The Parallel Factor Analysis and the Tucker Model: a Simulation Study

Zuzana Rošťáková, Roman Rosipal

Institute of Measurement Science
Slovak Academy of Sciences

ODAM 2019

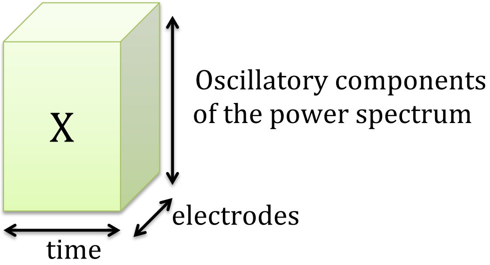
Motivation example

- real **EEG data**
  - mental activity, no real movements
  - **"time":** $\approx$ 5 minutes
  - **"space":** 10 electrodes
  - **"frequency":** oscillatory part of power spectrum in the range 4-25 Hz

$$\Rightarrow \text{a three–dimensional tensor } X \in \mathbb{R}^{I \times J \times K}$$

- **goal:** to detect hidden sources of neural activity

  → Parallel Factor Analysis was used in [Miwakeichi et al., 2004], [Rosipal et al., 2019],. . .



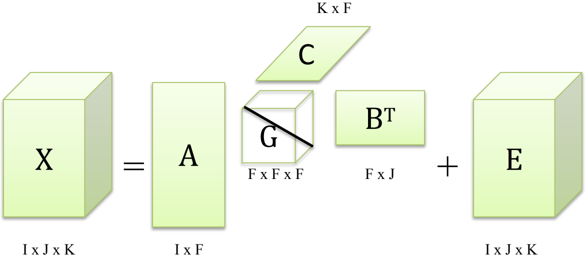
PARAllel FACtor Analysis (PARAFAC)

- [Harshman, 1970] (PARAFAC), [Carroll and Chang, 1970] (CANDECOMP)

$X \in \mathbb{R}^{\mathbb{I}\times\mathbb{J}\times\mathbb{K}}$:

$$X_{ijk} = \sum_{f=1}^{F} g_{fff}\, a_{if}\, b_{jf}\, c_{kf} + e_{ijk},$$

$$\|a_f\| = \|b_f\| = \|c_f\| = 1$$

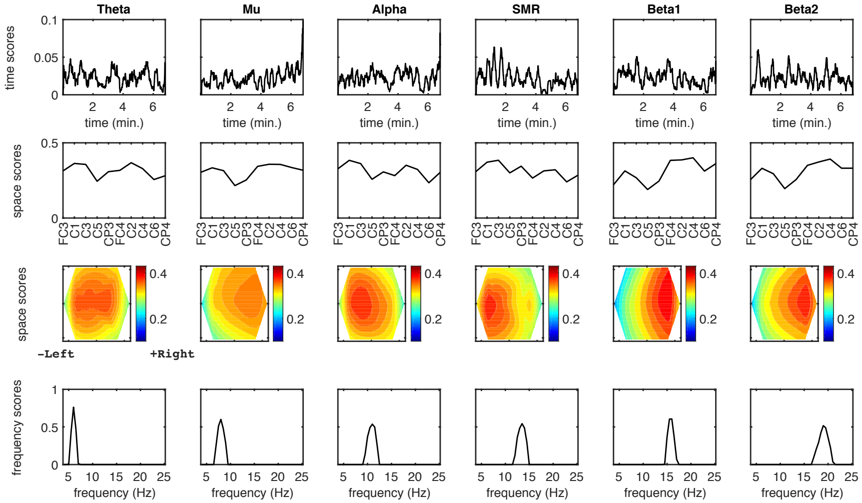


→ the same number F of components within each dimension

→ **restrictions (for EEG data):** nonnegativity; unimodality of columns in C

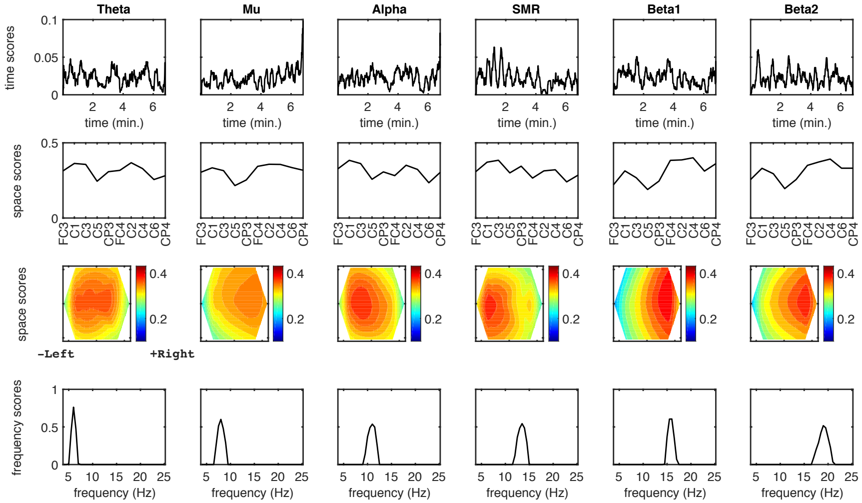
→ **method:** (nonnegative) alternating least squares

PARAFAC – results



→ lower number of space components ?

PARAFAC – results



→ lower number of space components ?

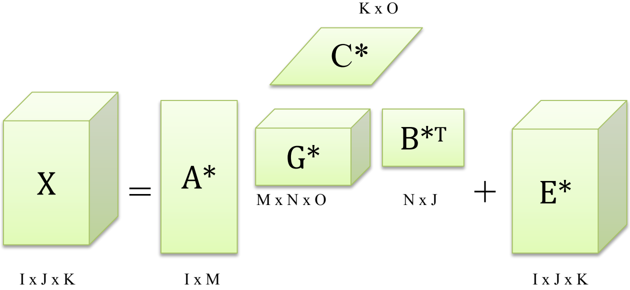
Tucker model

- [Tucker, 1966, Kroonenberg, 1983]

$X \in \mathbb{R}^{\mathbb{I}\times\mathbb{J}\times\mathbb{K}}$:

$$X_{ijk} = \sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{o=1}^{O} g^{\star}_{mno} a^{\star}_{im} b^{\star}_{jn} c^{\star}_{ko} + e^{\star}_{ijk},$$

$$\|a_m\| = \|b_n\| = \|c_o\| = 1$$



→ number of components can differ across dimensions

→ less restrictive than PARAFAC

Tucker model

goals:

- to find appropriate restrictions to the core tensor $G^\star$

  → interpretability

  → stability of the solution

- to detect situations where the Tucker model leads to a more parsimonious representation of EEG data, but with a comparable explanation of neural activity variability as the PARAFAC model

  → on different sets of simulated EEG data

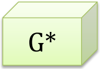
Tucker model – restrictions

- **matrices** $A^\star \in \mathbb{R}^{I\times M}, B^\star \in \mathbb{R}^{J\times N}, C^\star \in \mathbb{R}^{K\times O}$
  - nonnegativity; unimodality of columns in $C^\star$

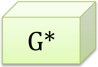
    $\rightarrow$ the same as in PARAFAC

- **the core tensor** $G^\star \in \mathbb{R}^{M\times N\times O}$

  i) **T_uncon:** unrestricted structure



  ii) **T_nonneg:** nonnegativity



> 0

  iii) **T_con:** nonnegativity, diagonal lateral slices

> 0

  - each factor has its own time and frequency scores

$$\Rightarrow M = O$$

  - factors can share the same space scores

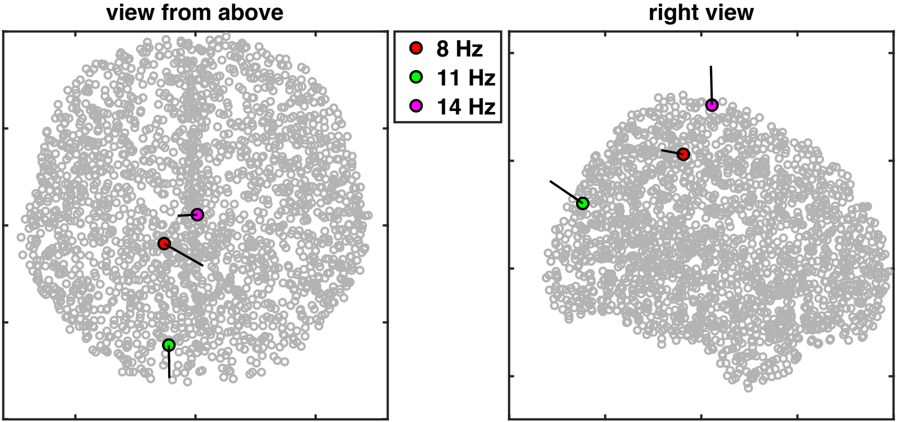
$$\Rightarrow N \leq M = O$$

Simulated EEG data

Simulated EEG data

- inspired by [Cohen, 2017]

1. **choose target dipoles and frequencies** $f_l, l = 1, \ldots, S$
   - $f_1 = 8$ and $f_2 = 14$ Hz in the central region
   - $f_3 = 11$ Hz in the occipital region



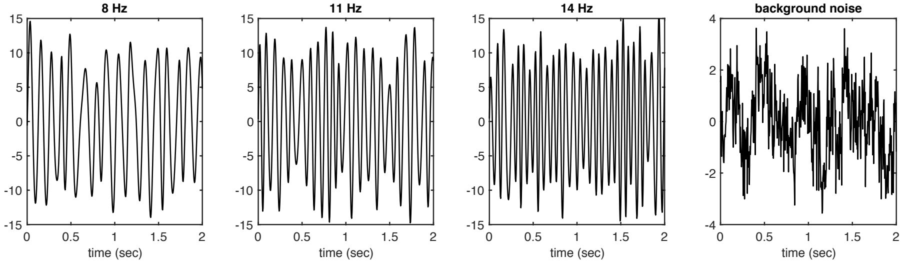
Simulated EEG data

2. **generate one minute of source signals on 2004 dipoles**
   - target dipole

$$x(t) = a(t)\sin\left(2\pi\left(f_l t + \frac{b(t)}{\text{sampling rate}}\right)\right)$$

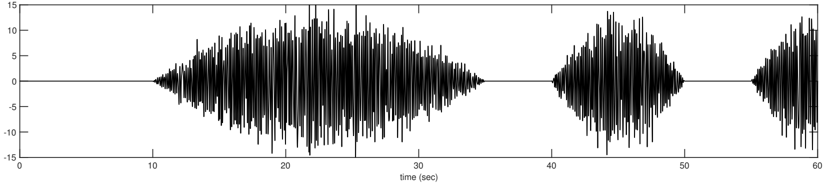
a, b – detrended and filtered series of random numbers

   - "non–target" dipole (background noise)
     - fractional Brownian motion with the Hurst exponent $H \in \{0.1, 0.5\}$



Simulated EEG data

3. generate random time intervals, where the target oscillations are active



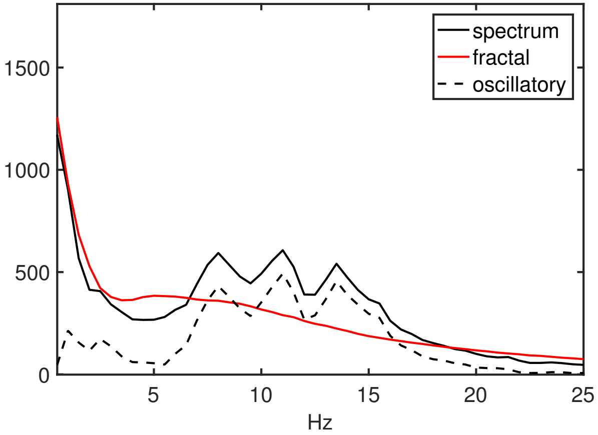
4. generate EEG data by using the forward model [Gramfort et al., 2010]

$$y_j^{EEG}(t) = \sum_{i=1}^{2004} c_{ji} x_i(t), \qquad j = 1, \ldots, 64 \text{ electrodes}$$

$C = \{c_{ji}\} \in \mathbb{R}^{64 \times 2004} \rightarrow$ Brainstorm toolbox in MATLAB [Tadel et al., 2011]

Oscillatory part of the EEG power spectrum

- **windowing** – 2 sec time intervals, overlapping period of 250 ms
- Irregular Resampling Auto–Spectral Analysis (IRASA) [Wen and Liu, 2016]

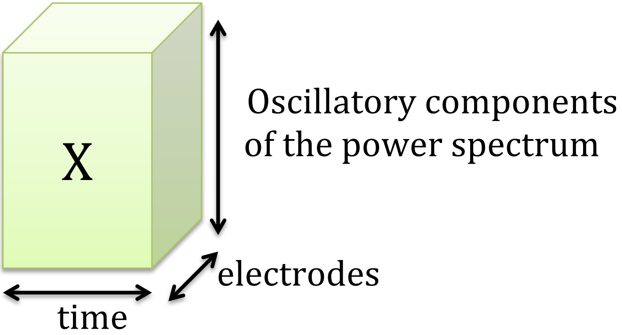


Centering

- centering in the first mode

$$X_{ijk}^{centered} = X_{ijk} - \frac{1}{I}\sum_{l=1}^{I} X_{ljk}$$

- final data



Simulated data analysis

Criteria

- **proportion of variance explained**

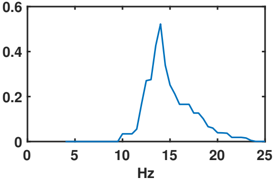
$$VarExpl = 100 \times \left(1 - \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(X_{ijk} - \widehat{X}_{ijk}\right)^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} X_{ijk}^2}\right)$$

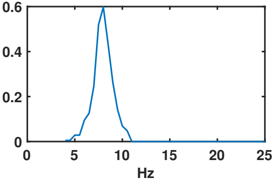
- **core consistency diagnostics** [Bro and Kiers, 2003]
  - for PARAFAC and Tucker models with constraints

$$CCD = 100 \times \left(1 - \frac{\sum_{m=1}^{F}\sum_{n=1}^{F}\sum_{o=1}^{F}(g_{mno} - g^{\star}_{mno})^2}{{g^{\star}_{mno}}^2}\right) \in (-\infty, 100]$$

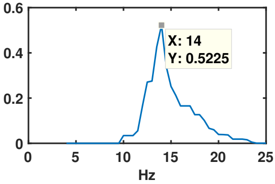
1. estimate A, B, C and G in PARAFAC/restricted Tucker model
2. estimate $G^{\star}$ in unrestricted Tucker model with A, B, C from step 1

Criteria

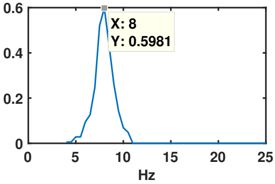




Criteria

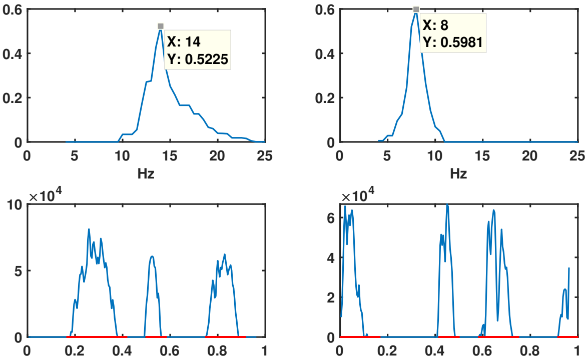


0.6
0.4
0.2
0
0
5
10
15
20
25
Hz
X: 14
Y: 0.5225



0.6
0.4
0.2
0
0
5
10
15
20
25
Hz
X: 8
Y: 0.5981

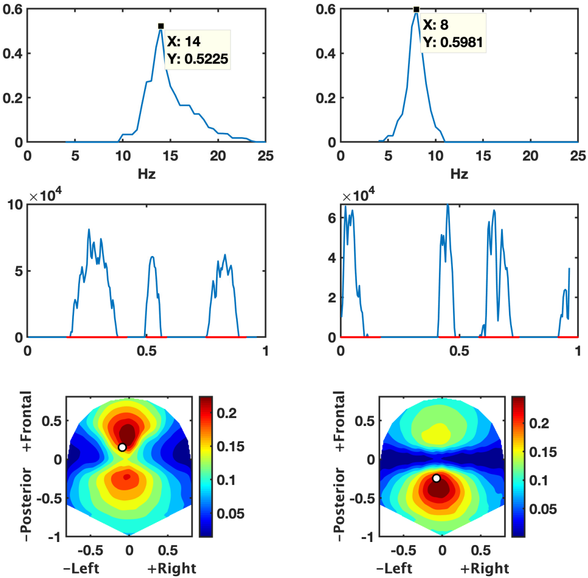
Criteria



- TI = number of target time intervals $Int_1, \ldots, Int_{TI}$
- μ_i, σ_i = mean and *sd* of time scores 3 sec **before** and **after** Int_i

$$Z = \frac{\sum_{i=1}^{TI} \left| \left\{ x \in Int_i : \frac{x-\mu_i}{\sigma_i} > \eta \right\} \right|}{\sum_{j=1}^{TI} |Int_j|}$$

Criteria



X: 14
Y: 0.5225
Hz
X: 8
Y: 0.5981
Hz
×10^4
×10^4
−Posterior +Frontal
−Left +Right
−Posterior +Frontal
−Left +Right

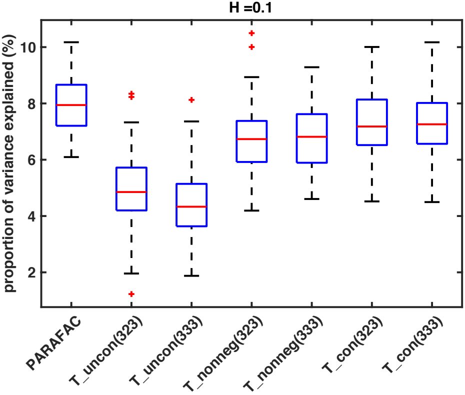
Case 1

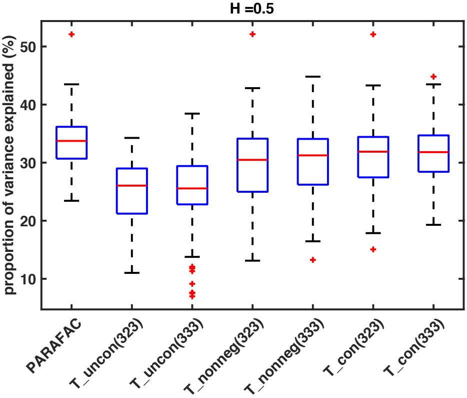
- 8 Hz and 14 Hz in the central region
- 11 Hz in the occipital region

- Hurst exponent $H \in \{0.1, 0.5\}$
- 100 datasets for each H
- 64 electrodes

- models:
  - PARAFAC with F = 3
  - T_uncon, T_nonneg, T_con with 3,2,3 and 3,3,3 factors

Variance explained

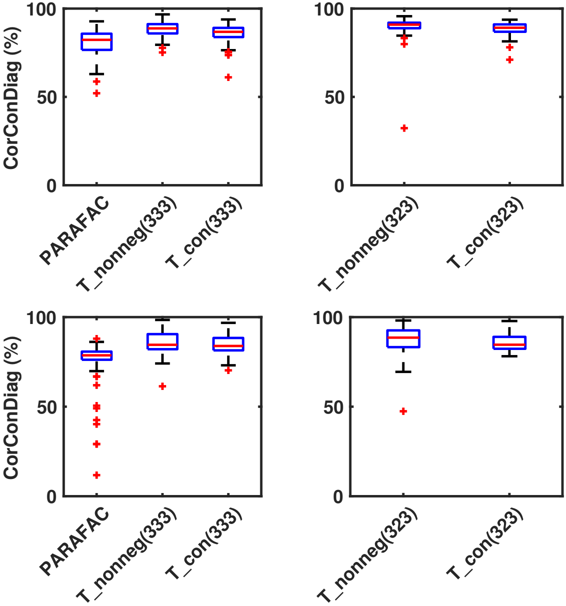




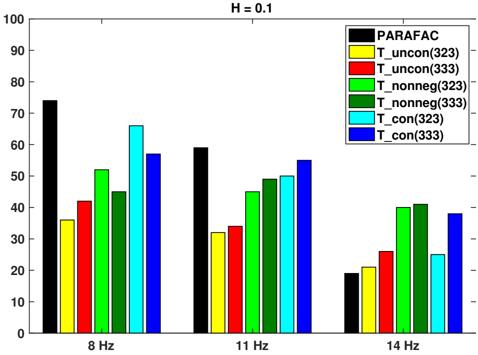
CorConDiag

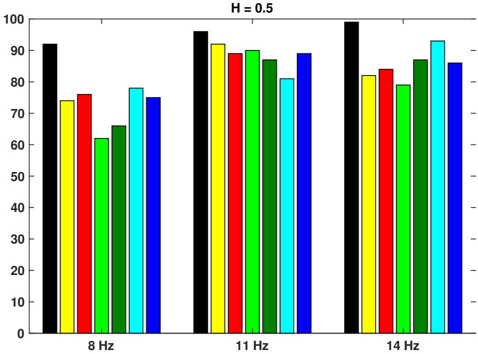
H = 0.1

H = 0.5

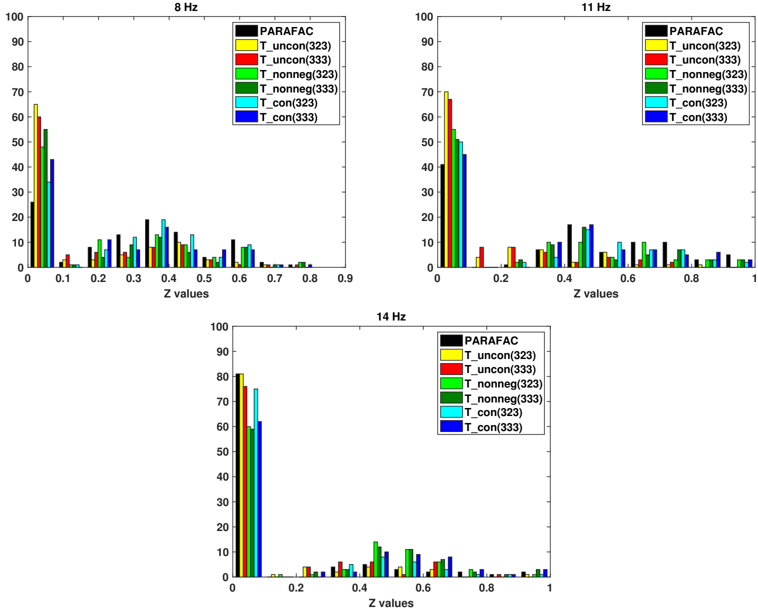


Frequency scores

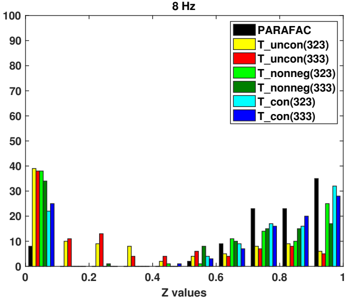


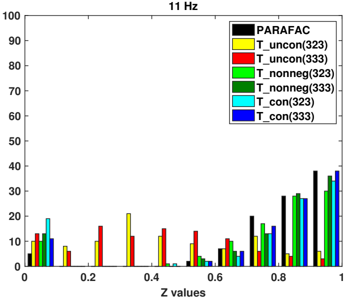


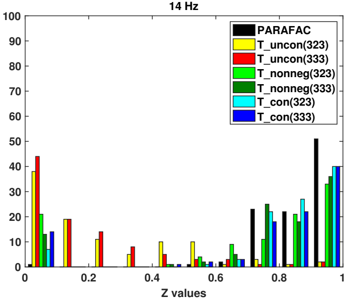
Time scores - H = 0.1



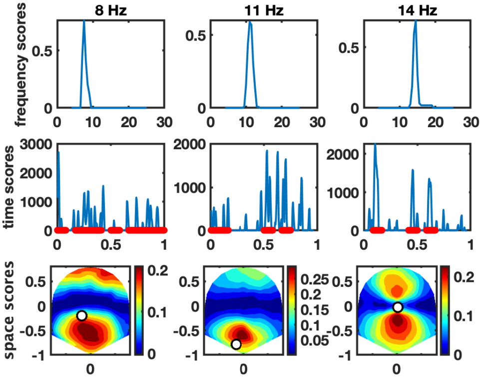
Time scores - H = 0.5



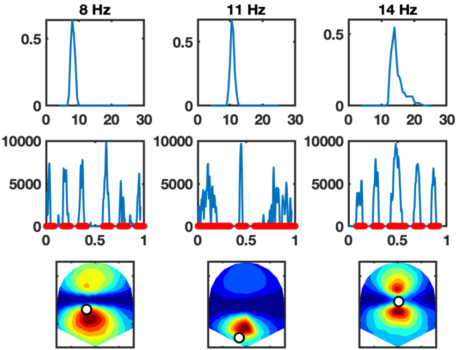




Small visual inspection....PARAFAC

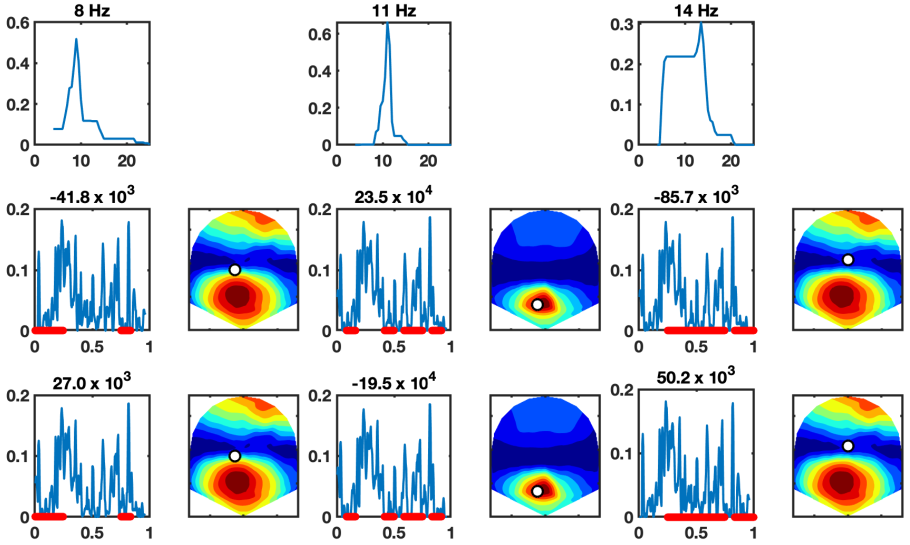


$\leftarrow$ H $=$ 0.1

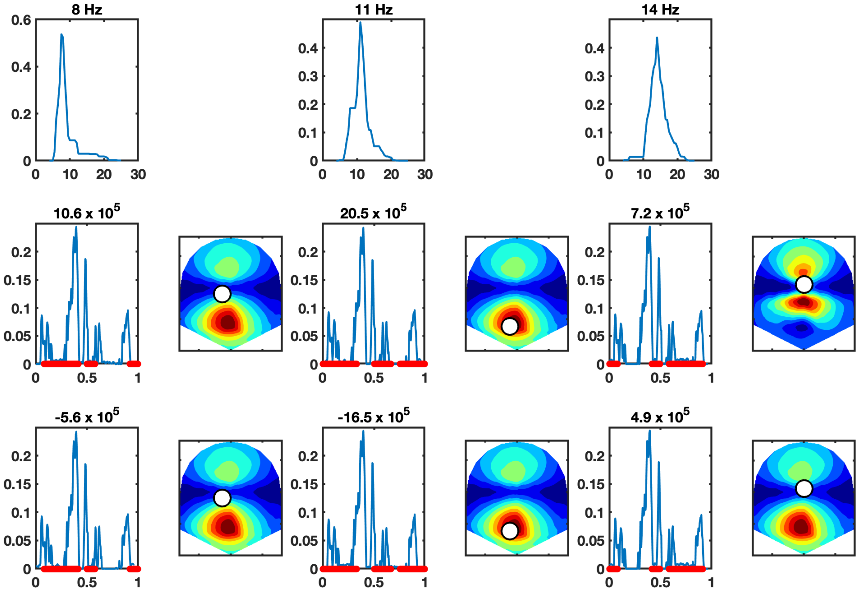


H $=$ 0.5 $\rightarrow$

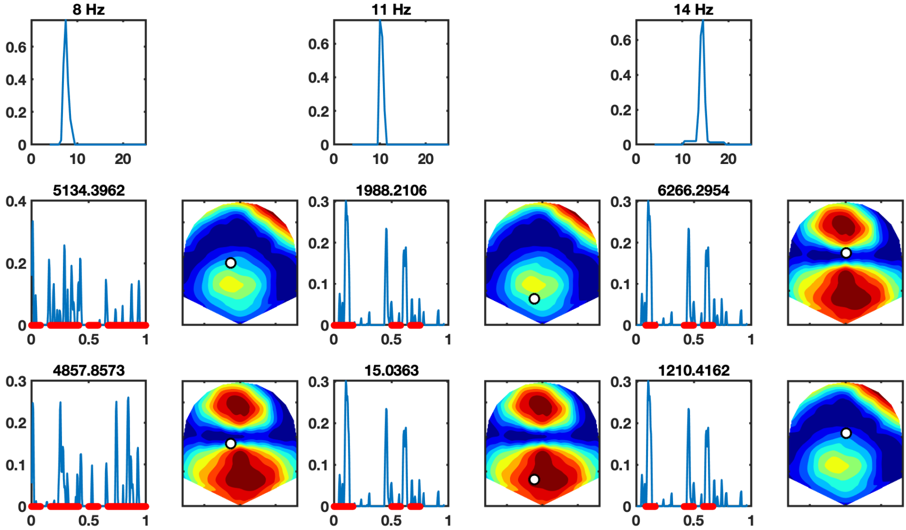
Small visual inspection....T_uncon, H = 0.1



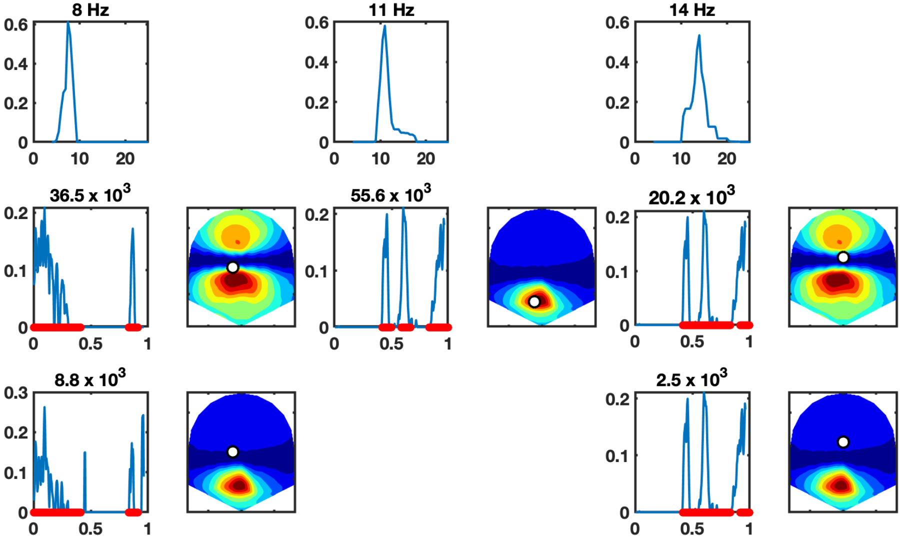
Small visual inspection....T_uncon, H = 0.5



Small visual inspection....T_nonneg, H = 0.1

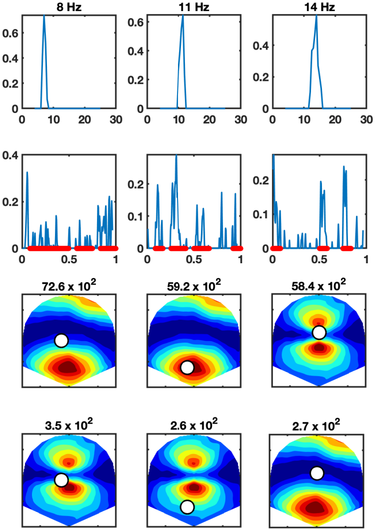


Small visual inspection....T_nonneg, H = 0.5

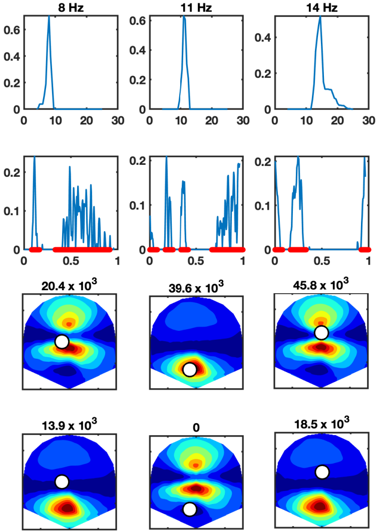


Small visual inspection....Tucker_con

H = 0.1



H = 0.5



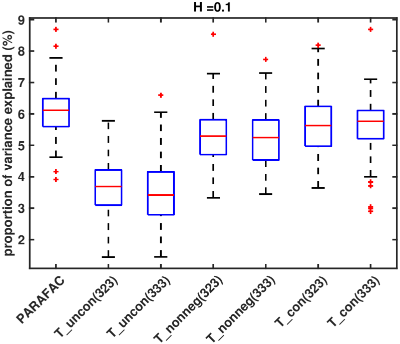
Case 2

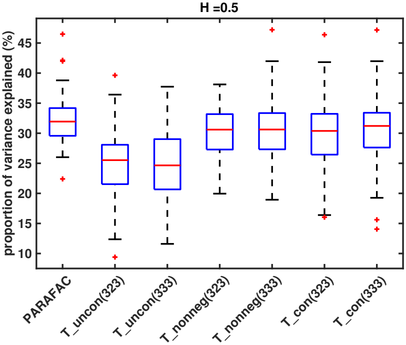
- 8 Hz and 14 Hz in the central region
- 11 Hz in the occipital region

- Hurst exponent $H \in \{0.1, 0.5\}$
- 100 datasets for each H
- 11 electrodes

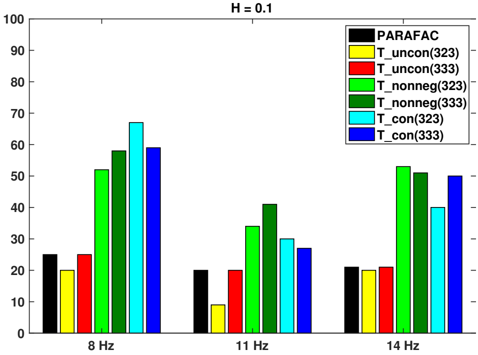
- models:
  - PARAFAC with F = 3
  - T_uncon, T_nonneg, T_con with 3,2,3 and 3,3,3 factors

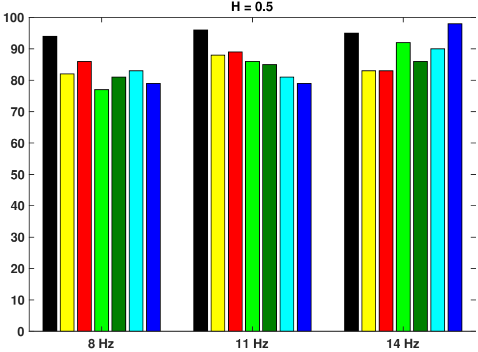
Variance explained



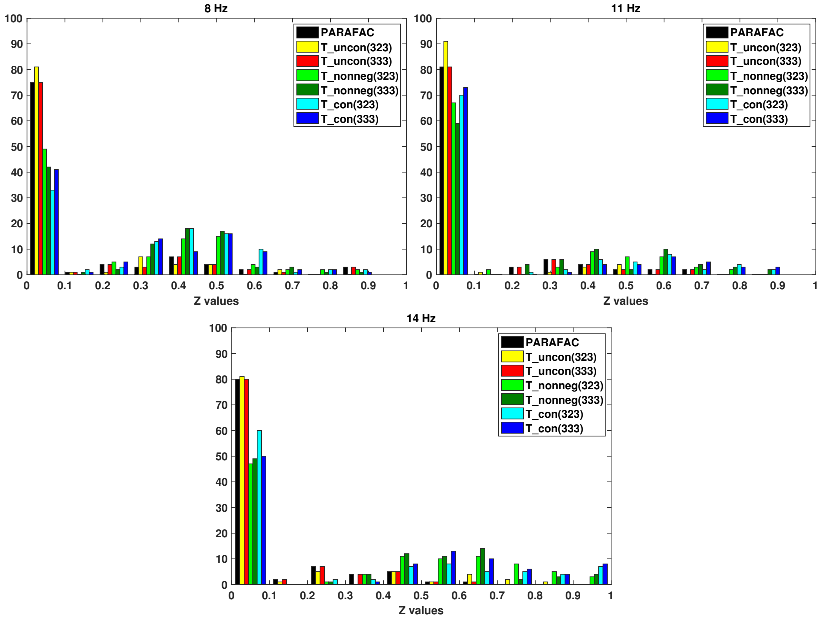


Frequency scores

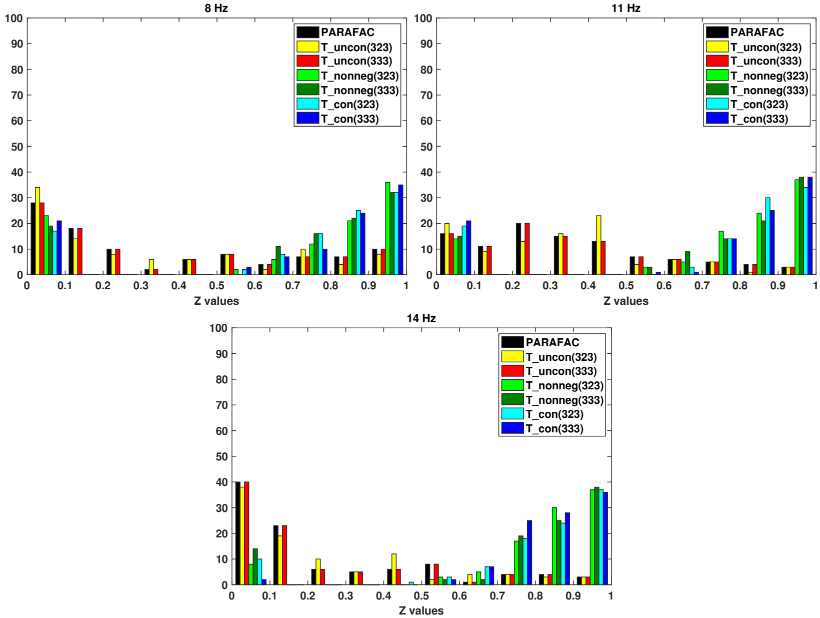




Time scores - H = 0.1



Time scores - H = 0.5



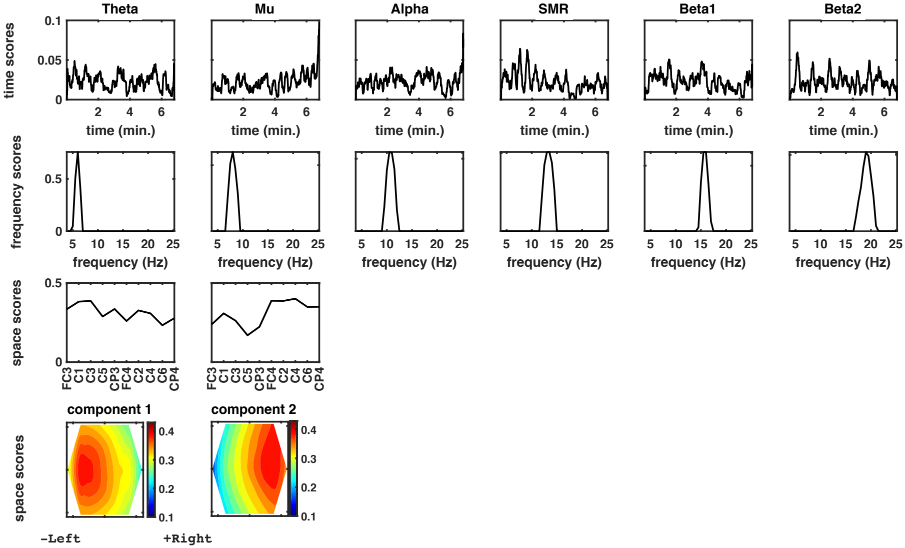
Conclusion

- T_uncon produced unsatisfactory results for this kind of data

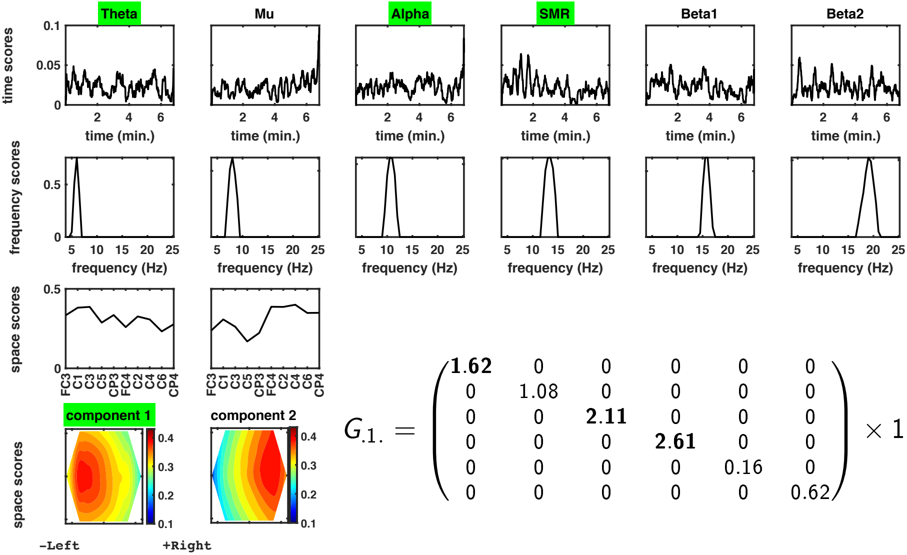
- **64 electrodes**
  - PARAFAC and T_nonneg, T_con produced results of similar quality

- **11 electrodes**
  - T_nonneg, T_con outperformed PARAFAC in the quality of time and frequency scores when a lower number of electrodes and higher level of background noise was considered

Tucker model – 4^{th} day

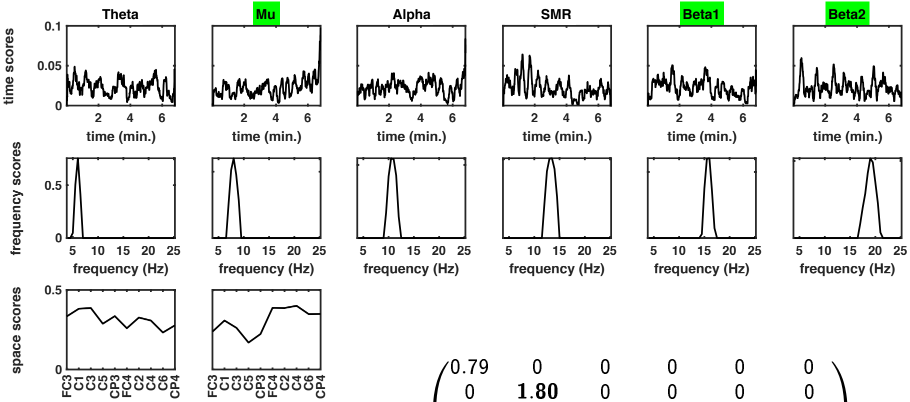


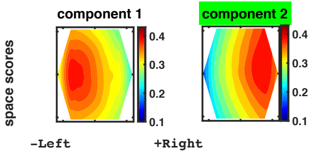
Tucker model – 4^{th} day



$$G_{1.} = \begin{pmatrix} \mathbf{1.62} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.08 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{2.11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{2.61} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.16 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.62 \end{pmatrix} \times 10^3$$

Tucker model – 4^{th} day





$$G_{2.} = \begin{pmatrix} 0.79 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1.80} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.48 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{2.20} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1.97} \end{pmatrix} \times 10^3$$

Literature I

Bro, R. and Kiers, H. A. L. (2003).
A new efficient method for determining the number of components in PARAFAC models.
Journal of Chemometrics, 17(5):274–286.

Carroll, J. D. and Chang, J.-J. (1970).
Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition.
Psychometrika, 35(3):283–319.

Cohen, M. X. (2017).
Comparison of linear spatial filters for identifying oscillatory activity in multichannel data.
Journal of Neuroscience Methods, 278:1 – 12.

Gramfort, A., Papadopoulo, T., Olivi, E., and Clerc, M. (2010).
Openmeeg: opensource software for quasistatic bioelectromagnetics.
BioMedical Engineering OnLine, 9(1):45.

Harshman, R. A. (1970).
Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis.
UCLA Working Papers in Phonetics, 16(1).

Kroonenberg, P. M. (1983).
Three-Mode Principal Component Analysis: Theory and Applications.
DSWO Press, Leiden.

Literature II

Miwakeichi, F., Martinéz-Montes, E., Valdeés-Sosa, P. A., Nishiyama, N., Mizuhara, H., and Yamaguchi, Y. (2004).
Decomposing eeg data into space-time-frequency components using parallel factor analysis.
NeuroImage, 22(3):1035 – 1045.

Rosipal, R., Porubcová, N., Barančok, P., Cimrová, B., Farkaš, I., and Trejo, L. J. (2019).
Effects of mirror-box therapy on modulation of sensorimotor EEG oscillatory rhythms: a single-case longitudinal study.
Journal of Neurophysiology, 121(2):620–633.

Tadel, F., Baillet, S., Mosher, J., Pantazis, D., and Leahy, R. (2011).
Brainstorm: A user-friendly application for meg/eeg analysis.
Computational Intelligence and Neuroscience, 2011:13 pages.

Tucker, L. R. (1966).
Some mathematical notes on three-mode factor analysis.
Psychometrika, 31(3):279–311.

Wen, H. and Liu, Z. (2016).
Separating fractal and oscillatory components in the power spectrum of neurophysiological signal.
Brain Topography, 29(1):13–26.