Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

The polyphonic brain

Sturm, Irene; Treder, Matthias S.; Dähne, Sven; Miklody, Daniel; Purwins, Hendrik; Blankertz, Benjamin; Curio, Gabriel

Publication date:
2014

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Sturm, I., Treder, M. S., Dähne, S., Miklody, D., Purwins, H., Blankertz, B., & Curio, G. (2014). The polyphonic brain: Extracting the neural representation of rhythmic structure for separate voices of polyphonic music using ERPs. Poster session presented at The Neurosciences and Music V, Dijon, France.

The polyphonic brain: Extracting the neural representation of rhythmic structure for separate voices of polyphonic music using ERPs.

Irene Sturm[1,2,3], Matthias Treder[2], Sven Dähne[2], Daniel Miklody[2], Hendrik Purwins[5], Benjamin Blankertz[2,4], Gabriel Curio[1,3,4]

[1]Berlin School of Mind and Brain
[2]Technische Universität Berlin, Neurotechnology Group
[3]Charité Medicine Berlin, Department of Neurology, Neurophysics Group
[4]Bernstein Focus Neurotechnology, Berlin
[5]Sound and Music Computing Group, Aalborg University Copenhagen

ERP responses to the onset structure of music

Rapid changes in the stimulus envelope (indicating tone onsets) elicit an N1-P2 ERP response, as has been shown for clicks and sine waves [1], musical tones [2] and for speech [3]. Canonical Correlation Analysis with temporal embedding (tkCCA), a multivariate correlation-based method, allows to extract brain responses to these changes in continuous auditory stimuli.

Here, we
- (1) probe, whether tkCCA can be applied to track changes in the stimulus envelope in the EEG of subjects who were presented with semi-artificial *monophonic* music clips of three instruments.
- (2) On *polyphonic* trials, composed of the same parts as in (1), we explore, whether the tkCCA-filters derived in (1) can recover a representation of each instrument's part from the EEG where subjects listened to the polyphonic stream.
- (3) We explore, whether, eventually, such a representation is influenced by focused attention.

Experiment (N=11), 64-channel EEG

Re-analysis of data from a 'Musical' brain computer interface application (MusicBCI, [4]): control of a user interface by shifting attention to one of three musical instruments (drums, keyboard, bass) that form a complex semi-naturalistic, polyphonic music stimulus.

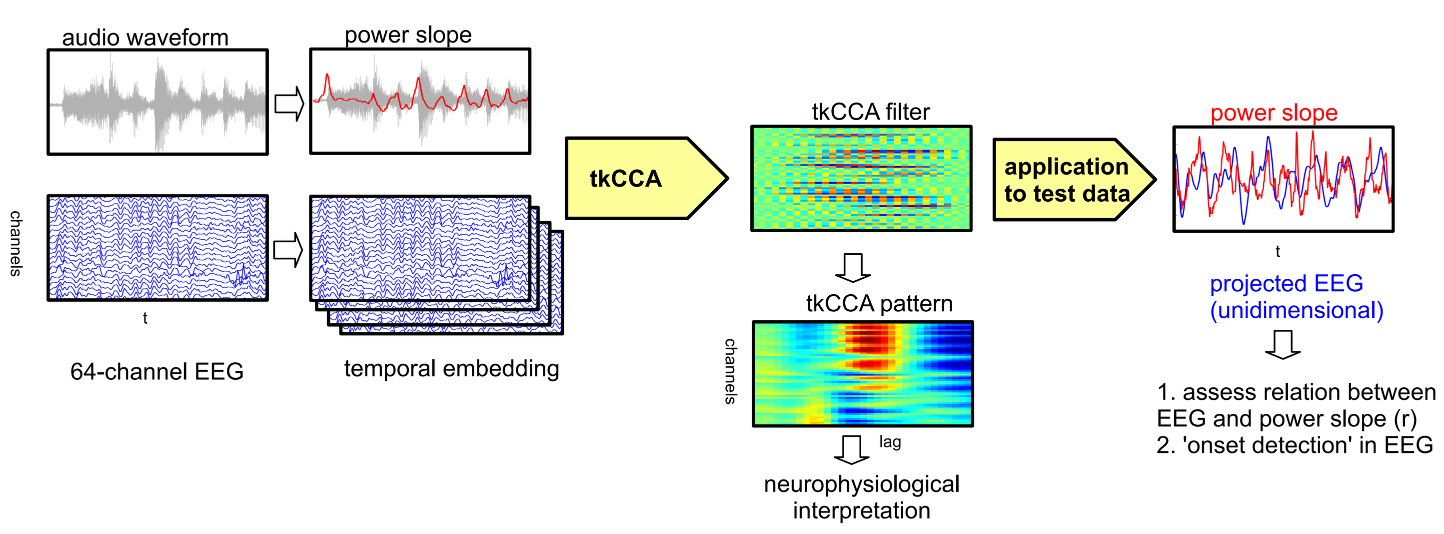
- Music clips of 40 s duration: sequences of a repetitive standard and a deviant pattern, resembling a minimalistic version of Depeche Mode's "Can't get enough" (1980s Electro Pop).
- 63 polyphonic trials, 14 monophonic trials for each instrument
- Task: count silently number of deviant patterns in target instrument.

Score of stimulus

Methods: Canonical Correlation Analysis with temporal embedding (tkCCA)

Idea:

- Design data-driven spatio-temporal filters [5] that maximize the correlation between the power slope of the music signal and EEG.
- Capture complex brain responses by integrate full set of electrodes and a range of time lags (50-250ms).
- Optimize and evaluate filter on separate portions of the data (cross-validation).



tkCCA Analysis pipeline

Steps of analysis:

- Pre-processing: Sampling frequency 100 Hz, HP filter at 1Hz
- Temporal embedding: Add to EEG data set $X_{1...n}$ additional dimensions that are copies of X, time-shifted by 5, ..., 25 data points $X_{2...n+1}, \ldots, X_{k...n+k}$. This allows to capture brain responses within a latency of 50 to 250ms.
- k-fold Leave-One-Out cross-validation: divide data set of k trials into training and test set, so that each trial is once the test set and the remaining clips are the training set.
- CCA (k folds): train spatio-temporal filter on training set, apply to test set. Correlate resulting uni-dimensional EEG projection with power slope of music signal.
- Non-parametric permutation-based testing: assess significance of derived correlation in permutation tests with surrogate data [6] that has the same spectral components, but random phases.
- Source reconstruction from tkCCA patterns: MUSIC algorithm ([7])

Results: Reconstruction of monophonic trials

- Significant correlation of tkCCA-filtered EEG (averaged for the 14 monophonic presentations of each instrument) with the respective power slope in 6/11 subjects for *drums*, in 9/11 subjects for *bass*, and in 1/11 subjects for *keyboard*.
- Grand Average: Significant correlation for *drums* (r=0.61, p=0.002) and *bass*(r=0.52, p=0.004).

subject	drums	bass	keyboard
S1	0.43	0.34	0.30
S2	0.08	0.42	0.25
S3	0.27	0.51	0.24
S4	0.50	0.34	0.17
S5	0.26	0.25	0.32
S6	0.20	0.15	0.07
S7	0.45	0.48	0.24
S8	0.42	0.56	0.22
S9	0.38	0.4	0.31
S10	0.3	0.22	0.28
S11	0.28	0.35	0.09
GA	0.61	0.52	0.53

Figure (1) Monophonic trials: Correlation of EEG projections (average of the 14 solo trials for each instrument) and power slope for each subject and instrument. Yellow shading indicates significance of correlation ($\alpha = 0.05$), determined by permutation tests.

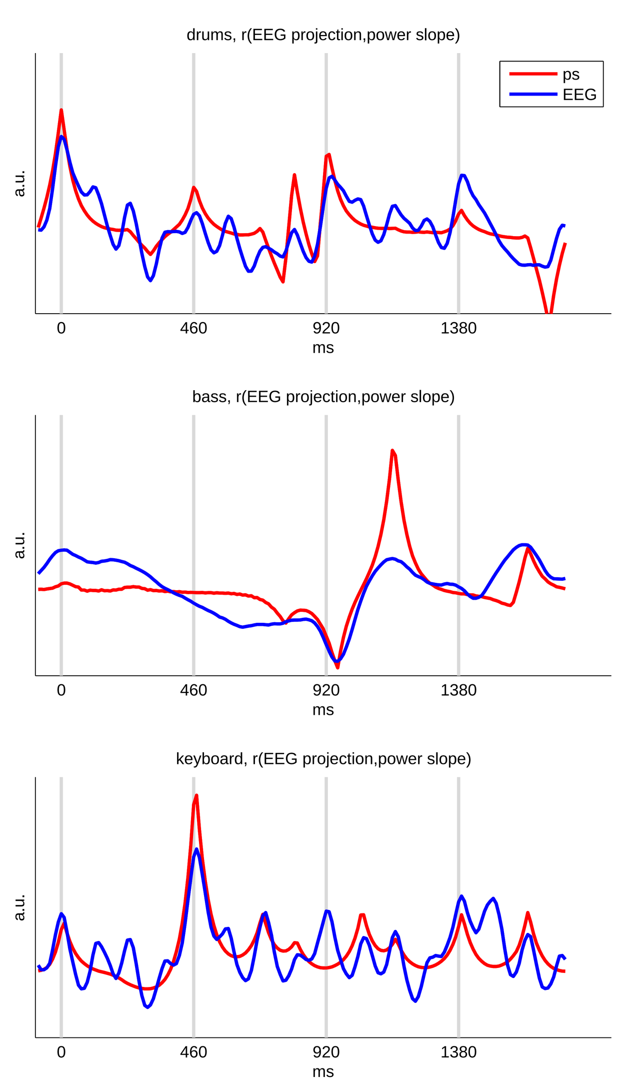


Figure (2) Right: Monophonic trials: Blue: Reconstructed power slope (EEG projection of tkCCA filter, averaged across trials, subjects and bars), Red: True power slope of audio signal

Results: Source reconstruction

- Decomposing the tkCCA patterns results in a fronto-central component, resembling the topography of the N1/P2 complex. The scalp pattern is consistent for the three instruments, the temporal evolution differs.

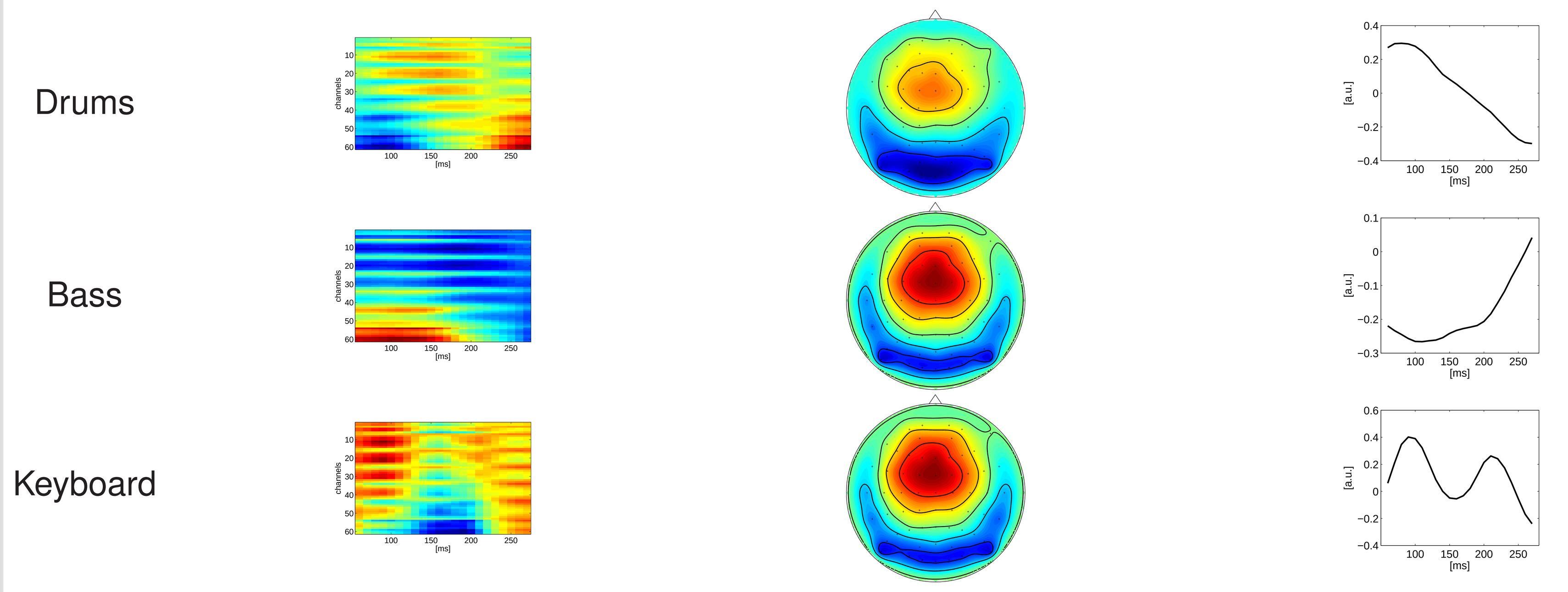


Figure (3) Subject S3: Left: tkCCA patterns, Middle: Spatial pattern of first component of the decomposed tkCCA pattern, Right: Temporal pattern of first component of the decomposed tkCCA pattern.

Results: Application of instrument-specific tkCCA filters on polyphonic trials

- Applying the tkCCA filters (trained on monophonic trials) on the polyphonic stimulus extracts the power slope of each single instrument with significant correlation in: 7/11 subjects for *drums*, in 3/11 subjects for *bass* in 7/11 subjects for *keyboard*.
- Grand Average: Significant correlation only for *keyboard* (r=0.42, p=0.001).
- In 7/11 subjects the power slopes of **two** instruments can be reconstructed in parallel from the (same) EEG where the polyphonic stimulus was presented.

subject	drums	bass	keyboard
S1	0.36	0.28	0.31
S2	0.11	0.05	0.33
S3	0.26	0.13	0.17
S4	0.12	0.07	0.29
S5	0.32	0.05	0.49
S6	0.03	0.34	0.21
S7	0.02	0.23	0.22
S8	0.06	0.08	0.09
S9	0.28	-0.02	0.33
S10	0.25	0.13	0.14
S11	0.13	0.15	0.10
GA	0.04	0.03	0.42

Figure (4) Polyphonic trials: Correlation of EEG projections (average of the 63 mixed trials for each instrument) derived by applying the tkCCA filters optimized on the monophonic trials to the polyphonic trials for each subject and instrument. Yellow shading indicates significance of correlation ($\alpha = 0.05$), determined by permutation tests.

subject	drums	bass	keyboard
GA	0.71	0.63	0.74

Figure (5) Behavioral performance (GA) counting task

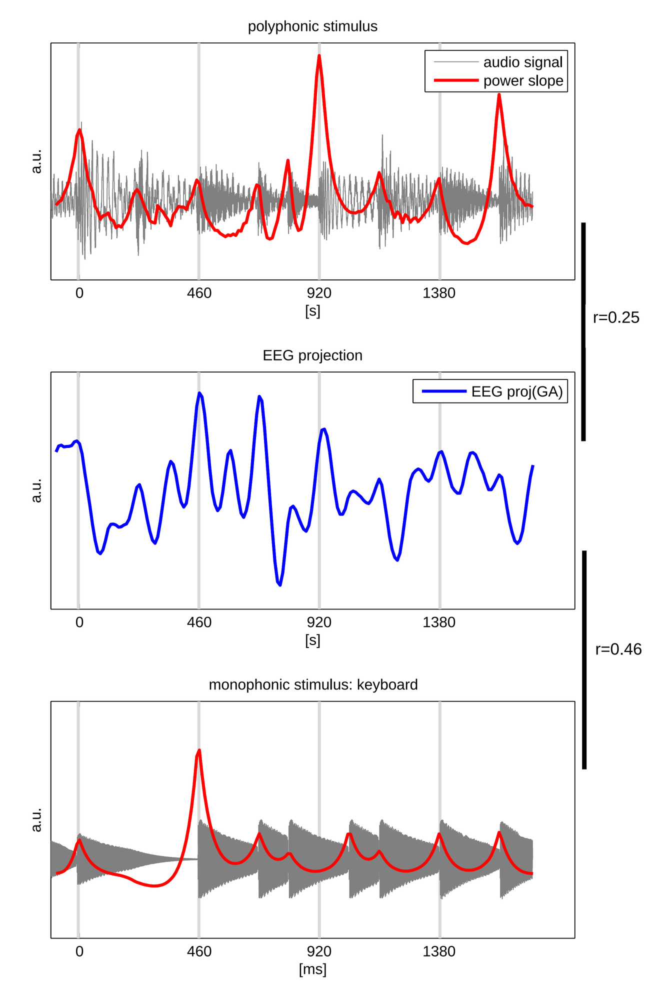


Figure (6) Top: Audio signal and power slope of polyphonic stimulus (one bar). Middle: Reconstructed *keyboard* power slope (EEG projection, derived by applying the tkCCA filter for *keyboard* to EEG of polyphonic stimulus, averaged for subjects and bars). Correlation with the power slope (r=0.46) is higher than with the power slope of the polyphonic stimulus (r=0.25). Bottom: Audio signal and power slope of *keyboard* (one bar).

Results: Effects of attention in polyphonic stimuli

Correlation between EEG projection and power slope of an instrument was significantly enhanced if this instrument was the target of attention for:
- Drums: S3, $p<0.007$
- Bass: S4, $p < 0.018$, S6, $p < 0.029$
- Keyboard: S7, $p < 0.002$, S9, $p < 0.002$ and S10, $p < 0.03$

Within the group of subjects a significant effect of attention was present for *bass* ($p < 0.046$) and for *keyboard* ($p < 0.012$). The behavioral performance differs for the three instruments, with highest performance for *keyboard*.

Discussion & Conclusion

- Canonical Correlation Analysis with temporal embedding (tkCCA) allows to extract neural responses to the onset structure of a continuous music stimulus.
- Extracted neural sources resemble a N1/P2 complex with a consistent scalp topography across instruments, but varying temporal characteristics. These may reflect the onset properties and the structural characteristics of each instrument's part.
- In 10/11 subjects instrument-specific tkCCA filters recovered the power slope of at least one instrument from the EEG of the polyphonic stimulus. This suggests that subtle characteristics of onsets and rhythmic structure that are 'learned' by the tkCCA filter could be effectual for the perceptual segregation of a polyphonic piece of music.
- The power slope of *keyboard* was reconstructed best and most consistently within the group of subjects. Tentatively, this may point to an enhanced representation of melodic instrument in the present N1-P2 response.
- Our preliminary results suggest, that attention may enhance the present brain response to the onset structure of music as suggested in [8] for the typical N1-P2 component. Thus, tkCCA may be a promising tool to investigate a listener's spontaneous fluctuation of attention in continuous music.

References

[1] T. W. Picton e. a., *Electroencephalography and Clin. Neurophysiology*, 36:179–190, 1974.
[2] N. Kaganovich, e. a., *European Journal of Neuroscience*, 37(7-8):1295–1307, 2013.
[3] I. Hertrich e. a., *Psychophysiology*, 49(3):322–334, 2012.
[4] M. Treder e. a., *Journal of Neural Engineering*, 11(1):026009, 2014.
[5] F. Biessmann e. a., *Machine Learning*, 79(1–2):5–27, 2010.
[6] J. Theiler e. a., *Physica D: Nonlinear Phenomena*, 58(1):77–94, 1992.
[7] J.C. Mosher e. a., *Biomedical Engineering, IEEE Tr. on*, 45(11):1342–1354, 1998.
[8] R. Näätänen e. a., *Psychophysiology*, 24(4):375-425, 1987.

Acknowledgments

This work is funded by the Berlin School of Mind and Brain, Humboldt Universität zu Berlin.

Contact: irene.sturm@charite.de