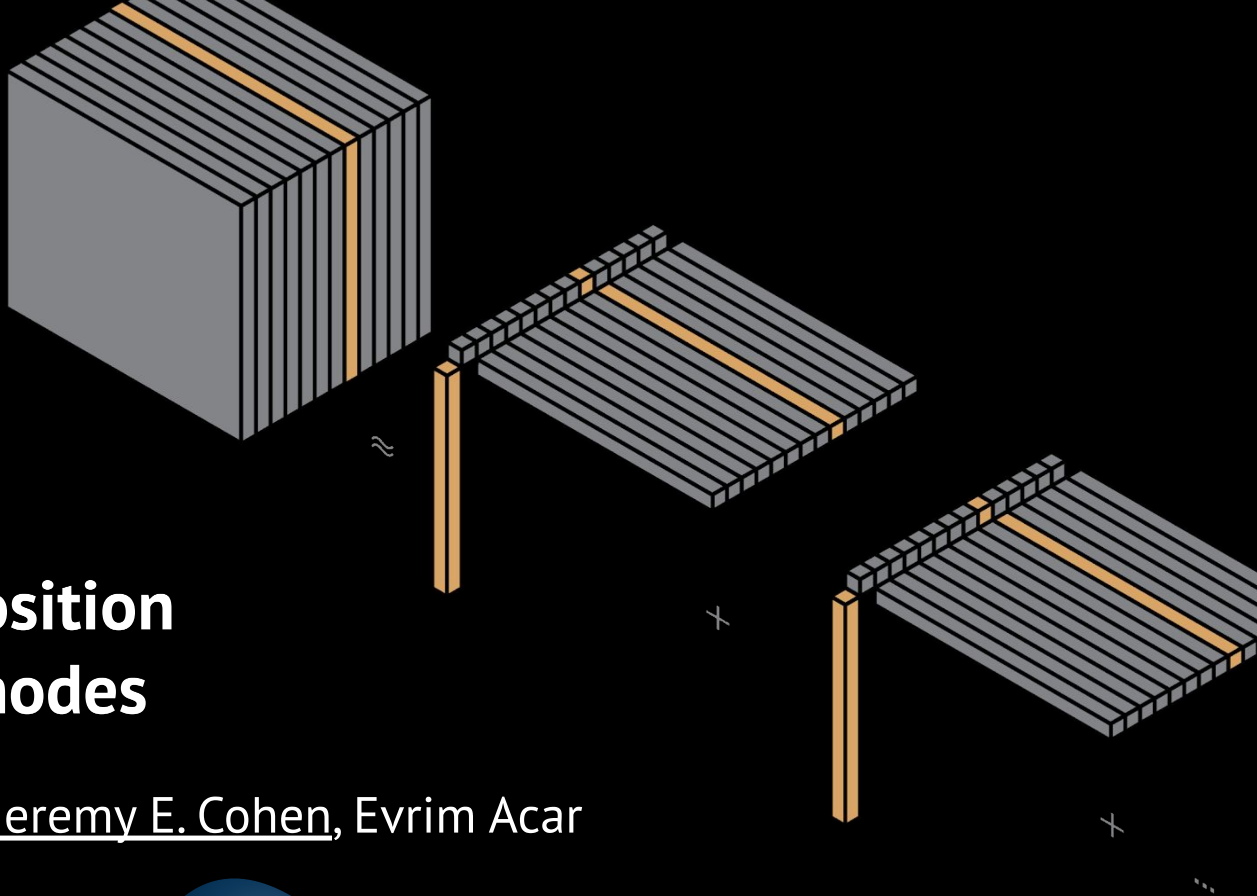
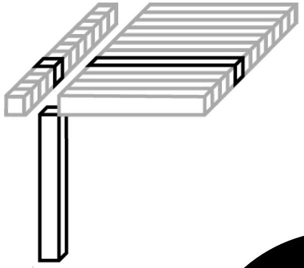


PARAFAC2 Decomposition constrained on all modes

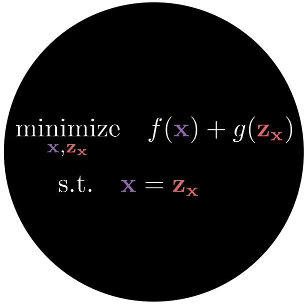
Marie Roald, Carla Schenker, Jeremy E. Cohen, Evrim Acar

GDR ISIS - 19.01.22

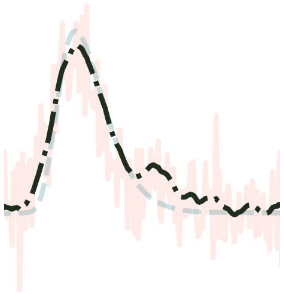
simulamet OSLOMET cnrs



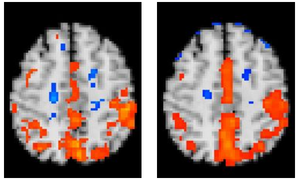
Background and motivation



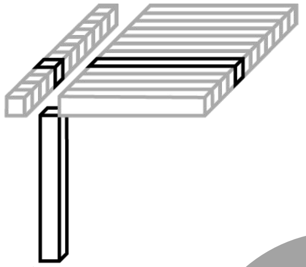
AO-ADMM for evolving components



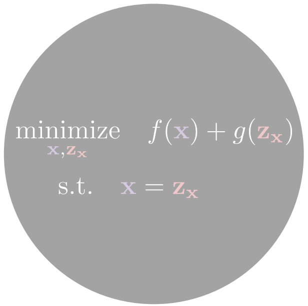
Numerical experiments



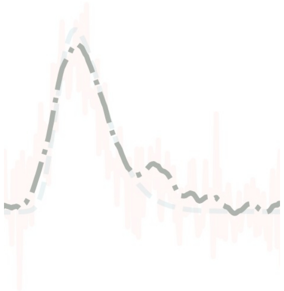
Regularised dynamic networks



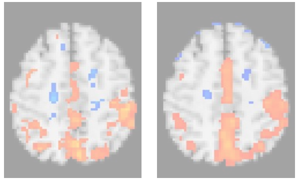
Background and motivation



AO-ADMM for evolving components

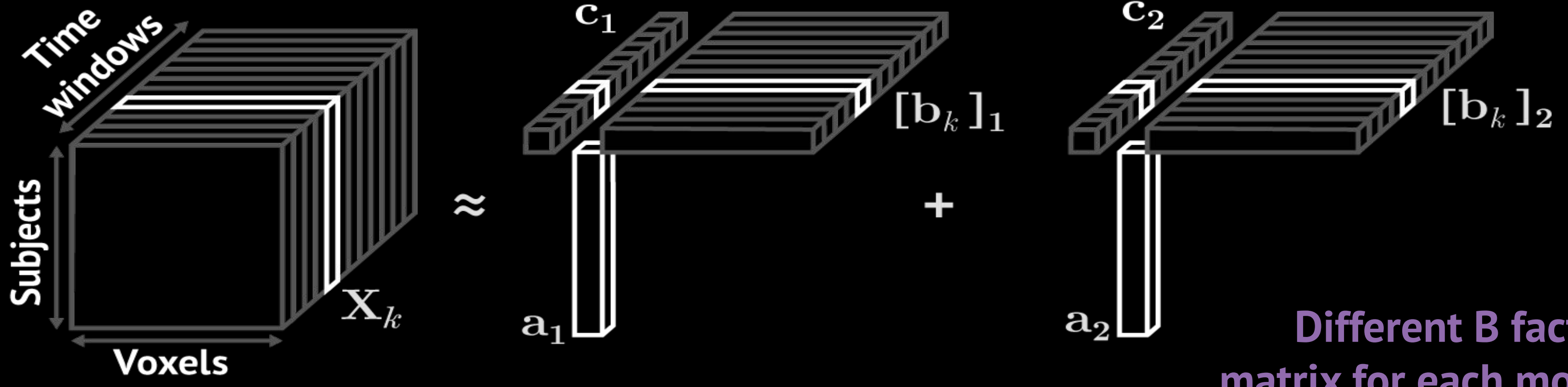


Numerical experiments



Regularised dynamic networks

PARAFAC2 is a tensor decomposition method that allows the $\mathbf{B}$ mode to have a different factor matrix for each frontal slice



Different B factor matrix for each mode

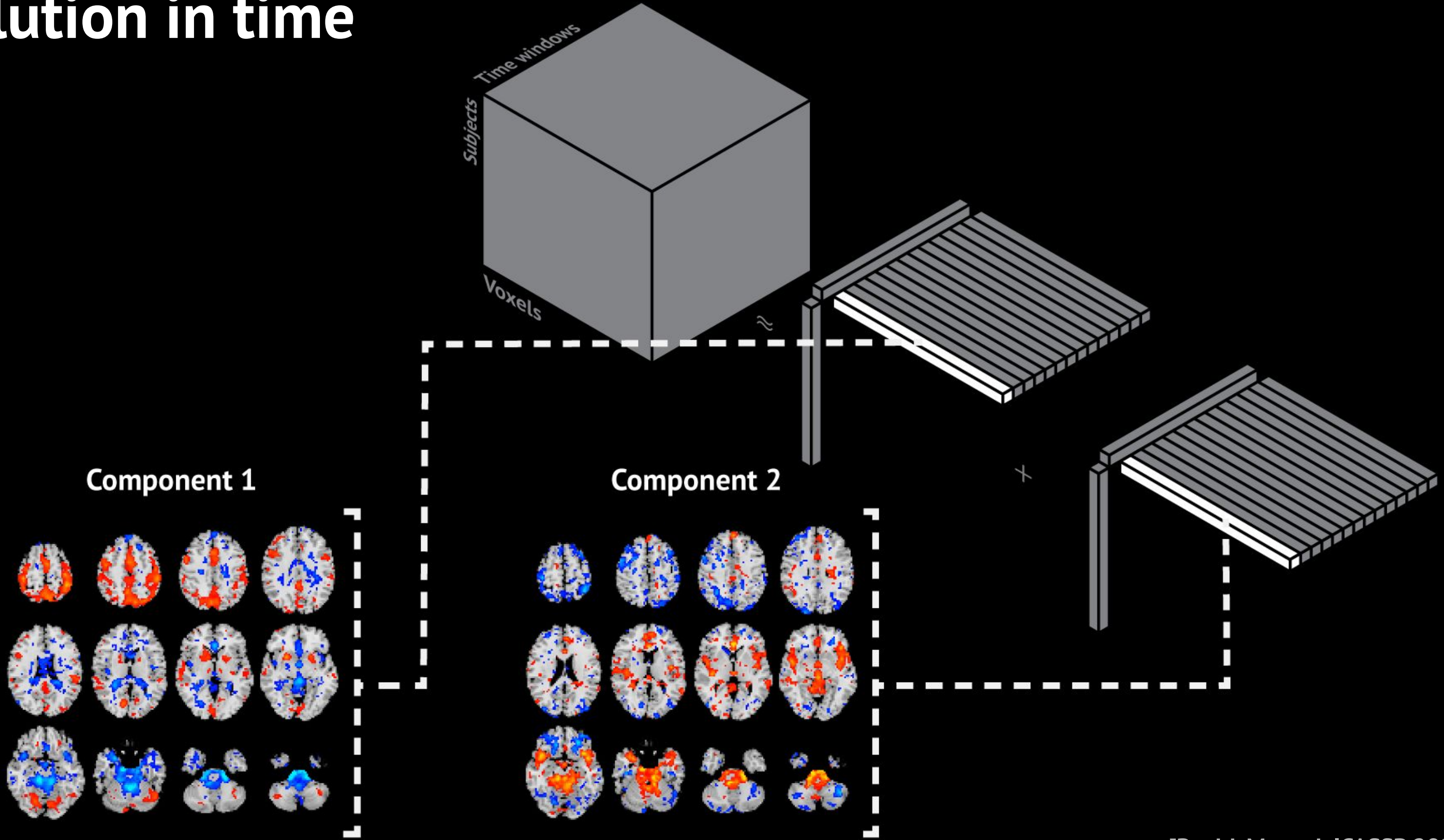
$$\mathbf{X}_k \approx \mathbf{A}\mathrm{diag}\left(\mathbf{c}_{k:}\right)\mathbf{B}_k^{\top}$$

$$\mathbf{B}_k^{\top}\mathbf{B}_k = \Phi$$

Constant cross product for each time step

[Harshman, RA. UCLA working papers in phonetics 1972]

PARAFAC2 captures both the meaningful components and their evolution in time



[Roald, M. et al. ICASSP 2020]

However, the PARAFAC2 model fits the noise more than the PARAFAC model and yields noisy components

Therefore we want to encourage smooth components through regularisation

To solve the PARAFAC2 problem (with ALS):

$$\underset{\mathbf{A},\{\mathbf{B}_k,\mathbf{D}_k\}_{k\leq K}}{\text{minimize}} \quad \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{B}_k^\mathsf{T}\mathbf{B}_k = \Phi \qquad \forall k$$

$\mathbf{D}_k = \text{diag}\left(\mathbf{c}_{k:}\right)$

[Kiers HAL. et al. J. Chemometrics 1999]

To solve the PARAFAC2 problem (with ALS):

$$\underset{\mathbf{A},\{\mathbf{B}_k,\mathbf{D}_k\}_{k\leq K}}{\text{minimize}} \quad \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\top - \mathbf{X}_k \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{B}_k^\top\mathbf{B}_k = \Phi \qquad \forall k$$

We reformulate it to this problem

$$\underset{\mathbf{A},\boldsymbol{\Delta}_\mathbf{B}\{\mathbf{P}_k,\mathbf{D}_k\}_{k\leq K}}{\text{minimize}} \quad \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k\boldsymbol{\Delta}_\mathbf{B}{}^\top\mathbf{P}_k^\top - \mathbf{X}_k \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{P}_k^\top\mathbf{P}_k = \mathbf{I} \qquad \forall k$$

$$\mathbf{D}_k = \text{diag}\left(\mathbf{c}_{k:}\right)$$

[Kiers HAL. et al. J. Chemometrics 1999]

To solve the PARAFAC2 problem (with ALS):

$$\underset{\mathbf{A},\{\mathbf{B}_k,\mathbf{D}_k\}_{k\leq K}}{\text{minimize}} \quad \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\top - \mathbf{X}_k \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{B}_k^\top \mathbf{B}_k = \Phi \qquad \forall k$$

We reformulate it to this problem

$$\underset{\mathbf{A},\boldsymbol{\Delta}_\mathbf{B}\{\mathbf{P}_k,\mathbf{D}_k\}_{k\leq K}}{\text{minimize}} \quad \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k \boldsymbol{\Delta}_\mathbf{B}{}^\top \mathbf{P}_k^\top - \mathbf{X}_k \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{P}_k^\top \mathbf{P}_k = \mathbf{I} \qquad \forall k$$

Constraint

Evolving components

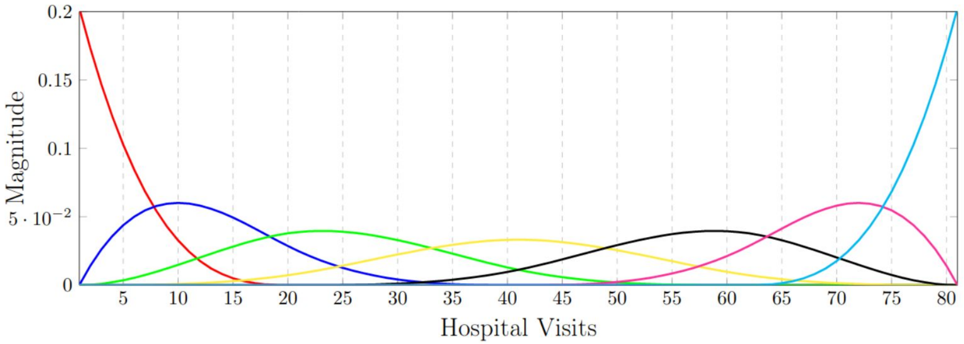
$\mathbf{D}_k = \text{diag}\left(\mathbf{c}_{k:}\right)$

[Kiers HAL. et al. J. Chemometrics 1999]

Previous work ensures smooth components by projecting the data onto a subspace of smooth data

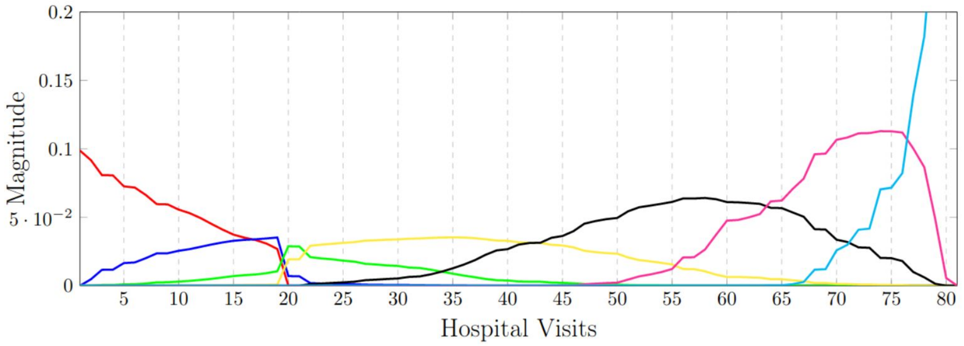
Standard B-spline basis vectors

[Helwig, N.E. Biometrical Journal 2017]

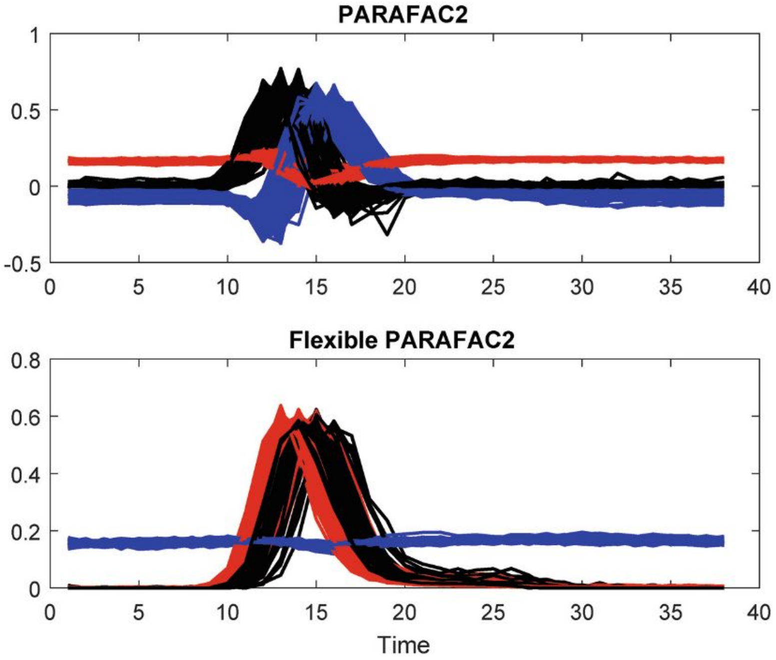


Informed M-spline basis vectors

[Afshar, A. et al. CIKM 2018]

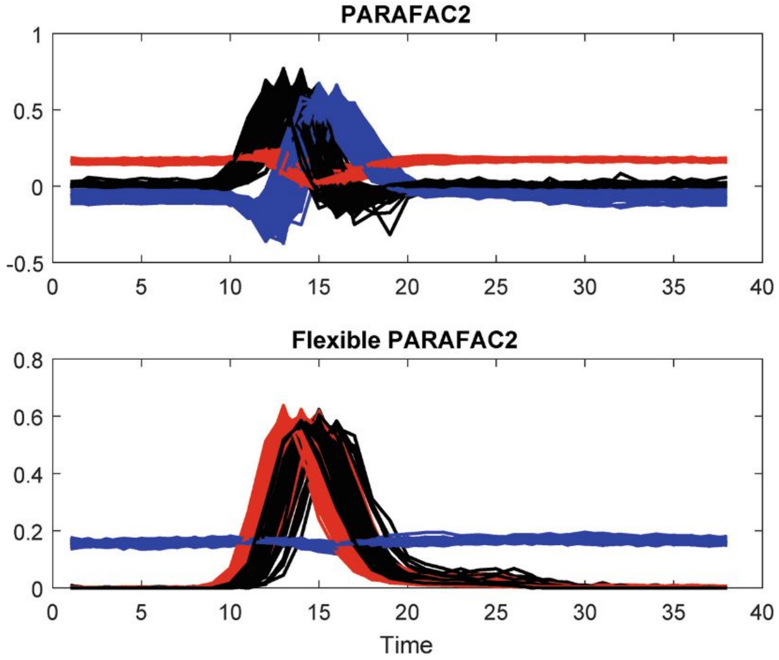


Non-negativity has been imposed via a flexible coupling approach and with a PARAFAC2 inspired regulariser



[Cohen, JE. Bro, R. LVA/ICA 2018]

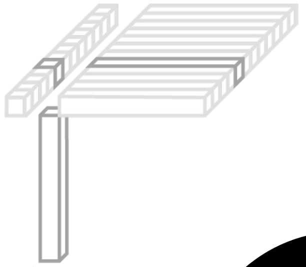
Non-negativity has been imposed via a flexible coupling approach and with a PARAFAC2 inspired regulariser



[Cohen, JE. Bro, R. LVA/ICA 2018]

$$\mathcal{R}_1 = \sum_{k=1}^{K} \frac{\mu}{2} \left\| \mathbf{U}_k{}^\top \mathbf{U}_k - \mathbf{\Phi} \right\|_F^2 ,$$

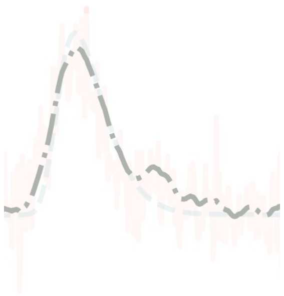
[Yin K. et al. KDD 2020]



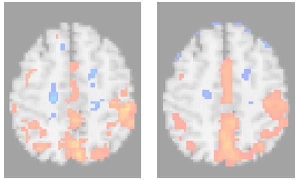
Background and motivation

$$\begin{aligned}\underset{\mathbf{x},\mathbf{z}_\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) + g(\mathbf{z}_\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} = \mathbf{z}_\mathbf{x}\end{aligned}$$

AO-ADMM for evolving components



Numerical experiments



Regularised dynamic networks

We propose using ADMM to update the $\mathbf{B}_k$-components

$$\underset{\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x}) + g(\ \mathbf{x}\ )$$

We propose using ADMM to update the $\mathbf{B}_k$-components

$$\underset{\mathbf{x},\mathbf{z}_\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x}) + g(\mathbf{z}_\mathbf{x})$$

Auxiliary variable for the regularisation

We propose using ADMM to update the $\mathbf{B}_k$-components

$$\underset{\mathbf{x},\mathbf{z}_\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x}) + g(\mathbf{z}_\mathbf{x})$$

$$\text{s.t.} \quad \mathbf{x} = \mathbf{z}_\mathbf{x}$$

Auxiliary variable for the regularisation

We used the following splitting scheme to find $\mathbf{B}_k$-matrices with regularisation

$$\begin{aligned}\underset{\{\mathbf{B}_k, \mathbf{Z}_{\mathbf{B}_k}, \mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}}{\text{minimize}} \quad & \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) \\ \text{s.t.} \quad \mathbf{B}_k &= \mathbf{Z}_{\mathbf{B}_k}, \qquad \forall k \\ \mathbf{B}_k &= \mathbf{Y}_{\mathbf{B}_k}, \qquad \forall k \\ \mathbf{Y}_{\mathbf{B}_k}^\mathsf{T}\mathbf{Y}_{\mathbf{B}_k} &= \Phi, \qquad \forall k\end{aligned}$$

We used the following splitting scheme to find $\mathbf{B}_k$-matrices with regularisation

$$
\begin{aligned}
\underset{\{\mathbf{B}_k, \mathbf{Z}_{\mathbf{B}_k}, \mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}}{\text{minimize}} \quad & \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) \\
\text{s.t.} \quad & \mathbf{B}_k = \mathbf{Z}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{B}_k = \mathbf{Y}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{Y}_{\mathbf{B}_k}^\mathsf{T} \mathbf{Y}_{\mathbf{B}_k} = \Phi, \qquad \forall k
\end{aligned}
$$

We used the following splitting scheme to find $\mathbf{B}_k$-matrices with regularisation

$$
\begin{aligned}
\underset{\{\mathbf{B}_k, \mathbf{Z}_{\mathbf{B}_k}, \mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}}{\text{minimize}} \quad & \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) \\
\text{s.t.} \quad & \mathbf{B}_k = \mathbf{Z}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{B}_k = \mathbf{Y}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{Y}_{\mathbf{B}_k}^\mathsf{T} \mathbf{Y}_{\mathbf{B}_k} = \Phi, \qquad \forall k
\end{aligned}
$$

We used the following splitting scheme to find $\mathbf{B}_k$-matrices with regularisation

$$
\begin{aligned}
\underset{\{\mathbf{B}_k, \mathbf{Z}_{\mathbf{B}_k}, \mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}}{\text{minimize}} \quad & \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) \\
\text{s.t.} \quad & \mathbf{B}_k = \mathbf{Z}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{B}_k = \mathbf{Y}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{Y}_{\mathbf{B}_k}^\mathsf{T} \mathbf{Y}_{\mathbf{B}_k} = \Phi, \qquad \forall k
\end{aligned}
$$

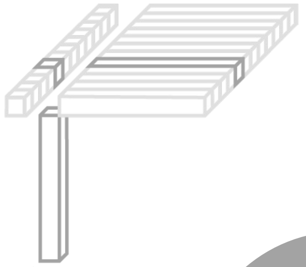
We used the following splitting scheme to find $\mathbf{B}_k$-matrices with regularisation

$$
\begin{aligned}
\underset{\{\mathbf{B}_k, \mathbf{Z}_{\mathbf{B}_k}, \mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}}{\text{minimize}} \quad & \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^\top - \mathbf{X}_k \right\|_F^2 + g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) \\
\text{s.t.} \quad & \mathbf{B}_k = \mathbf{Z}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{B}_k = \mathbf{Y}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{Y}_{\mathbf{B}_k}^\top \mathbf{Y}_{\mathbf{B}_k} = \Phi, \qquad \forall k
\end{aligned}
$$

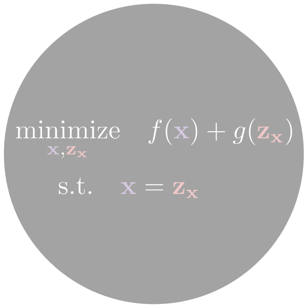
To obtain a problem that can be solved by ADMM, we use an implicit constraint instead of an explicit constraint for $\mathbf{Y}_{\mathbf{B}_k}$

$$
\begin{aligned}
\underset{\{\mathbf{B}_k, \mathbf{Z}_{\mathbf{B}_k}, \mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}}{\text{minimize}} \quad & \sum_{k=1}^{K} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \iota_{\text{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) \\
\text{s.t.} \quad & \mathbf{B}_k = \mathbf{Z}_{\mathbf{B}_k}, \qquad \forall k \\
& \mathbf{B}_k = \mathbf{Y}_{\mathbf{B}_k}, \qquad \forall k
\end{aligned}
$$

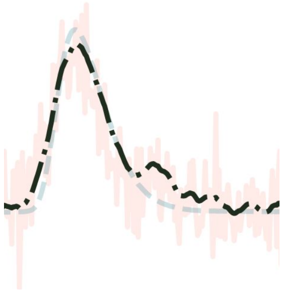
$$
\iota_{\text{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) = \begin{cases} 0, & \text{if } \mathbf{Y}_{\mathbf{B}_k}^\mathsf{T} \mathbf{Y}_{\mathbf{B}_k} = \Phi \quad \forall k \\ \infty, & \text{otherwise} \end{cases}
$$



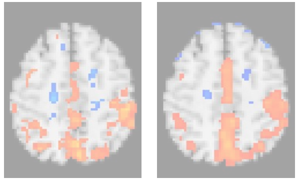
Background and motivation



AO-ADMM for evolving components



Numerical experiments



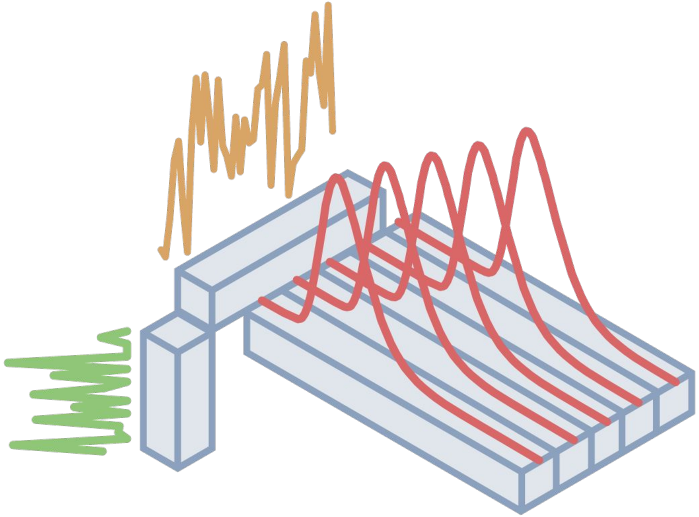
Regularised dynamic networks

To evaluate the performance for different constraints we used the following setup

A - Truncated normal distribution (I=20)

B - Created based on the regularisation/constraints and shifted for each timestep (J={30, 201, 200})

C - Uniform between 0.1 and 1.1 (K={20, 40})

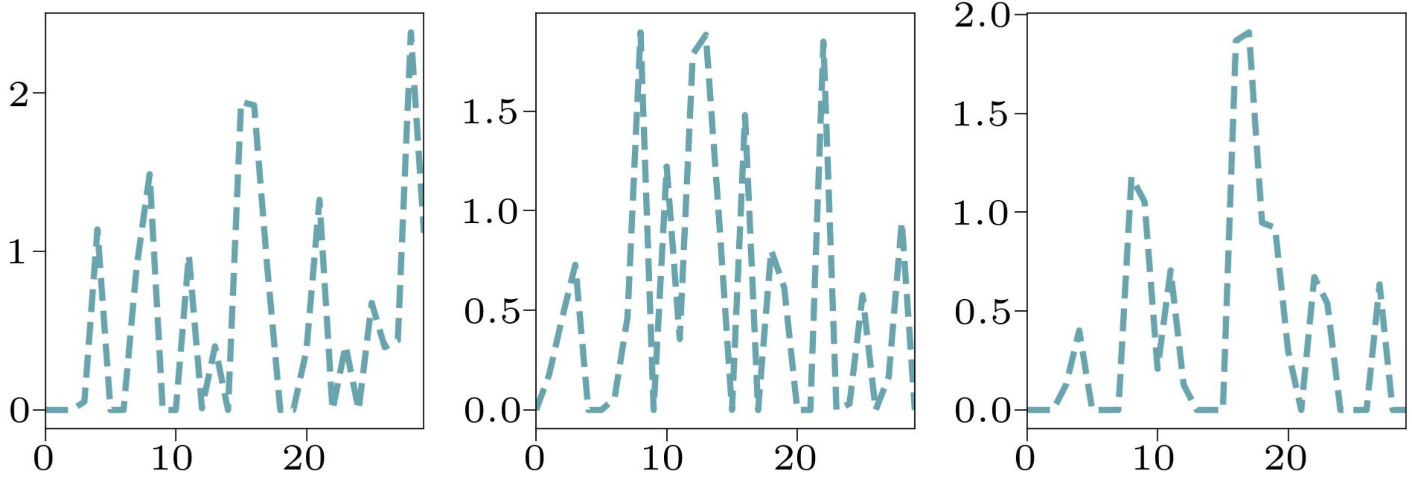


$$\mathcal{X}_{\text{noise}} = \mathcal{X} + \eta \mathcal{E} \frac{||\mathcal{X}||_F}{||\mathcal{E}||_F} \quad \mathcal{E}_{ijk} \sim \mathcal{N}(0, 1)$$

50 different datasets each setup, decomposed with **5** random initialisations for all models, selected model with lowest SSE.

[Roald M. et al. EUSIPCO 2021]

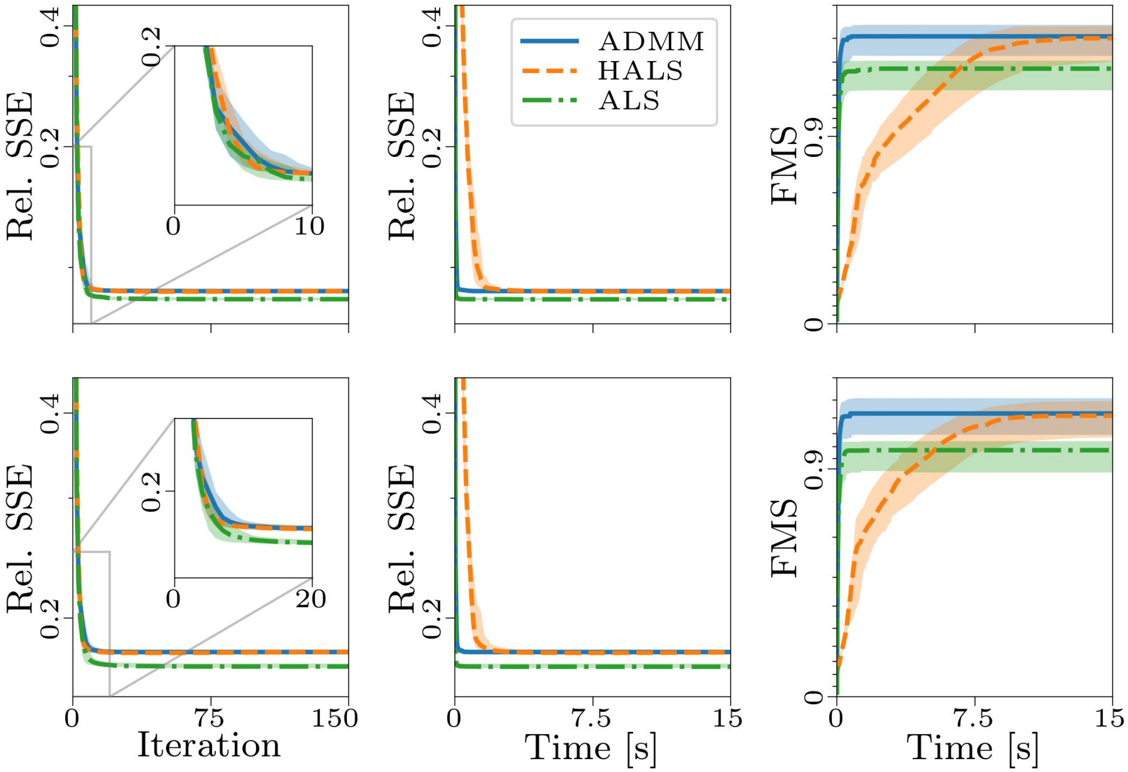
To compare performance of non-negative PARAFAC2, we used $\mathbf{B}_k$ matrices with elements from a truncated normal distribution



AO-ADMM fits the data quicker than flexible, ALS has no constraints on B mode and overfits to the noise

Noise level 0.33

Noise level 0.5



$$\mathrm{FMS} = \frac{1}{R}\sum_{r=1}^{R} \mathbf{a}_r^\mathsf{T}\hat{\mathbf{a}}_r \mathbf{b}_r^\mathsf{T}\hat{\mathbf{b}}_r \mathbf{c}_r^\mathsf{T}\hat{\mathbf{c}}_r,$$

HALS: [Cohen, JE. Bro, R. LVA/ICA 2018]
ADMM : [Roald M. et al. EUSIPCO 2021]

There are also a variety of *structure imposing* regularisation penalties

Encode known similarities of features

$$\mathbf{x}^\top \mathbf{L}\mathbf{x} = \sum_{n=1}^{N-1} (x_{n+1} - x_n)^2$$

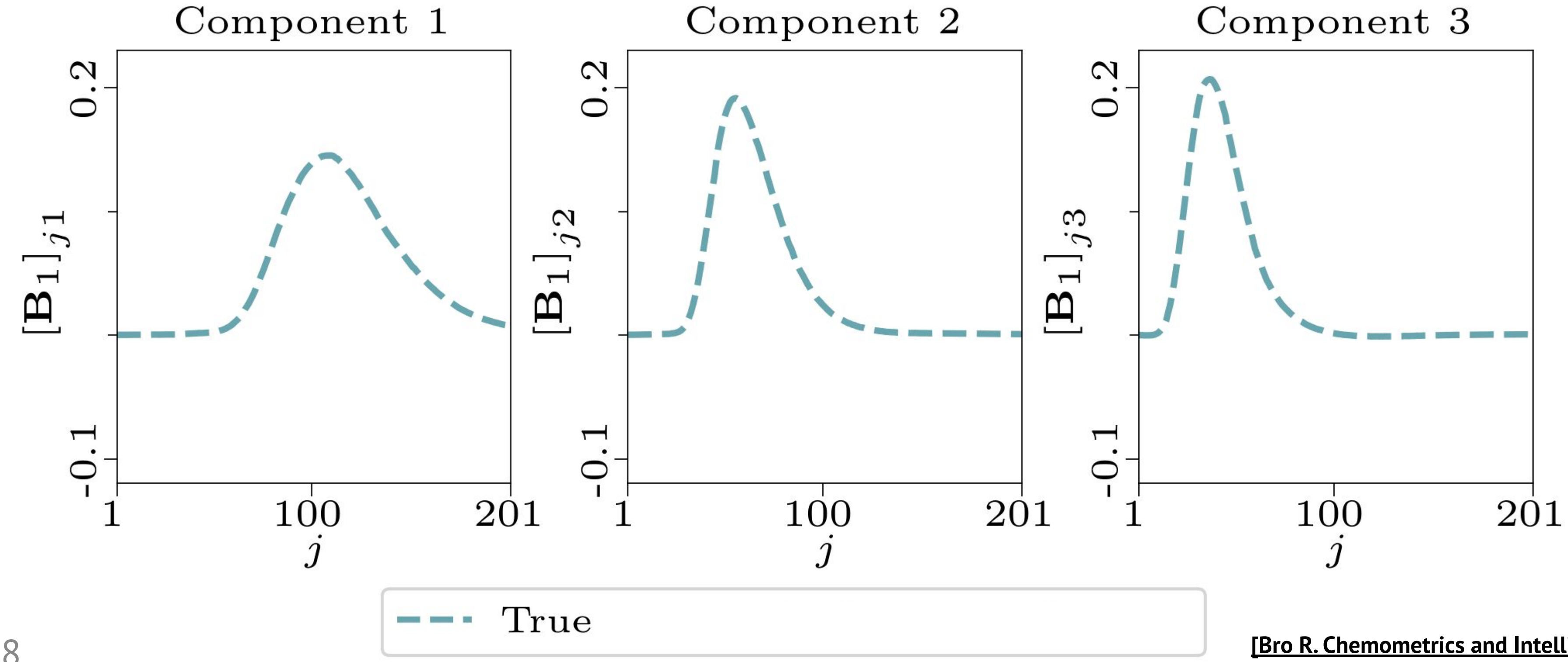
Encode piecewise constant components

$$\|\mathbf{x}\|_{TV} = \sum_{n=1}^{N-1} |x_{n+1} - x_n|$$

Encode bounded components

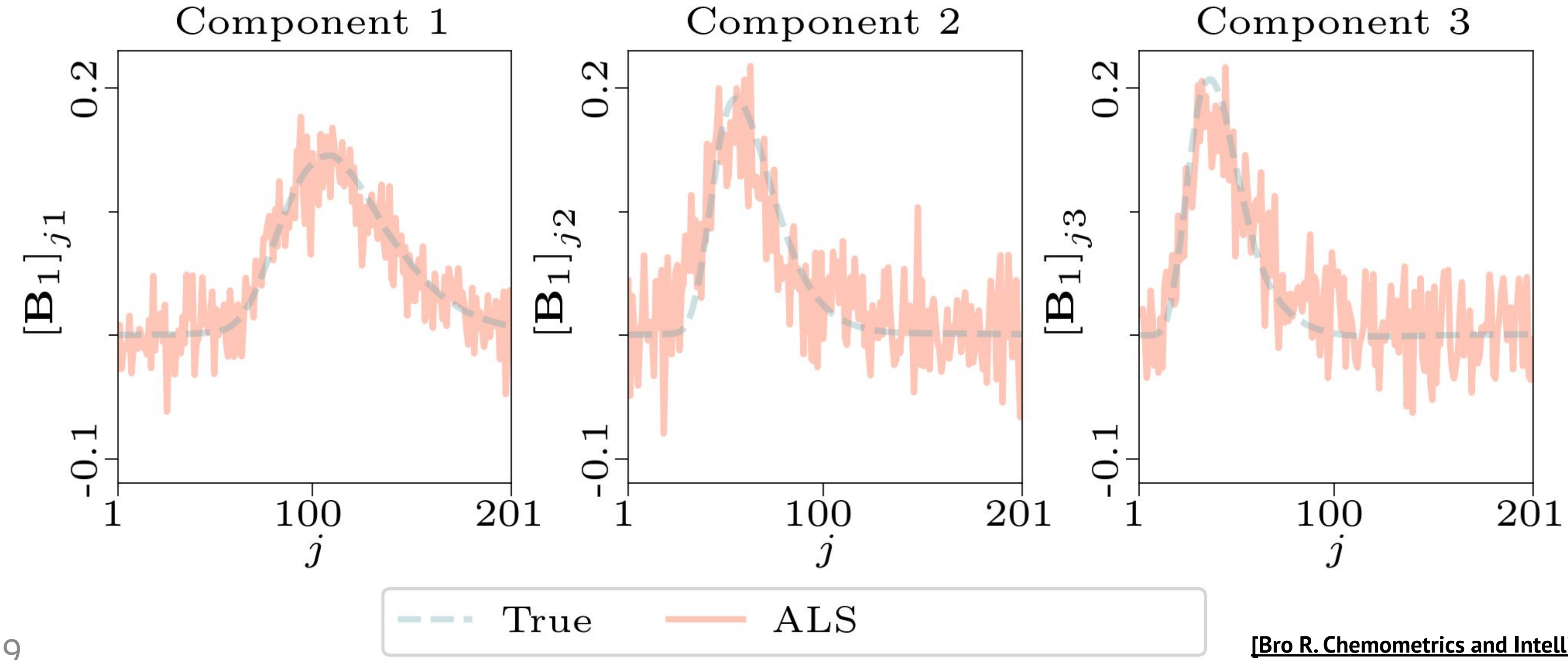
$$\|\mathbf{x}\|^2 = \sum_{n=1}^{N} x_n^2$$

To evaluate the performance of smoothness regularisation we used components from fluorescence spectroscopy



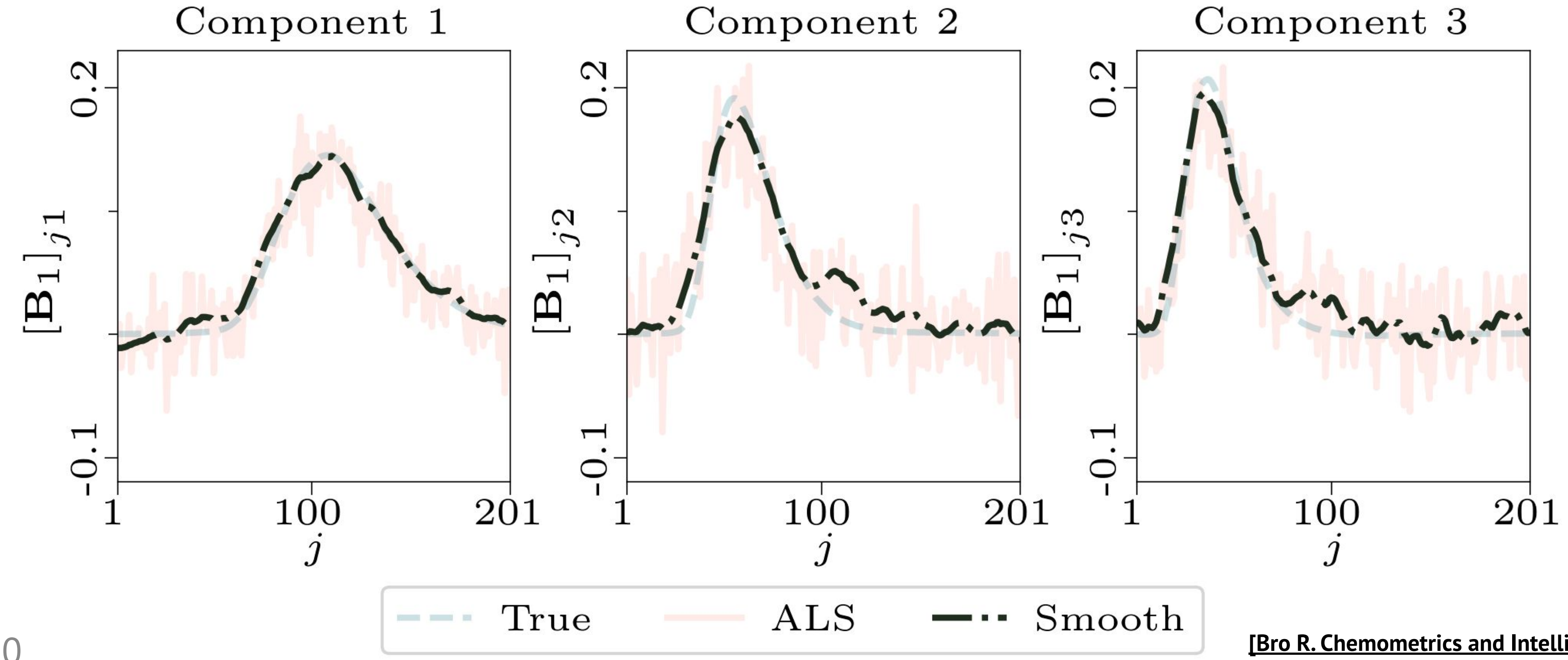
[Bro R. Chemometrics and Intelligent Laboratory Systems 1997]

Standard unregularised PARAFAC2 (ALS) finds noisy components with the right overall shape



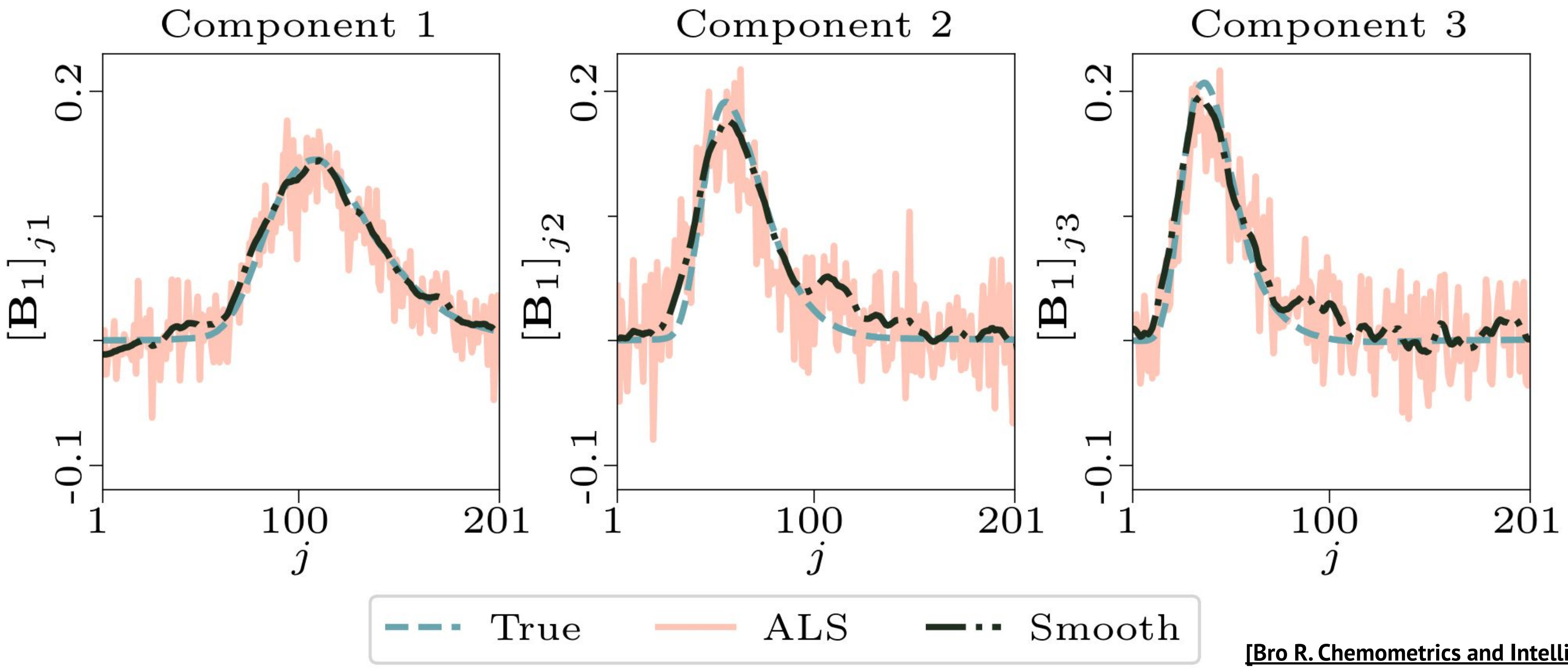
[Bro R. Chemometrics and Intelligent Laboratory Systems 1997]

Graph laplacian (smoothness)-based regularisation finds components that are closer to the truth



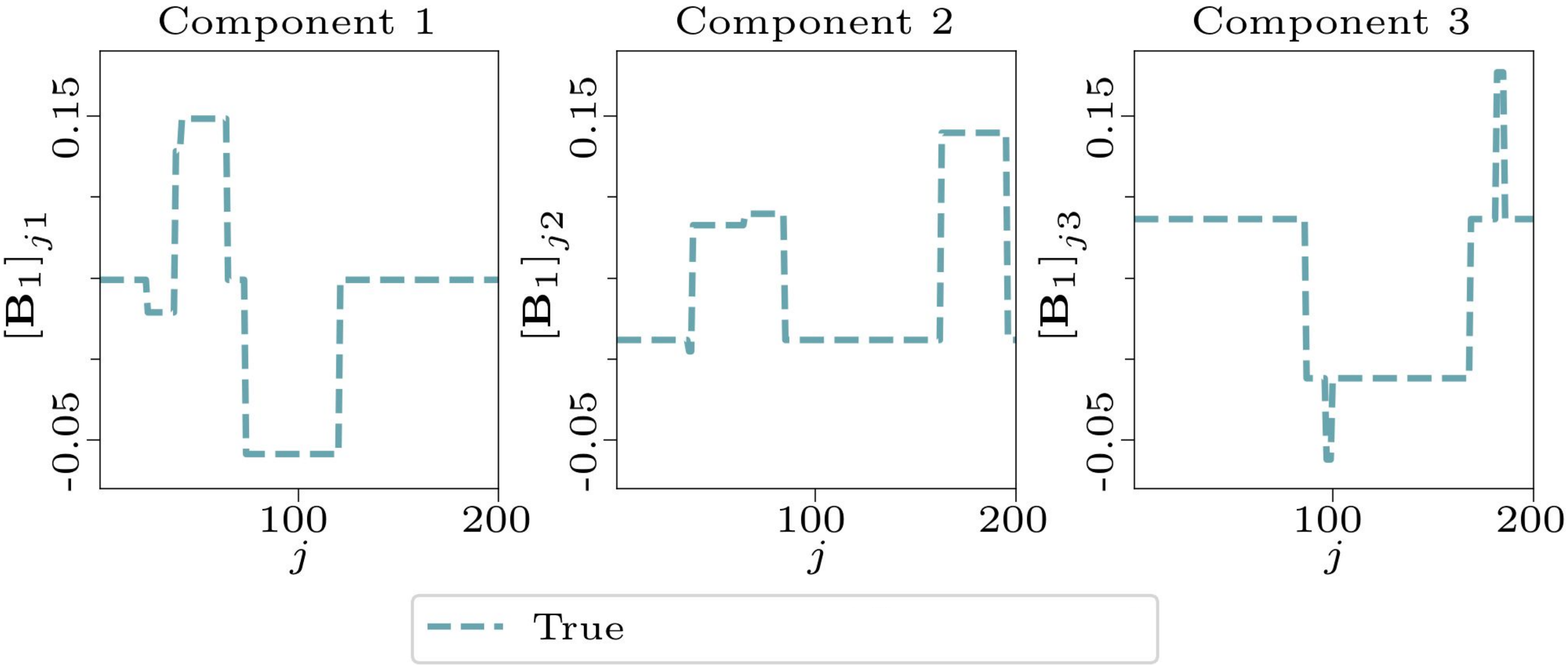
[Bro R. Chemometrics and Intelligent Laboratory Systems 1997]

Graph laplacian (smoothness)-based regularisation finds components that are closer to the truth

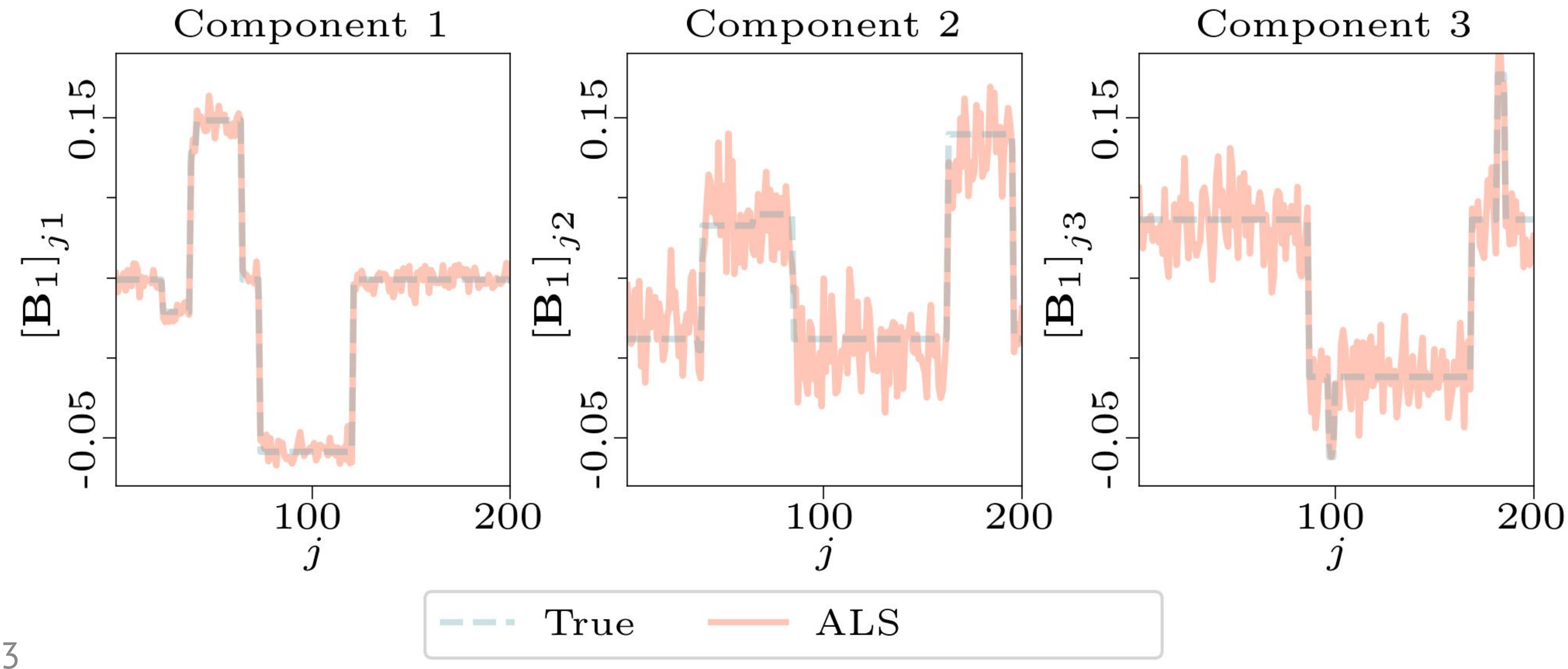


[Bro R. Chemometrics and Intelligent Laboratory Systems 1997]

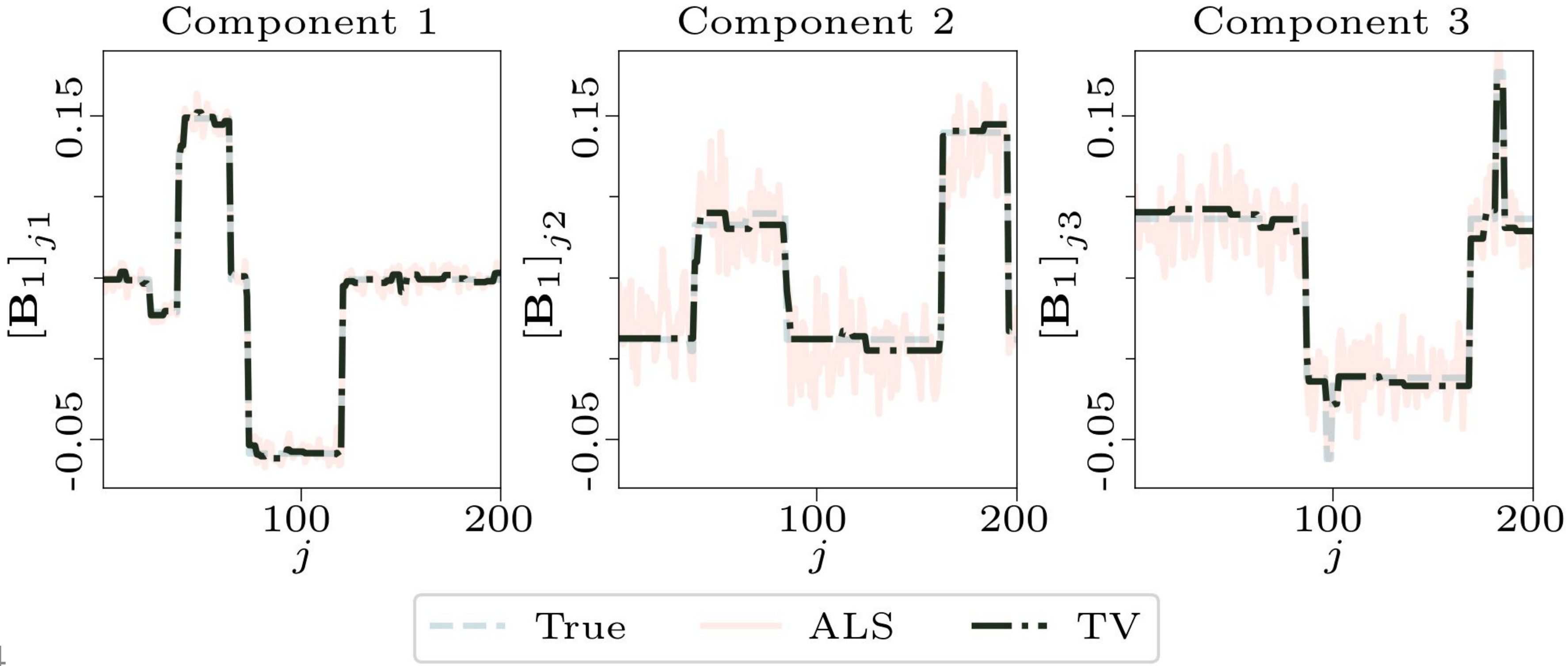
To evaluate the performance with TV regularisation, we simulated piecewise constant components with 6 jumps



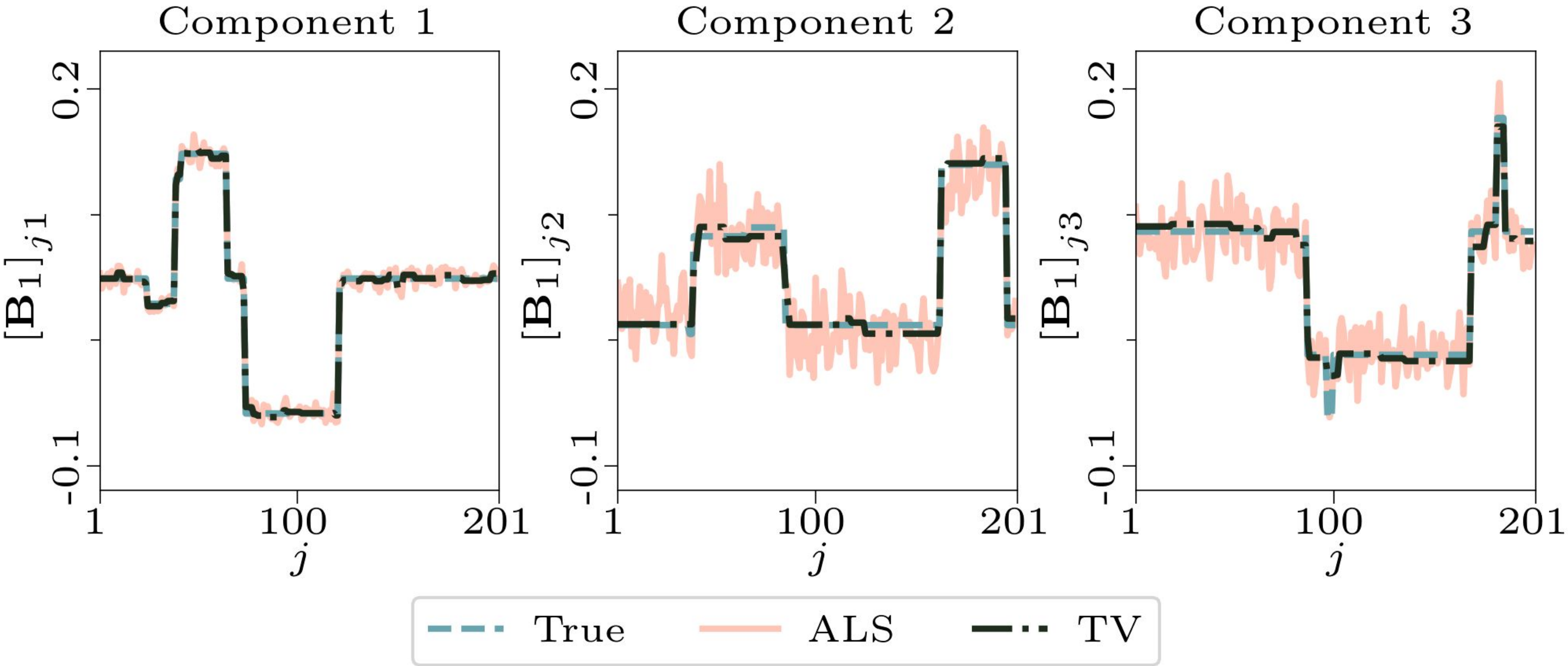
Again, the standard ALS algorithm yields noisy components



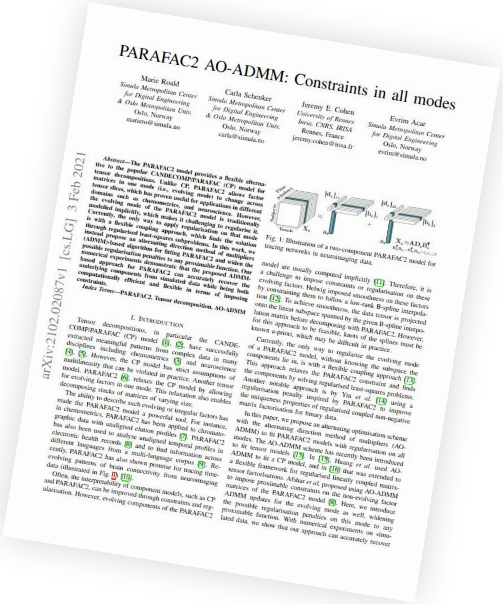
The components obtained with TV regularisation are closer to the ground truth



The components obtained with TV regularisation are closer to the ground truth



More details about the experiment setup and the results are available in the papers



PARAFAC2 AO-ADMM: Constraints in all modes

Marie Roald, Carla Schenker, Jeremy E. Cohen, Evrim Acar

arXiv:2102.02087v1 [cs.LG] 3 Feb 2021

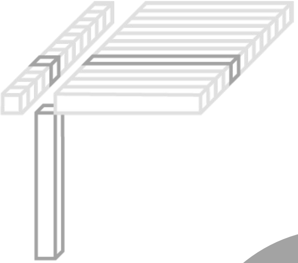
An AO-ADMM approach to constraining PARAFAC2 on all modes*

Marie Roald, Carla Schenker, Rasmus Bro, Jeremy E. Cohen, Evrim Acar

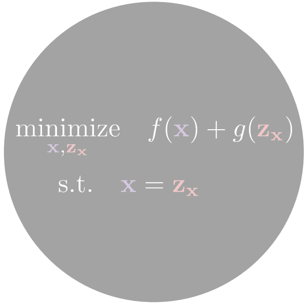
arXiv:2110.01278v1 [cs.LG] 4 Oct 2021

M. Roald, C. Schenker, J. E. Cohen, E. Acar.
PARAFAC2 AO-ADMM: Constraints in all modes.
EUSIPCO2021
Submitted to SIMODS, arxiv preprint available

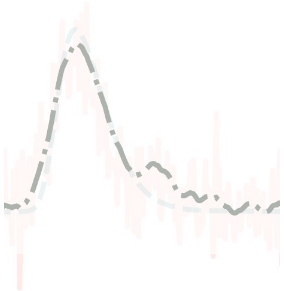
Additional details and the code is available
github.com/MarieRoald/PARAFAC2-AOADMM-EUSIPCO21
github.com/MarieRoald/PARAFAC2-AOADMM-SIMODS



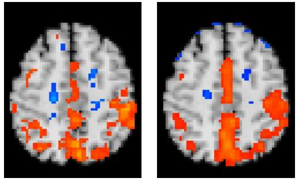
Background and motivation



AO-ADMM for evolving components

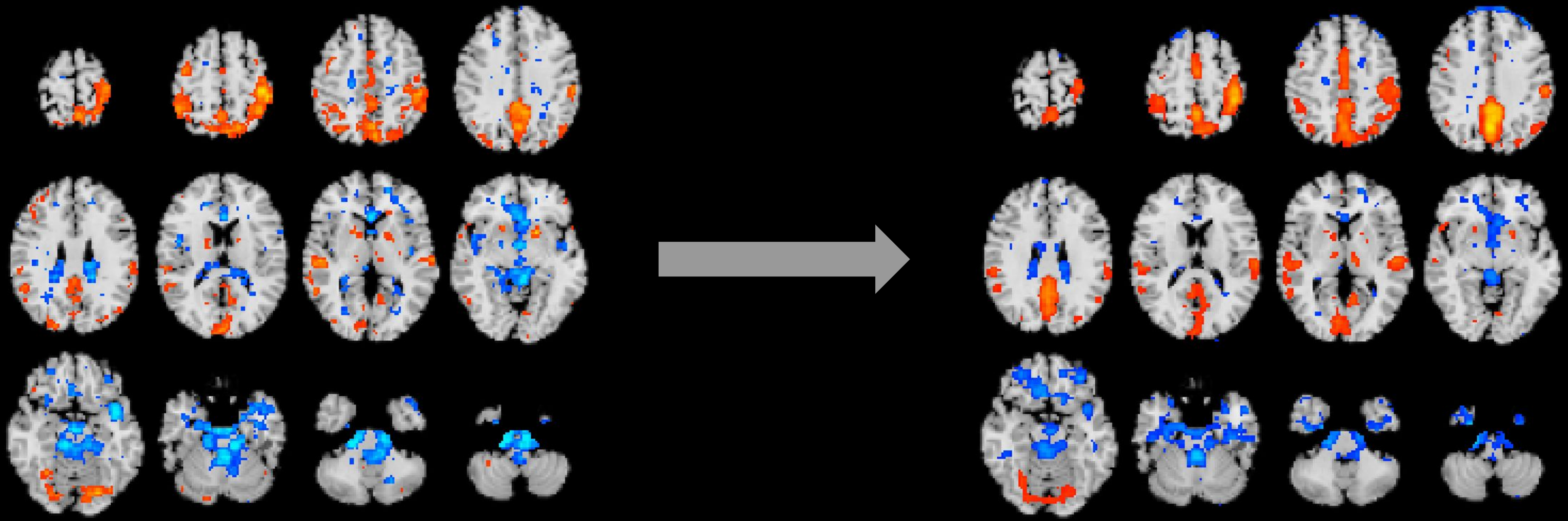


Numerical experiments

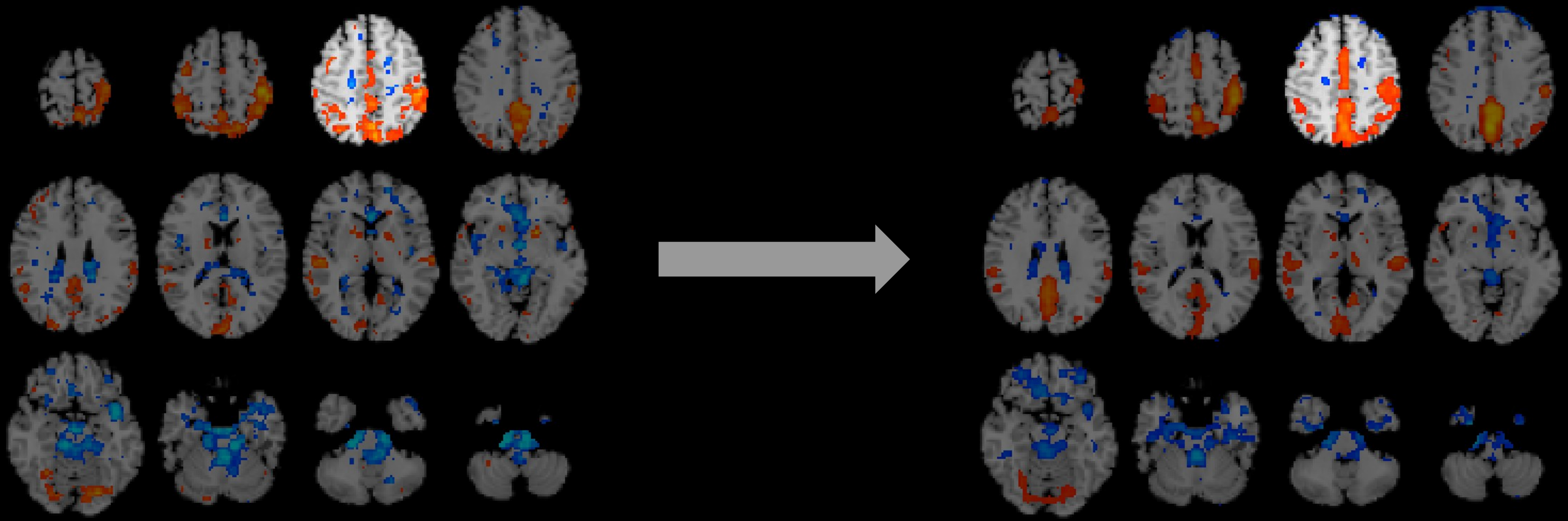


Regularised dynamic networks

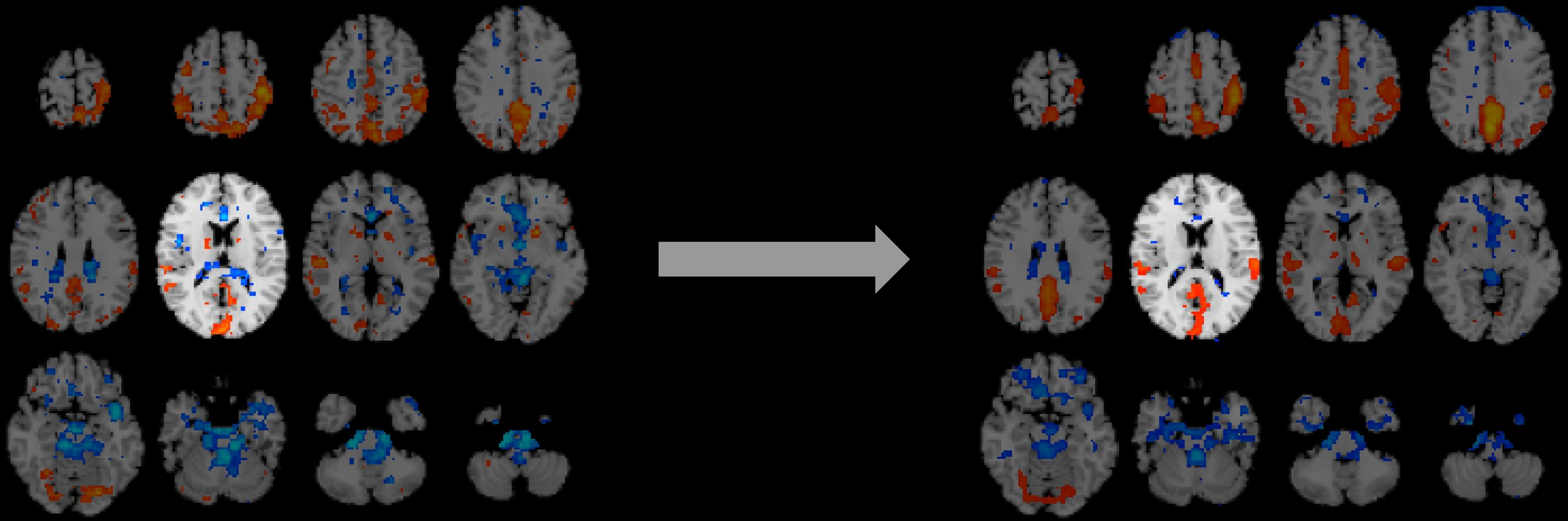
Smoothness regularisation leads to less noisy brain-activation maps when applied to fMRI data



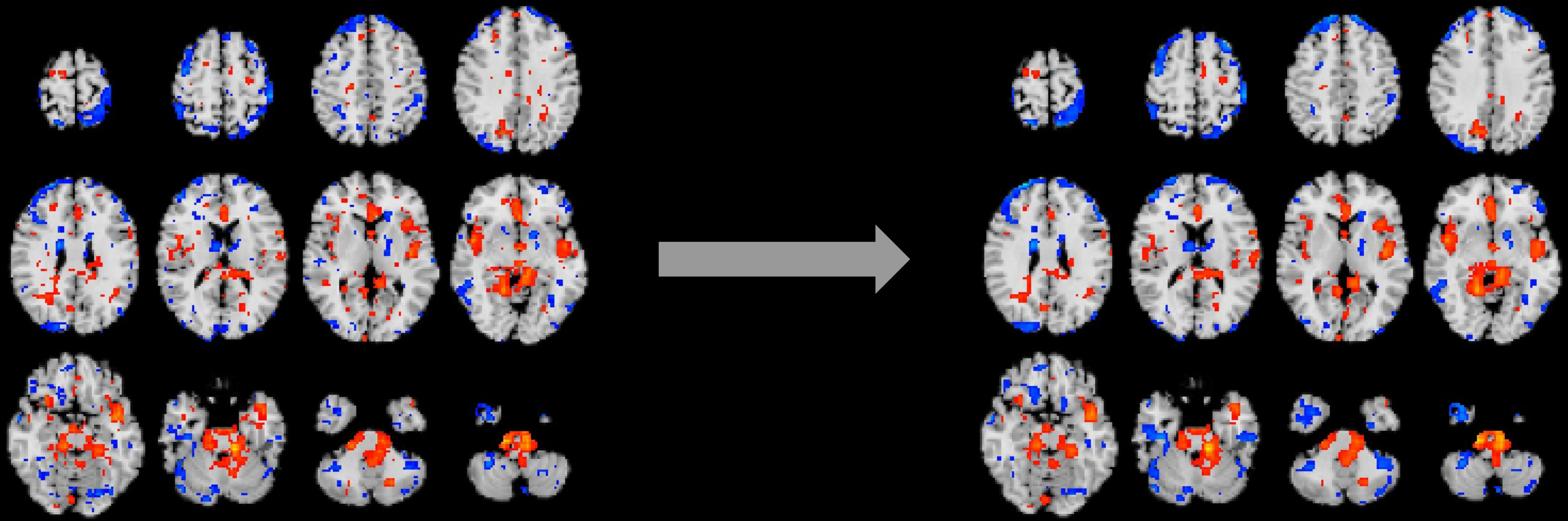
Smoothness regularisation leads to less noisy brain-activation maps when applied to fMRI data



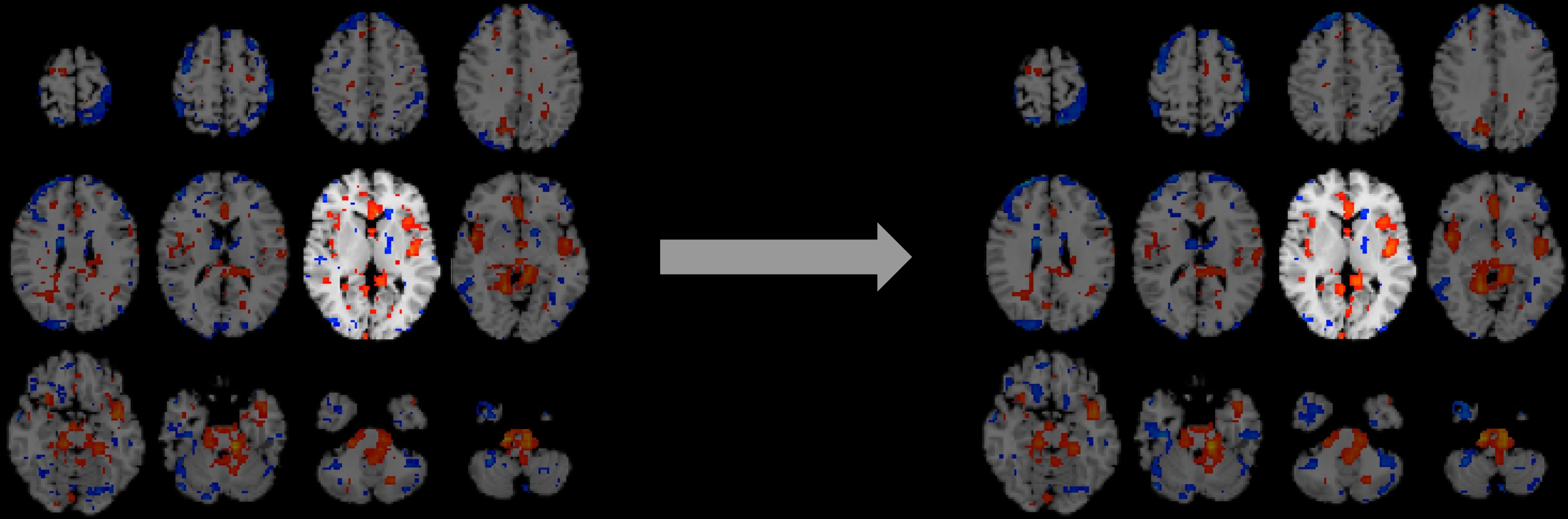
Smoothness regularisation leads to less noisy brain-activation maps when applied to fMRI data



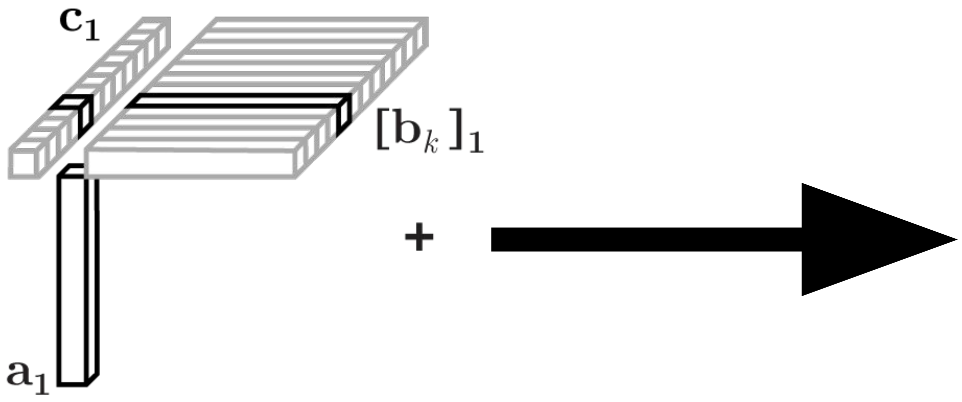
Smoothness regularisation leads to less noisy brain-activation maps when applied to fMRI data

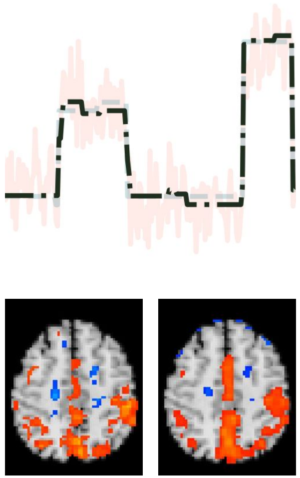


Smoothness regularisation leads to less noisy brain-activation maps when applied to fMRI data



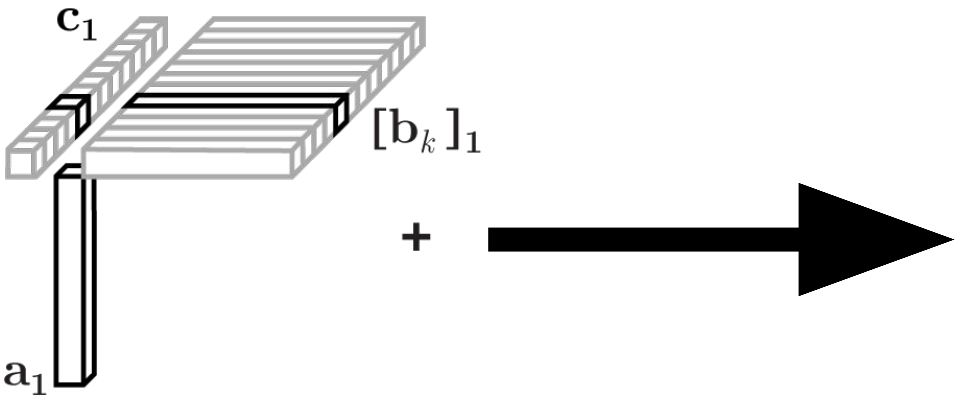
In summary, PARAFAC2 is a promising model for tracing evolving networks, and with PARAFAC2 AO-ADMM, we can improve model interpretability through meaningful constraints

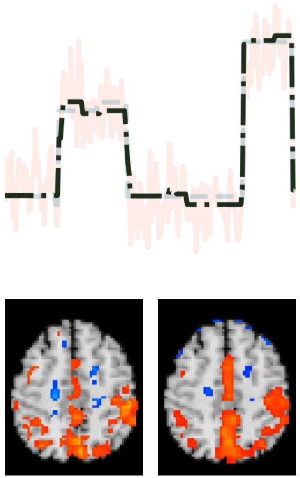




[Roald, M. et al. ICASSP 2020]
[Roald M. et al. EUSIPCO 2021]

In summary, PARAFAC2 is a promising model for tracing evolving networks, and with PARAFAC2 AO-ADMM, we can improve model interpretability through meaningful constraints





Questions?

simulamet OSLOMET cnrs

[Roald, M. et al. ICASSP 2020]
[Roald M. et al. EUSIPCO 2021]

For the AO-ADMM scheme, we fit the modes alternatingly and solve the regularised subproblems with ADMM

Until convergence:

  Update $\mathbf{A}$ matrix

  Update $\mathbf{B}_k$ matrices

  Update $\mathbf{C}$ matrix ($\mathbf{D}_k$ matrices)

[Huang, K. et al. IEEE Trans. Signal Proc. 2015]

The ADMM updates for the A and C matrix are well known, so we focus on how to update the $\mathbf{B}_k$ matrices with regularisation

Until convergence:

Update $\mathbf{A}$ matrix

Update $\mathbf{B}_k$ matrices

Update $\mathbf{C}$ matrix ($\mathbf{D}_k$ matrices)

Using ADMM, we obtain the following update steps:

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{c} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2}\left\| \mathbf{B}_k - (\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 + \\ \frac{\rho_k}{2}\left\| \mathbf{B}_k - (\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\Delta_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 \end{array} \right\}$$

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \frac{\rho_k}{2}\left\|\mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}\right)\right\|_F^2$$

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k\leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2}\left\|\mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\Delta_{\mathbf{B}_k}}^{(t)}\right)\right\|_F^2$$

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

$$\mu_{\Delta_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\Delta_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

Update the components to fit the data well, while still being close to the auxiliary variables

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{c} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 \end{array} \right\}$$

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \frac{\rho_k}{2}\left\|\mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}\right)\right\|_F^2$$

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k\leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2}\left\|\mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t)}\right)\right\|_F^2$$

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

$$\mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{l} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 \end{array} \right\}$$

Update **first auxiliary variable** to follow regularisation while being close to the **components**

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \frac{\rho_k}{2} \left\| \mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

$$\mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{l} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\Delta_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 \end{array} \right\}$$

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} \; g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \frac{\rho_k}{2} \left\| \mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

Update second auxiliary variable to follow the PF2 constraint while being close to the components

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\Delta_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

$$\mu_{\Delta_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\Delta_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{l} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^\mathsf{T} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\Delta_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 \end{array} \right\}$$

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \frac{\rho_k}{2} \left\| \mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\Delta_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

Update the first scaled dual variable to correct the regularisation coupling

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

$$\mu_{\Delta_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\Delta_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{l} \left\| \mathbf{A}\mathbf{D}_k\mathbf{B}_k^{\mathsf{T}} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - (\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t)}) \right\|_F^2 \end{array} \right\}$$

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k}\right) + \frac{\rho_k}{2} \left\| \mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t)}\right) \right\|_F^2$$

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

Update the second scaled dual variable to correct the constraint coupling

$$\mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{\Delta}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

We repeat these steps N times or until convergence for every outer iteration

$$\mathbf{B}_k^{(t+1)} \leftarrow \min_{\mathbf{B}_k} \left\{ \begin{array}{c} \left\| \mathbf{A}\mathbf{D}_k \mathbf{B}_k^{\mathsf{T}} - \mathbf{X}_k \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - \left(\mathbf{Z}_{\mathbf{B}_k}^{(t)} - \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} \right) \right\|_F^2 + \\ \frac{\rho_k}{2} \left\| \mathbf{B}_k - \left(\mathbf{Y}_{\mathbf{B}_k}^{(t)} - \mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t)} \right) \right\|_F^2 \end{array} \right\}$$

$$\mathbf{Z}_{\mathbf{B}_k}^{(t+1)} \leftarrow \min_{\mathbf{Z}_{\mathbf{B}_k}} g_{\mathbf{B}}\left(\mathbf{Z}_{\mathbf{B}_k} \right) + \frac{\rho_k}{2} \left\| \mathbf{Z}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} \right) \right\|_F^2$$

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\text{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K} \right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t)} \right) \right\|_F^2$$

$$\mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\mathbf{Z}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Z}_{\mathbf{B}_k}^{(t+1)}$$

$$\mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t+1)} \leftarrow \mu_{\boldsymbol{\Delta}_{\mathbf{B}_k}}^{(t)} + \mathbf{B}_k^{(t+1)} - \mathbf{Y}_{\mathbf{B}_k}^{(t+1)}$$

To compute the prox for the PARAFAC2 constraint, we use the $\mathbf{Y}_{\mathbf{B}_k} = \mathbf{P}_k \mathbf{\Delta}_{\mathbf{B}}$ reformulation used for unconstrained PARAFAC2

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k\leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}} \iota_{\text{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}\,k}}^{(t)}\right)\right\|_F^2$$

$$\mathbf{P}_k^{(t+1)} \leftarrow \mathbf{U}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T}, \qquad \mathbf{U}_k^{(t+1)} \mathbf{S}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T} = \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}\,k}}^{(t)}\right) {\mathbf{\Delta}_{\mathbf{B}}}^{(t)\mathsf{T}}$$

$${\mathbf{\Delta}_{\mathbf{B}}}^{(t+1)} \leftarrow \frac{1}{\sum_{k=1}^{K} \rho_{\mathbf{B}_k}} \sum_{k=1}^{K} \rho_{\mathbf{B}_k} {\mathbf{P}_k^{(t+1)}}^\mathsf{T} \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}\,k}}^{(t)}\right)$$

To compute the prox for the PARAFAC2 constraint, we use the $\mathbf{Y}_{\mathbf{B}_k} = \mathbf{P}_k \mathbf{\Delta}_\mathbf{B}$ reformulation used for unconstrained PARAFAC2

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}k}}^{(t)}\right) \right\|_F^2$$

$$\mathbf{P}_k^{(t+1)} \leftarrow \mathbf{U}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T}, \qquad \mathbf{U}_k^{(t+1)} \mathbf{S}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T} = \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}k}}^{(t)}\right) {\mathbf{\Delta}_\mathbf{B}}^{(t)\mathsf{T}}$$

$$\mathbf{\Delta}_\mathbf{B}^{(t+1)} \leftarrow \frac{1}{\sum_{k=1}^{K} \rho_{\mathbf{B}_k}} \sum_{k=1}^{K} \rho_{\mathbf{B}_k} {\mathbf{P}_k^{(t+1)}}^\mathsf{T} \left(\mathbf{B}_k^{(t+1)} + \mu_{\mathbf{\Delta}_{\mathbf{B}k}}^{(t)}\right)$$

To compute the prox for the PARAFAC2 constraint, we use the $\mathbf{Y}_{\mathbf{B}_k} = \mathbf{P}_k \boldsymbol{\Delta}_{\mathbf{B}}$ reformulation used for unconstrained PARAFAC2

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k\leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}} \iota_{\mathrm{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k\leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \boldsymbol{\mu}_{\boldsymbol{\Delta}_{\mathbf{B}k}}^{(t)}\right)\right\|_F^2$$

$$\mathbf{P}_k^{(t+1)} \leftarrow \mathbf{U}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T}, \qquad \mathbf{U}_k^{(t+1)} \mathbf{S}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T} = \left(\mathbf{B}_k^{(t+1)} + \boldsymbol{\mu}_{\boldsymbol{\Delta}_{\mathbf{B}k}}^{(t)}\right) {\boldsymbol{\Delta}_{\mathbf{B}}}^{(t)\mathsf{T}}$$

$${\boldsymbol{\Delta}_{\mathbf{B}}}^{(t+1)} \leftarrow \frac{1}{\sum_{k=1}^{K} \rho_{\mathbf{B}_k}} \sum_{k=1}^{K} \rho_{\mathbf{B}_k} {\mathbf{P}_k^{(t+1)}}^\mathsf{T} \left(\mathbf{B}_k^{(t+1)} + \boldsymbol{\mu}_{\boldsymbol{\Delta}_{\mathbf{B}k}}^{(t)}\right)$$

To compute the prox for the PARAFAC2 constraint, we use the $\mathbf{Y}_{\mathbf{B}_k} = \mathbf{P}_k \boldsymbol{\Delta}_{\mathbf{B}}$ reformulation used for unconstrained PARAFAC2

$$\{\mathbf{Y}_{\mathbf{B}_k}^{(t+1)}\}_{k \leq K} \leftarrow \min_{\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}} \iota_{\text{PF2}}\left(\{\mathbf{Y}_{\mathbf{B}_k}\}_{k \leq K}\right) + \sum_{k=1}^{K} \frac{\rho_k}{2} \left\| \mathbf{Y}_{\mathbf{B}_k} - \left(\mathbf{B}_k^{(t+1)} + \boldsymbol{\mu}_{\boldsymbol{\Delta}_{\mathbf{B}k}}^{(t)}\right) \right\|_F^2$$

$$\mathbf{P}_k^{(t+1)} \leftarrow \mathbf{U}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T}, \qquad \mathbf{U}_k^{(t+1)} \mathbf{S}_k^{(t+1)} {\mathbf{V}_k^{(t+1)}}^\mathsf{T} = \left(\mathbf{B}_k^{(t+1)} + \boldsymbol{\mu}_{\boldsymbol{\Delta}_{\mathbf{B}k}}^{(t)}\right) {\boldsymbol{\Delta}_{\mathbf{B}}^{(t)}}^\mathsf{T}$$

$$\boldsymbol{\Delta}_{\mathbf{B}}^{(t+1)} \leftarrow \frac{1}{\sum_{k=1}^{K} \rho_{\mathbf{B}_k}} \sum_{k=1}^{K} \rho_{\mathbf{B}_k} {\mathbf{P}_k^{(t+1)}}^\mathsf{T} \left(\mathbf{B}_k^{(t+1)} + \boldsymbol{\mu}_{\boldsymbol{\Delta}_{\mathbf{B}k}}^{(t)}\right)$$