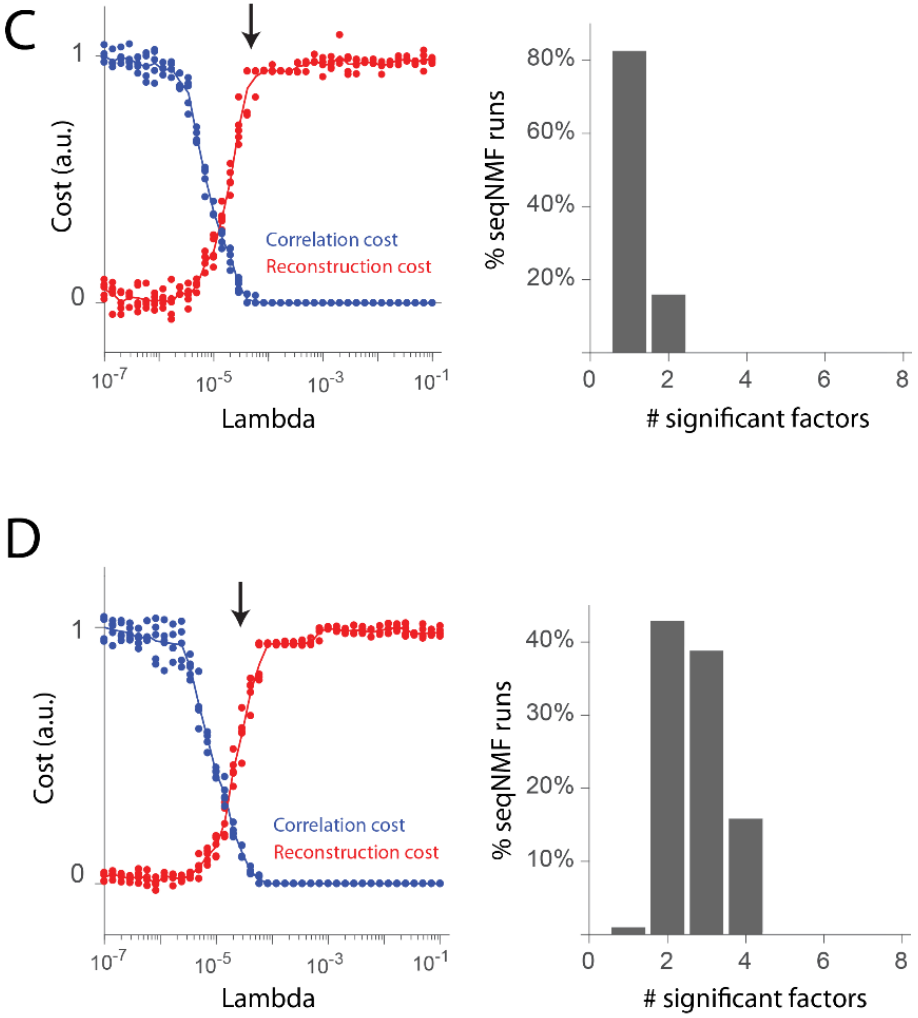
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on hippocampal sequences



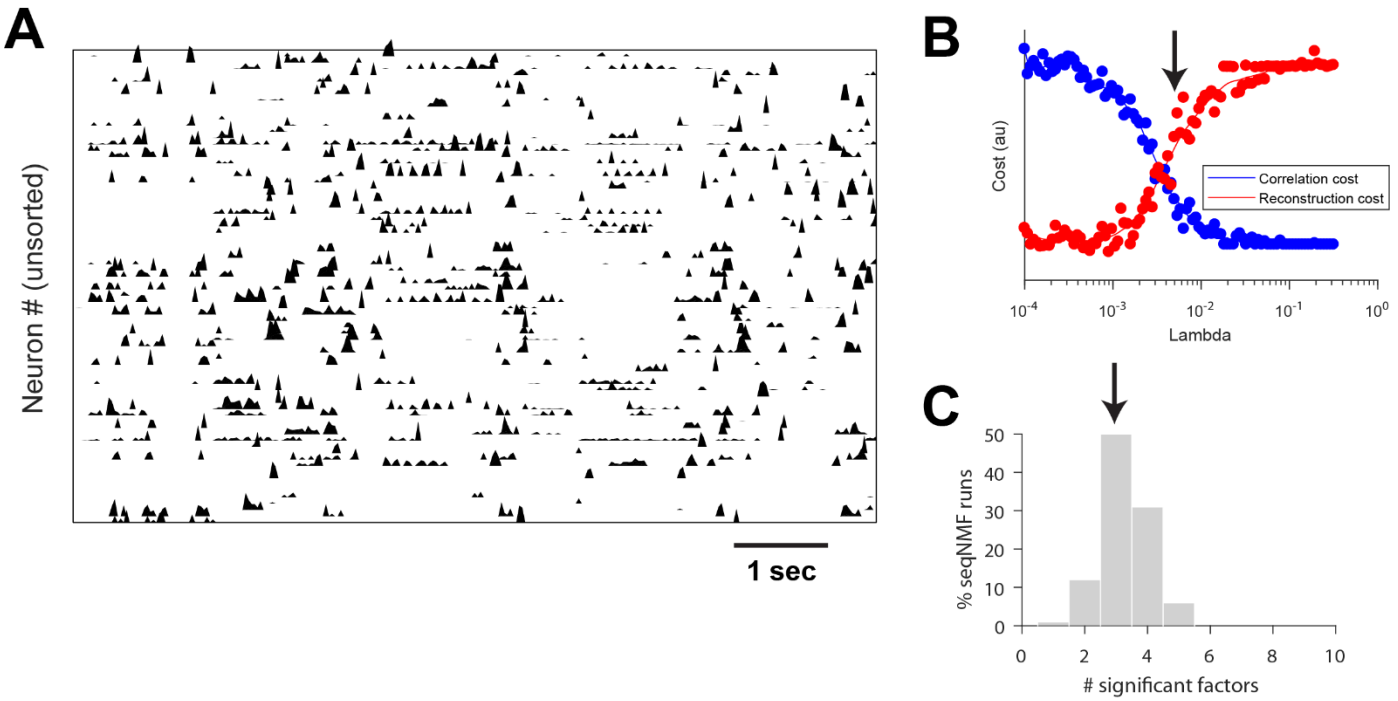
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on hippocampal sequences



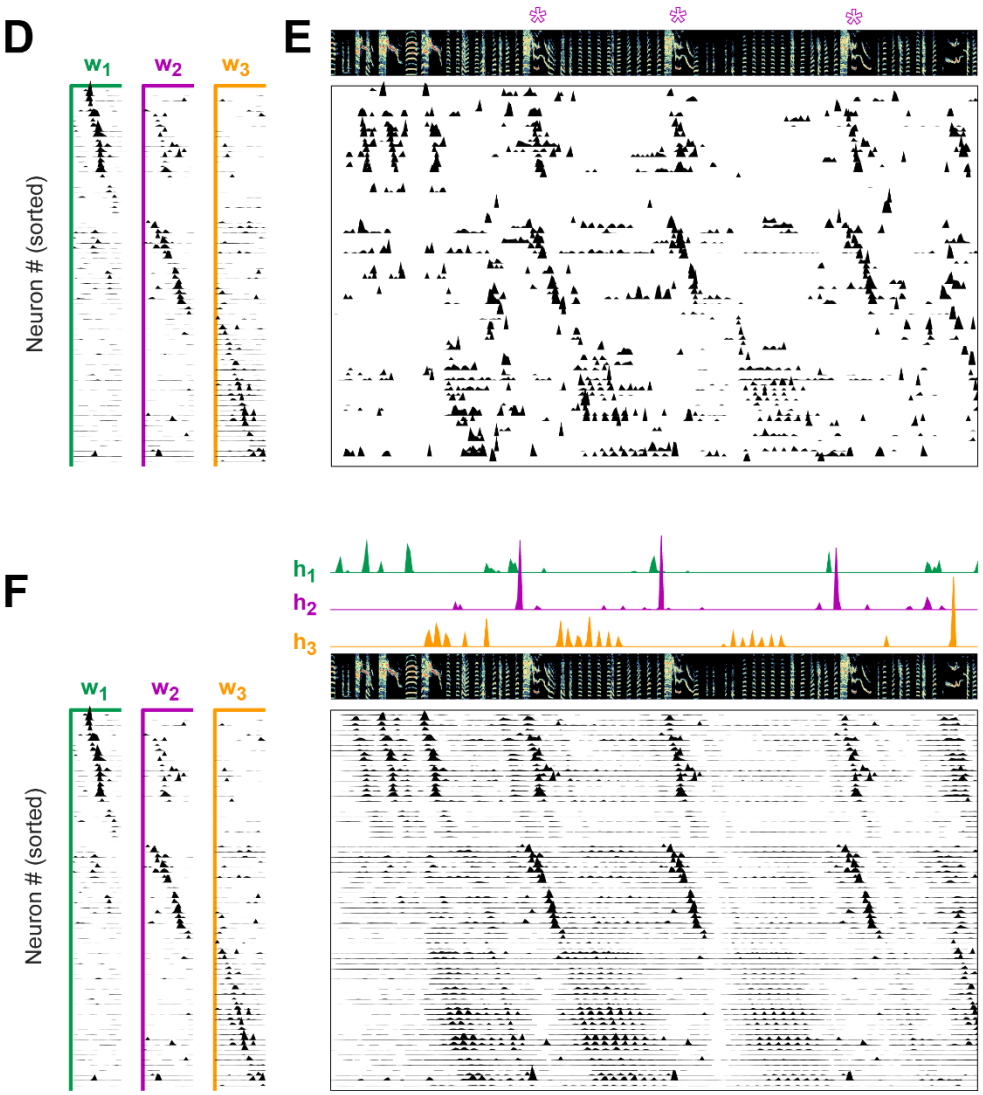
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on songbird HVC



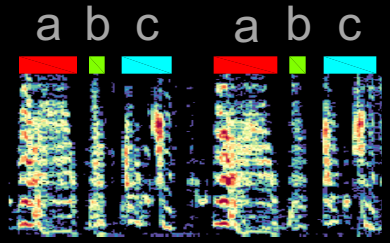
Mackevicius and Bahle, et al., 2018

Testing SeqNMF on songbird HVC



Mackevicius and Bahle, et al., 2018

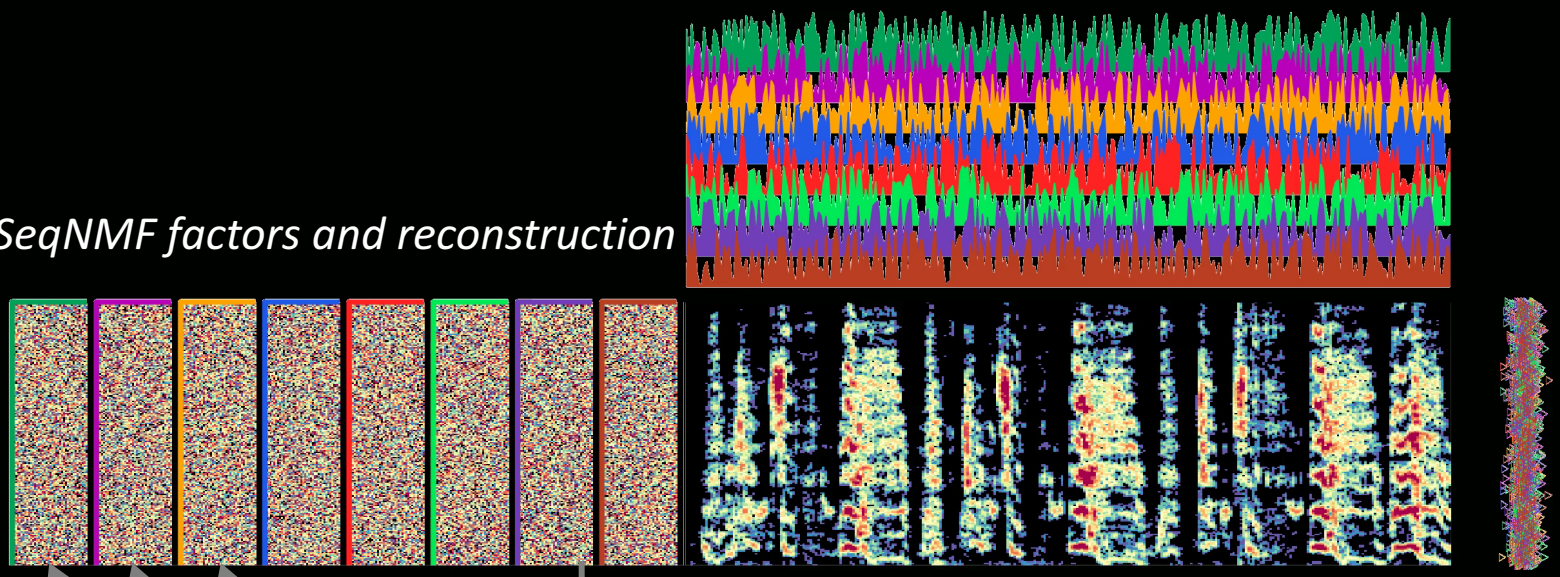
SeqNMF is broadly applicable to many high-dimensional datasets

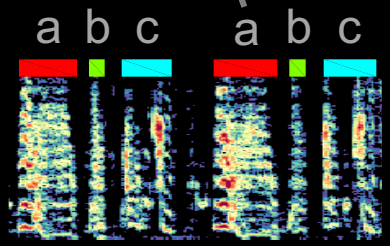


Hand-labeled song syllables, Okubo 2015

SeqNMF is broadly applicable to many high-dimensional datasets

SeqNMF factors and reconstruction





. . .

Hand-labeled song syllables, Okubo 2015

github.com/Feelab/seqNMF
BioRxiv

Agenda

- We hope you now have an idea of whether seqNMF might work on your data
- Extra Tools
  - Cross-validation procedure to estimate the number of sequences
  - Choosing between “parts-based” vs “events-based” factorizations
  - Denoising with convolutional NMF
  - Exercises/link to code

Cross-validation procedure for NMF

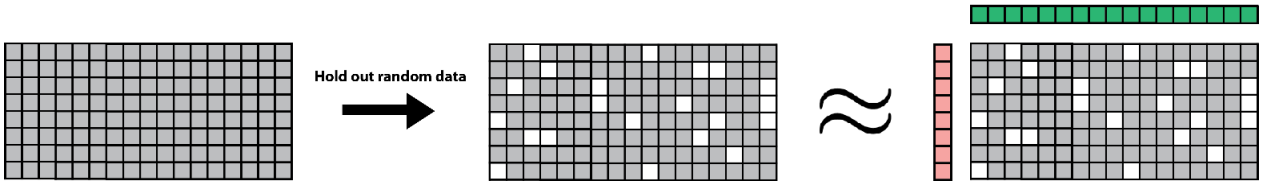


See http://alexhwilliams.info/itsneuronalblog/2018/02/26/crossval/

Cross-validation procedure for NMF

Minimize error only in non-masked elements

$$arg\min_{W,H} ||M \circ (WH - X)||_2^2$$



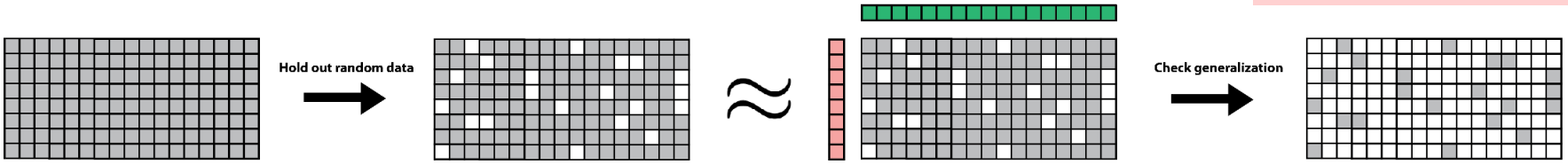
See http://alexhwilliams.info/itsneuronalblog/2018/02/26/crossval/

Cross-validation procedure for NMF

Check generalization

$$arg\min_{W,H} ||M \circ (WH - X)||_2^2$$

$$|| \sim M \circ (WH - X)||_2^2$$

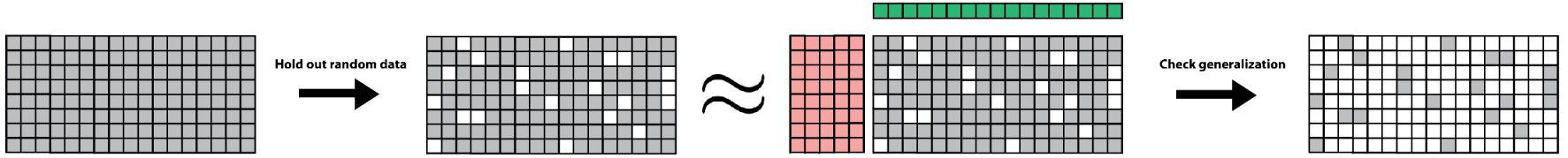


See http://alexhwilliams.info/itsneuronalblog/2018/02/26/crossval/

Cross-validation procedure for convolutional NMF

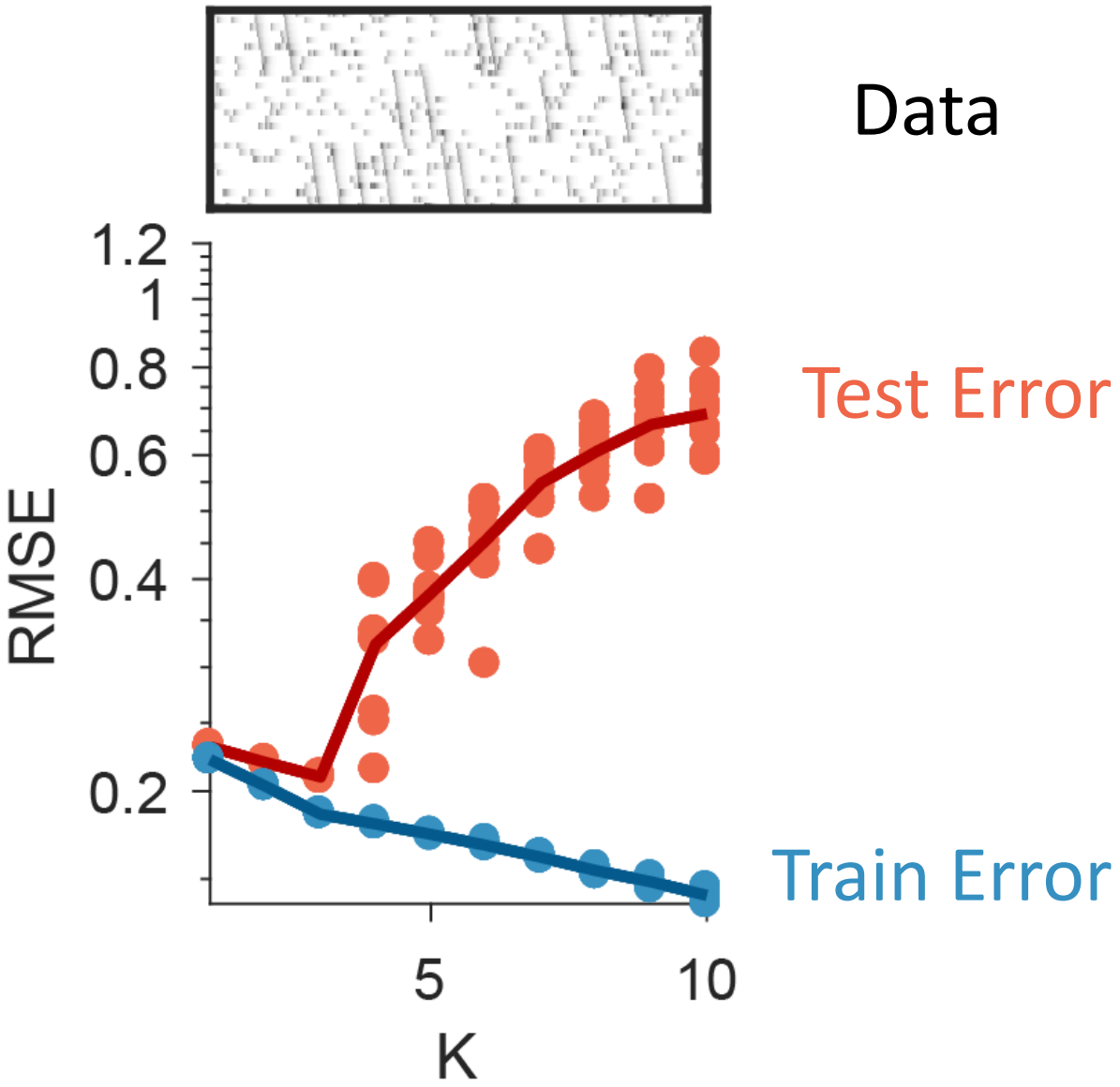
$$arg\min_{W,H} ||M \circ (W \circledast H - X)||_2^2$$

$$|| \sim M \circ (WH - X)||_2^2$$

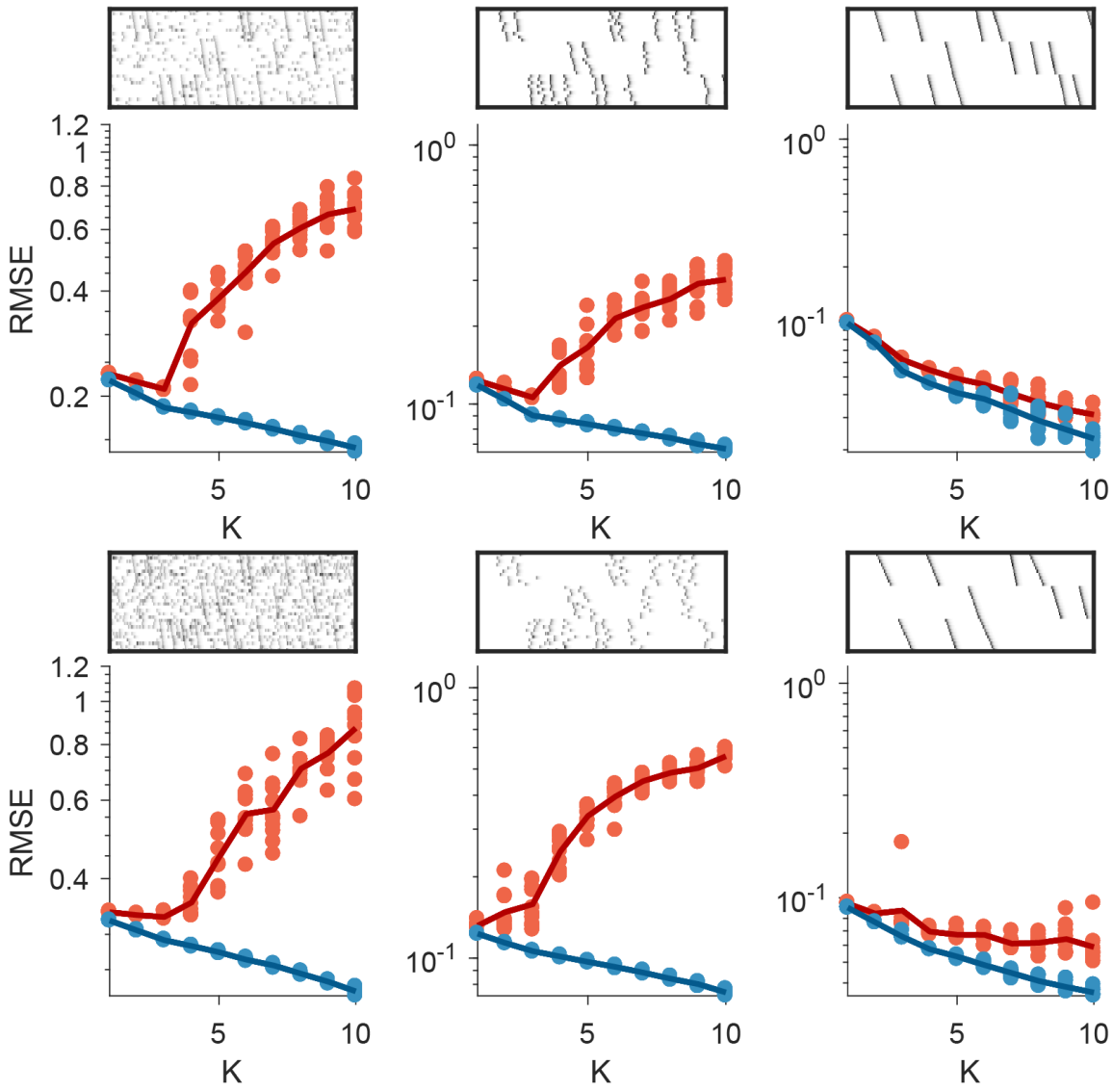


Mackevicius and Bahle, et al., in preparation

Generalization as a function of model components

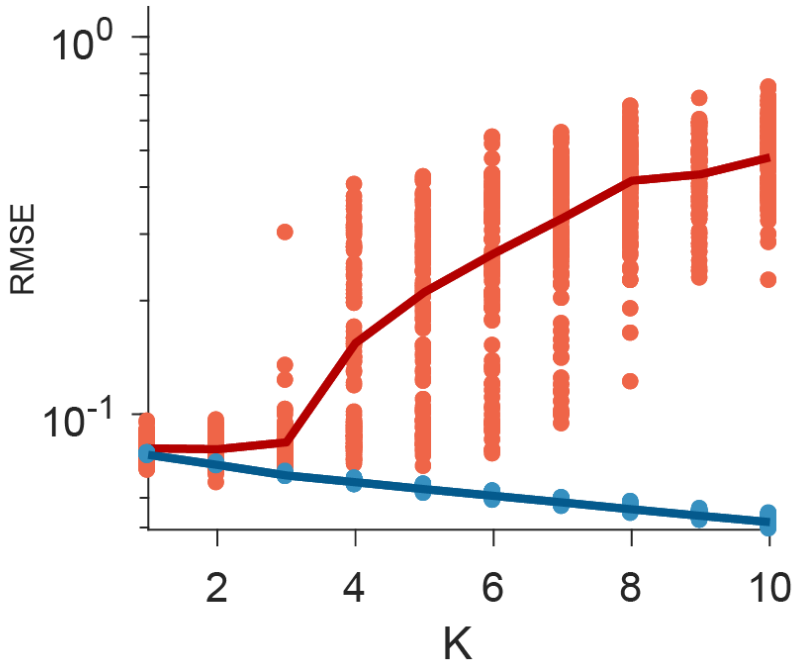


Mackevicius and Bahle, et al., in preparation

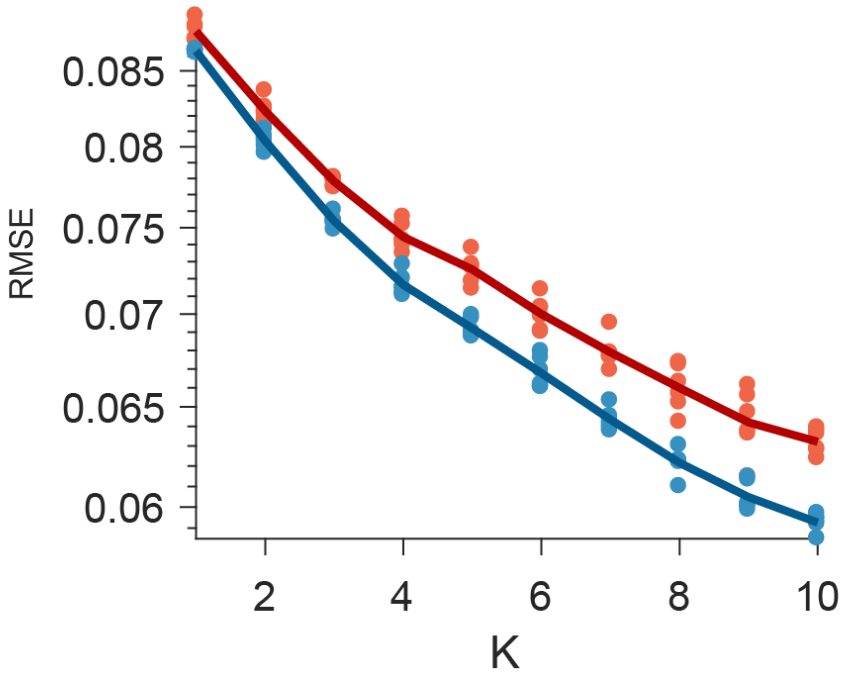


Mackevicius and Bahle, et al., in preparation

Songbird HVC



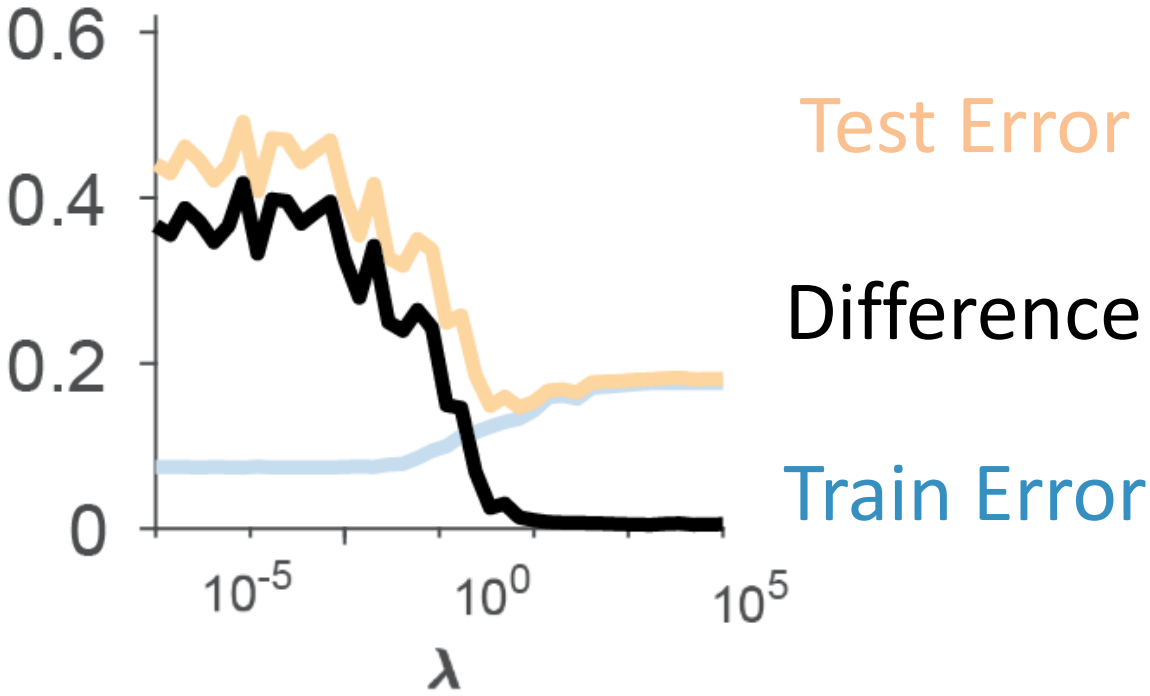
Rat Hippocampus



Mackevicius and Bahle, et al., in preparation

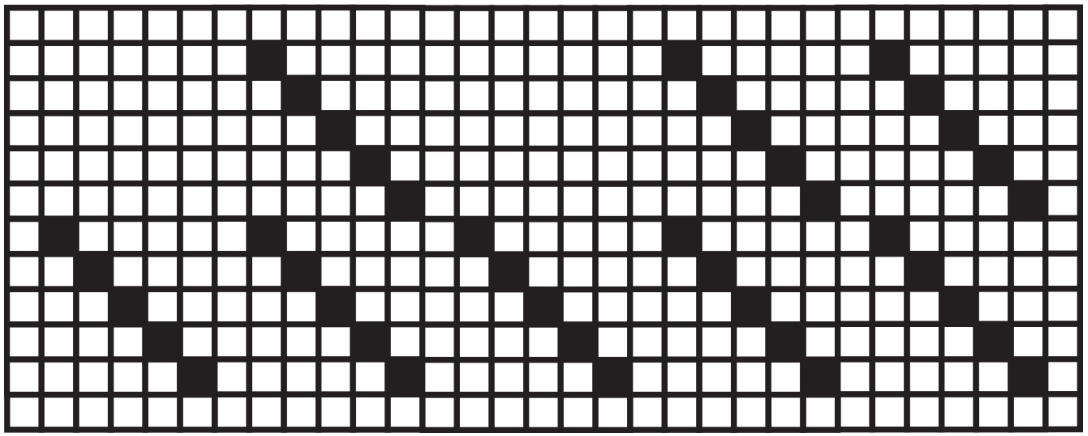
Generalization as a function of penalty





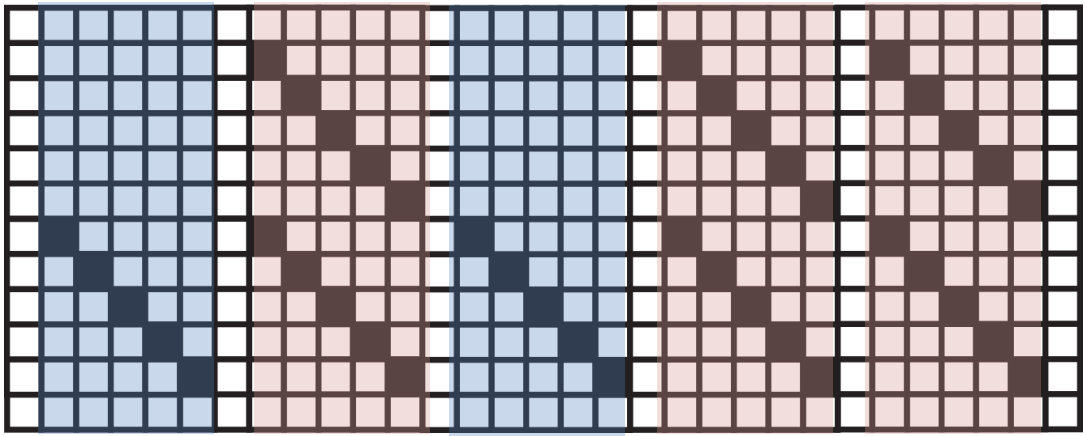
Mackevicius and Bahle, et al., in preparation

Choosing between “parts-based” vs “events-based” factorizations



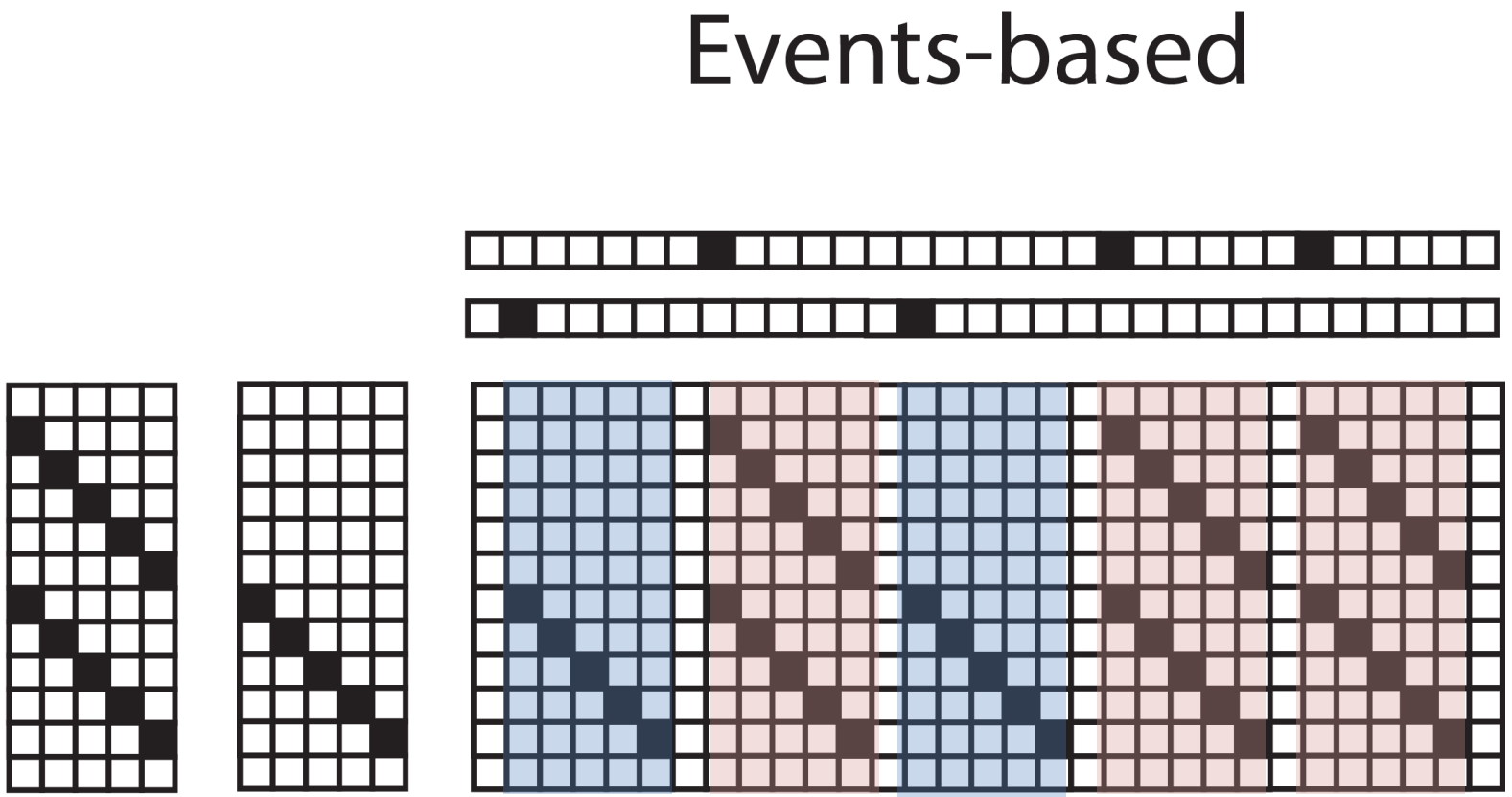
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



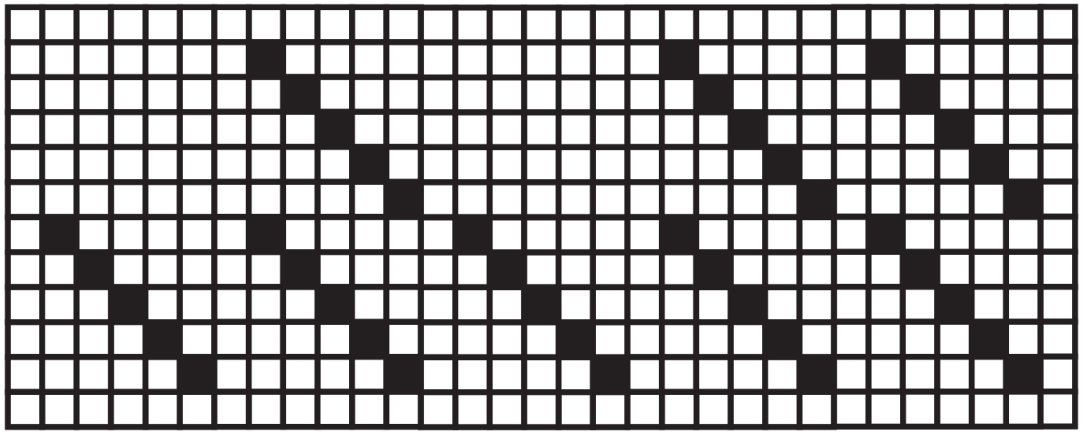
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



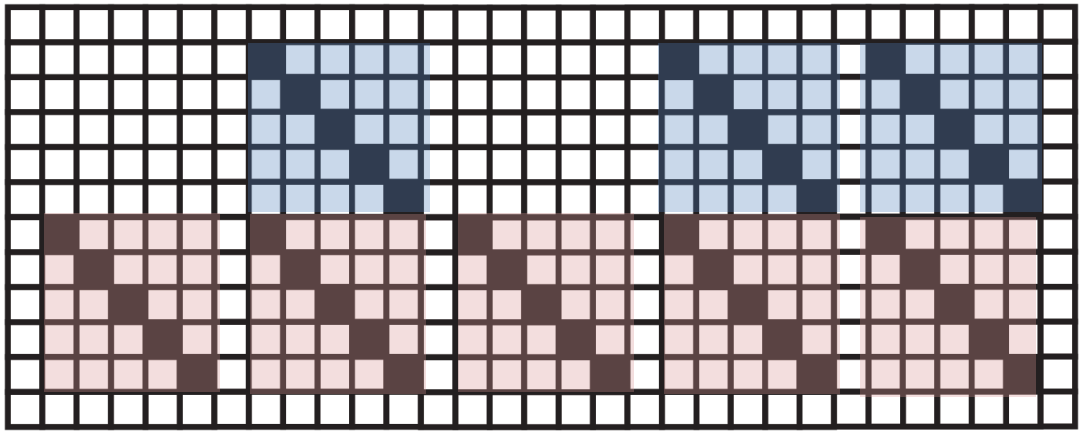
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



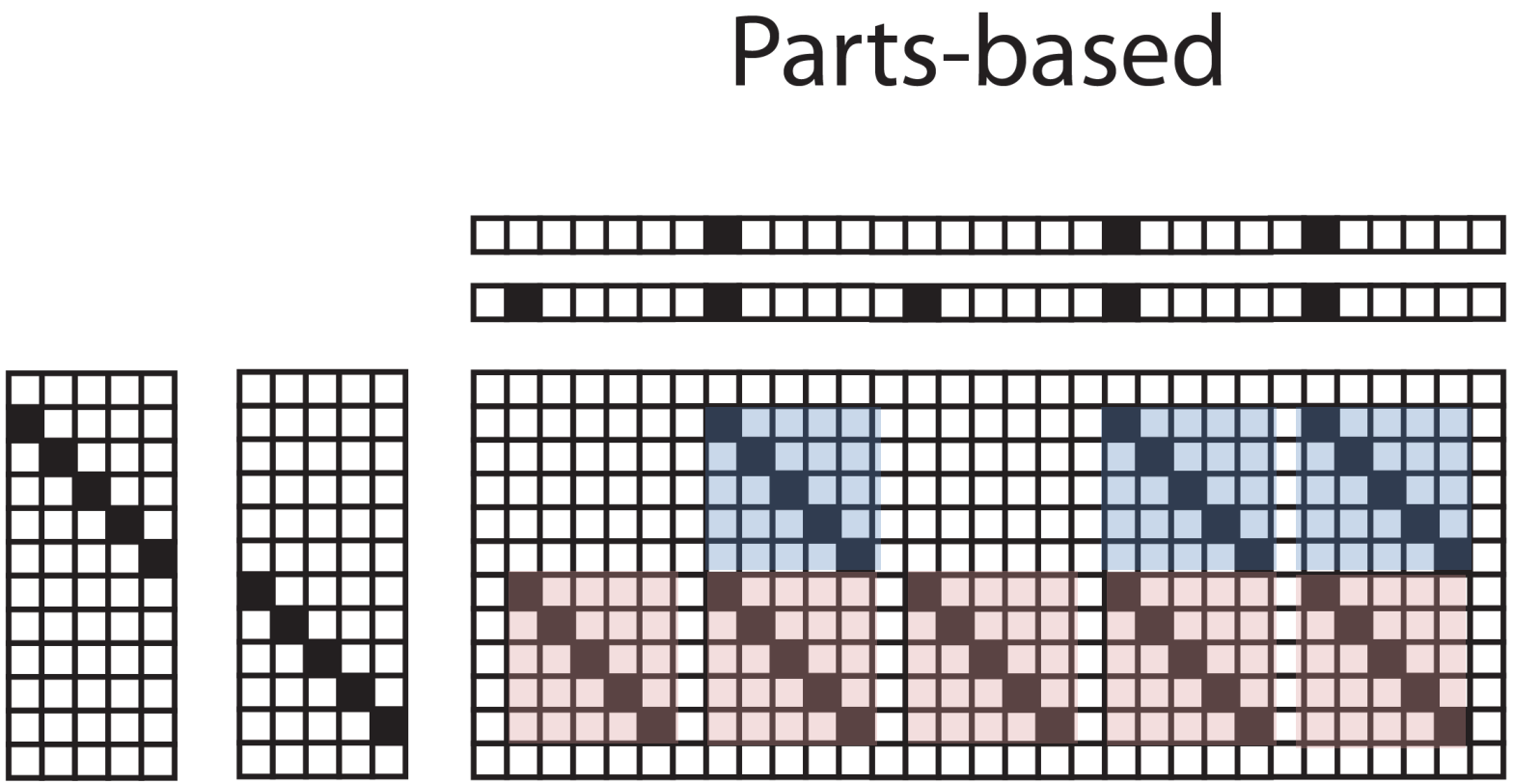
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



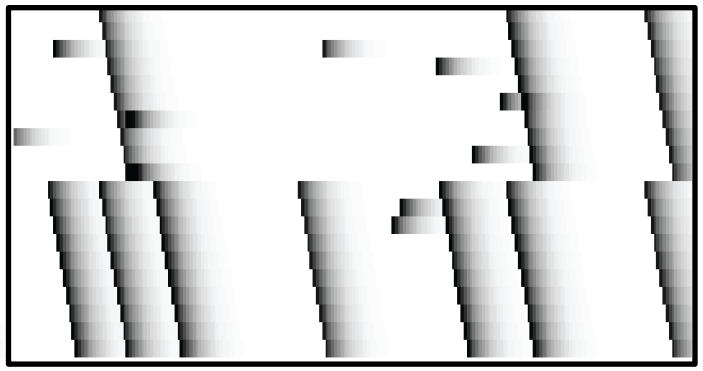
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



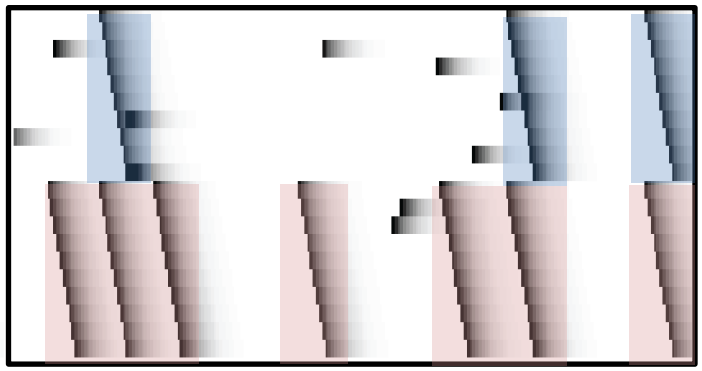
Mackevicius and Bahle, et al., 2018

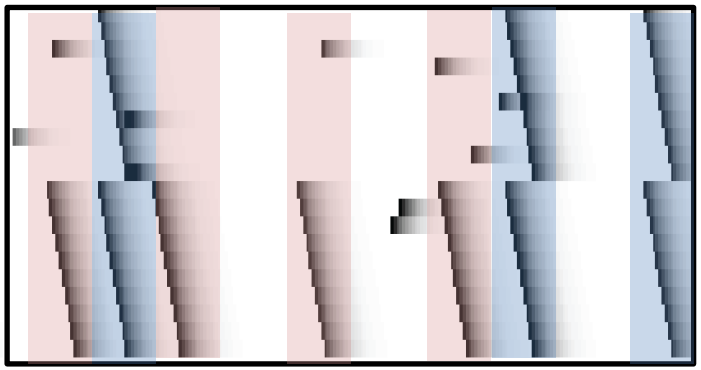
Choosing between “parts-based” vs “events-based” factorizations



Mackevicius and Bahle, et al., 2018

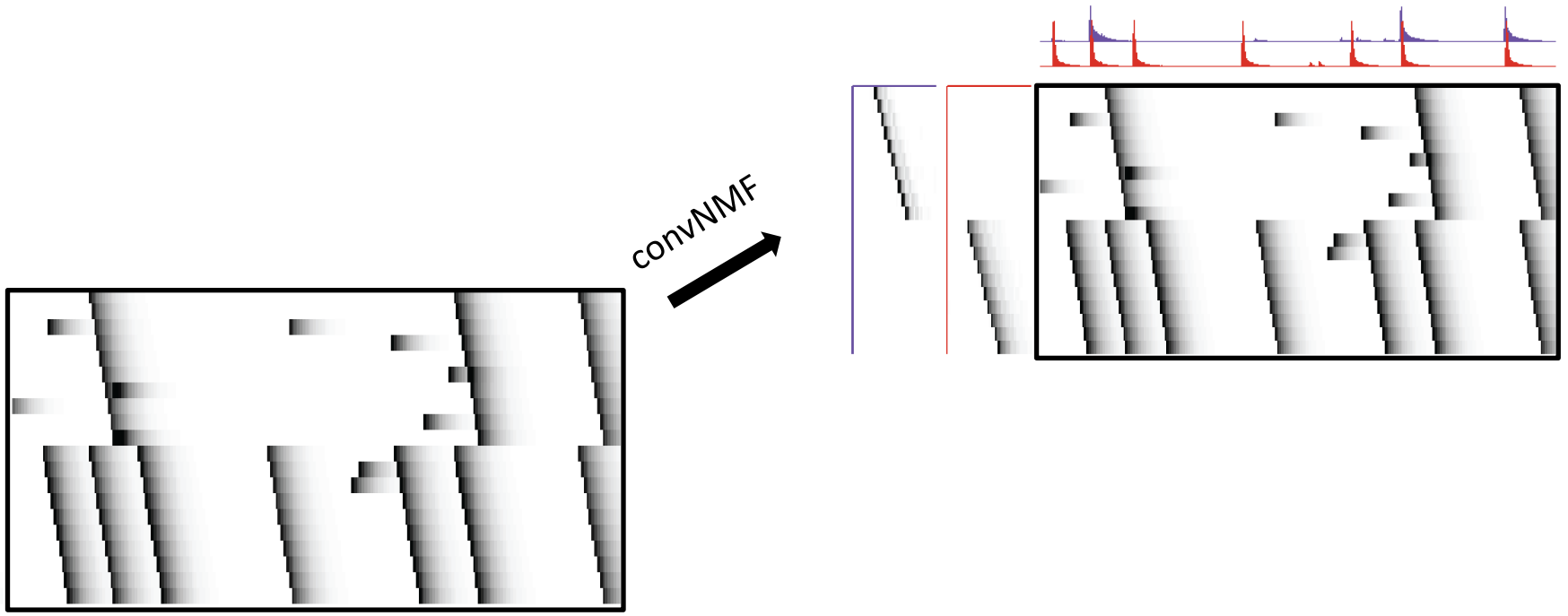
Choosing between “parts-based” vs “events-based” factorizations





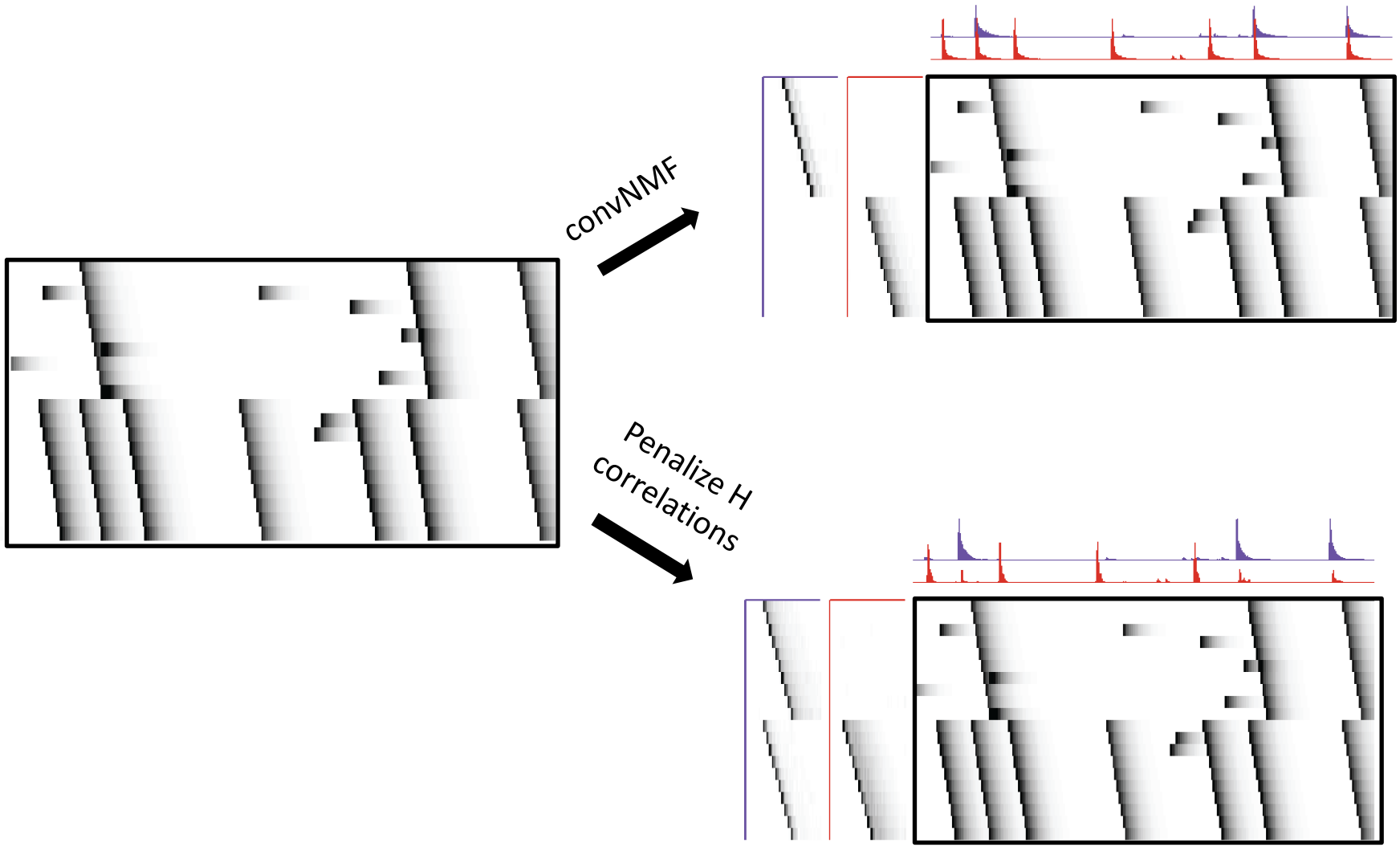
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



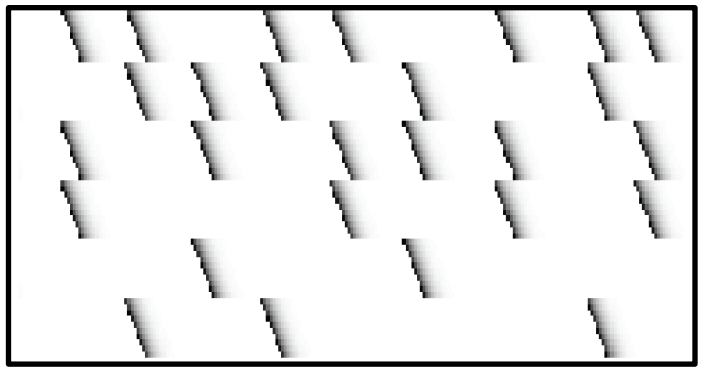
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



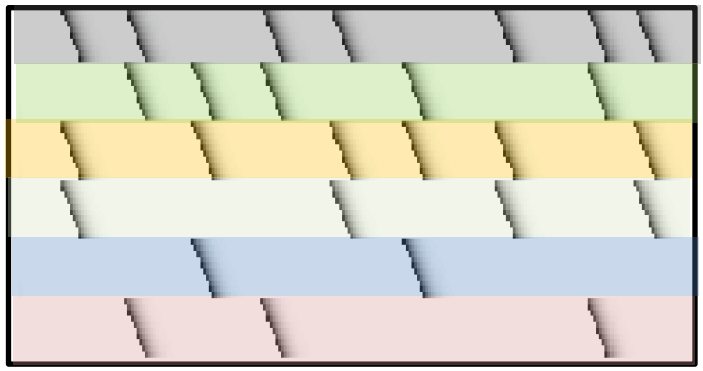
Mackevicius and Bahle, et al., 2018

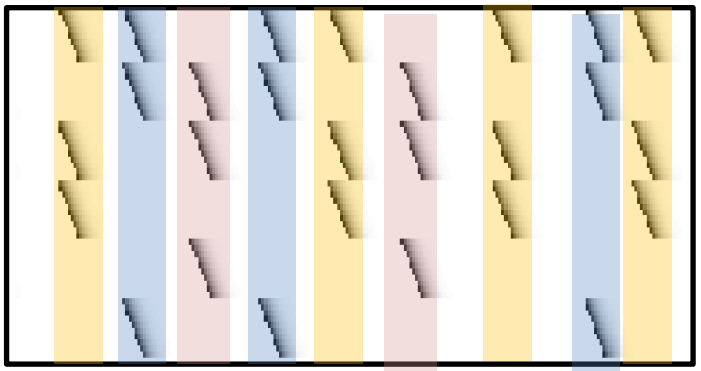
Choosing between “parts-based” vs “events-based” factorizations



Mackevicius and Bahle, et al., 2018

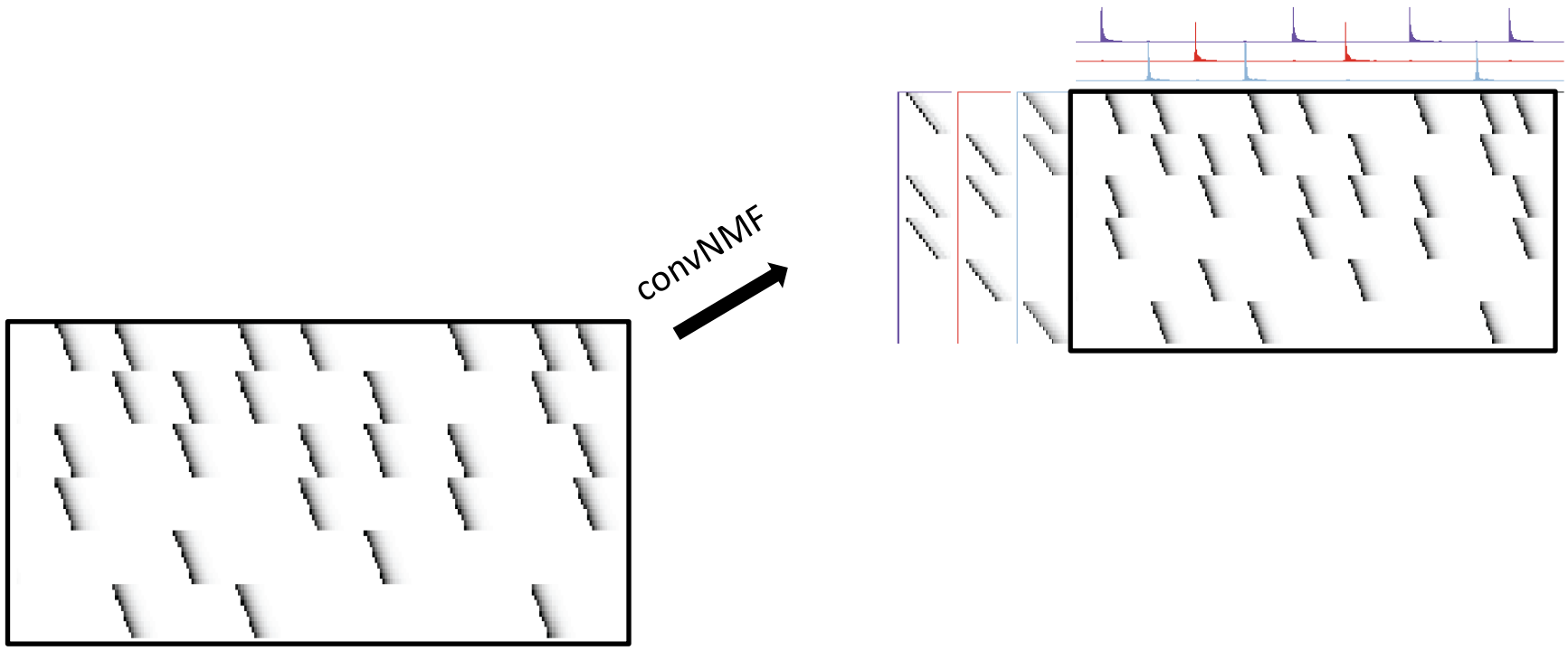
Choosing between “parts-based” vs “events-based” factorizations





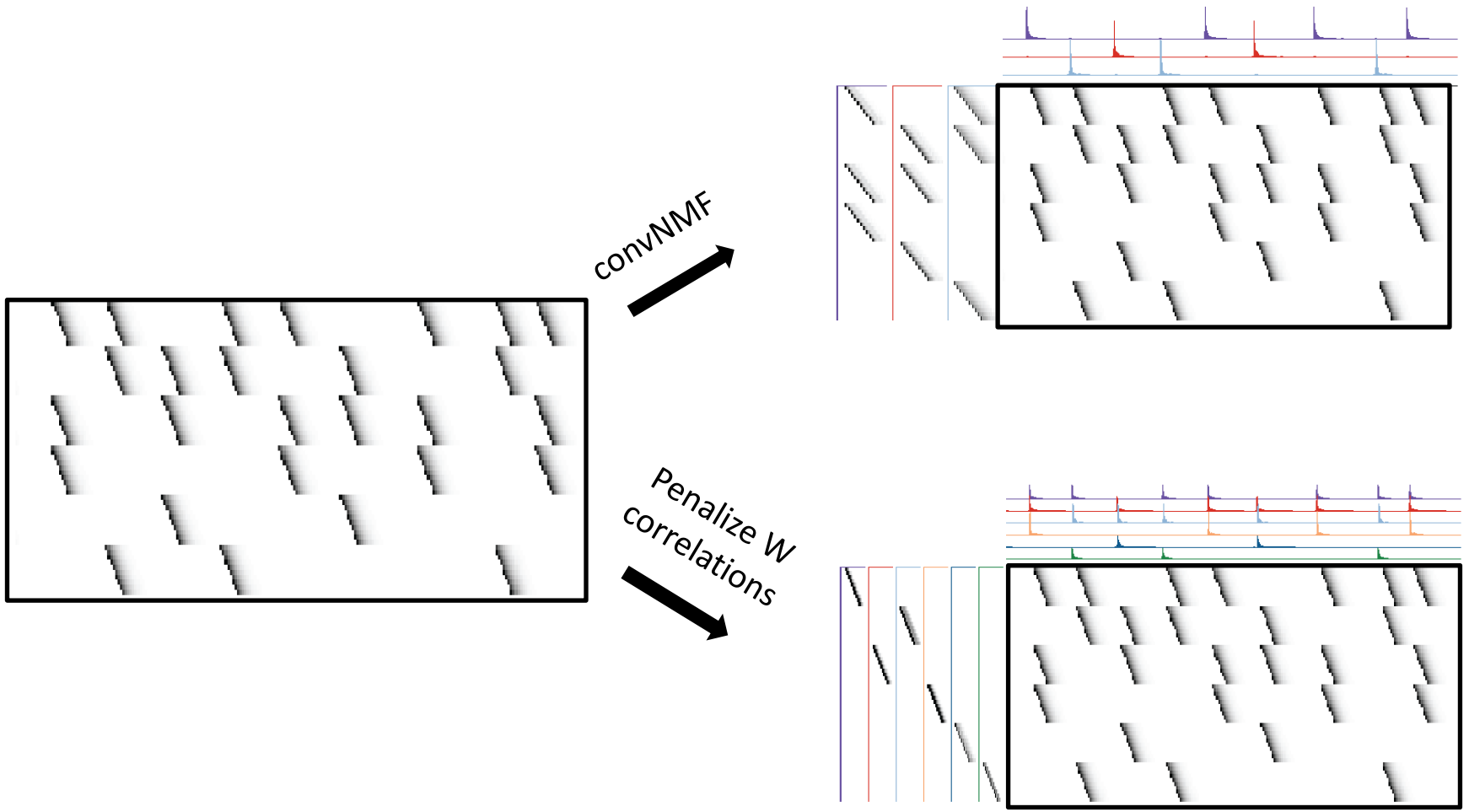
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations



Mackevicius and Bahle, et al., 2018

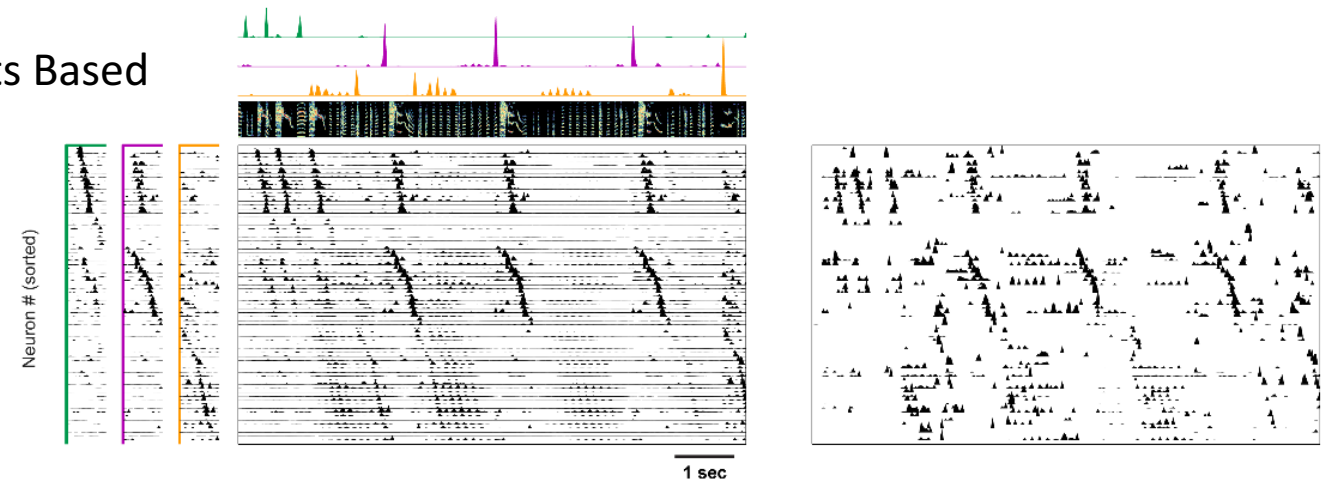
Choosing between “parts-based” vs “events-based” factorizations



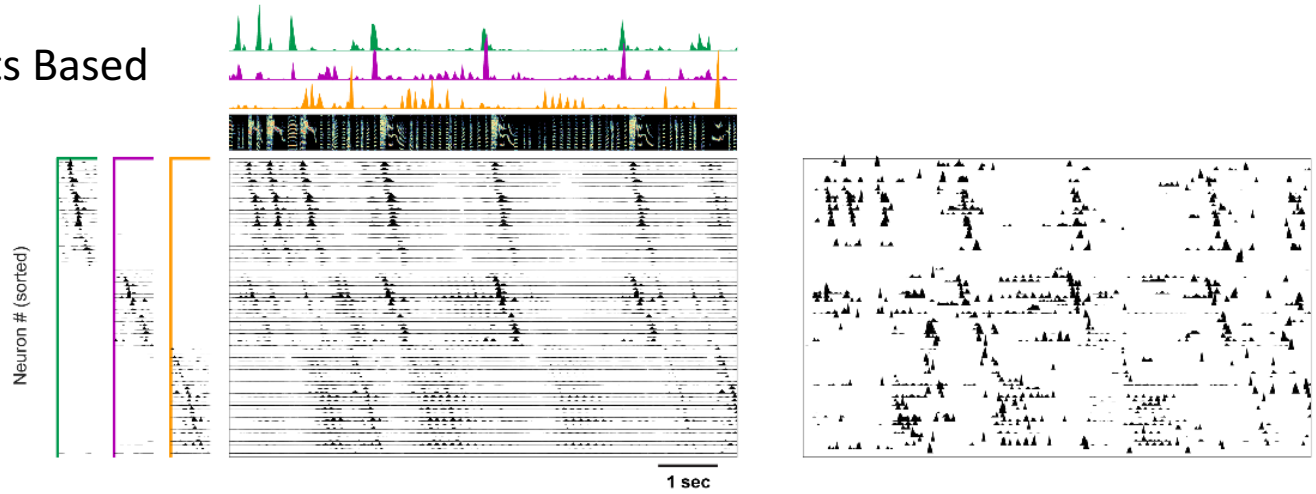
Mackevicius and Bahle, et al., 2018

Choosing between “parts-based” vs “events-based” factorizations

Events Based

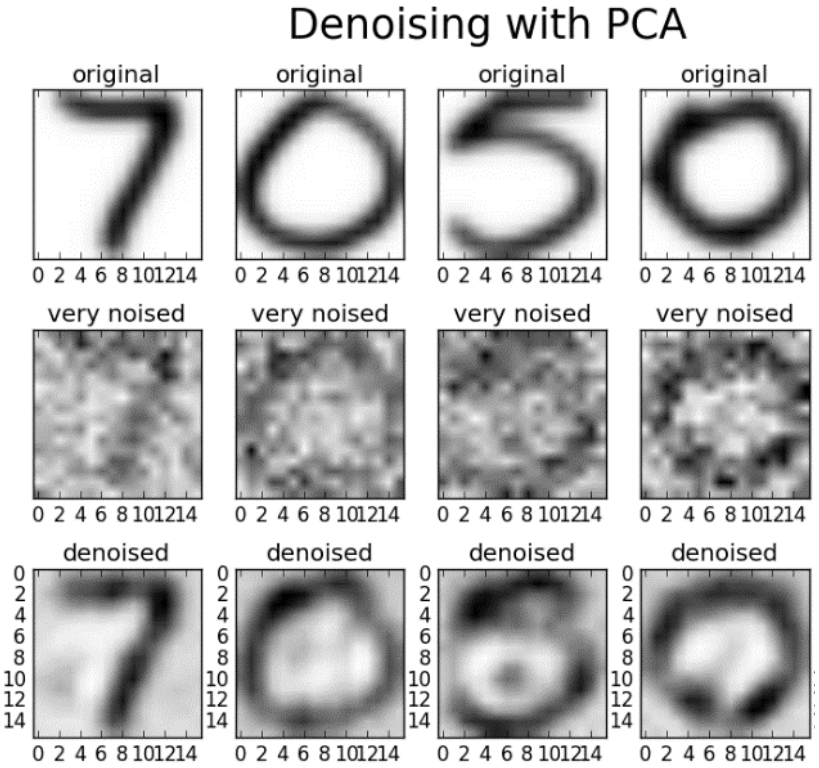


Parts Based

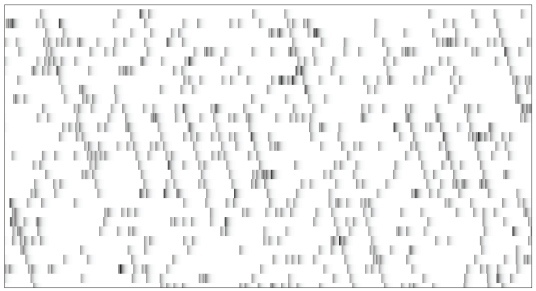


Mackevicius and Bahle, et al., 2018

Denoising with convolutional NMF

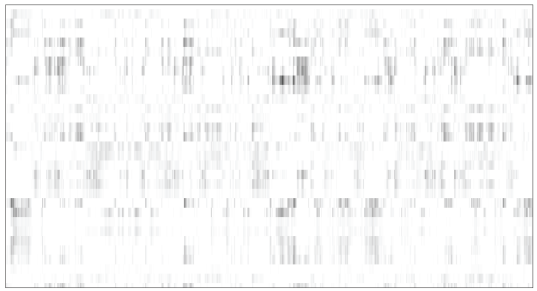


Data

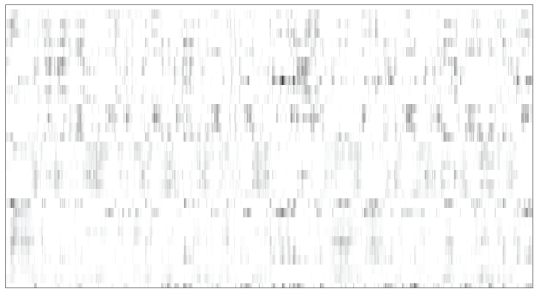


SVD (PCA)

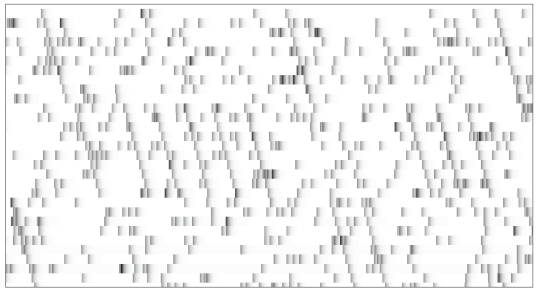
3 components (21090 parameters)



10 components (70300 parameters)

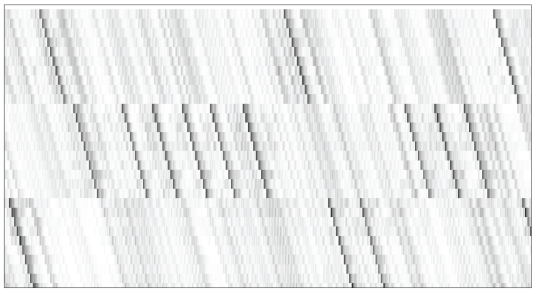


30 components (210900 parameters)

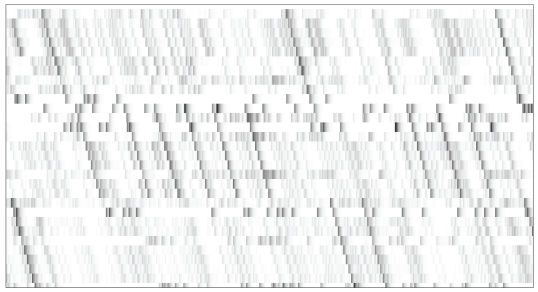


convNMF

3 components (34500 parameters)



10 components (115000 parameters)



Exercises

- Download and try the demo code: github.com/Feelab/seqNMF
  - Fit seqNMF
  - Choose lambda
  - Test significance
  - Switch between parts-based and events-based factorizations
  - Try seqNMF on a new dataset

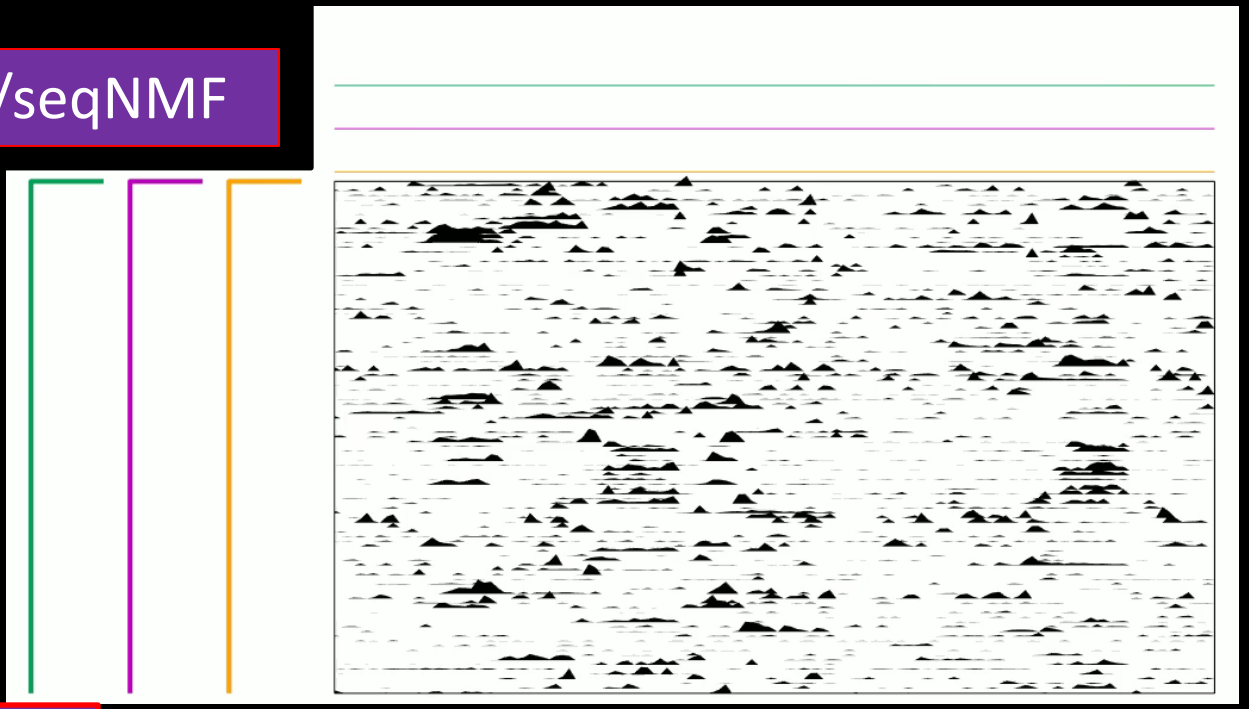
SeqNMF: GO FIND SEQUENCES!

github.com/Feelab/seqNMF

BioRxiv

elm@mit.edu

abahle@mit.edu





Emily Mackevicius



Andrew Bahle



Alex Williams



Shijie Gu



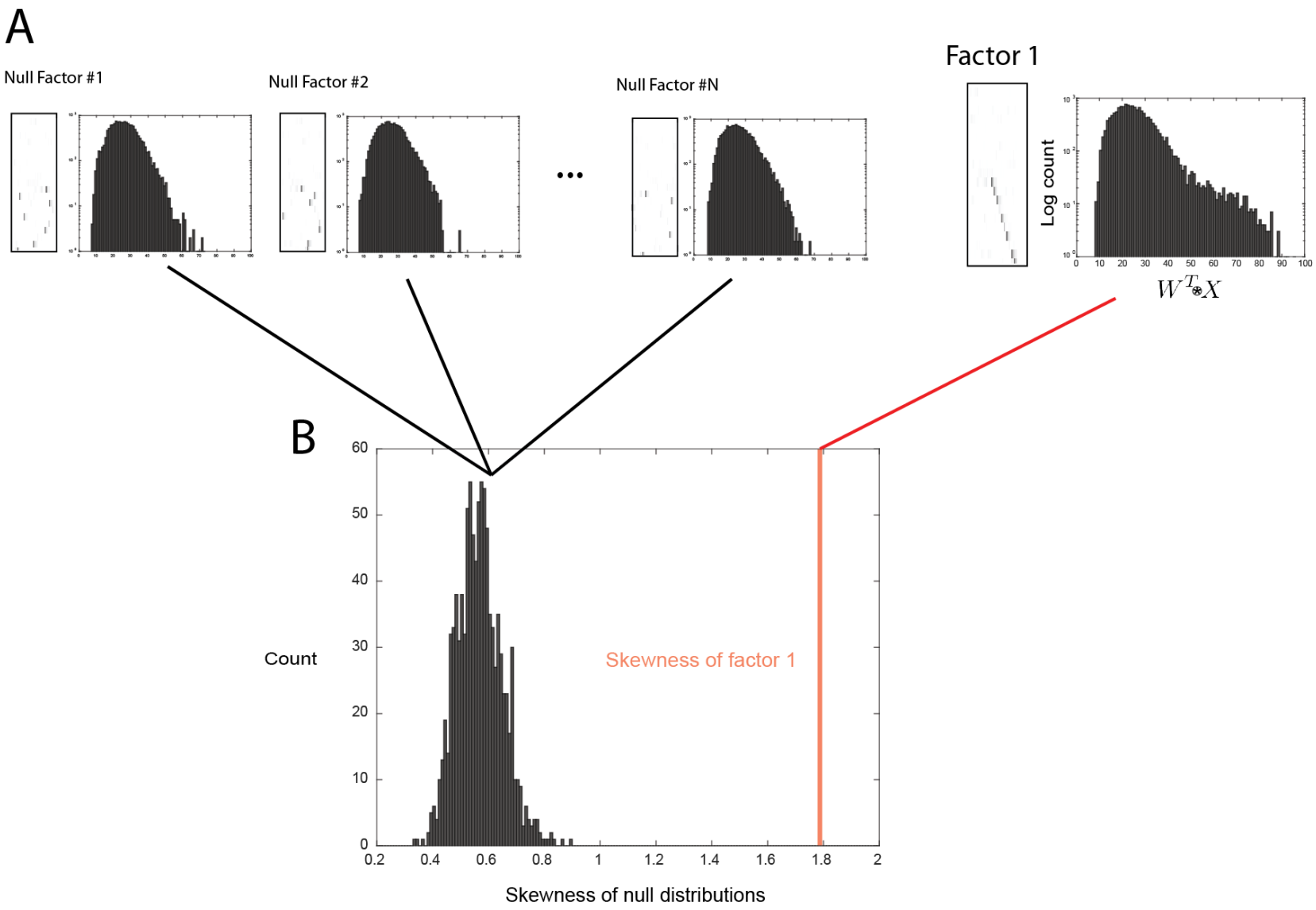
Natasha Denissenko



Mark Goldman



Michale Fee


A
Null Factor #1
Null Factor #2
Null Factor #N
Factor 1
Log count
$W^T X$
B
Count
Skewness of factor 1
Skewness of null distributions