University of Dundee

ToFFi - Toolbox for Frequency-based Fingerprinting of Brain Signals

Komorowski, Michal; Rykaczewski, Krzysztof ; Piotrowski, Tomasz ; Jurewicz, Katarzyna ; Wojciechowski, Jakub ; Keitel, Anne

ToFFi - Toolbox for Frequency-based Fingerprinting of Brain Signals

Michał K. Komorowski[a,b], Krzysztof Rykaczewski[b], Tomasz Piotrowski[a,b],
Katarzyna Jurewicz[c,d], Jakub Wojciechowski[e,c], Anne Keitel[f],
Joanna Dreszer[g], Włodzisław Duch[a,b]

[a]*Department of Informatics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University, Toruń, Poland*
[b]*Neurocognitive Laboratory, Centre for Modern Interdisciplinary Technologies, Nicolaus Copernicus University, Toruń, Poland*
[c]*Nencki Institute of Experimental Biology, Polish Academy of Sciences, Warsaw, Poland*
[d]*Department of Neurosciences, Faculté de médecine, Université de Montréal, Montréal, QC, Canada*
[e]*Bioimaging Research Center, Institute of Physiology and Pathology of Hearing, Kajetany, Poland*
[f]*Psychology, University of Dundee, Scrymgeour Building, Dundee, United Kingdom*
[g]*Faculty of Philosophy and Social Sciences, Institute of Psychology, Nicolaus Copernicus University in Toruń, Poland*

Abstract

Spectral fingerprints (SFs) are unique power spectra signatures of human brain regions of interest (ROIs, Keitel & Gross, 2016). SFs allow for accurate ROI identification and can serve as biomarkers of differences exhibited by non-neurotypical groups. At present, there are no open-source, versatile tools to calculate spectral fingerprints. We have filled this gap by creating a modular, highly-configurable MATLAB Toolbox for Frequency-based Fingerprinting (ToFFi). It can transform MEG/EEG signals into unique spectral representations using ROIs provided by anatomical (AAL, Desikan-Killiany), functional (Schaefer), or other custom volumetric brain parcellations. Toolbox design supports reproducibility and parallel computations.

Email address: `michak@is.umk.pl` (Michał K. Komorowski)

1. Introduction

Brain dynamics and brain oscillations are among the most important topics in neuroscience. Different methods proved to be useful for studying robust whole-brain, as well as regionally-specific patterns of activity, called brain fingerprints. They can serve as signatures for mental states during task execution or rest [1], [2], [3], [4]. Frequency of oscillations turned out to be one of the key features in many studies describing particular regions of interest (ROIs) [5], [6], [7] and large-scale brain networks [8], [9], [10], [11], [12]. Spectral fingerprints (Fig. 1) play a role as biomarkers that are sufficiently specific to permit the successful identification of brain regions using their spectral characteristics. Moreover, spectral profiles' peaks that correspond to the natural frequencies of ROIs [12], [13], are consistently modulated by specific tasks, neurological or mental disorders. They can be generalized across groups of participants [14], [7]. In this paper, we introduce a novel implementation of the Spectral Fingerprinting technique, in a highly configurable MATLAB toolbox.

2. Problems and Background

There are many open software packages available to analyze neural data. The Fieldtrip Toolbox[1] [15] was designed to perform analysis both on sensor and source level of EEG/MEG/iEEG/NIRS data. EEGLAB[2] [16] helps with processing continuous and event-related electrophysiological data implementing many analytic methods (ICA, time/frequency analysis, artifact rejection, event-related statistics, microstates analysis) and several useful routines for visualization. To simulate brain dynamics, perform connectivity analyses, and solve forward/inverse problems, the supFunSim[3] toolbox [17] and the Virtual Brain[4] system [18] are among suitable choices. However, there is no open software for analyzing spectral fingerprints, and our work attempts

---

[1]`https://www.fieldtriptoolbox.org/` ; accessed: 14.10.2021
[2]`https://eeglab.org/` ; accessed: 14.10.2021
[3]`https://github.com/nikadon/supFunSim` ; accessed: 14.10.2021
[4]`https://www.thevirtualbrain.org/` ; accessed: 14.10.2021

to fill this gap. We designed the Toolbox for Frequency-based Fingerprinting (ToFFi, `https://github.com/micholeodon/ToFFi_Toolbox`) for analysis of MEG, EEG, and other multichannel data. Users can configure many parameters for each stage of processing, including the selection of the brain parcellation, and decide which of them will run in parallel (cluster computations are supported). Results of the calculations are reproducible thanks to the implemented control using pseudo-random number generators and visualization scripts.

3. Software Framework

3.1. Software Architecture

ToFFi is a modular piece of software that consists of five components: I. Data Preparation, II. Spectral Fingerprinting, III. Analysis, IV. Presentation, and V. Maintenance (Fig. 2). The Data Preparation module (I) is responsible for arranging sensor time series signals, spatial filters, and brain parcellation data, for processing by the second step routines. The Spectral Fingerprinting (II) module transforms MEG/EEG multichannel array of signals, through a series of five stages, into spatially localized power spectrum-driven representations called spectral fingerprints (Fig. 3). Fourier Transform, source reconstruction (beamforming), and Gaussian Mixture Modeling algorithms are used to compute spectral fingerprints both at the individual and the group level. The third component (III) consists of additional routines that can analyze particular output files from component II. Currently, we have implemented group-level brain regions identification, individual-level brain regions identification, and regional clustering (network analysis) - all based on the concept of modeling brain activity as spectral fingerprints. The Presentation component (IV) is a collection of auxiliary scripts used to visualize particular results of performed computations for easier interpretation. Maintenance routines (V) are used to automate some parts of the workflow, e.g.: manage configuration files, manage output data files, etc. A more detailed description of how the data are transformed can be found in Appendix A (6. METHODS), Fig. 4, Fig. 5, and Fig. 6, which summarize the whole pipeline.

3.2. Software Functionalities

Spectral Fingerprinting can be performed on multichannel time series data (e.g. MEG, EEG) of arbitrary size, with any sampling-frequency ad-

justed to the desired frequency resolution, acquired from a single multiple subjects, and divided into segments of selected, equal duration (e.g. 1000 ms). These segments may contain non-overlapping pieces of a continuous recording (e.g. resting-state) or trials with brain responses for several repetitions of the same experimental condition (event-related paradigm). For individual-level analysis, the toolbox offers a reconstruction of voxel-wise time series power spectra with beamforming using precomputed spatial filters (e.g. LCMV, [19], [20]) and multichannel empirical sensor signals or artificial white Gaussian noise signals. Power spectra of ROIs can be estimated both at individual and group level, using different brain parcellations (anatomical: AAL [21], Desikan-Killiany [22]; functional: Schaefer [23]), and optionally normalized. These spectra can be clustered (currently, only the k-means algorithm is implemented) with arbitrary distance metric and subsequently modeled as a regularized Gaussian mixture of regional spectra with a fixed or optimal number of clusters to construct group-level fingerprints. The user can also estimate the accuracy of identification of brain regions from their spectral fingerprints using cross-validation. Hierarchical clustering of spectral fingerprints (network analysis) was also implemented. The scope of selected brain regions of interest (ROIs) and set of subjects of choice can be limited if desired. For the majority of stages, one can perform computations in parallel on a single computer with multiple cores, or on a grid of multiple-core machines orchestrated via workload manager (currently, only SLURM manager is supported). Interpretation of the outputs of the software components II and III are supported with visualization routines. For reproducibility, data maintenance routines and pseudo-random generator control are implemented as well.

4. Implementation and Empirical Results

4.1. Implementation

The toolbox can be operated under Linux, macOS, and Windows systems. Maintenance scripts (V) are coded mostly in Bash, which is accessible both for Linux (as default) and Windows (using Cygwin, cmder, or other shell emulator). All calculations are carried out entirely in MATLAB (version R2021a or newer is recommended) with Signal Processing Toolbox, Statistics and Machine Learning Toolbox, Parallel Computing Toolbox, and open-source Fieldtrip Toolbox (version 20210816 or newer is preferred, [15]).

4

Additionally, `vline.m` and `hline.m` functions[5] by Brandon Kuczenski are used for plotting, and `HZmvntest.m` function[6] by Antonio Trujillo-Ortiz for multivariate normality testing. If the user wishes to enable cluster computations, the toolbox is prepared to work in coordination with the SLURM workload manager[7]. To the best of our knowledge, there is no other software for Spectral Fingerprinting available, apart from the illustrative beta-version script referenced by the authors of [7].

4.2. Empirical results

Keitel and Gross [7] showed that rendering regional brain activity as a combination of spectra via Spectral Fingerprinting allows for the identification of ROIs with high accuracy. They noticed that clustering of the brain areas according to the similarity of spectral profiles shows patterns similar to macroscale organization of the human brain cortex. Auditory spectral profiles turned out to be modulated during auditory processing. Lubinus and colleagues [14] have discovered that visual deprivation is reflected in the modulation of spectral fingerprints, indicating possible correspondence with the structural and functional adaptation of the human brain. Likewise, Mellem with collaborators [6] demonstrated via a similar method that there is a mix of lower and higher frequency peaks across the brain and it does not follow a simple lower order-higher order processing hierarchy.

5. Illustrative Example

Please consult the following parts of Appendix A to run Illustrative Example smoothly: *Chapter 2. Conventions* - to learn notation used throughout the documentation; *Chapter 3. Installation* - to set up a computational environment properly, *Sections 5.3* and *5.4* - to get the input data.

After installation, one is advised to follow the instructions in *Chapter 4. Illustrative Example* in Appendix A to complete the illustrative example using the Human Connectome Project MEG dataset (HCP Reference Manual, [24]). We have selected N=10 subjects with the MEG resting-state

---

[5]`https://www.mathworks.com/matlabcentral/fileexchange/`
`1039-hline-and-vline` ; accessed: 14.10.2021

[6]`https://www.mathworks.com/matlabcentral/fileexchange/17931-hzmvntest` ;
accessed: 14.10.2021

[7]`https://slurm.schedmd.com/` ; accessed: 14.10.2021

cleaned signal acquired via a 248 channel array in three subsequent runs, approximately 3 min each.

Spectral Fingerprinting routines were configured to optimize the frequency resolution for the lower frequencies, thus accounting for the $1/f$ power trend present in the typical electrophysiological activity of the human brain. The proper number of clusters to be constructed was estimated using the Silhouette optimality criterion. Choosing cosine dissimilarity as a distance measure helped to compose frequency clusters of power spectra similar in shape, diminishing the influence of the power spectra amplitude. To speed up computations, the number of CPU cores was set to two.

Group-level fingerprints in the 1–40 Hz frequency interval from 8 distant regions of the human brain (ROIs; Fig. 7) were found (Fig. 8). The similarity of fingerprints was assessed (Fig. 9), together with the accuracy of how well one can identify them (Fig. 10).

The proposed method allowed for discrimination between different modes of operation for a range of brain areas. Dominant and supportive group-level oscillation profiles were recognized and separated. Functional similarity between homologue areas was confirmed using hierarchical clustering analysis. Recognition of the brain areas based on their spectral fingerprints turned out to be challenging among homologue areas due to their functional similarity, yet remaining informative.

6. Conclusions

Spectral Fingerprinting allows for the discovery of meaningful oscillatory patterns from electrophysiological time series that can show task-induced modulations or serve as a signature of the brain's regional activity in the particular parcellation. Our novel Toolbox for Frequency-based Fingerprinting (ToFFi) provides researchers with a modular, highly configurable tool for computing regional source-reconstructed power spectra, finding optimal prototypes common for a group of subjects via individual- and group-level clustering algorithms, together with testing their properties using additional analytical routines. The efficiency is boosted with parallel computation support, and reproducibility is controlled with pseudo-random number generator parameters. An in-depth understanding of the underlying algorithms is facilitated by the function reference (Appendix B) and the toolbox manual (Appendix A). Presented software is compatible with various modern tools

used by the neuroscientific community and allows easy adaptation of its modular structure to specific tasks.

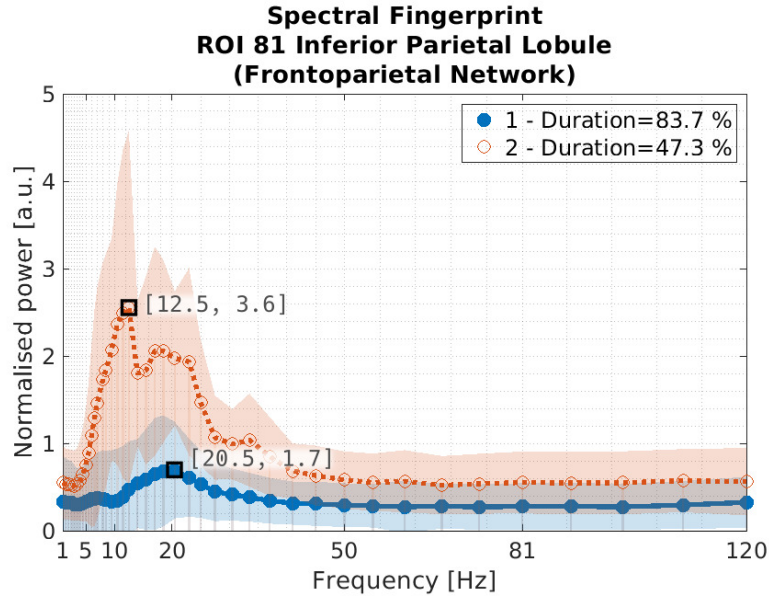## Acknowledgments

Figures

Figure 1: A spectral fingerprint of the inferior parietal lobule. For this particular region, it consists of two spectral modes. It is formed by clustering power spectra segments (normalized, i.e. spectral power in comparison to the whole brain) first on the individual subjects level and then clustered again on the group level. Each mode corresponds to one of the centroids found by the clustering algorithm. Shaded regions depict the standard deviation (1σ) estimated from the covariance matrix of the Gaussian Mixture Model component corresponding to the given spectral mode. The first mode peaks at 12.5 Hz, and the second mode peaks at 20.5 Hz. The frequency axis resolution can be set to logarithmic to optimize spectral analysis resolution of lower frequencies. Duration is shown as a percentage of time segments in which each spectral mode was present on average during recording.
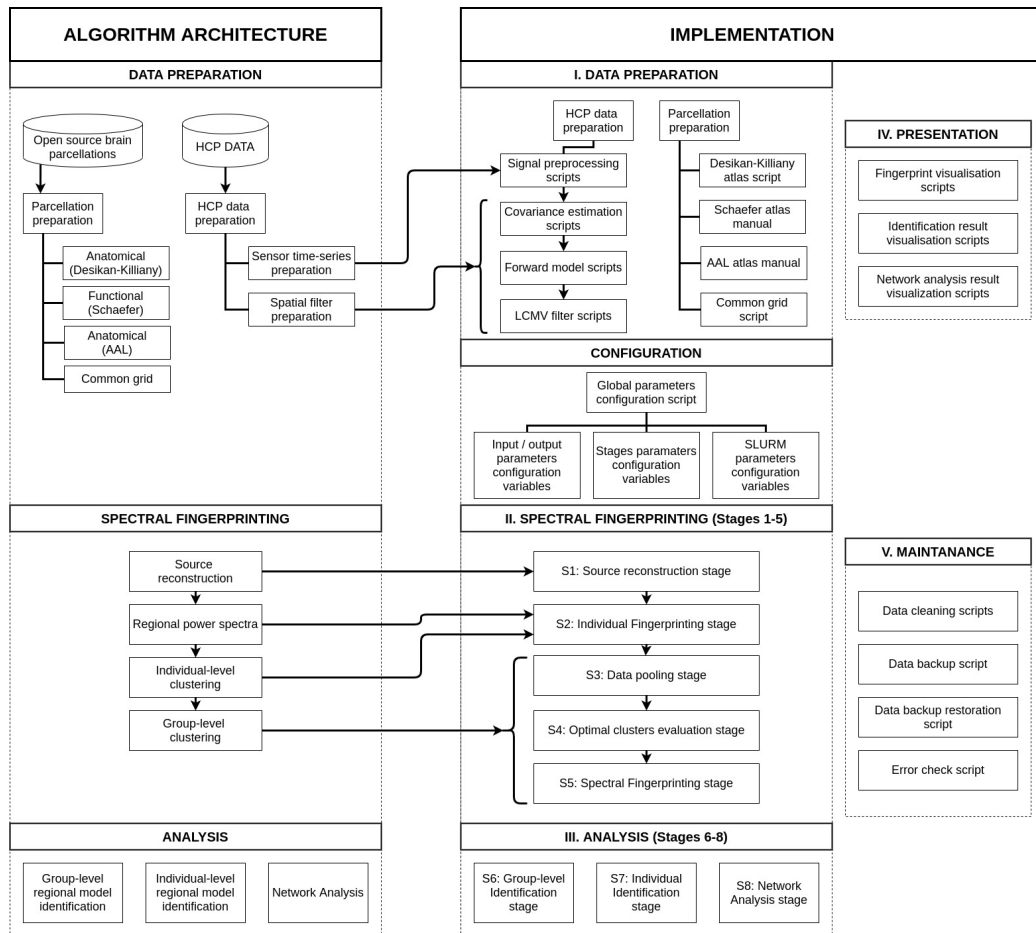
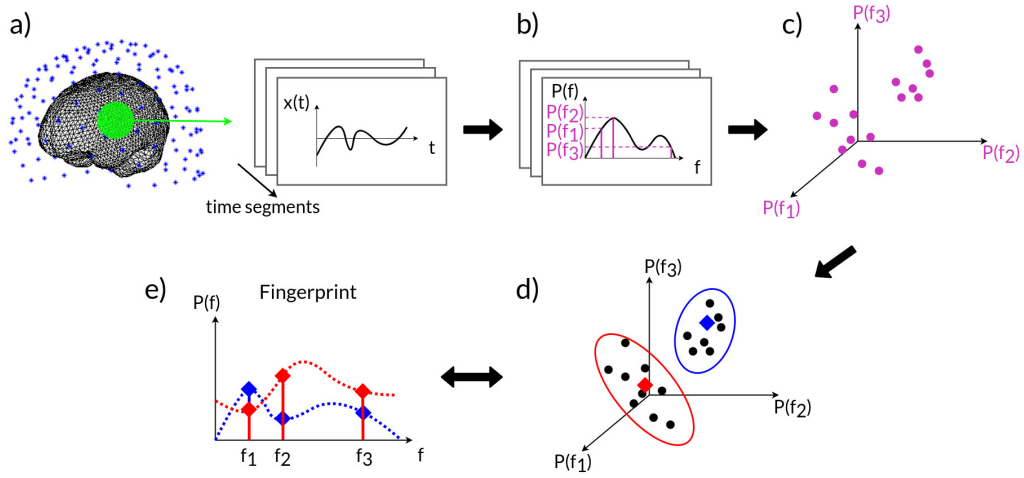Figure 2: ToFFi toolbox architecture.

Figure 3: Illustration of the Spectral Fingerprinting algorithm [7]. a) Mean-field, electrical activity of the brain is recorded via an array of sensors (blue dots). A beamformer [20] solves so-called *inverse problem*, thus enabling reconstruction of the source activity in selected voxels constituting chosen brain regions of interest (green area). Source-level time series are cut into segments of equal length. b) Power within each segment is estimated for a given set of frequencies of interest (here three frequencies, f_1, f_2, f_3, are shown for readability) and then averaged across locations inside the region of interest. c) Power spectrum of each segment is represented as a point in a n-dimensional frequency space, where power of the selected frequencies provides coordinates of the point. d) Segments are clustered together. Each centroid is equivalent to one *spectral mode* of a given brain region and depicted as an interpolant curve spanned between the frequencies of interest in the resulting fingerprint (interpolation type is arbitrary, as it serves visualization purpose only).
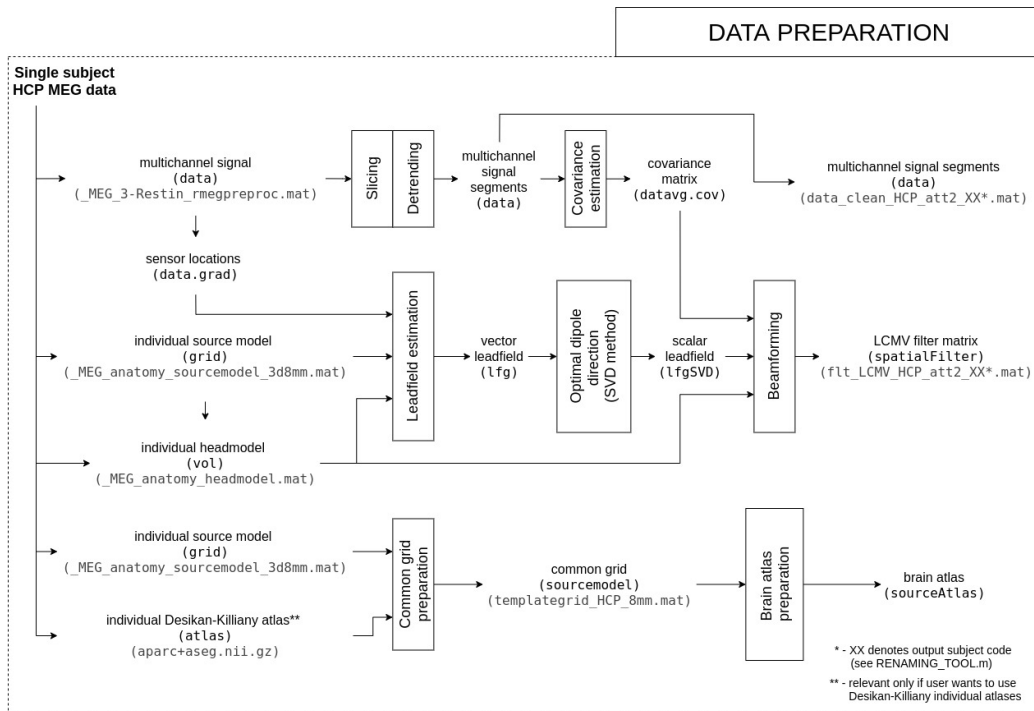
Figure 4: Diagram of how the data are processed by the DATA PREPARATION component (see Fig. 2). Each depicted piece of data is endowed with its name, MATLAB workspace variable name (middle parentheses), and a file name (bottom parentheses) if it is read or written to the disk during processing.
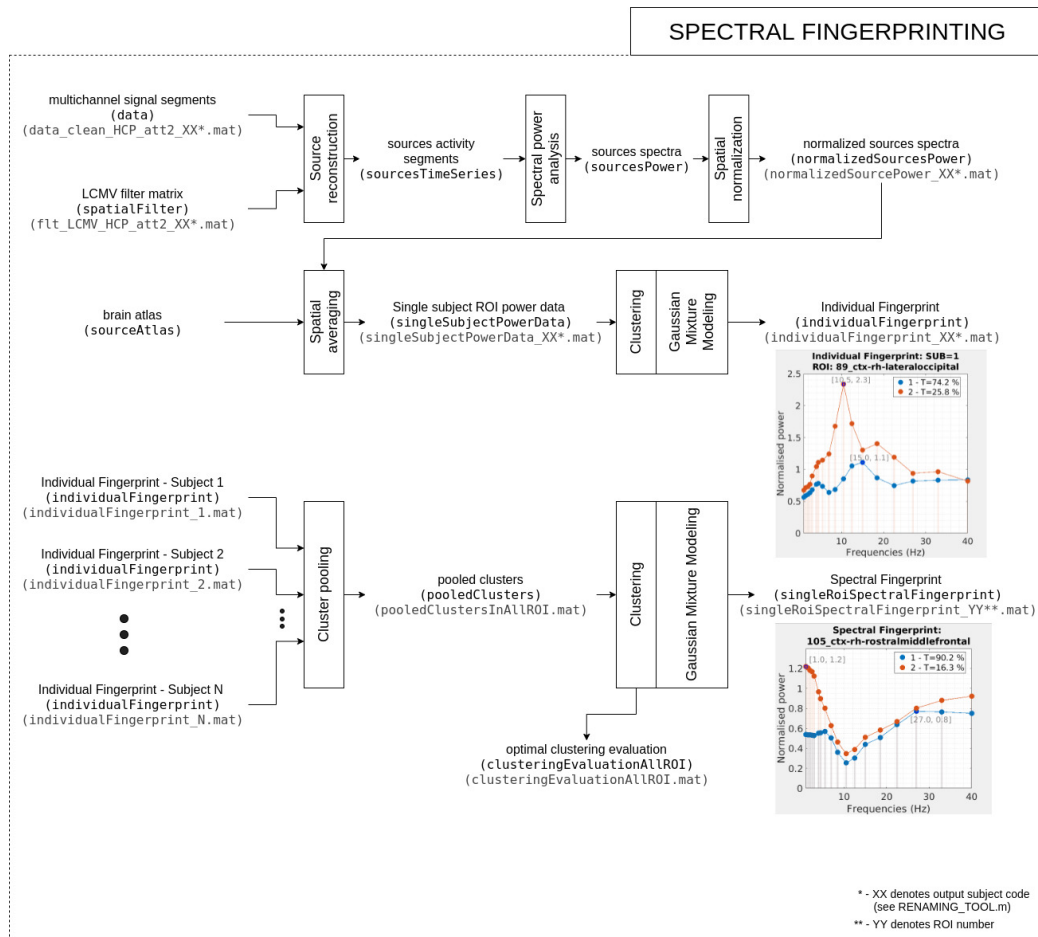
Figure 5: Diagram of how the data are processed by the SPECTRAL FINGERPRINTING component (see Fig. 2). Each depicted piece of data is endowed with its name, MATLAB workspace variable name (middle parentheses), and a file name (bottom parentheses) if it is read or written to the disk during processing.
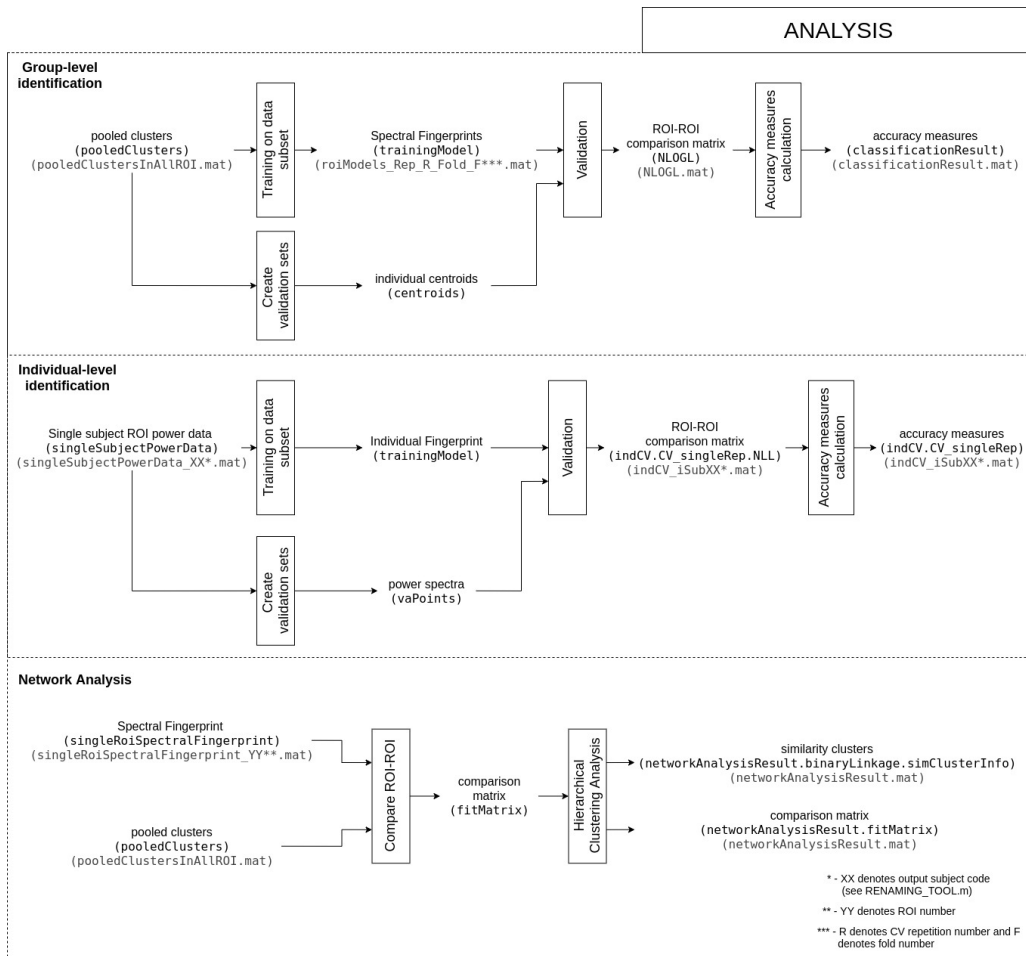
Figure 6: Diagram of how the data are processed by the ANALYSIS component (see Fig. 2). Each depicted piece of data is endowed with its name, MATLAB workspace variable name (middle parentheses), and a file name (bottom parentheses) if it is read or written to the disk during processing.
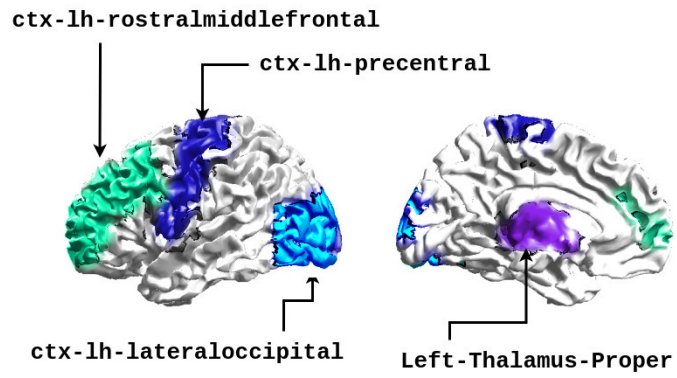
Figure 7: Brain regions chosen from the Desikan-Killiany atlas for the purpose of the illustrative example. Here only the left counterpart is depicted, whereas right hemisphere homologues were chosen as well.
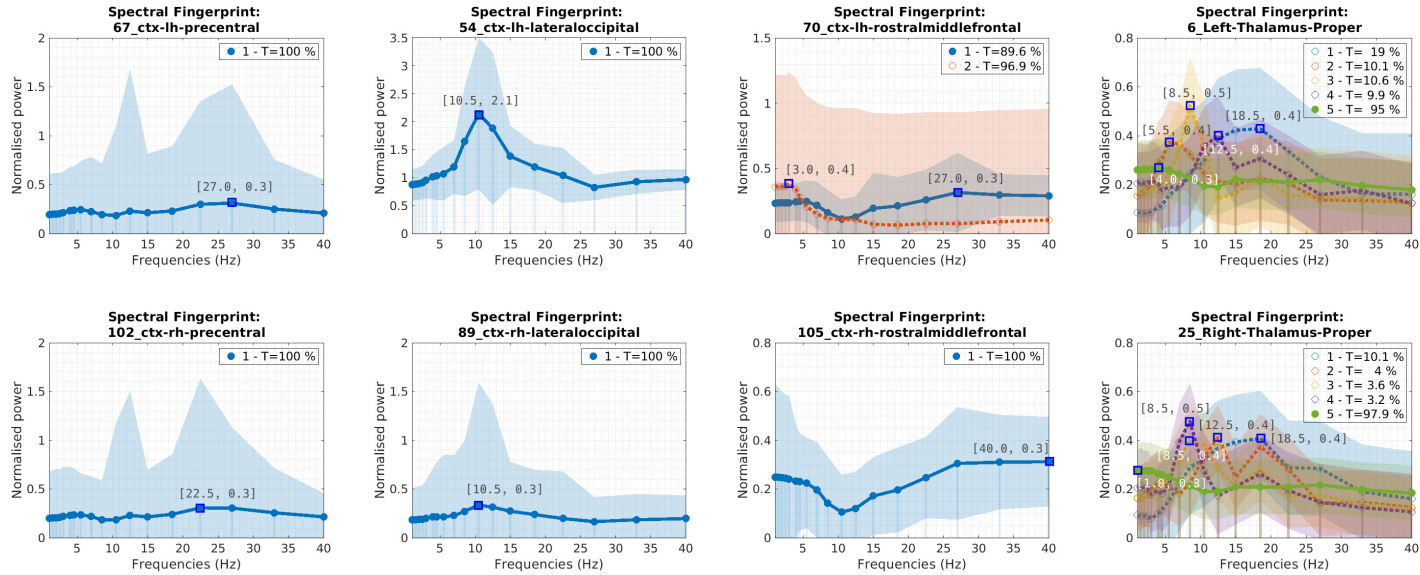
Figure 8: (in color) Resting-state spectral fingerprints for Desikan-Killiany atlas in the 1–40 Hz frequency interval. Each column shows two homologue brain areas. Legends show the corresponding duration of each spectral mode (i.e., the percentage of trials in which each spectrum was present on average during recording) and whether the mode was present for at least five subjects (filled dot) or not (empty dot). The frequency axis was configured to be logarithmic in order to optimize the lower frequencies resolution. Y-axis depicts the power normalized in relation to the average spectrum of the whole brain. Shaded regions depict the standard deviation (1σ) of the corresponding spectral mode. For i-th of total F frequencies of interest, standard deviation was estimated as $\sqrt{\Sigma_{i,i}}$, where $\Sigma_{i,i}$ is the i-th diagonal entry of the covariance matrix of the Gaussian Mixture Model component corresponding to the given spectral mode. Standard deviations have relatively large values due to the small number of subjects used in the illustrative example. Homologue areas have very similar fingerprints.
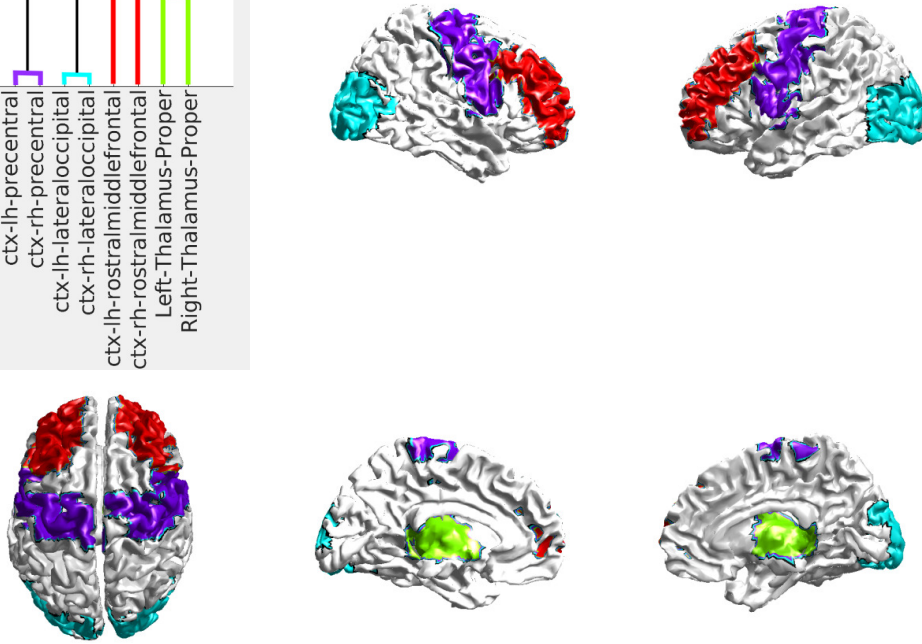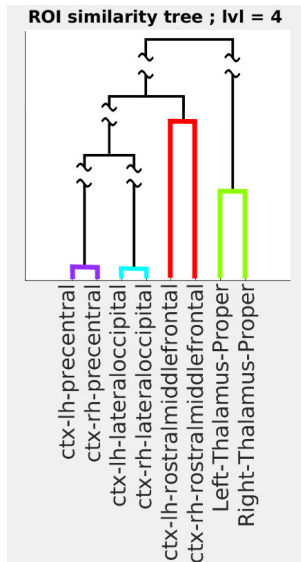
Figure 9: Result of the hierarchical agglomerative clustering of the spectral fingerprints presented in Fig. 8. Homologue areas were automatically matched together according to the similarity of their fingerprints. The similarity tree has disproportionately long branches that were broken for clarity (waved lines).
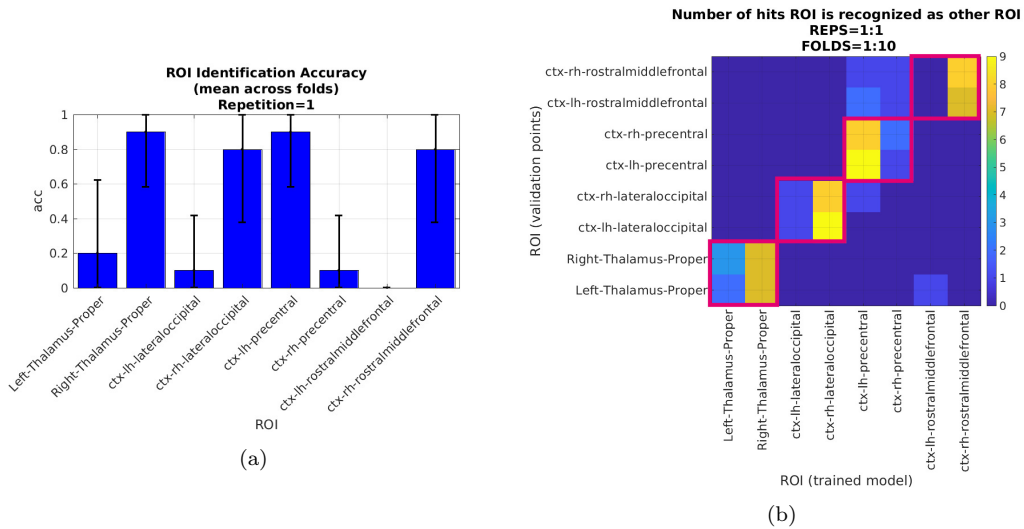
Figure 10: Group-level identification accuracy: a) bar plot showing the average identification accuracy across cross-validation iterations (leave-one-out), b) confusion matrix showing in each row a distribution of "votes" for each ROI. Each ROI was tested ten times (model trained on nine subjects versus one validation subject). For ideal identification, this matrix would have a value of 10 for the diagonal elements and zeros elsewhere. Confusion happens mostly between homologue areas (2x2 red boxes). Left hemisphere ROIs are recognized as the right hemisphere homologue areas.

Appendix A. ToFFi Toolbox Manual

Link: `https://github.com/micholeodon/ToFFi_Toolbox/tree/master/ToFFi_Toolbox-20211013/docs/ToFFi_Manual.pdf`

Appendix B. Functions reference

Functions reference documents most important M-File Functions of the ToFFi Toolbox.

It can be accessed **after** downloading/cloning the ToFFi Toolbox repository (`https://github.com/micholeodon/ToFFi_Toolbox`).

Functions reference can be found here:
`ToFFi_Toolbox-YYYYMMDD/docs/FUNCTIONS_REFERENCE.html`, where `YYYYMMDD` stands for the toolbox revision number.

References

[1] M. Bola, B. A. Sabel, Dynamic reorganization of brain functional networks during cognition, NeuroImage 114 (2015) 398–413. `doi:10.1016/j.neuroimage.2015.03.057`.

[2] F. M. Krienen, B. T. T. Yeo, R. L. Buckner, Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture, Philosophical Transactions of the Royal Society B: Biological Sciences 369 (1653) (2014) 20130526. `doi:10.1098/rstb.2013.0526`.

[3] R. Ciric, J. S. Nomi, L. Q. Uddin, A. B. Satpute, Contextual connectivity: A framework for understanding the intrinsic dynamic architecture of large-scale functional brain networks, Scientific Reports 7 (1) (2017) 6537. `doi:10.1038/s41598-017-06866-w`.

[4] J. N. Keynan, Y. Meir-Hasson, G. Gilam, A. Cohen, G. Jackont, S. Kinreich, L. Ikar, A. Or-Borichev, A. Etkin, A. Gyurak, I. Klovatch, N. Intrator, T. Hendler, Limbic Activity Modulation Guided by Functional Magnetic Resonance Imaging–Inspired Electroencephalography Improves Implicit Emotion Regulation, Biological Psychiatry 80 (6) (2016) 490–496. `doi:10.1016/j.biopsych.2015.12.024`.

[5] M. Siegel, T. H. Donner, A. K. Engel, Spectral fingerprints of large-scale neuronal interactions, Nature Reviews Neuroscience 13 (2) (2012) 121–134. `doi:10.1038/nrn3137`.

[6] M. S. Mellem, S. Wohltjen, S. J. Gotts, A. S. Ghuman, A. Martin, Intrinsic frequency biases and profiles across human cortex, Journal of Neurophysiology 118 (5) (2017) 2853–2864. `doi:10.1152/jn.00061.2017`.

[7] A. Keitel, J. Gross, Individual human brain areas can be identified from their characteristic spectral activation fingerprints, PLoS Biol 14 (6) (2016) e1002498. `doi:https://doi.org/10.1371/journal.pbio.1002498`.

[8] K. Mahjoory, J.-M. Schoffelen, A. Keitel, J. Gross, The frequency gradient of human resting-state brain oscillations follows cortical hierarchies, eLife 9 (2020) e53715. `doi:10.7554/eLife.53715`.

[9] J. Samogin, Q. Liu, M. Marino, N. Wenderoth, D. Mantini, Shared and connection-specific intrinsic interactions in the default mode network, NeuroImage 200 (2019) 474–481. `doi:10.1016/j.neuroimage.2019.07.007`.

[10] M. Marino, Q. Liu, J. Samogin, F. Tecchio, C. Cottone, D. Mantini, C. Porcaro, Neuronal dynamics enable the functional differentiation of resting state networks in the human brain, Human Brain Mapping 40 (5) (2019) 1445–1457. `doi:10.1002/hbm.24458`.

[11] C. D. Hacker, A. Z. Snyder, M. Pahwa, M. Corbetta, E. C. Leuthardt, Frequency-specific electrophysiologic correlates of resting state fMRI networks, NeuroImage 149 (2017) 446–457. `doi:10.1016/j.neuroimage.2017.01.054`.

[12] M. Rosanova, A. Casali, V. Bellina, F. Resta, M. Mariotti, M. Massimini, Natural Frequencies of Human Corticothalamic Circuits, Journal of Neuroscience 29 (24) (2009) 7679–7685. `doi:10.1523/JNEUROSCI.0445-09.2009`.

[13] F. Ferrarelli, S. Sarasso, Y. Guller, B. A. Riedner, M. J. Peterson, M. Bellesi, M. Massimini, B. R. Postle, G. Tononi, Reduced natural oscillatory frequency of frontal thalamocortical circuits in schizophrenia, Archives of General Psychiatry 69 (8) (2012) 766–774. `doi:10.1001/archgenpsychiatry.2012.147`.

[14] C. Lubinus, J. Orpella, A. Keitel, H. Gudi-Mindermann, A. K. Engel, B. Roeder, J. M. Rimmele, Data-Driven Classification of Spectral Profiles Reveals Brain Region-Specific Plasticity in Blindness, Cerebral Cortex 31 (5) (2021) 2505–2522. `doi:10.1093/cercor/bhaa370`.

[15] R. Oostenveld, P. Fries, E. Maris, J.-M. Schoffelen, FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data, Computational Intelligence and Neuroscience 2011 (2010) e156869. `doi:10.1155/2011/156869`.

[16] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, Journal of Neuroscience Methods 134 (1) (2004) 9–21. `doi:10.1016/j.jneumeth.2003.10.009`.

[17] K. Rykaczewski, J. Nikadon, W. Duch, T. Piotrowski, supFunSim: Spatial Filtering Toolbox for EEG, Neuroinformatics 19 (1) (2021) 107–125. `doi:10.1007/s12021-020-09464-w`.

[18] P. Sanz Leon, S. A. Knock, M. M. Woodman, L. Domide, J. Mersmann, A. R. McIntosh, V. Jirsa, The Virtual Brain: a simulator of primate brain network dynamics, Frontiers in Neuroinformatics 7. `doi:10.3389/fninf.2013.00010`.

[19] B. Van Veen, W. Van Drongelen, M. Yuchtman, A. Suzuki, Localization of brain electrical activity via linearly constrained minimum variance spatial filtering, IEEE Transactions on Biomedical Engineering 44 (9) (1997) 867–880. `doi:10.1109/10.623056`.

[20] K. Sekihara, S. S. Nagarajan, Adaptive Spatial Filters for Electromagnetic Brain Imaging, Series in Biomedical Engineering, Springer-Verlag, Berlin Heidelberg, 2008. `doi:10.1007/978-3-540-79370-0`.

[21] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, NeuroImage 15 (1) (2002) 273–289. `doi:10.1006/nimg.2001.0978`.

[22] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T.

Hyman, M. S. Albert, R. J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, NeuroImage 31 (3) (2006) 968–980. `doi:10.1016/j.neuroimage.2006.01.021`.

[23] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, B. T. T. Yeo, Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI, Cerebral Cortex (New York, N.Y.: 1991) 28 (9) (2018) 3095–3114. `doi:10.1093/cercor/bhx179`.

[24] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, The WU-Minn Human Connectome Project: An Overview, NeuroImage 80 (2013) 62–79. `doi:10.1016/j.neuroimage.2013.05.041`.