IMPROVED MOTOR IMAGERY DECODING WITH SPATIOTEMPORAL FILTERING BASED ON BETA BURST KERNELS

S. Papadopoulos[1,2,3], L. Darmet[1,2,3], M.J. Szul[1,3], M. Congedo[4], J.J. Bonaiuto[1,3,†], J. Mattout[1,2,†]

[1] University Lyon 1, Lyon, France

[2] Lyon Neuroscience Research Center, CRNL, INSERM, U1028, CNRS, UMR 5292, Lyon, France

[3] Institut des Sciences Cognitives Marc Jeannerod, CNRS, UMR 5229, Lyon, France

[4] GIPSA-lab, University Grenoble Alpes, CNRS, Grenoble-INP, Grenoble, France

[†] These authors contributed equally

E-mail: sotirios.papadopoulos@univ-lyon1.fr

ABSTRACT: The description of the event-related desynchronization and synchronization phenomena in the mu and beta frequency bands has to a significant extent shaped our understanding of motor-related brain processes. Accordingly, Brain-Computer Interface applications leveraging attempted or imagined movements usually depend on spatially- and band-limited power changes as the brain markers of interest. Yet, converging neuroscience evidence question the idea that signal power best describes the movement-related modulation of brain activity. On a single-trial level, beta band activity is characterized by short, transient and heterogeneous events termed bursts rather than sustained oscillations. In a recent study we demonstrated that a beta burst analysis of hand motor imagery binary classification tasks is often superior to beta power in terms of classification score. Here we expand upon this idea proposing a comparable to state-of-the-art algorithm. We confirm our previous results by using convolution kernels extracted from beta bursts. Moreover, we show that these kernels can effectively be used in inter-session transfer learning strategies.

INTRODUCTION

Three decades ago, a number of seminal studies in motor neuroscience revealed for the first time, time-locked changes in induced power within specific frequency bands [1-3]. These studies described a relative-to-baseline gradual reduction in the power of brain signals recorded during an ongoing movement or motor imagery (MI) task in the mu (~8-12 Hz) [3-6] and beta (~13-30 Hz) [3, 5] frequency bands, termed event-related desynchronization (ERD). They, also, demonstrated a relative increase in power in the beta band shortly after the end of the task, known as event-related synchronization (ERS) [6-8]. The ERD is considered to be a high-level indication of brain processes pertaining to movement preparation and execution, and is particularly prominent in the contralateral sensorimotor cortex [2, 9-12].

Because of this spatial and frequency specificity, ERD is the main marker of interest for motor-related and especially MI-based, non-invasive BCI applications [13, 14]. Typically, signals recorded during MI are transformed in the time-frequency domain (TF) [15–17], and are then spatially filtered using the common spatial pattern algorithm (CSP) [18-20]. This results in an increase of signal-to-noise ratio and extracts the signal power in specific time windows and frequency bands of interest, while also maximizing the spatial disparity among different MI classes (e.g."left" or "right" hand).

The hypothesized reliability and reproducibility of these signal characteristics has also served as the basis for a range of transfer-learning attempts. Transfer learning refers to the exploitation of specific signal features extracted from past recording sessions, different subjects and/or experiments to guide decoding during future sessions [21].

Despite the fact that the ERD and ERS are widely observed, their nature is not well-understood and converging neurophysiology evidence puts these phenomena into question. The ERD and ERS are typically revealed by averaging signal power in the TF domain over multiple trials, especially in the beta frequency band [11, 22], under the assumption of sustained oscillations. However, this evidence points out that, on the contrary, on a single trial level, beta band activity occurs in short events, termed bursts [11, 22–26]. The rate of these beta bursts is more behaviorally relevant in motor processes [11, 24, 27-30] than averaged beta band power. Moreover, it has been shown that beta bursts comprise heterogeneous events [29] with different functions, alluded to by their differential modulation during different task conditions [31, 32] or phases [29, 30].

In a recent study we showed that the analysis of beta

bursts from channels C3 and C4 during hand motor imagery can be advantageous to beta power in terms of classification, confirming the hypothesis that on the single-trial level beta burst rate modulations are more behaviorally relevant than beta band power changes [31]. In this article we expand upon that study. We introduce an algorithm that exploits beta bursts in order to transform brain signals into measures of waveform-resolved burst rate. Moreover, this algorithm can take advantage of an arbitrary number of recorded signals while being computationally efficient, thus constructing decoding features that are comparable to state-of-the-art in BCI. We analyze the activity during "left" and "right" hand MI of three open EEG datasets and show that the use of beta bursts instead of beta band power can improve classification results. Finally, we adopt a transfer learning approach and show that beta bursts detected in one recorded session can be exploited to guide the decoding in other sessions.

MATERIALS AND METHODS

Datasets: We analyzed three open EEG MI datasets: BNCI 2014-001 [13], BNCI 2014-004 [33] and Zhou 2016 [34], all available through the MOABB project [14] (Table 1). These datasets are composed of recordings of numerous subjects who were required to perform sustained MI following the appearance of a visual cue on a screen. For our analysis we only considered trials corresponding to the 'left hand' or 'right hand' classes even though two experimental paradigms consisted of more MI classes. A brief account of the tasks is described in [31].

Pre-processing: Each subject's epoched recordings were loaded with the MOABB python package (v0.4.6) LeftRightImagery class, and were filtered with a low pass cutoff of 120 Hz (zero-phase FIR filter designed with the windowed approach and transition bandwidth of 25% of the low pass frequency). Then, we rejected trials using the autoreject python package [35] (v0.4.0, function get_rejection_threshold) (Table 1). We refer the reader to [31] for more details regarding the pre-processing.

Burst-detection: Following pre-processing, a subset of channels above the sensorimotor cortex was defined ('C3', 'Cz', 'C4', and 'FC3', 'FCz', 'FC4', 'CP3', 'Cpz', 'CP4' when available; as in [31]). The corresponding recordings were first transformed in the time-frequency (TF) domain from 1 to 43 Hz using the superlets algorithm [36] (parameters: o_{min} = 1, o_{max} = 40, c = 4) with a frequency resolution of 0.5 Hz. Then, we identified bursts within the beta frequency range (15-30Hz) from each TF matrix using a previously published procedure that allowed us to extract the waveforms of the beta bursts within a fixed time-window (see [29] for more information on the algorithm and [31] on how it was specifically applied to these datasets).

Kernel selection: For each dataset and subject we randomly sampled 10% of the recording trials (or all

Table 1. Dataset attributes

Dataset	BNCI 2014-001	BNCI 2014-004	Zhou 2016
# Subjects	9	9	4
# Channels	3	22	64
# Sessions	2	5	3
Min - Max # Trials across subjects	288	680-760	290-319
Min - Max # Trials (after trial rejection)	217-287	269-621	114-280
Sampling freq. (Hz)	250	250	250
Trial duration (s)	4.0	4.5	5.0
Reference	[13]	[33]	[34]

trials corresponding to a recording session when assessing transfer learning) of each participant after trial rejection (Table 1) ensuring class balance, in order to create a large sample of beta burst waveforms while restricting the number of trials per participant we excluded from classification (see *Feature Selection* and *Classification*). We aggregated all detected bursts within these trials in a matrix (irrespective of the trial class, i.e. 'left hand' or 'right hand') after robust scaling (scikit-learn package [37], v1.0.2). Then, we used principal component analysis (PCA) [38] (scikit-learn package, v1.0.2) in order to reduce the time dimension of the waveforms. We defined an index of lateralized modulation of the average-per-axis PCA score I_m (a metric of the difference between any burst waveform and the average shape) from the baseline to the trial periods of the recordings as using electrodes C3 and C4:

$$I_m = |(u_{ipsi}^{C3} - u_{contr}^{C4}) - (u_{ipsi}^{C4} - u_{contr}^{C3})|, m \in \{2,...,9\}$$

$$u = |\widehat{score}_{trial\ period} - \widehat{score}_{baseline}|$$

where *ipsi (contr)* refers to bursts recorded from channels C3 / C4 during a left / right (right / left) hand MI. Using this index, we identified the three PCA axes that maximized I_m among components 2 to 9. We did not consider the first PCA component as it simply describes the temporal skew of the burst waveforms [29, 31]. Finally, based on our previous study [31], we split the computed score range of each of the three selected axes in seven equally spaced groups. We considered the two groups that lie further away from the origin (score equal to 0), grouped similarly shaped bursts together and defined 2 kernels per PCA axis by computing the average waveform of all bursts within these two groups. As a result, we ended up with 6 kernels that describe the
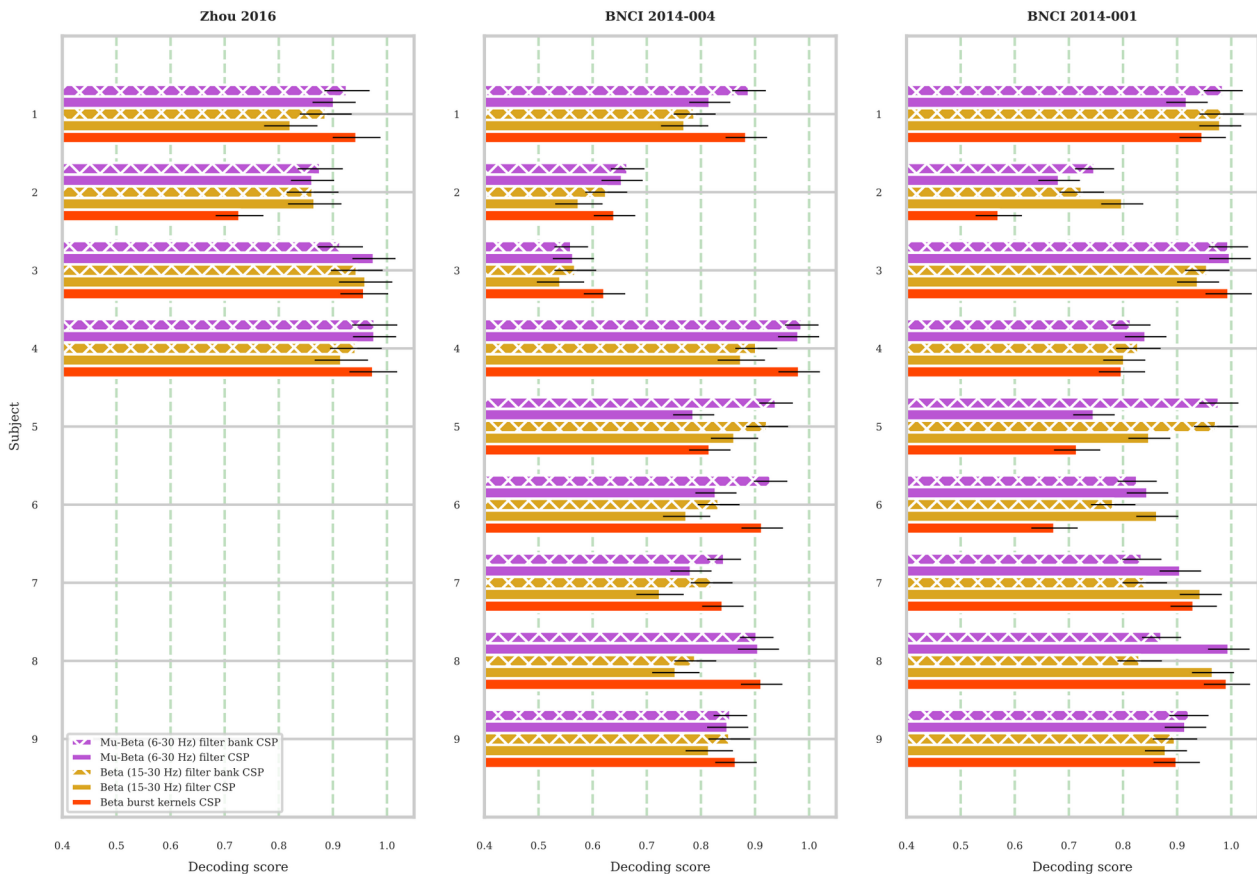
Figure 1: Average decoding score (area under the curve of the receiver operator characteristic) and standard deviation per dataset, subject and classification feature.

burst waveforms that were expected to be maximally rate-modulated during the task, compared to baseline, for each participant of each dataset separately.

Feature selection: We used these kernels in order to transform the available EEG recordings of all initially available channels, not considering the subset of trials that were used for defining the kernels (the random trials sample or a given recording session). Specifically, a copy of each subject's epoched data was convolved with one kernel, resulting in a proxy of the waveform-resolved burst rate. Then, each of the temporally filtered epoched data was spatially filtered using the CSP algorithm (MNE package [39], v 1.5.1, function CSP, parameters: n_components = 4, transform_into = "average_power").

We also created band-limited spatial features with a standard filtering technique. Specifically, we computed the envelope of the epoched data Hilbert transform (MNE package, v 1.5.1, function apply_hilbert) after independently applying a single filter in the beta frequency band (15-30 Hz) and a wider frequency range encompassing both the mu and beta bands (6-30 Hz). Then, these signals served as inputs to the CSP algorithm (we kept the parameters unchanged).

In order to assess whether the number of spatial features used for classification (24 for the beta burst kernel approach versus 4 for the Hilbert power approach)

affected the observed results, we also adopted a filter bank approach. We split either frequency range in non-overlapping filter banks of 3 Hz range; this resulted in 5 filters for the beta band (15-18 Hz, 18-21 Hz, 21-24 Hz, 24-27 Hz, 27-30 Hz) and 8 filters for the mu-beta band (6-9 Hz, 9-12 Hz, 12-15 Hz, 15-18 Hz, 18-21 Hz, 21-24 Hz, 24-27 Hz, 27-30 Hz) corresponding to 20 and 32 spatial features respectively.

We also estimated kernel-specific CSP filters on one session and used these filters in the remaining sessions. For comparison, we estimated the CSP filters of a session and used them to transform the signals of the other sessions for each of the previously described filtering techniques.

Classification: The resulting spatially filtered data were concatenated in a single matrix, once per subject. Using a repeated (n=10), 5-fold cross validation procedure we shuffled the remaining trials (Table 1) and estimated the decoding score using LDA (scikit-learn, v1.0.2) as a classifier using the whole trial recordings. The decoding scores were based on the area under the curve (AUC) of the receiver operating characteristic (scikit-learn, v1.0.2). All numeric computations were based on the numpy python package (v1.21.6; [40]) and an environment running python (v3.10).

Statistical analysis: To estimate, at the population level, any statistical difference between methods, we compa-

Figure 2: Average difference in decoding score (area under the curve of the receiver operator characteristic) between classification score of a transfer learning session and classification score obtained using all available trials per subject. A value of 0 indicates no difference in decoding between the two approaches. Positive values or negative values indicate a performance gain or loss respectively when adopting a transfer learning approach.

red classification results of the waveform-resolved burst features against multiple band-limited power features. We used a generalized linear mixed model with a binomial distribution and logit link function with across-trial average classification score as the dependent variable setting the number of trials as prior weights, the type of classification feature as a fixed effect, and the subject nested within the dataset as random intercepts. We also compared the across-session average difference in classification score between a transfer learning approach and an across-session, global classification, using a similar model but with a Gaussian distribution. Statistical analyses were conducted using R (v4.1.2) and lme4 (v1.1-31; [41]). Fixed effects were assessed using type II Wald X 2 tests using car (v3.1-1; [42]). Pairwise Tukey- or Sidak-corrected follow-up tests were carried out using estimated marginal means from the emmeans package (v.1,8,7; [43]).

RESULTS

We estimated the across-session decoding score per subject of each dataset using the waveform-resolved approach. In order to assess the significance of the waveform-resolved burst rate features we also computed classification features based on signal power and compared the different decoding results. We used a filter bank approach to assess whether the number of spatial features used for classification affected the decoding scores (Fig. 1).

Across all datasets, the waveform-resolved burst rate features resulted in statistically significant improvement in decoding performance compared to beta band power estimated using a single filter (X^2 (5) = 97.081, p < 0.001), non-significant differences compared to the beta band filter bank and mu-beta band single filter approaches (p = 0,1047 and p = 0,9987 respectively), and significant decoding score decrease compared to mu-beta band filter bank approach (p = 0.004).

Moreover, we assessed whether the beta burst kernels can be exploited in a transfer learning approach. To do so, after trial rejection we iteratively estimated the kernels based on the recordings of one session and used them for transforming the recordings of the remaining sessions. We also examined whether we can use kernel-specific CSP filters for transfer learning. We compared these approaches to transfer learning of CSP filters after applying the standard filtering techniques.

The difference between the results depicted in Fig 1. and the transfer learning approach results per subject of

each dataset are depicted in Fig. 2. Negative values indicate greater decoding performance when estimating the results based on all available recordings, whereas positive values indicate an improvement in decoding score using transfer learning. Without surprise, transfer learning based on the beta burst kernels slightly reduced the overall decoding score, but was, interestingly, the only method that had the potential to improve decoding for some of the subjects. The two kernel-based transfer learning approaches yielded statistically non-significant discrepancies in decoding score difference (X^2 (6) = 28.8, p = 0.93). Transfer learning based only on the kernels yielded significantly smaller reduction in decoding score difference compared to all filtering methods (p = 0.013 vs beta band filter; p < 0.001 vs beta band filter bank; p = 0.0237 vs mu-beta band filter; p = 0.0011 vs mu-beta band filter bank). Transfer learning based on kernel-specific CSP patterns resulted in significantly smaller decoding score differences in all comparisons except for filtering in the mu-beta band using a single filter (p = 0.0392 vs beta band filter; p = 0.0161 vs beta band filter bank; p = 0.3061 vs mu-beta band filter; p = 0.034 vs mu-beta band filter bank).

DISCUSSION

Recently, the fields of neurophysiology and systems neuroscience have been experiencing a surge in the development of novel methods for analyzing neural recordings. An increasing number of articles are concerned with unveiling traditionally disregarded signal characteristics [11, 22-26, 29, 30] with several important implications for BCI. Particularly, the description of beta bursts has put into question the importance of band-limited power modulations in motor or MI tasks.

BCI applications often rely on signal power under the assumption of sustained and/or oscillatory signals. However, the description of waveform-specific beta burst modulations opens up new possibilities and holds the potential of improving decoding [31]. In this article we verified that waveform-resolved burst rate features can be informative markers of the underlying brain activity during MI tasks. In line with previous results we showed that indeed waveform-resolved burst rate features can be more informative than beta band power alone. We also showed that the information content of these features is comparable to that of filtering within a wider frequency band encompassing the mu band, although the mu-beta filter bank approach still yielded the best decoding results. A possible explanation for this finding is that beta burst kernels also capture slower modulations of the underlying activity and, thus, by adapting the waveform-resolved burst rate features to the mu band characteristics [45, 46] we may be able to further improve decoding.

Additionally, we adopted an inter-session transfer learning approach, and showed that the waveform-resolved features are relatively stable over recording sessions. We demonstrated that on the dataset level

reduction in classification score can be minimal, and that these features can even contribute in improving decoding for subjects with low classification score. This finding is important because it proves that, despite the difficulty of exploiting spatial features learned during previous sessions, there is potential in improving how previously acquired data can be leveraged during offline calibration or even online decoding sessions. In the future, an assessment of the across-subject kernel similarity may open the path for more reliable inter-subject transfer learning.

CONCLUSION

The waveform-resolved burst rate analysis is a promising, neurophysiology-grounded alternative to classic descriptions of beta band activity in motor and MI tasks. The results of this work reaffirm that classification features based on beta bursts can efficiently decode MI binary classification tasks and suggest that beta bursts kernels are stable across recording sessions, thus potentially serving as an interesting feature inter-session transfer learning.

REFERENCES

[1] Pfurtscheller G 1981 Central beta rhythm during sensorimotor activities in man Electroencephalogr. Clin. Neurophysiol. 51 253–64

[2] Pfurtscheller G and Berghold A 1989 Patterns of cortical activation during planning of voluntary movement Electroencephalogr. Clin. Neurophysiol. 72 250–8

[3] Pfurtscheller G and Da Silva F L 1999 Event related EEG/MEG synchronization and desynchronization: basic principles Clin. Neurophysiol. 110 1842–57

[4] Pfurtscheller G, Brunner C, Schlögl A and Da Silva F L 2006 Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks NeuroImage 31 153–9

[5] Pfurtscheller G, Neuper C, Flotzinger D and Pregenzer M 1997 EEG-based discrimination between imagination of right and left hand movement Electroencephalogr. Clin. Neurophysiol. 103 642–51

[6] Neuper C, Wörtz M and Pfurtscheller G 2006 ERD/ERS patterns reflecting sensorimotor activation and deactivation Progress in Brain Research vol 159 ch 14, pp 211–22

[7] Pfurtscheller G, Stancák A and Neuper C 1996 Post-movement beta synchronization. A correlate of an idling motor area? Electroencephalogr. Clin. Neurophysiol. 98 281–93

[8] Alayrangues J, Torrecillos F, Jahani A and Malfait N 2019 Error-related modulations of the sensorimotor post-movement and foreperiod beta-band activities arise from distinct neural substrates and do not reflect efferent signal processing NeuroImage 184 10–24

[9] Makeig S, Enghoff S, Jung T P and Sejnowski T J

2000 A natural basis for efficient brain-actuated control IEEE Trans. Rehabil. Eng. 8 208–11

[10] Pfurtscheller G and Neuper C 1997 Motor imagery activates primary sensorimotor area in humans Neurosci. Lett. 239 65–68

[11] Little S, Bonaiuto J, Barnes G and Bestmann S 2019 Human motor cortical beta bursts relate to movement planning and response errors PLoS Biol. 17 1–30

[12] Zich C, Quinn A J, Bonaiuto J J, O'Neill G, Mardell L C, Ward N S et al. 2023 Spatiotemporal organization of human sensorimotor beta burst activity elife 12 e80160

[13] Tangermann M et al 2012 Review of the BCI competition IV Front. Neurosci. 6 1–31

[14] B. Aristimunha, "Mother of all BCI Benchmarks". Zenodo, Oct. 23, 2023. doi: 10.5281/zenodo. 10034224.

[15] Bruns A 2004 Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? J. Neurosci. Methods 137 321–32

[16] Herman P, Prasad G, McGinnity T M and Coyle D 2008 Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification IEEE Trans. Neural Syst. Rehabil. Eng. 16 317–26

[17] Brodu N, Lotte F and Lécuyer A 2011 Comparative study of band-power extraction techniques for motor imagery classification IEEE SSCI 2011—Symp. Ser. Comput. Intell.—CCMB 2011 2011 IEEE Symp. Comput. Intell. Cogn. Algorithms, Mind, Brain pp 95–100

[18] Koles Z J 1991 The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG Electroencephalogr. Clin. Neurophysiol. 79 440–7

[19] Blankertz B, Kawanabe M, Tomioka R, Hohlefeld F U, Nikulin V and Müller K R 2008 Invariant common spatial patterns: alleviating non-stationarities in brain-computer interfacing Advances in Neural Information Processing Systems 20—Proc. 2007 Conf. Pp 1–8

[20] Müller-Gerking J, Pfurtscheller G and Flyvbjerg H 1999 Designing optimal spatial filters for single-trial EEG classification in a movement task Clin. Neurophysiol. 110 787–98

[21] Lotte F, Bougrain L, Cichocki A, Clerc M, Congedo M, Rakotomamonjy A et al. 2018 A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update J. Neural Eng. 15 031005

[22] Jones S R 2016 When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning Curr. Opin. Neurobiol. 40 72–80

[23] Lundqvist M, Rose J, Herman P, Brincat S, Buschman T and Miller E 2016 Gamma and beta bursts underlie working memory Neuron 90 152–64

[24] Wessel J R 2020 β-bursts reveal the trial-to-trial dynamics of movement initiation and cancellation J. Neurosci. 40 411–23

[25] Shin H, Law R, Tsutsui S, Moore C I and Jones S R 2017 The rate of transient beta frequency events predicts impaired function across tasks and species elife 6:e29086

[26] Torrecillos F et al 2018 Modulation of beta bursts in the subthalamic nucleus predicts motor performance J. Neurosci. 38 8905–17

[27] Hannah R, Muralidharan V, Sundby K K and Aron A R 2020 Temporally-precise disruption of prefrontal cortex informed by the timing of beta bursts impairs human action-stopping NeuroImage 222 117222

[28] Enz N, Ruddy K L, Rueda-Delgado L M and Whelan R 2021 Volume of β-bursts, but not their rate, predicts successful response inhibition J. Neurosci. 41 5069–79

[29] Szul M J, Papadopoulos S, Alavizadeh S, Daligaut S, Schwartz D, Mattout J et al. 2023 Diverse beta burst waveform motifs characterize movement-related cortical dynamics Prog. Neurobiol. 165187

[30] Rayson H, Szul M J, El-Khoueiry P, Debnath R, Gautier-Martins M, Ferrari P F, et al. 2023 Bursting with potential: how sensorimotor beta bursts develop from infancy to adulthood J. Neurosci. 43 8487–503

[31] Papadopoulos S, Szul M J, Congedo M, Bonaiuto J J and Mattout J, Beta bursts question the ruling power for brain-computer interfaces J. Neural Eng., 2024, 21 (1), pp.016010. https://doi.org/10.1088/ 1741-2552/ad19ea.

[32] Langford Z D, Procyk E and Wilson C R E 2023 Frontal oscillatory beta bursts have rhythmically distinct regimes with differing functional relevance bioRxiv 1–19

[33] Leeb R, Lee F, Keinrath C, Scherer R, Bischof H and Pfurtscheller G 2007 Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment IEEE Trans. Neural Syst. Rehabil. Eng. A 15 473–82

[33] Zhou B, Wu X, Lv Z, Zhang L and Guo X 2016 A fully automated trial selection method for optimization of motor imagery based brain-computer interface PLoS One 11 1–20

[35] Jas M, Engemann D, Bekhti Y, Raimondo F and Gramfort A 2017 Autoreject: Automated artifact rejection for MEG and EEG data NeuroImage 159 417–129

[36] Moca V V, Bârzan H, Nagy-Dăbâcan A and Muresan R C 2021 Time-frequency super-resolution with superlets Nat. Commun. 12 1–18

[37] Pedregosa F et al 2011 Scikit-learn: machine learning in Python Fabian J. Mach. Learn. Res. 12 2825–30

[38] Shlens J 2014 A tutorial on principal component analysis (arXiv:1404.1100)

[39] Gramfort A, Luessi M, Larson E, Engemann D A, Strohmeier D, Brodbeck C, et al. 2013 MEG and EEG data analysis with MNE-Python Front. Neurosci.

[40] Harris C R et al 2020 Array programming with NumPy Nature 585 357–62

[41] Bates D, Mächler M, Bolker B M and Walker S C 2015 Fitting linear mixed-effects models using lme4 J. Stat. Softw. 67 1–48

[42] Fox J and Weisberg S 2019 An R Companion to Applied Regression (Sage)

[43] Lenth R V 2023 emmeans: estimated marginal means, aka least-squares means

[44] Gerster M, Waterstraat G, Litvak V, Lehnertz K, Schnitzler A, Florin E, et al. 2022 Separating Neural Oscillations from Aperiodic 1/f Activity: Challenges and Recommendations Neuroinformatics 20 991–1012

[45] Vigué-Guix I and Soto-Faraco S 2022 Using occipital α-bursts to modulate behaviour in real-time BioRxiv

[46] Chen Y Y, Lambert K J M, Madan C R and Singhal A 2021 Mu oscillations and motor imagery performance: a reflection of intra-individual success, not inter-individual ability Hum. Mov. Sci. 78 1–12