Single-trial characterization of neural rhythms: potentials and challenges

Julian Q. Kosciessa[1,2,*], Thomas H. Grandy[2], Douglas D. Garrett[1,2], Markus Werkle-Bergner[2,*]

[1]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin/London; [2]Center for Lifespan Psychology, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

[*] Corresponding authors:
kosciessa@mpib-berlin.mpg.de; werkle@mpib-berlin.mpg.de

Abstract

The average power of rhythmic neural responses as captured by M/EEG/LFP recordings is a prevalent index of human brain function. Increasing evidence questions the utility of trial-/group averaged power estimates, as seemingly sustained activity patterns may be brought about by time-varying transient signals in each single trial. Hence, it is crucial to accurately describe rhythmic and arrhythmic neural responses on the single trial-level. However, it is less clear how well this can be achieved in empirical M/EEG/LFP recordings. Here, we extend an existing rhythm detection algorithm ("eBOSC") to systematically investigate boundary conditions for estimating neural rhythms at the single-trial level. Using simulations and resting and task-based EEG recordings from a micro-longitudinal assessment, we show that rhythms can be successfully captured at the single-trial level with high specificity, but that the quality of single-trial estimates varies greatly between subjects. Importantly, our analyses suggest that rhythmic estimates at the single-trial level are reliable within-subject markers, but are not consistently valid descriptors of the individual rhythmic process. Finally, we discuss the utility and potential of rhythm detection, and various implications for single-trial analyses of neural rhythms in electrophysiological recordings.

Keywords: rhythm detection; abundance; alpha power; inter-individual differences; single-trial estimates

Highlights

- Extension of a state-of-the-art rhythm detection method (eBOSC).
- Rhythm detection can offer specific indices of single-trial rhythmicity.
- Arrhythmic duration systematically biases rhythmic power estimates.
- Power- and phase-based definitions of rhythmicity derive similar estimates of rhythmic duration.
- Surface EEG recordings exhibit stable inter-individual differences in α-rhythmicity.
- Low levels of rhythmic strength constrain single-trial characterization of neural rhythms.

1.1 Towards a single-trial characterization of neural rhythms

Episodes of rhythmic neural activity in electrophysiological recordings are of prime interest for research on neural representations and computations across multiple scales of measurement (e.g. Buzsáki, 2006; Wang, 2010). At the macroscopic level, the study of rhythmic neural signals has a long heritage, dating back to Hans Berger's classic investigations into the Alpha rhythm (Berger, 1938). Since then, advances in recording and processing techniques have facilitated large-scale spectral analysis schemes (e.g. Gross, 2014) that were not available to the pioneers of electrophysiological research, who often depended on the manual analysis of single time series to indicate the presence and magnitude of rhythmic events. Interestingly, improvements in analytic methods still do not capture all information extracted by manual inspection. For example, current analysis techniques are largely naïve to the specific temporal presence of rhythms in the continuous recordings, as they often employ windowing of condition- or group-based averages to extract putative rhythm-related characteristics (Cohen, 2014). However, the underlying assumption of stationary, sustained rhythms within the temporal window of interest might not consistently be met (Jones, 2016; Stokes & Spaak, 2016), thus challenging the appropriateness of the averaging model (i.e. the ergodicity assumption (Molenaar & Campbell, 2009). Furthermore, in certain situations, single-trial characterizations become necessary to derive unbiased individual estimates of neural rhythms (Cohen, 2017). For example, this issue becomes important when asking whether rhythms appear as bursts or in sustained form (van Ede, Quinn, Woolrich, & Nobre, 2018), or when only single-shot acquisitions are feasible (i.e. resting state).

1.2 Duration as a powerful index of rhythmicity

The presence of rhythmicity is a necessary prerequisite for the accurate interpretation of measures of amplitude, power, and phase (Aru et al., 2015; Jones, 2016; Muthukumaraswamy & Singh, 2011). This is exemplified by the bias that arrhythmic periods exert on rhythmic power estimates. Most current time-frequency decomposition methods of neurophysiological signals (such as the electroencephalogram (EEG)) are based on the Fourier transform (Gross, 2014). Following Parceval's theorem (e.g. Hansen, 2014), the Fast Fourier Transform (FFT) decomposes an arbitrary time series into a sum of sinusoids at different frequencies. Importantly, FFT-derived power estimates do not differentiate between high

amplitude transients and low amplitude sustained signals. In the case of FFT power, this is a direct result of the violated assumption of stationarity in the presence of a transient signal. Short-time FFT and wavelet techniques alleviate (but do not eliminate) this problem by analyzing shorter epochs, during which stationarity is more likely to be obtained. However, whenever spectral power is averaged across these episodes, both high-amplitude rhythmic and low-amplitude arrhythmic signal components may once again become intermixed. In the presence of arrhythmic content (often referred to as the "signal background," or "noise"), this results in a reduced amplitude estimate of the underlying rhythm, the extent of which relates to the duration of the rhythmic episode relative to the length of the analyzed segment (which we will refer to as 'abundance') (see Figure 1A for a schematic). Therefore, integration across epochs that contain a mixture of rhythmic and arrhythmic signals results in an inherent ambiguity between the strength of the rhythmic activity (as indexed by power/amplitude) and its duration (as indexed by the abundance of the rhythmic episode within the segment).

Crucially, the strength and duration of rhythmic activity differ in their neurophysiological interpretation. Rhythmic power most readily indexes the magnitude of synchronized changes in membrane potentials within a network (Buzsáki, Anastassiou, & Koch, 2012), and is thus related to the size of the participating neural population. The duration of a rhythmic episode, by contrast, tracks how long population synchrony is upheld. Notably, measures of rhythm duration have recently gained interest as they may provide additional information regarding the biophysical mechanisms that give rise to the recorded signals (Peterson & Voytek, 2017; Sherman et al., 2016), for example, by differentiating between burst-like and sustained rhythmic events (van Ede et al., 2018).

1.3. Single-trial rhythm detection as a methodological challenge

In general, the accurate estimation of process parameters depends on a sufficiently strong signal in the neurophysiological recordings under investigation. Especially for scalp-level M/EEG recordings it remains elusive whether neural rhythms are sufficiently strong to be clearly detected in single trials. Here, a large neural population has to be synchronously active to give rise to potentials that are visible at the scalp surface. This problem intensifies further by signal attenuation through the skull (in the case of EEG) and the superposition of signals from diverse sources of no interest both in- and outside the brain (da Silva, 2018). In sum, these considerations lead to the proposal that the signal-to-noise ratio (SNR), here operationally

defined as the ratio of rhythmic to arrhythmic variance, may fundamentally constrain the accurate characterization of single-trial rhythms.

Recently, different methods have been proposed to characterize rhythmicity at the single-trial-level: the power-based Better OSCillation Detection (BOSC; Caplan, Madsen, Raghavachari, & Kahana, 2001; Whitten, Hughes, Dickson, & Caplan, 2011) and the phase-based lagged coherence index (Fransen, van Ede, & Maris, 2015). Notably, both proposed algorithms make different assumptions regarding the definition of rhythmicity: BOSC assumes that rhythms are defined as spectral peaks on top of an arrhythmic 1/f background, whereas lagged coherence defines rhythms based on the predictability of phase estimates at a lag defined by the rhythm's period. Following those considerations, we set out to answer the following hypotheses and questions: (1) A precise differentiation between rhythmic and arrhythmic timepoints increases rhythmic power estimates through disambiguating the strength and the duration of the rhythmic process. (2) The feasibility of single-trial rhythm detection depends on the individual rhythmic representation, which may vary across subjects. To what extent does the single-trial rhythm representation in empirical data allow for an accurate single-trial characterization? (3) To what extent do power- and phase-based definitions of rhythmicity converge in rhythmicity estimates?

Here, we apply an extended version of the BOSC algorithm (eBOSC) to resting- and task-state data from a micro-longitudinal dataset to systematically investigate the feasibility to derive reliable and valid indices of neural rhythmicity from single-trial scalp EEG data. We calculate lagged coherence during the resting state to probe the convergence between rhythm definitions. Furthermore, we use computational simulations to derive rhythm detection benchmarks and to probe its boundary conditions. Finally, we showcase eBOSC's ability to characterize multiple rhythmic features. We focus on alpha rhythms (~8-15 Hz; defined here based on individual FFT-peaks) due to (a) their high amplitude in human EEG recordings, (b) the pervasive focus on alpha rhythms in the detection literature (Caplan, Bottomley, Kang, & Dixon, 2015; Fransen et al., 2015; Whitten et al., 2011), and (c) their importance for human cognition (Grandy et al., 2013; Klimesch, 2012; Sadaghiani & Kleinschmidt, 2016).

2. Methods

2.1 Study design

Resting state and task data were collected in the context of a larger assessment, consisting of eight sessions in which an adapted Sternberg short-term memory task (Sternberg, 1966) and three additional cognitive tasks were repeatedly administered. Resting state data are from the first session, task data are from sessions one, seven and eight, during which EEG data were acquired. Sessions one through seven were completed on consecutive days (excluding Sundays) with session seven completed seven days after session one by all but one participant (eight days due to a two-day break). Session eight was conducted approximately one week after session seven (M = 7.3 days, SD = 1.4) to estimate the stability of the behavioral practice effects. The reported EEG sessions lasted approximately three and a half to four hours, including approximately one and a half hours of EEG preparation. For further details on the study protocol and results of the behavioural tasks see (Grandy, Lindenberger, & Werkle-Bergner, 2017).

2.2 Participants

The sample contained 32 young adults (mean age = 23.3 years, SD = 2.0, range 19.6 to 26.8 years; 17 women; 28 university students) recruited from the participant database of the Max Planck Institute for Human Development, Berlin, Germany (MPIB). Participants were right-handed, as assessed with a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal vision, as assessed with the Freiburg Visual Acuity test (Bach, 1996; 2007). Participants reported to be in good health with no known history of neurological or psychiatric incidences and were paid for their participation (8.08 € per hour, 25.00 € for completing the study within 16 days, and a performance-dependent bonus of 28.00 €; see below). All participants gave written informed consent according to the institutional guidelines of the ethics committee of the MPIB, which approved the study.

2.3 Procedure

Participants were seated at a distance of 80 cm in front of a 60 Hz LCD monitor in an acoustically and electrically shielded chamber. A resting state assessment was conducted prior to the initial performance of the adapted Sternberg task. Two resting state periods were used: the first encompassed a duration of two minutes of continuous eyes open (EO1) and eyes closed (EC1) periods, respectively; the second resting state was comprised of two 80 second runs,

totalling 16 repetitions of 5 seconds interleaved eyes open (EO2) – eyes closed (EC2) periods. An auditory beep indicated to the subjects when to open and close their eyes.

Following the resting assessments, participants performed an adapted version of the Sternberg task. Digits were presented in white on a black background and subtended ~2.5° of visual angle in the vertical and ~1.8° of visual angle in the horizontal direction. Stimulus presentation and recording of behavioral responses were controlled with E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, USA). The task design followed the original report (Sternberg, 1966). Participants started each trial by pressing the left and right response key with their respective index fingers to ensure correct finger placement and to enable fast responding. An instruction to blink was given, followed by the sequential presentation of 2, 4 or 6 digits from zero to nine. On each trial, the memory set size (i.e. load) varied randomly between trials, and participants were not informed about the upcoming condition. Also, the single digits constituting a given memory set were randomly selected in each trial. Each stimulus was presented for 200 ms, followed by a fixed 1000 ms blank inter-stimulus interval (ISI). The offset of the last stimulus coincided with the onset of a 3000 ms blank retention interval, which concluded with the presentation of a probe item that was either contained in the presented stimulus set (*positive probe*) or not (*negative probe*). Probe presentation lasted 200 ms, followed by a blank screen for 2000 ms, during which the participant's response was recorded. A beep tone indicated the end of the trial. The task lasted about 50 minutes.

For each combination of load x probe type, 31 trials were conducted, cumulating in 186 trials per session. Combinations were randomly distributed across four blocks (block one: 48 trials; blocks two through four: 46 trials). Summary feedback of the overall mean RT and accuracy within the current session was shown at the end of each block. At the beginning of session one, 24 practice trials were conducted to familiarize participants with the varying set sizes and probe types. To sustain high motivation throughout the study, participants were paid a 28 € bonus if their current session's mean RT was faster or equal to the overall mean RT during the preceding session, while sustaining accuracy above 90%. Only correct trials were included in the analyses.

2.4 EEG recordings and pre-processing

EEG was continuously recorded from 64 Ag/AgCl electrodes using BrainAmp amplifiers (Brain Products GmbH, Gilching, Germany). Sixty scalp electrodes were arranged within an elastic cap (EASYCAP GmbH, Herrsching, Germany) according to the 10% system

(cf. Oostenveld, Fries, Maris, & Schoffelen, 2011) with the ground placed at AFz. To monitor eye movements, two electrodes were placed on the outer canthi (horizontal EOG) and one electrode below the left eye (vertical EOG). During recording, all electrodes were referenced to the right mastoid electrode, while the left mastoid electrode was recorded as an additional channel. Prior to recording, electrode impedances were retained below 5 kΩ. Online, signals were recorded with an analog pass-band of 0.1 to 250 Hz and digitized at a sampling rate of 1 kHz.

Preprocessing and analysis of EEG data were conducted with the FieldTrip toolbox (Oostenveld et al., 2011) and using custom-written MATLAB (The MathWorks Inc., Natick, MA, USA) code. Offline, EEG data were filtered using a 4th order Butterworth filter with a pass-band of 0.5 to 100 Hz, and were linearly detrended. Task data were segmented to 21s epochs ranging from -9s to +12s with regard to the onset of the 3s retention interval. Resting data were segmented to their respective recording interval. For the interleaved data, the first and last trial for each condition were removed, resulting in an effective trial number of 14 trials per condition. Blink, movement and heart-beat artifacts were identified using Independent Component Analysis (ICA; Bell & Sejnowski, 1995) and removed from the signal. Subsequently, data were downsampled to 250 Hz and all channels were re-referenced to mathematically averaged mastoids. Artifact-contaminated channels (determined across epochs) were automatically detected (a) using the FASTER algorithm (Nolan, Whelan, & Reilly, 2010) and (b) by detecting outliers exceeding three standard deviations of the kurtosis of the distribution of power values in each epoch within low (0.2-2 Hz) or high (30-100 Hz) frequency bands, respectively. Rejected channels were interpolated using spherical splines (Perrin, Pernier, Bertrand, & Echallier, 1989). Subsequently, noisy epochs were likewise excluded based on FASTER and recursive outlier detection, resulting in the rejection of approximately 13% of trials. To prevent trial rejection due to artifacts outside the signal of interest, artifact detection was restricted to epochs that included 2.4s of additional signal around the on- and offset of the retention interval, corresponding to the longest effective segment that was used in the analyses. A further 2.65% of incorrectly answered trials from the task were subsequently excluded.

2.5 Rhythm-detection using extended BOSC

We applied an extended version of the Better OSCillation detection method (eBOSC; cf. Caplan et al., 2001; Whitten et al., 2011) to automatically separate rhythmic from arrhythmic

episodes. The BOSC method reliably identifies rhythms using data-driven thresholds based on theoretical assumptions of the signal characteristics. Briefly, the method defines rhythms as temporal episodes during which wavelet-derived power at a particular frequency exceeds a *power threshold* based on an estimate of the arrhythmic signal background for a minimum duration of cycles (*duration threshold*) at the particular frequency. The theoretical *duration threshold* excludes high amplitude transients. Previous applications of the BOSC method focused on the analysis of resting-state data or long data epochs, where reliable detection has been established regardless of specific parameter setups (Caplan et al., 2001; 2015; Whitten et al., 2011). We introduce the following adaptations here (for details on the implementation of the adaptations, please see the Supplementary Methods): (1) we use a form of robust regression in place of linear regression following removal of the alpha peak; (2) we combine detected time points into continuous rhythmic episodes and (3) we reduce the impact of wavelet convolution on abundance estimates (see Supplementary Methods 1 & 2). We benchmarked the algorithm and compared it to standard BOSC using simulations (see Supplementary Methods 3).

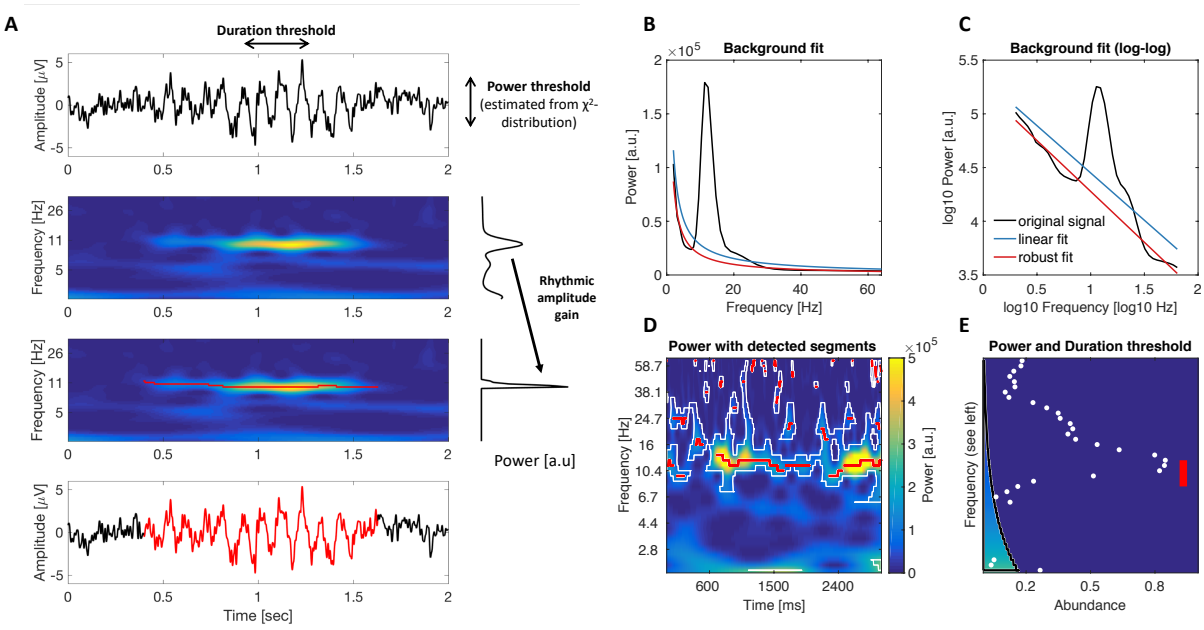


Figure 1: Schematic illustration of rhythm detection. (A) Average amplitude estimates (right) increase with the constraint on rhythmic episodes within the averaged time interval. The left plots show simulated time series and the corresponding time-frequency power. Superimposed red traces indicate rhythmic time points. The upper right plot shows the average power spectrum averaged across the entire epoch, the lower plot presents amplitudes averaged exclusively across rhythmic time points. An amplitude gain is observed due to the exclusion of arrhythmic

low amplitude time points. (B-E) Comparison of standard and extended oscillation detection. (B+C) Rhythms were detected based on a power threshold estimated from the arrhythmic background spectrum. Standard BOSC applies a linear fit in log-log space to define the background power, which may overestimate the background at the frequencies of interest in the case of data with large rhythmic peaks. Robust regression following peak removal alleviates this problem. (D) Example of episode detection. White borders circumfuse time frequency points, at which standard BOSC indicated rhythmic content. Red traces represent the continuous rhythmic episodes that result from the extended post-processing. (E) Applied thresholds and detected rhythmic abundance. The black border denotes the duration threshold at each frequency (corresponding to D), i.e. for how long the power threshold needed to be exceeded to count as a rhythmic period. Note that this threshold can be set to zero for a post-hoc characterization of the duration of episodes (see Methods 2.12). The color scaling within the demarcated area indicates the power threshold at each frequency. Abundance corresponds to the relative length of the segment on the same time scale as presented in D. White dots correspond to the standard BOSC measure of rhythmic abundance at each frequency (termed Pepisode). Red lines indicate the abundance measure used here, which is defined as the proportion of sample points at which a rhythmic episode between 8-15 Hz was indicated (shown as red traces in D).

2.7 Definition of amplitude, abundance and rhythmic probability

The ***abundance*** of alpha rhythms denotes the relative duration of rhythmic episodes with a mean frequency in the alpha range (8 to 15 Hz). This frequency range was selected as individual resting state spectra showed clear peaks within this range (Supplementary Figure 5). ***Rhythmic amplitudes*** were extracted within these rhythmic episodes. If no alpha episode was indicated, abundance was set to zero, and amplitude was set to missing. To compare this rhythm-specific measure to the standard measure of window-averaged amplitudes, ***overall amplitudes*** were computed by averaging across the entire segment at its alpha peak frequency. Unless indicated otherwise, both amplitude measures were normalized by subtracting the amplitude estimate of the fitted background spectrum. This step represents a parameterization of rhythmic power (cf. Haller et al., 2018) and is conceptually similar to baseline normalization, without requiring an explicit baseline segment. We further define ***rhythmic probability*** as the probability of a detected rhythmic episode across trials within the alpha frequency range. Unless stated differently, subject-, and condition-specific amplitude and abundance values were

averaged within and across trials, and across posterior-occipital channels (P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2), in which alpha power was maximal (Figure 2A).

In addition, we calculated an ***empirical signal-to-noise ratio (empSNR)*** as the ratio of the individual alpha frequency (IAF) amplitude to the background fit: empSNR = $\frac{OverallPower}{Background}$. In addition, we defined ***'effective rhythmic SNR' (effSNR)*** as the background-normalized rhythmic power above the scale free background as an estimate for the representation of the rhythmic peak: effSNR = $\frac{RhythmicPower-Background}{Background}$. Note that the empirical SNR is used to investigate the potential influence of SNR on detection performance, as it may influence the power threshold and hence which parts of the signal are eligible to be defined as rhythmic. In contrast, while the effective rhythmic SNR offers a less arrhythmically biased measure (as arrhythmic periods do not contribute to the 'signal'; see Figure 1A, Figure 4, Supplementary Figure 2), it is already influenced by the detection procedure. Hence, the empirical SNR may constrain, whereas effective empirical SNR partially results from detection efficacy. Due to this definition, empirical SNR constitutes a lower bound on effective rhythmic SNR.

2.9 Calculation of phase-based lagged coherence

To investigate the convergence between the power-based duration estimate (abundance) and a phase-based alternative, we calculated lagged coherence at linearly scaled frequencies in the range of 1 to 40 Hz for each resting-state condition. Lagged coherence assesses the consistency of phase clustering at a single sensor for a chosen cycle lag (see Fransen et al., 2015 for formulas). Instantaneous power and phase were estimated via 3-cycle wavelets. Data were segmented to be comparable to the output of rhythm detection (e.g. same removal of signal shoulders as described above). In reference to the duration threshold for power-based rhythmicity, we calculated the averaged lagged coherence using two adjacent epochs à three cycles. We computed an index of alpha rhythmicity by averaging values across epochs and posterior-occipital channels, finally extracting the value at the maximum lagged coherence peak in the 8 to 15 Hz range.

2.10 Dynamics of rhythmic probability and rhythmic power during task performance

To investigate the detection properties in the task data, we analysed the temporal dynamics of rhythmic probability and power in the alpha band. We created time-frequency representations as described above (section 2.5) and extracted the power time series at the frequency with maximal power in the range of 8-15 Hz (alpha), separately for each person, condition, channel and trial. Subsequently, we z-scored the power time series to accentuate signal dynamics and attenuate between-subject SNR differences. These time series were averaged within subject to create individual averages of rhythm dynamics. These were in turn averaged across subjects to derive Grand Average rhythm dynamics. At the single-trial level, values were allocated to rhythmic vs. arrhythmic time points according to whether a rhythmic episode with mean frequency in the respective range was indicated by eBOSC (Figure 2B; Figure 3C).

2.11 Rhythmic frequency variability during rest

As a characteristic of rhythmicity, we assessed the stability of IAF estimates across trials of the task. Trial-wise rhythmic IAF variability (Figure 8A) was calculated as the standard deviation between trials of within-trial frequency averages of designated rhythmic episodes. Whole-trial IAF variability (Figure 8B) was calculated as the standard deviation of the mean frequency across the whole trial (i.e. encompassing segments both designated as rhythmic and arrhythmic). Equivalent calculations were made using the different amplitude and abundance simulations (see Supplementary Methods 3).

2.12 Post-hoc characterization of sustained rhythms vs. transients

Instead of exclusively relying on an *a priori* duration threshold as done in standard BOSC, eBOSC also allows for a post-hoc separation of rhythms and transients based on the duration of frequency episodes that exceed the power threshold. This is afforded by our extended post-processing that results in a more specific identification of the duration and frequency of the detected rhythmic episodes (see Supplementary Methods 1 & 2). For this analysis (Section 3.5, Figure 9), we set the *a priori* duration threshold to zero and separated the resulting episodes post-hoc based on the amount of cycles that they encompassed at their mean frequency.

3. Results

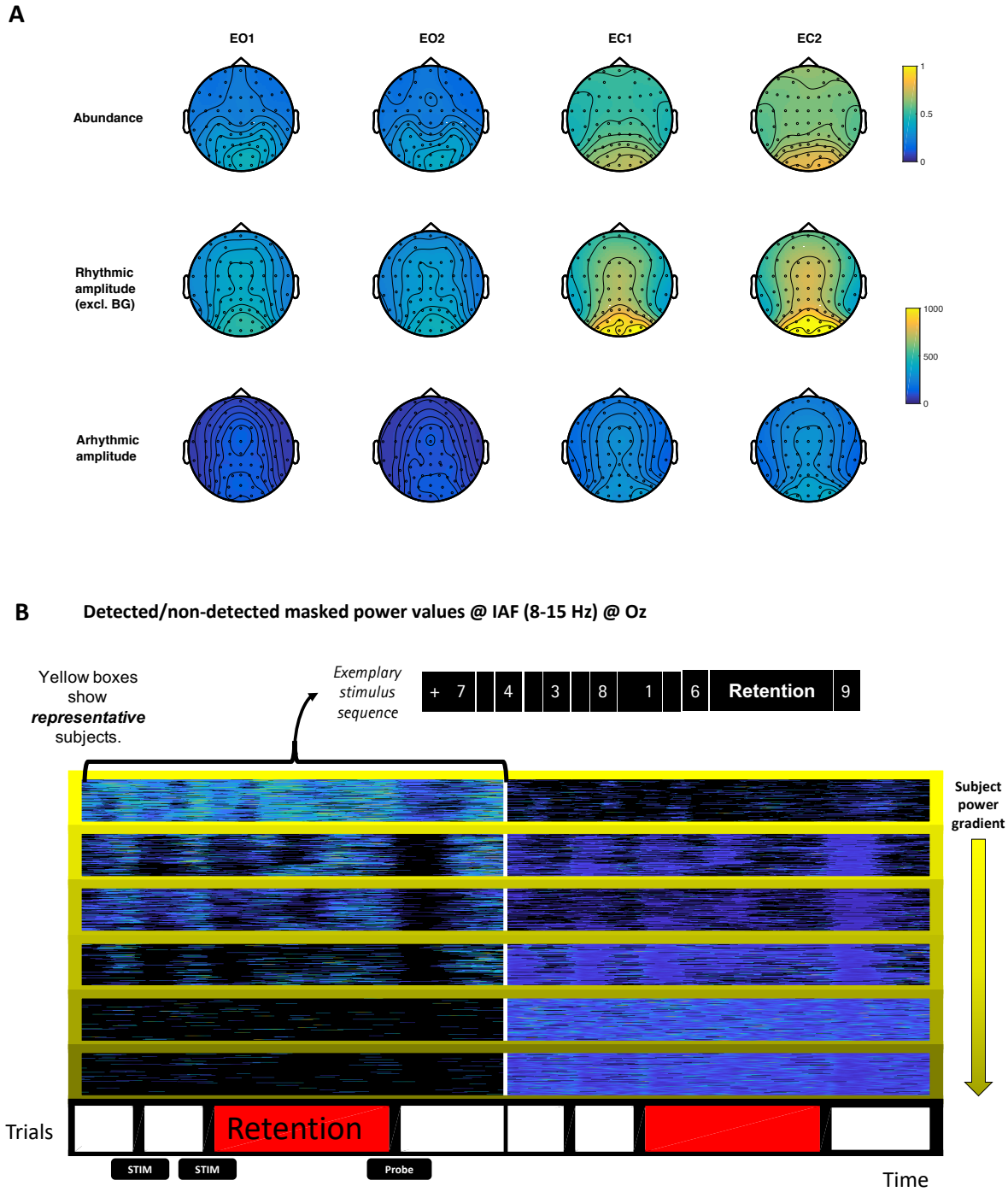


Figure 2: Extended BOSC successfully identifies single-trial alpha rhythms during rest (**A**) and task (**B**) states. (**A**) Individual topographies of abundance (top), and amplitude across time points designated as rhythmic (center) or arrhythmic (bottom) in the alpha range (8-15 Hz; bottom row), split by resting condition (rows): EO1: continuous eyes open; EO2: interleaved eyes open; EC1: continuous eyes closed; EO2: interleaved eyes closed. eBOSC identifies high occipital alpha power and abundance. (**B**) Task-related alpha dynamics are captured by eBOSC

to the degree to which they are visible in single trials. Each box displays individual trial-wise z-standardized alpha power at the individual alpha frequency, separately for rhythmic (left) and non-rhythmic (right) time points. The y-axis displays single trials across load levels, the timing on the x-axis covers the last two stimulus & encoding periods, the retention interval as well as the probe presentation and response interval of the adapted Sternberg task. The subplots' frame colour indicates the subjects' raw power maximum (i.e. the data scaling). Data are from channel Oz during the first session across load conditions.

3.1 eBOSC detects single-trial alpha rhythms during rest and task states

Individual power spectra show clear rhythmic peaks in the alpha range for every participant during both the eyes closed and the eyes open resting state (Supplementary Figure 5), indicating the general presence of alpha rhythms. eBOSC achieves high specificity for alpha rhythms in simulated (see Supplementary Results 1) and in empirical data from rest and task periods. In line with a putative source in visual cortex, rhythmic time-points exhibit increased alpha power compared with time points that were identified as arrhythmic especially at occipito-parietal channels (Figure 2A).

Task EEG recordings were characterized by stereotypic, design-locked, alpha power dynamics at encoding, retention and probe presentation at average and single-trial levels (Figure 2B, 3C). Those dynamics were detected at satisfying levels in single trials when alpha power was visibly elevated from the background (i.e. when the signal-to-noise ratio was high; upper plots in Figure 2B; note that plots are sorted by descending power (see frame colour) of the depicted subjects). However, strong inter-individual differences were apparent in the expression of single-trial alpha power dynamics and the associated detection results, with little detected rhythmicity when individual alpha power was low, and hence less rhythmic signal was available to detect in single trials (bottom plots in Figure 2B). Notably, individual abundance estimates were stable across multiple sessions, suggesting that they are indicative of trait-like characteristics (Figure 3A). Note that it is unlikely that differences in detection are due to misfit of the background spectrum. Simulations suggest that compared to the linear background fit that is implemented in standard BOSC, the robust fit with alpha peak removal successfully removes the bias of rhythmic alpha power on background estimates (Supplementary Figure 4B), while individual fits look reasonable (Supplementary Figure 5).

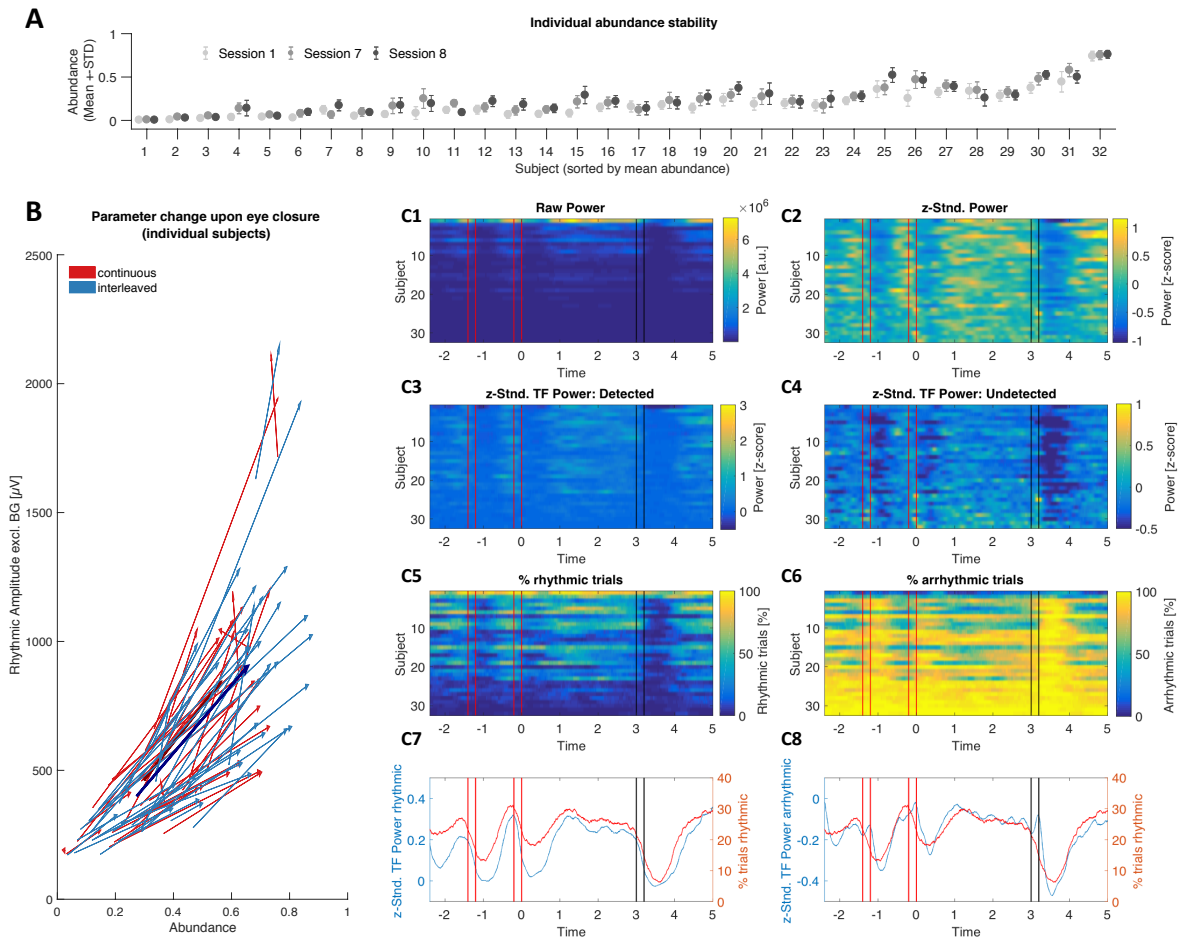


Figure 3: Characterization of detected rhythms during rest and task states. (**A**) Individual abundance estimates are stable across sessions. Data were averaged across posterior-occipital channels and high (i.e. 6) item load trials. (**B**) Detected rhythmic amplitude and abundance are co-modulated by eye closure within almost all subjects. Arrows indicate the direction and magnitude of parameter change upon eye closure for each subject. Red arrows indicate data during continuous eyes closed/eyes open intervals, blue arrows represent data from the interleaved acquisition. Thick arrows indicate the condition mean across subjects. (**C**) Individual alpha power and rhythm presence at electrode Oz under high memory load (6 items). (C7) Rhythmic power (blue) and rhythm probability (red) exhibit stereotypic temporal dynamics during encoding (red bars), retention (0 to 3 s) and retrieval (black bars). While rhythmic and arrhythmic power show similar temporal dynamics, the latter is strongly reduced in power (C4 vs. C3; see scales in C7 and C8). The arrhythmic power dynamics are characterized by additional transient increases following stimulus presentations. Data are from the first session. Power was extracted at the individual alpha peak frequency.

Supplementing the detection results, alpha power and abundance exhibited state modulations during both rest and task. As one of the earliest findings in cognitive electrophysiology (Berger, 1938), alpha amplitudes increase in magnitude upon eye closure. Here, eye closure was reflected by a joint shift towards higher amplitudes and abundances for almost all participants (Figure 3B). During the Sternberg task, the rhythmic probability closely tracks rhythmic amplitude dynamics (Figure 3C), with a notable exception. During stimulus onsets, time points designated as 'arrhythmic' revealed transient power increases relative to standardized rhythmic power (Figure 3C8 vs. 3C7), suggesting a transient engagement of encoding processes. Given the thresholds of the eBOSC algorithm, short burst-like alpha increases would be considered arrhythmic if their duration was smaller than the duration threshold and an increase in transient events is indeed observed without an *a priori* duration threshold (see section 3.5). In sum, these results suggest the suitability of eBOSC for specific and fine-grained detection of rhythmic neural characteristics, both at the group and individual level.

3.2 Presence of arrhythmic signal systematically biases overall power estimates

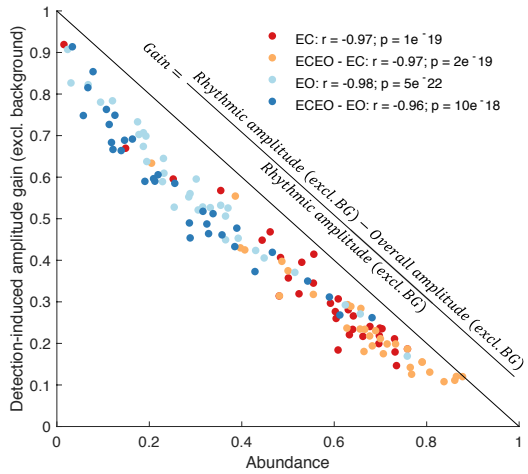


Figure 4: The duration of arrhythmic periods biases power estimates proportionally. The relative rhythm-specific amplitude gain (see schematic in Figure 1A) decreases with the duration of rhythmicity in the investigated period. (A) Amplitude gains by episode detection scale inversely linear to rhythmic abundance. Dots represent individual condition averages during the resting state.

A central prediction of the present work is that 'overall' amplitudes that were estimated across periods with mixed rhythmic and arrhythmic episodes underestimate rhythmic power due to violations of the Fourier theorem's stationarity assumption. Hence, the extent of amplitude underestimation should scale with the presence of arrhythmic signal. Stated differently, if most of the signal contains rhythmicity, the difference between overall amplitude estimates and estimates derived from rhythmic episodes should be minimal. Conversely, if large parts of the signal are arrhythmic, rhythm-specific amplitude estimates should strongly increase from their unspecific counterparts (see Figure 1A for a schematic, Supplementary Figure 2 for simulations). In line with these expectations, we observe an inverse relationship between the amplitude gain arising from focusing on rhythmic alpha episodes vs. 'overall' whole-trial-averages and the abundance of alpha rhythms, suggesting a successful exclusion of low-amplitude arrhythmic signal components (Figure 4). This relationship was highly linear, with constant relative increases in rhythmic power due to specificity increases.

3.3 Sufficient rhythmic SNR is a prerequisite for unbiased duration estimates

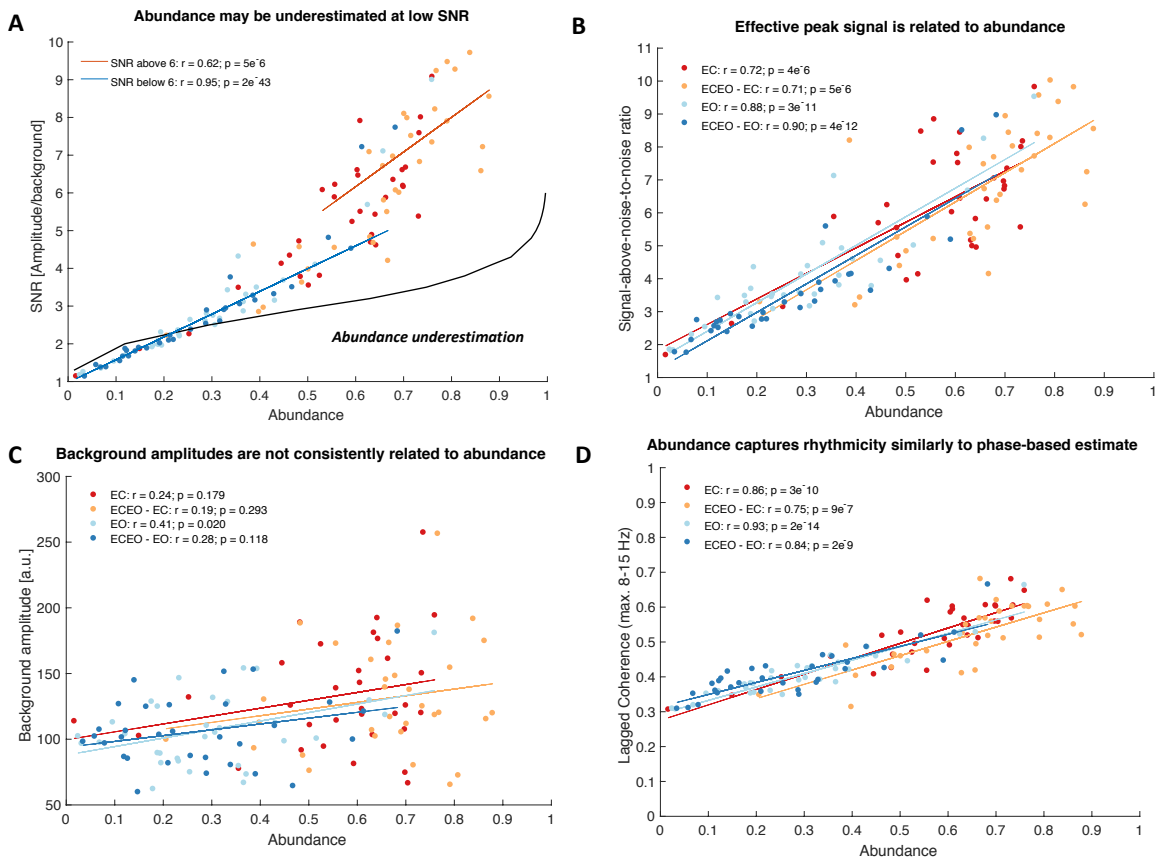


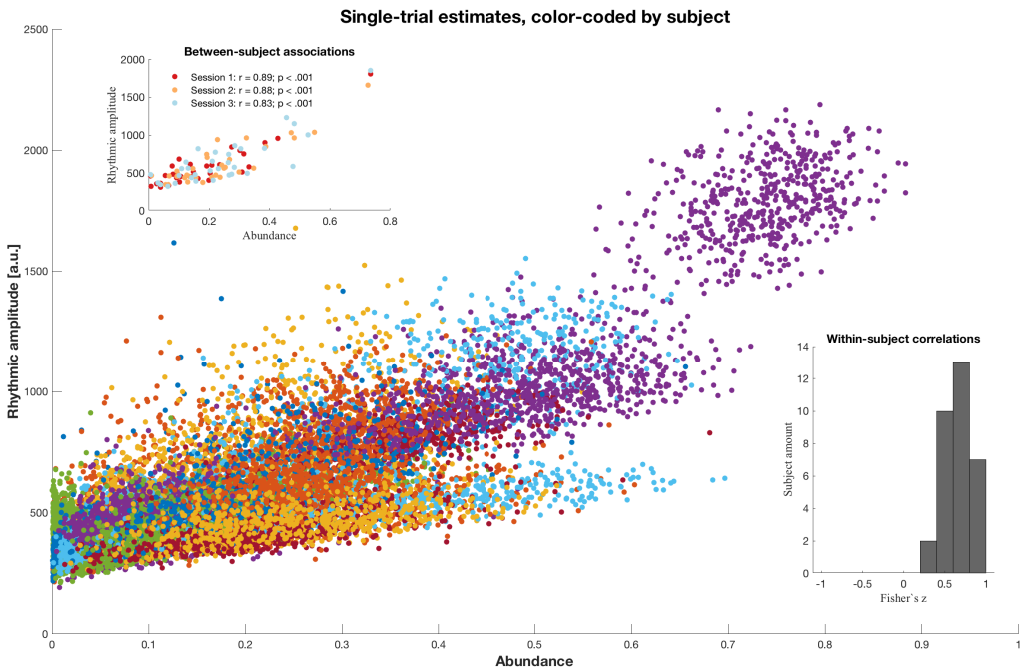
Figure 5: Abundance characterizes a recording's effective rhythmicity. (**A**) Individual abundance estimates are strongly related to the empirical signal-to-noise ratio (SNR) of the alpha peak. This relationship is also observed when only considering individual data within the SNR range for which simulation analyses indicated an unbiased abundance estimation. (**B**) The effective rhythmic signal can be conceptualized as the background-normalized rhythmic amplitude above the background estimate. This proxy for signal clarity is inter-individually linked to abundance estimates. (**C**) Background estimates are not consistently related to abundance. This implies that the relationship between amplitude and abundance is mainly driven by the signal, but not background amplitude (i.e. the effective signal 'clarity') and that associations do not arise from a misfit of the background. (**D**) Rhythmicity estimates translate cohere between power- and phase-based definition of rhythmicity. Lagged coherence data represent the individual peak value of the average lagged coherence spectrum in the 8-15 Hz range. This indicates that the BOSC-detected rhythmic spectral peak above the 1/f spectrum contains the rhythmic information that is captured by phase-based duration estimates. All data are from the resting state.

While rhythm detection with eBOSC successfully separates rhythmic from arrhythmic episodes, we also observed strong inter-individual variability in the abundance of detected rhythmicity. This may imply actual differences in the duration of rhythmic engagement. However, we also observed a severe underestimation of abundance as a function of rhythmic SNR in simulations (Supplementary Figure 1). Notably, the empirical data suggested a dependence of the abundance estimates on the empirical SNR (i.e., the IAF amplitude divided by the noise background), approximating a deterministic form in the range where simulations indicate an underestimation of abundance (blue line in Figure 5A. The black line indicates simulation-based estimates for stationary alpha rhythms at different empirical SNR levels; see Supplementary Methods 3). However, this association also persisted in ranges where simulations suggest sufficient SNR for unbiased abundance estimates (orange line in Figure 5A). Together these observations implicate sufficient SNR as a prerequisite for unbiased duration estimates.

As eBOSC defines rhythmicity as the spectral peak exceeding the background (i.e. the *signal* in SNR), the rhythmic peak itself may serve as an even better predictor of abundance. This derives from the power threshold that constrains eligible rhythmic periods (and thus the inferred rhythmic abundance) to those time points when power falls within the spectral peak. In line with this consideration, the 'effective peak' signal exhibits a strong linear relationship

to abundance (Figure 5B). Importantly, the background estimate is not consistently related to the indicated abundance (Figure 5C), emphasizing that it is the 'signal' and not the 'noise' component of SNR that constrains detection.

It remains unclear however whether this underestimation is exclusive to such a 'power thresholding'-approach or a more general constraint for single-trial rhythm characterization. To probe this possibility, we calculated a phase-based measure of rhythmicity, termed 'lagged coherence' (Fransen et al., 2015), which assesses the stability of phase clustering at a single sensor for a chosen cycle lag. Here, 3 cycles were chosen to improve comparability with the eBOSC duration threshold. Crucially, this definition of rhythmicity led to highly concordant estimates with eBOSC's abundance measure[1] (Figure 5D), suggesting that power-based rhythm detection above the scale-free background overlaps to a large extent with the rhythmic information captured in the phase-based lagged-coherence measures and that duration estimation is generally more challenging at low empirical SNR.



[1] The eBOSC duration measure was further strongly correlated with the traditional Pepisode measure (estimated at the IAF) that results from the standard BOSC algorithm (EC: r = .96, p = 2e^-18; EC2: r=.94, p = 2e^-15; EO: r = .97, p = 3e^-20; EO2 = .97, p = 2e^-20), suggesting that both measures are similarly sensitive in our empirical data and reflect to a large extent overlapping information.

Figure 6: Rhythmic amplitude and abundance are positively associated between and within subjects. Amplitude-abundance association within subjects in the Sternberg task (all sessions and trials). Dots represent single trial estimates, color-coded by subject. (Inlay 1) Amplitude-abundance association between subjects. Dots represent condition averages for each subject. (Inlay 2) Histogram of within-subject Fisher's z-coefficients of amplitude-abundance associations. Relationships are exclusively positive.

The constraints of alpha power on duration estimates also offers an explanation for the previously observed association between amplitude and abundance of rhythmic alpha episodes (Caplan et al., 2015; Whitten et al., 2011). In the present data, we observed between- and within-subject associations of alpha amplitude and abundance during both rest and task states (Figure 6). These relationships were also present in simulations, which suggests that inter-individual rhythmic power differences can account for the between-person association. The positive within-person coupling, on the other hand, may equally stem from either varying rhythmic abundance or power (Supplementary Figure 4A). Figure 7 schematically shows how amplitude-abundance coupling may be reflected in single-trials as a function of single-trial rhythmicity (rhythmic SNR).

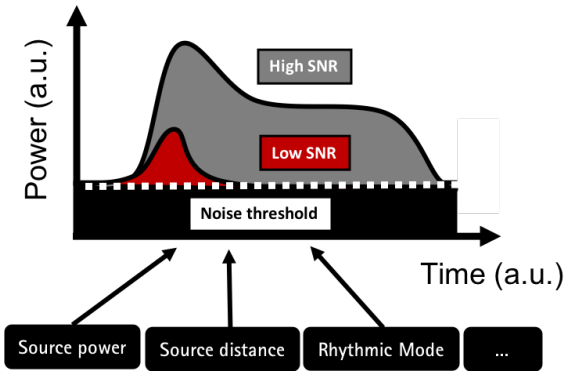


Figure 7: Schematic of the potential interdependence of single-trial rhythmic power and abundance. Low SNR may cause the detection of shorter peaks with constrained amplitude ranges, whereas prolonged periods may exceed the threshold when the rhythmic signal is clearly separated from the background.

Finally, we investigated how the differences in single-trial rhythmicity affect the estimation of another characteristic, namely the individual alpha frequency (IAF) that generally shows high temporal stability (i.e. trait-qualities) within person (Grandy et al., 2013). We

observed a strong negative association between the estimated rhythmicity and fluctuations in the rhythmic IAF between trials (Figure 8A). That is, for subjects with pervasive alpha rhythms, IAF estimates were reliably stable across trials, whereas frequency variations were induced when rhythmicity was low. Notably a qualitatively and quantitatively similar association was observed in simulations. As lower abundance implies a smaller number of samples from which the IAF is estimated, this effect could amount to a sampling confound. However, we observed a similar link between empirical alpha SNR and IAF variability when the latter was estimated across all timepoints in a trial (Figure 8B). Again, simulations with stationary 10 Hz rhythms gave rise to similar results, suggesting that estimated frequency fluctuations can arise (at least in part) from low SNR and the associated absence of clear single-trial rhythmicity. Hence, even when the IAF is intra-individually stable, its moment-to-moment estimation may induce variability when the rhythms are not clearly present.

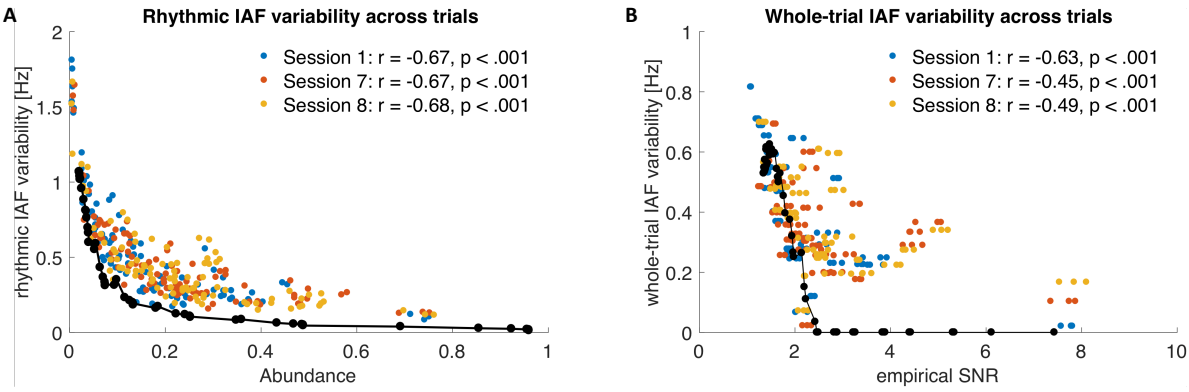


Figure 8: Trial-by-trial IAF variability is associated with sparse single-trial rhythmicity. (**A**) Individual alpha frequency (IAF) precision across trials is related to abundance. Lower individual abundance estimates are associated with increased across-trial IAF variability. (**B**) This relationship also exists when considering overall alpha SNR and IAF estimates from across the whole trial. Superimposed black lines show the 6th order polynomial fit for simulation results encompassing varying rhythm durations and amplitudes. Empirically estimated frequency variability is quantitatively similar to the bias observed at low SNRs in the simulated data.

Combined, these results suggest that the efficacy of an accurate single-trial characterization of neural rhythms relies on sufficient individual rhythmicity and not only constrains the validity of power and duration estimates, but broadly affects a range of rhythm characteristics that can be inferred at the single-trial level.

3.4 Characterizing sustained rhythms and transients

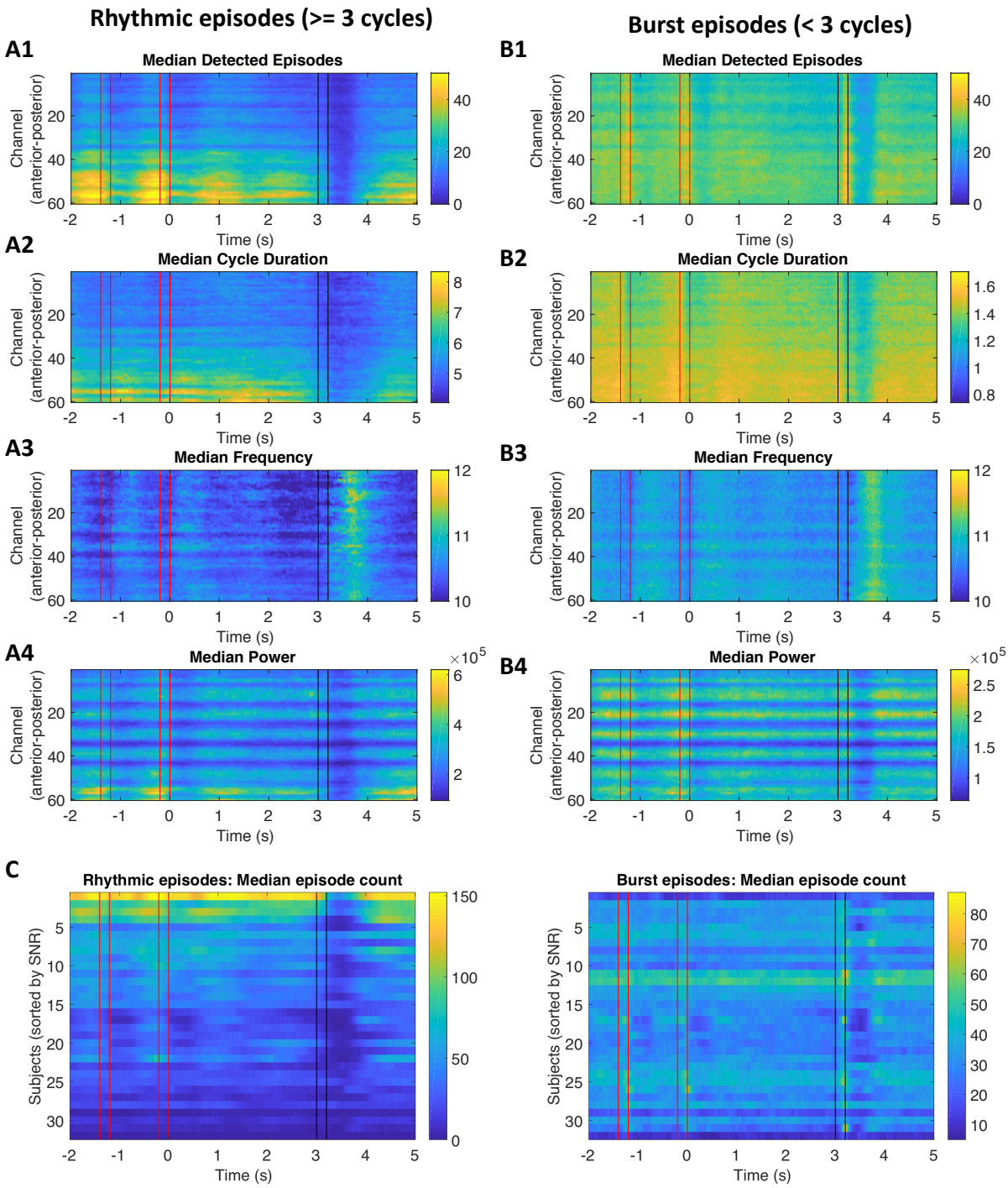


Figure 9: eBOSC provides a varied characterization of duration-specific frequency content, hence separating sustained rhythmicity (**A**) from transients (**B**). Here, episodes with mean frequency between 8 and 15 Hz were post-hoc sorted by falling below or above a post-hoc 3-cycle duration threshold. For each index, estimates were averaged across all episodes at any

time point, followed by averaging across sessions and subjects. Note that all indices are based on episodes that fulfil the power threshold for rhythmicity. There are notable differences (e.g. an increased prevalence of bursts upon stimulus onset: B1 vs. A1). Furthermore, we observe frequency increases during the response period, which may relate to motor suppression. (**C**) Whereas SNR posed a major constraint on the identification of sustained rhythmicity, it did not constrain the number of detected transients, suggesting separable sources.

From the joint investigation of detection performance in simulated and empirical data, it follows that low rhythmic SNR conditions (approx. below an SNR of 6) constitute a severe challenge for single trial rhythm detection and characterization. However, given sufficiently rhythmic signals, eBOSC nonetheless provides a specific identification of frequency-specific content based on its extensive post-processing (see Supplementary Methods 1 & 2). This affords a characterization of episodes that adhere to the power assumptions of rhythmicity in the absence of an *a priori* duration threshold, and enables a post-hoc investigation of sustained and transient rhythmic episodes. Using the traditional 3-cycle threshold as a post-hoc criterion, we observe differences in the temporal prevalence of bursts and sustained rhythms, with a larger number of transient events following stimulus onsets, in line with the observations made for rhythmic vs. arrhythmic power (Figure 3C8; section 3.1). In addition, these episodes provide further characteristics of the neural dynamics, such as the average cycle duration (Figure 9A2, Figure 9B2) and the event-specific frequency (Figure 9A3, Figure 9B3). The latter exhibits transient increases around the response period, likely related to motor inhibition. Notably, while empirical SNR constrains the detection of sustained rhythmicity (e.g. Figure 5A), the same was not observed for the number of transient episodes (Figure 9C), thereby suggesting differential origins of these signal contributions.

4. Discussion

In the present manuscript, we explored the feasibility of characterizing neural rhythms at the level of single trials. To achieve this goal, we extended a previously published rhythm detection method, BOSC (Whitten et al., 2011). Based on simulations we demonstrate that our extended BOSC (eBOSC) algorithm performs well and increases detection specificity. Crucially, the reliance on robust regression in conjunction with removal of the rhythmic power band effectively decoupled estimation of the noise background from the rhythmic signal component (as reflected in the divergent associations with rhythmicity estimates). In real data,

we can successfully separate rhythmic and arrhythmic, sometimes burst-like components. In total, single-trial characterization of neural rhythms appears promising for improving a mechanistic understanding of rhythmic processing modes during rest and task.

However, the simulations also reveal challenges for accurate rhythm characterization in that the abundance estimates clearly depend on a participant's empirical SNR. The comparison to a phase-based rhythm detection further suggests that this a general limitation independent of the chosen detection algorithm. Below, we will discuss potentials and challenges of single-trial rhythm detection in more detail.

4.1 The utility and potential of rhythm detection

Single-trial analyses are rapidly gaining importance (Jones, 2016; Stokes & Spaak, 2016), in part due to a debate regarding sustained vs. transient nature of neural rhythms that cannot be resolved at the level of data averages (Jones, 2016; van Ede et al., 2018). In short, due to the non-negative nature of power estimates, time-varying transient power increases may be represented as sustained power upon averaging. Importantly, sustained and transient events may differ in their neurobiological origin (Sherman et al., 2016), indicating high theoretical relevance for the differentiation of sustained rhythms from bursts. Moreover, many analysis procedures, such as phase-based functional connectivity, assume that estimates are directly linked to the presence of rhythmicity, therefore leading to interpretational difficulties when it is unclear whether this condition is met (Aru et al., 2015; Muthukumaraswamy & Singh, 2011). Clear identification of rhythmic time periods based on single-trial data is necessary to resolve these issues. In the current study, we extended a state-of-the-art rhythm detection algorithm, and systematically investigated its ability to characterize neural alpha rhythms at the single-trial level in scalp EEG recordings.

Across simulations and empirical data, we observed that eBOSC detects rhythms with high specificity, albeit with slightly impaired sensitivity at low SNRs compared to standard BOSC (Supplementary Results 1). However, we consider its higher specificity to be of particular importance with regard to deriving rhythm-specific estimates. Furthermore, due to the highly concordant estimates of individual rhythmicity between power- and phase-based methods (Figure 5C), eBOSC's sensitivity decreases are unlikely to be a major constraint in the current analyses. In contrast, eBOSC offers multiple advantages compared to extant tools. First, by excluding the rhythmic peak prior to fitting the arrhythmic background, eBOSC removes a potential bias of rhythmic power on the background fit (Supplementary Figure 4B), thereby

effectively uncoupling the estimated background amplitude from the indicated rhythmicity (Figure 5C).[2] Second, the post-processing of detected segments affords a more specific characterization of neural rhythms compared to standard BOSC (Supplementary Figures 1 & 2). In particular, accounting for the temporal extension of the wavelet increases the temporal specificity of indicated rhythms. In line with this, our simulations indicate a better adherence to an *a priori* duration threshold by reducing detection of transients, which are not excluded by standard BOSC (Supplementary Figure 2). This further allows for a more specific post-hoc characterization of episode duration without the requirement of an *a priori* duration threshold. eBOSC thereby provides a fine-grained temporal indexing of rhythms and bursts (Figure 9) that can be used to improve the characterization of neural dynamics. Such characterization includes multiple parameters, such as the frequency of rhythmic episodes, their duration, their amplitude and other indices not considered here (e.g. instantaneous phase, appearance in the time domain). Crucially, a focus on rhythmic periods can resolve biases that arise when arrhythmic periods are included in the segments of interest. For example, refining rhythmic amplitudes based on knowledge about the temporal occurrence of rhythms effectively removes the adverse impact of arrhythmic time points. In line with our hypotheses, simulations (Supplementary Figure 2) and empirical data (Figure 4) indicate that arrhythmic episodes in the analysed segment bias rhythmic power estimates relative to the extent of their duration. Conversely, a focus on rhythmic periods induces the most pronounced amplitude gains when rhythmic periods are sparse.

Moreover, by allowing a post-hoc duration classification of high-power episodes, eBOSC can disentangle transient and sustained events in a principled manner. This may provide new insights and alternative interpretations about biophysical signal generators (Sherman et al., 2016). Here, we observed an increased number of detected alpha power transients following stimulus onsets, and more sustained rhythms when no stimulus was presented. In line with these observations, (Peterson & Voytek, 2017) recently proposed alpha bursts to increase visual gain during stimulus onsets and contrasted this role with decreased cortical processing during sustained alpha rhythms. Our data supports such a distinction between sustained and transient events, although the Peterson & Voytek (2017) interpretation remains to be investigated. Note that the reported duration of 'burst' events in the literature is still diverse (Peterson & Voytek, 2017) exceeding the 3-cycle threshold used here. In contrast to eBOSC however, previous work

[2] While a positive association may indicate an overestimation of the background amplitude in the presence of the rhythmic peak, it does not provide sufficient evidence for fitting problems as the two parameters may be naturally correlated.

has not accounted for the impact of wavelet duration. It is thus conceivable that power transients that were previously characterized as 3 cycles or longer are actually shorter after correcting for the impact of wavelet convolution, as is done in the current eBOSC implementation. Due to its flexible post-hoc classification and temporally precise indication of rhythmic periods, eBOSC provides a tool to further probe the duration of neural dynamics. In total, single-trial characterization of neural rhythms appears promising for improving a mechanistic understanding of rhythmic processing modes during rest and task.

4.2 Single-trial detection of rhythms: rhythmic SNR as a central challenge

The empirical analyses revealed gradual and stable inter-individual differences of the estimated duration of rhythmicity. Crucially however, abundance was related to rhythmic SNR, most strongly in ranges where simulations indicated an underestimation of rhythmic duration (see Figure 5A black line). It is therefore uncertain whether low rhythmic duration estimates for these subjects indicate temporally constrained rhythms or underestimation due to low rhythmicity, as was true in the simulations.

Multiple results suggest that duration underestimation at low SNRs does not arise from idiosyncrasies of our algorithm. Notably, the duration measure obtained with the eBOSC pipeline was strongly correlated to standard BOSC's Pepisode measure (Whitten et al., 2011) as well as the phase-based lagged coherence index (Fransen et al., 2015), thus showing high convergence with different state-of-the-art techniques (Figure 5D). Furthermore, detection performance was visually satisfying given observable task-locked rhythm dynamics (Figure 2B) and the characterization of arrhythmic time points as low-amplitude power (Figure 2A). Furthermore, the observed relationship between amplitude gain and abundance suggests a successful exclusion of (low-power) arrhythmic episodes at the individual level (Figure 4). These observations indicate that eBOSC rhythm detection performed well, but that low rhythmic SNR conditions present a fundamental challenge to single-trial characterization, suggesting that resulting indices have to be interpreted carefully when rhythmicity is low (see section 4.4).

It is crucial to highlight the single-trial perspective, as the signal properties of one-shot signals may vary from those at the average level. A common argument for the utility of trial averages is an improved clarity of a stationary signal due to the removal of superimposed, randomly distributed noise. Here, most subjects exhibited similar mean alpha power dynamics (Figure 3C2) despite large differences in single-trial rhythmicity (Figure 3A). However,

identical average dynamics may arise both from sustained rhythmic periods and transient bursts at the target frequency in single trials (Jones, 2016; Sherman et al., 2016). Given similar average dynamics, the observed inter-individual differences in abundance may thus be explained by two scenarios: (1) Subjects have engaged a similar rhythmic process, but the representation of this process in single trials is not sufficient (e.g. due to differences in external noise levels). In this case, low abundance estimates indicate an underestimation of rhythmic duration. (2) The current sample is heterogeneous with regard to the expression of extended rhythmicity vs. alpha bursts. In this case, abundance estimates may vary inter-individually as a function of sustained vs. rhythmic engagement.

While the current results do not provide a definitive answer, we suggest that the first interpretation might be more likely due to the following reasons: (a) low-abundance subjects were in an SNR range in which simulations of rhythmicity indicated a similar underestimation of abundance (Figure 5A); (b) subjects with decreased rhythmicity did not show increased amounts of bursts (Figure 9C) during periods where mean dynamics indicated power increases, and; (c) increased bursts were event-locked to visual onsets for both high- and low-rhythmic subjects, with no systematic differences in burst probability as a function of rhythmicity (Figure 9C). Taken together, we tentatively infer that subjects with low rhythmicity may share similar rhythm dynamics, but that these are not sufficiently represented in their continuous data and can only be recovered upon averaging. Importantly, the stability at follow-ups a week later (Figure 3A) suggests an intrinsic determinant of the recordings' rhythmicity (compared to biases induced by e.g., the processing parameters or technical setup).

The present data are agnostic as to whether stable inter-individual differences in rhythmic power arise from variable characteristics of the originating neural processes (e.g. the magnitude of neural synchronization) or factors related to sensitivity differences at the sensor level (e.g. distance between source and sensor; head shape; skull thickness). One speculative factor of the latter category is the exact anatomical source of the rhythms and its distance to the sensor. Unlike MEG, EEG is generally sensitive to signals originating from both tangential and radial sources (hence showing little dependency of sensitivity on dipole orientation; (Ahlfors, Han, Belliveau, & Hämäläinen, 2010)), and spatial differences in the source location may induce strong differences in SNR (Goldenholz et al., 2009) with decreased sensitivity for signals from deeper sources. In contrast, it is also likely that differences in alpha power reflect real differences in the extent of neural synchronization of a similarly positioned source, in line with a large literature on between-person differences in alpha power that relate to cognitive

outcomes. In this case, many subjects' rhythmic synchronization may be too faint compared to the scalp sensors' sensitivity to allow for accurate single-trial depictions.

4.4. General implications for single-trial rhythm indices at low SNR

Low rhythmic SNR (whether due to a low propensity to synchronize or measurement noise) does not exclusively challenge single-trial detection, but more generally affects single-trial rhythm characterization. This may at first appear trivial when considering that a suitable process has to be present and recognized to be further characterized. It is however important to stress that stable inter-individual differences in rhythmic power and SNR can affect a variety of indices, resulting in reliable, yet potentially invalid conclusions when power differences are not considered. This concerns all 'meta'-indices (like phase, frequency, duration) whose estimation accuracy relies on apparent rhythmicity.

Our simulations suggest that as the rhythmic SNR decreases, estimates of the duration (Figure 5A) and frequency stationarity (Figure 8) increasingly deviate from the simulated parameters. Changes in instantaneous alpha frequency as a function of cognitive demands have been theorized and reported in the literature (Haegens, Cousijn, Wallis, Harrison, & Nobre, 2014; Herrmann, Murray, Ionta, Hutt, & Lefebvre, 2016; Mierau, Klimesch, & Lefebvre, 2017; Samaha & Postle, 2015; Wutz, Melcher, & Samaha, 2018), with varying degrees of control for power differences between conditions and individuals. Our empirical analyses suggest an increased trial-by-trial variability of individual alpha frequency estimates as alpha SNR decreases (Figure 8). Meanwhile, simulations suggest that such increased variance - both estimated within indicated rhythmic periods and across whole trials – may represent an effect of low SNR. While our results do not negate the possibility of real frequency changes of the alpha rhythm under task load, they emphasize the importance of controlling for the presence of rhythms, mirroring considerations for the interpretation of phase estimates (Muthukumaraswamy & Singh, 2011) and amplitudes (section 4.1).

Furthermore, we observed a strong correlation between the strength and duration of alpha rhythms both between and within subjects, in line with previous reports of high overlap between these indices (Caplan et al., 2015; Fransen et al., 2015). Whereas such coupling takes on a deterministic form in a low SNR range, it is notable that a positive coupling between amplitude and abundance was also observed between subjects with high alpha power. Furthermore, we exclusively observed positive within-subject relationships at the between-trial level (Figure 6). Accordingly, simulations suggest that positive relationships may arise even at

high levels of SNR in the presence of fluctuations in the strength and duration of rhythms (Supplementary Figure 4B), although no association is simulated given the stationarity of one parameter. Hence, these observed associations are likely artificial, rather than intrinsic couplings between rhythmic strength and duration. In cases where e.g. abundance is only used as a sensitive and/or specific replacement for power (Caplan et al., 2015; Fransen et al., 2015) such coupling may not be problematic, but care has to be taken in interpreting available duration indices as power-independent characteristics of rhythmic episodes.

A proper duration index becomes increasingly important however to assess whether rhythms are stationary or burst-like. For this purpose, both amplitude thresholding and phase-progression criteria have been proposed (Cole & Voytek, 2018; Peterson & Voytek, 2017; Sherman et al., 2016; van Ede et al., 2018; Vidaurre, Myers, Stokes, Nobre, & Woolrich, 2018). Here, we show that both methods arrive at similar conclusions regarding individual rhythmic duration and that the above challenges are therefore applicable to both approaches. As an alternative to threshold-based methods, Van Ede et al. (2018) propose methods based on e.g. Hidden Markov Models (Vidaurre et al., 2018; 2016) for the estimation of rhythmic duration. These approaches are interesting; the definition of states to be inferred in single trials is based on individual (or group) averages, while the multivariate nature of the signals across channels is also taken into account. It is a viable question whether such approaches can characterize rhythmicity in scenarios where the present methods fail due to insufficient rhythmicity.

Likewise, single-trial properties are gaining relevance in decoding analyses that traditionally operate with few if any trial averages. Depending on whether the relevant feature vectors include neural rhythms, differences in rhythmicity may therefore also affect decoding feasibility. Recently, large inter-individual differences in decoding performance have been observed (Westner, Dalal, Hanslmayr, & Staudigl, 2018), and it remains an intriguing question whether such decoding efficacy covaries with the extent of rhythmicity. By characterizing a recording's rhythmicity, eBOSC provides a tool to investigate such putative links.

4.5 Limitations of eBOSC rhythm detection

While the eBOSC approach is theoretically agnostic to the derivation of power estimates, the currently employed post-processing steps require wavelet estimates. This results from model-based assumptions about the time-frequency extension of the wavelet that are used for refining detected episodes (see Supplementary Methods 1 & 2). This is an optional step in the eBOSC toolbox however, which should allow for different power derivations to be used, as

long as the derived time series adhere to the other method assumptions (e.g. 1/f background). In addition, there is now an surge of interest in characterizing neural dynamics by their waveform shape, circumventing power analyses entirely (Cole & Voytek, 2018). While such an approach is intriguing, further work is needed to show which analysis sequence is more fruitful: (a) identifying events first and then describing the associated waveform shape (e.g. eBOSC) or (b) identifying events based on a description of their waveform shape (e.g. cycle-by-cycle analysis). As both procedures operate on the basis of single-trials, similar challenges are likely to apply to both approaches.

Finally, while the goal is to derive single-trial metrics, the power threshold is calculated based on condition-averages, as individual trial spectra do not consistently follow a clear 1/f structure in the presence of rhythmic content. In addition, a pre-removal of power in passbands with large rhythmic power peaks is necessary to remove rhythmic biases on background estimates (Supplementary Figure 4B). Recently, (see also Haller et al., 2018) have proposed a principled approach to the removal of power peaks, which may afford rhythm-unbiased background estimates without requiring priors regarding target frequencies. It may thus represent a useful pre-processing step for further applications that aim to detect rhythmicity across frequencies. For our application, we anticipate no qualitative changes compared to our alpha exclusion approach as (a) simulations and empirical data did not show an association between background and rhythmicity estimates (Figure 5C. Supplementary Figure 4B), and the signal was dominated by an alpha frequency peak, with visually satisfying individual background fits (Supplementary Figure 5).

5. Conclusion

We extended a state-of-the-art rhythm detection method and characterized alpha rhythms in simulated, resting and task data at the single trial level. By using simulations, we show that rhythm detection can be employed to derive specific estimates of rhythmicity, with fine-grained control over its definition, and to reduce the bias of rhythm duration on amplitude estimates that commonly exists in standard analysis procedures. However, we also observe striking inter-individual differences in the indicated duration of rhythmicity, which for subjects with low alpha power may be due to insufficient single-trial rhythmicity. We further show that this may lead to biased estimates, in particular underestimated duration and increased variability of rhythmic frequency. Given these constraints, we have provided examples of eBOSC's efficacy to characterize rhythms that may prove useful for investigating the origin and

functional role of neural rhythms in health and disease, and in turn, the current study works to establish the foundation for ideographic analyses of neural rhythms.

Data availability

The scripts implementing the eBOSC pipelines are available at github.com/jkosciessa/eBOSC alongside the simulation scripts that were used to assess eBOSC's detection properties.

Funding

This study was conducted within the project 'Cognitive and Neuronal Dynamics of Memory across the Lifespan (CONMEM)' at the Center for Lifespan Psychology, Max Planck Institute for Human Development (MPIB). MW-B's work was supported by grants from the German Research Foundation (DFG, WE 4269/3-1 and WE 4269/5-1) as well as an Early Career Research Fellowship 2017 – 2019 awarded by the Jacobs Foundation. The study was conducted in partial fulfillment of the doctoral dissertation of JQK.

References

Ahlfors, S. P., Han, J., Belliveau, J. W., & Hämäläinen, M. S. (2010). Sensitivity of MEG and EEG to source orientation. *Brain Topography*, *23*(3), 227–232. http://doi.org/10.1007/s10548-010-0154-x

Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., et al. (2015). Untangling cross-frequency coupling in neuroscience., *31*, 51–61. http://doi.org/10.1016/j.conb.2014.08.002

Bach, M. (1996). The Freiburg Visual Acuity test--automatic measurement of visual acuity. *Optometry & Vision Science*, *73*(1), 49–53.

Bach, M. (2007). The Freiburg Visual Acuity Test-variability unchanged by post-hoc re-analysis, *245*(7), 965–971. http://doi.org/10.1007/s00417-006-0474-4

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159.

Berger, H. (1938). Über das Elektrenkephalogramm des Menschen. *Archiv Für Psychiatrie Und Nervenkrankheiten*, *108*(3), 407–431. http://doi.org/10.1007/BF01824101

Buzsáki, G. (2006). Rhythms of the Brain. New York: Oxford University Press.

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, *13*(6), 1–14. http://doi.org/10.1038/nrn3241

Caplan, J. B., Bottomley, M., Kang, P., & Dixon, R. A. (2015). Distinguishing rhythmic from non-rhythmic brain activity during rest in healthy neurocognitive aging. *NeuroImage*, *112*, 341–352. http://doi.org/10.1016/j.neuroimage.2015.03.001

Caplan, J. B., Madsen, J. R., Raghavachari, S., & Kahana, M. J. (2001). Distinct patterns of brain oscillations underlie two basic parameters of human maze learning. *Journal of Neurophysiology*, *86*(1), 368–380.

Cohen, M. X. (2014). Analyzing neural time series data: theory and practice.

Cohen, M. X. (2017). Where Does EEG Come From and What Does It Mean? *Trends in Neurosciences*, *40*(4), 208–218. http://doi.org/10.1016/j.tins.2017.02.004

Cole, S. R., & Voytek, B. (2018). Cycle-by-cycle analysis of neural oscillations. *bioRxiv*, 302000. http://doi.org/10.1101/302000

da Silva, F. H. L. (2018). Niedermeyer's Electroencephalography. Oxford University Press.

Fransen, A. M. M., van Ede, F., & Maris, E. (2015). Identifying neuronal oscillations using rhythmicity. *NeuroImage*, *118*(C), 256–267. http://doi.org/10.1016/j.neuroimage.2015.06.003

Goldenholz, D. M., Ahlfors, S. P., H m l inen, M. S., Sharon, D., Ishitobi, M., Vaina, L. M., & Stufflebeam, S. M. (2009). Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Human Brain Mapping*, *30*(4), 1077–1086. http://doi.org/10.1002/hbm.20571

Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2013). Individual alpha peak frequency is related to latent factors of general cognitive abilities. *NeuroImage*, *79*(C), 10–18. http://doi.org/10.1016/j.neuroimage.2013.04.059

Grandy, T., Lindenberger, U., & Werkle-Bergner, M. (2017). When group means fail: Can one size fit all? *bioRxiv*. http://doi.org/10.1101/126490

Gross, J. (2014). Analytical methods and experimental approaches for electrophysiological studies of brain oscillations. *Journal of Neuroscience Methods*, *228*, 57–66. http://doi.org/10.1016/j.jneumeth.2014.03.007

Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., & Nobre, A. C. (2014). Inter- and intra-individual variability in alpha peak frequency. *NeuroImage*, *92*(C), 46–55. http://doi.org/10.1016/j.neuroimage.2014.01.049

Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., et al. (2018). Parameterizing neural power spectra. *bioRxiv*, 1–16. http://doi.org/10.1101/299859

Hansen, E. W. (2014). DFT Properties and Theorems. In *Fourier transforms. Principles and applications* (p. 128). Hoboken, New Jersey: John Wiley & Sons.

Herrmann, C. S., Murray, M. M., Ionta, S., Hutt, A., & Lefebvre, J. (2016). Shaping Intrinsic Neural Oscillations with Periodic Stimulation. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *36*(19), 5328–5337. http://doi.org/10.1523/JNEUROSCI.0236-16.2016

Jones, S. R. (2016). When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Current Opinion in Neurobiology*, *40*, 72–80. http://doi.org/10.1016/j.conb.2016.06.010

Klimesch, W. (2012). alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. http://doi.org/10.1016/j.tics.2012.10.007

Mierau, A., Klimesch, W., & Lefebvre, J. (2017). State-dependent alpha peak frequency shifts: Experimental evidence, potential mechanisms and functional implications. *Neuroscience*, *360*, 146–154. http://doi.org/10.1016/j.neuroscience.2017.07.037

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112–117. http://doi.org/10.1111/j.1467-8721.2009.01619.x

Muthukumaraswamy, S. D., & Singh, K. D. (2011). A cautionary note on the interpretation of phase-locking estimates with concurrent changes in power. *Clinical Neurophysiology*, *122*(11), 2324–2325. http://doi.org/10.1016/j.clinph.2011.04.003

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods*, *192*(1), 152–162. http://doi.org/10.1016/j.jneumeth.2010.07.015

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. http://doi.org/10.1016/0028-3932(71)90067-4

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*(1), 156869–9. http://doi.org/10.1155/2011/156869

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184–187.

Peterson, E. J., & Voytek, B. (2017). Alpha oscillations control cortical gain by modulating excitatory-inhibitory background activity. *bioRxiv*, 185074. http://doi.org/10.1101/185074

Sadaghiani, S., & Kleinschmidt, A. (2016). Brain Networks and α-Oscillations: Structural and Functional Foundations of Cognitive Control. *Trends in Cognitive Sciences*, *20*(11), 805–817. http://doi.org/10.1016/j.tics.2016.09.004

Samaha, J., & Postle, B. R. (2015). The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. *Current Biology*, *25*(22), 2985–2990. http://doi.org/10.1016/j.cub.2015.10.007

Sherman, M. A., Lee, S., Law, R., Haegens, S., Thorn, C. A., Hämäläinen, M. S., et al. (2016). Neural mechanisms of transient neocortical beta rhythms: Converging evidence

from humans, computational modeling, monkeys, and mice. *Proceedings of the National Academy of Sciences*, *113*(33), E4885–E4894. http://doi.org/10.1073/pnas.1604135113

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654.

Stokes, M., & Spaak, E. (2016). The Importance of Single-Trial Analyses in Cognitive Neuroscience. *Trends in Cognitive Sciences*, *20*(7), 483–486. http://doi.org/10.1016/j.tics.2016.05.008

van Ede, F., Quinn, A. J., Woolrich, M. W., & Nobre, A. C. (2018). Neural Oscillations: Sustained Rhythms or Transient Burst- Events? *Trends in Neurosciences*, 1–3. http://doi.org/10.1016/j.tins.2018.04.004

Vidaurre, D., Myers, N., Stokes, M., Nobre, A. C., & Woolrich, M. W. (2018). Temporally unconstrained decoding reveals consistent but time-varying stages of stimulus processing, 1–23. http://doi.org/10.1101/260943

Vidaurre, D., Quinn, A. J., Baker, A. P., Dupret, D., Tejero-Cantero, A., & Woolrich, M. W. (2016). Spectrally resolved fast transient brain states in electrophysiological data. *NeuroImage*, *126*(C), 81–95. http://doi.org/10.1016/j.neuroimage.2015.11.047

Wang, X. J. (2010). Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews*, *90*(3), 1195–1268. http://doi.org/10.1152/physrev.00035.2008

Westner, B. U., Dalal, S. S., Hanslmayr, S., & Staudigl, T. (2018). Across-subjects classification of stimulus modality from human MEG high frequency activity. *PLoS Computational Biology*, *14*(3), e1005938. http://doi.org/10.1371/journal.pcbi.1005938

Whitten, T. A., Hughes, A. M., Dickson, C. T., & Caplan, J. B. (2011). A better oscillation detection method robustly extracts EEG rhythms across brain state changes: The human alpha rhythm as a test case. *NeuroImage*, *54*(2), 860–874. http://doi.org/10.1016/j.neuroimage.2010.08.064

Wutz, A., Melcher, D., & Samaha, J. (2018). Frequency modulation of neural oscillations according to visual task demands. *Proceedings of the National Academy of Sciences*, *115*(6), 1346–1351. http://doi.org/10.1073/pnas.1713318115