Standard multiscale entropy reflects spectral power at mismatched temporal scales: What's signal irregularity got to do with it?

Julian Q. Kosciessa[123]*, Niels A. Kloosterman[12], and Douglas D. Garrett[12]*

[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; [3]Department of Psychology, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany

* Corresponding authors: kosciessa@mpib-berlin.mpg.de; garrett@mpib-berlin.mpg.de

Abstract:

The (ir)regularity of neural time series patterns as assessed via Multiscale Sample Entropy (MSE; e.g., Costa et al., 2002) has been proposed as a complementary measure to signal variance, but the con- and divergence between these measures often remains unclear in applications. Importantly, the estimation of sample entropy is referenced to the magnitude of fluctuations, leading to a trade-off between variance and entropy that questions unique entropy modulations. This problem deepens in multi-scale implementations that aim to characterize signal irregularity at distinct timescales. Here, the normalization parameter is traditionally estimated in a scale-invariant manner that is dominated by slow fluctuations. These issues question the validity of the assumption that entropy estimated at finer/coarser time scales reflects signal irregularity at those same scales. While accurate scale-wise mapping is critical for valid inference regarding signal entropy, systematic analyses have been largely absent to date. Here, we first simulate the relations between spectral power (i.e., frequency-specific signal variance) and MSE, highlighting a diffuse reflection of rhythms in entropy time scales. Second, we replicate known cross-sectional age differences in EEG data, while highlighting how timescale-specific results depend on the spectral content of the analyzed signal. In particular, we note that the presence of both low- and high-frequency dynamics leads to the reflection of power spectral density slopes in finer time scales. This association co-occurs with previously reported age differences in both measures, suggesting a common, power-based origin. Furthermore, we highlight that age differences in high frequency power can account for observed entropy differences at coarser scales via the traditional normalization procedure. By systematically assessing the impact of spectral signal content and normalization choice, our findings highlight fundamental biases in traditional MSE implementations. We make multiple recommendations for future work to validly interpret estimates of signal irregularity at time scales of interest.

Highlights

- Multiscale sample entropy (MSE) links to spectral power via an internal similarity criterion.
- Counterintuitively, traditional MSE implementations lead to slow-frequency reflections in fine-scale entropy, and high-frequency biases on coarse-scale entropy.
- Fine-scale entropy reflects power spectral density slopes, a multi-scale property.
- Narrowband sample entropy indexes (non-stationary) rhythm (ir)regularity at matching time scales.

Keywords: multiscale sample entropy; time scale bias; resting state EEG; age differences; rhythms

1 Introduction

1.1 Entropy as a measure of signal (ir)regularity

Neural times series exhibit a wealth of dynamic patterns that may be tightly linked to neural computations. While some of these patterns consist of stereotypical deflections (e.g., periodic neural rhythms; Buzsaki & Draguhn, 2004; X. J. Wang, 2010), others have a more complex appearance that may still be equally relevant for characterizing neural function (S. R. Cole & Voytek, 2017; Diaz, Bassi, Coolen, Vivaldi, & Letelier, 2018). Multiscale entropy (MSE) (Costa, Goldberger, & Peng, 2002, 2005) has been proposed as an information-theoretic metric that estimates the temporal irregularity in a signal (in theory providing information above and beyond traditional spectral metrics), while accommodating that neural dynamics occur across multiple spatiotemporal scales. In tandem, dynamic perspectives on brain function in the framework of nonlinear dynamics and complex systems have gained traction (Breakspear, 2017; Stam, 2005; Vakorin & McIntosh, 2012), suggesting that optimal computations in the brain may be characterized by metastable states that afford flexible movement between distinct attractor states. Following this conceptual framework, MSE has been increasingly applied to characterize the apparent "irregularity" (or non-linearity) of neural dynamics of different brain states, across the lifespan and in relation to health and disease (Bruce, Bruce, & Vennelaganti, 2009; Jaworska et al., 2018; McIntosh et al., 2014; Miskovic, MacDonald, Rhodes, & Cote, 2019; Sleimen-Malkoun et al., 2015; Takahashi et al., 2010; H. Wang, McIntosh, Kovacevic, Karachalios, & Protzner, 2016; Werkle-Bergner et al., 2014; Yang et al., 2013). With its novel focus on non-linear dynamics, MSE has thus become an attractive measure to gain new perspectives into brain function. However, its relation to extant, linear signal characteristics (e.g. spectral power) is considered complex in its own right (Courtiol et al., 2016; Nikulin & Brismar, 2004; Vakorin & McIntosh, 2012). Many applications highlight a joint modulation of both entropy and spectral power, although the specifics of their potential association (e.g., regarding their time scales) are not always clear. Given the apparent sensitivity of MSE in many applications, we argue that a better understanding of the relation of MSE to established linear signal characteristics such as spectral power (Buzsaki & Draguhn, 2004; Buzsaki & Mizuseki, 2014; Lopes da Silva, 2013) is critical. In particular, work on the interpretation of entropy time scales remains sparse. At best, this limits any temporally-specific interpretation of observed effects. Here, we probe two potential challenges to traditional interpretations of MSE estimates: (a) the validity of unique inferences regarding pattern irregularity of a neural signal vs. its variance, and; (b) the validity of the time-scale at which effects are observed.

1.2 The influence of variance on entropy challenges measurement validity

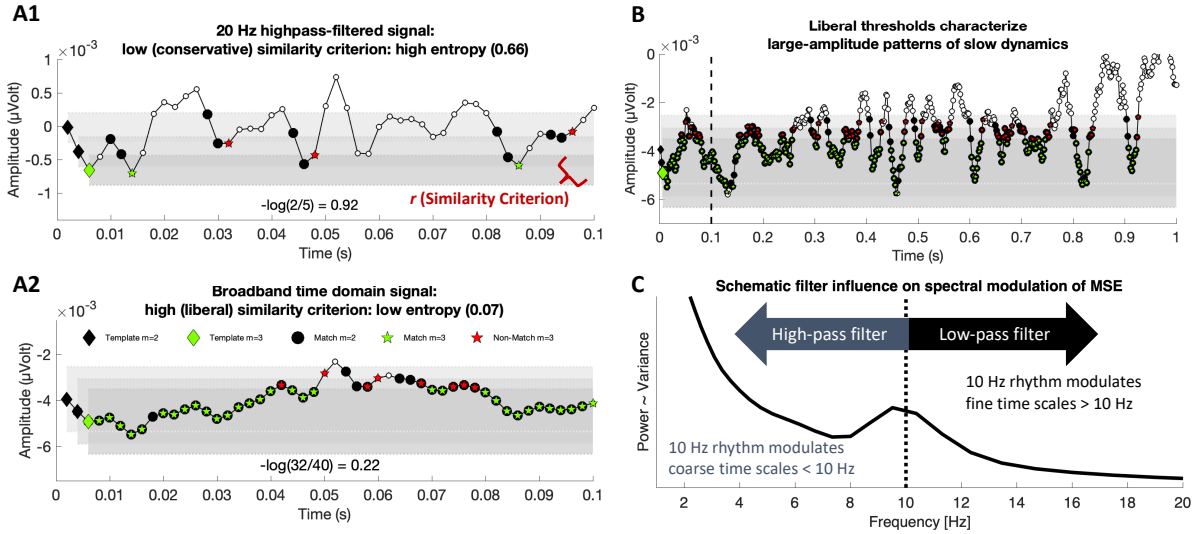

Figure 1: Scale-dependent entropy estimates are linked to spectral power via the similarity criterion (*r* parameter) and the regularity introduced by spectral events (e.g. rhythmicity). During the entropy calculation, template patterns of length m (here m = 2) are compared to the rest of the time series. Matches are detected when m consecutive samples fall within the templates' similarity bounds as indicated by the grey shading. Entropy is based on the ratio of m+1 vs. m target matches and increases with a disproportional number of patterns of length m that do not remain similar at length m+1 (non-matches). This procedure is iteratively repeated across samples, deriving the entropy for each template in time. (A) Sample entropy varies as a function of the variance-dependent similarity criterion *r* that in turn relies on the signal's spectral variance. Empirical example of fine-scale entropy estimation in identical high-frequency (A1) and broadband (A2) signals. The superimposed formula exemplifies the sample entropy calculation for the current template. When the same signal is constrained to high frequency content (A1), its variance and the associated similarity criterion reflect a conservative criterion for pattern similarity. This results in high sample entropy estimates that accurately reflect high frequency pattern irregularity. (A2), In contrast, broadband signals are typically characterized by strong low-frequency fluctuations that lead to large similarity criteria at fine scales (A2), which are more appropriate for characterizing the large-amplitude fluctuations of slow dynamics (B; note different x-axis scaling). (C) Scale-wise estimates may not reflect the irregularity of spectral events at matching time scales depending on filter choices. In addition to influencing the similarity criterion, added spectral systematicity also modulates entropy estimates at varying time scales as a function of filter choice. The schematic shows an exemplary power spectrum with a characteristic 1/f shape, i.e., dominance of power/variance at low frequencies and a prominent alpha frequency peak. Low-pass filtering leads to slow dynamics dominating fine time scales, whereas high-pass filtering leads to reflections of rhythmicity at coarse time scales.

Sample entropy is an information theoretic metric that indexes the pattern irregularity (or "complexity") of time series as the conditional probability that two sequences remain similar when another sample is included in the sequence (for a visual example see Figure 1A). Hence, sample entropy compares the relative rate of similar to dissimilar time domain patterns. Whereas signals with a similar/repetitive structure (like rhythmic fluctuations) are assigned low entropy, less predictable/dissimilar (or random) signals are characterized as having higher entropy. We presume that a necessary condition for valid non-linear interpretations of sample entropy is that "the degree of irregularity of a complex signal […] cannot be entirely captured by the SD [i.e., standard deviation]" (Costa, Goldberger, & Peng, 2004, p. 1; i.e., square root of variance), a linear characteristic (Al-Nashash et al., 2009). For this reason, sample entropy is traditionally assessed relative to the standard deviation of the broadband signal to intuitively normalize the estimation of irregularity for overall distributional width (Richman & Moorman, 2000). In particular, the *similarity parameter r* directly reflects the tolerance against which temporal patterns are labelled as being similar or different (for an example, see Figure 1A; for

details see methods). In particular, for each point in the time series, a repeating pattern is identified by falling within a range that is defined by the standard deviation of the signal (see Figure A1). However, contrary to the assumption that "[d]efining *r* as a fraction of the standard deviation eliminates the dependence of [sample entropy] on signal amplitude" (Bruce et al., 2009, p. 259; see also Costa et al., 2004), it is rather plausible that this procedure in itself introduces dependencies between signal variance and entropy. Specifically, as the magnitude of signal fluctuations increases, the threshold for pattern similarity becomes more liberal as more pattern are identified as similar (see Figure A2), thereby reducing estimated entropy and leading to a general anti-correlation between signal variance and entropy (Nikulin & Brismar, 2004; Richman & Moorman, 2000; Shafiei et al., 2019). Hence, contrary to common belief, the use of a variance-based normalization criterion may invoke rather than remove dependencies between entropy estimates and signal variance (see Hypothesis A in section 1.5).

This problem is compounded in the case of multiscale sample entropy (MSE), which aims to describe entropy at different time scales – from fast dynamics at fine (also referred to as 'short) time scales to slow fluctuations at coarse (or 'long) time scales. To characterize coarser time scales during the MSE calculation, signals are traditionally low-pass filtered, whereas the similarity criterion typically remains scale-invariant, and set relative to the original broadband signal ('Original' implementation). In turn, progressive time scale coarsening successively removes high frequency content from the signal, yet a fixed broadband criterion still retains the excluded frequencies; as a result, the increasingly mismatched criterion becomes a liberally biased threshold for pattern similarity, effectively reducing entropy estimates. This is most clearly illustrated by the observation that white noise signals, which should be characterized as equally random at each time scale, exhibit decreasing entropy values towards coarser scales when scale-invariant *r* parameters are used (Courtiol et al., 2016; Miskovic, Owens, Kuntzelman, & Gibb, 2016; Nikulin & Brismar, 2004). Hence, the use of scale-invariant similarity criteria renders links between signal variance and signal entropy ambiguous in standard applications (Nikulin & Brismar, 2004). This prior observation provided a rationale for scale-dependent computations of the *r* parameter (Valencia et al., 2009). This procedure adheres to the initial idea of normalizing the scale-dependent signal via its variance, without making estimates at coarser scales dependent on the variance of frequencies that have already been removed from the signal. However, the use of scale-invariant thresholds remains dominant in neuroscientific applications and in previous validation work (Courtiol et al., 2016), thus requiring an emphasis of the divergence between results from fixed and scale-varying thresholds.

While fixed similarity criteria present a general challenge to the validity of entropy estimation, a scale-specific re-estimation of normalization parameters does not by itself guarantee unique, variance-independent entropy estimates. In contrast, sample entropy remains conditional on signal variance due to the (scale-dependent) broadband variance normalization. It is well appreciated that the broadband signal represents the mixture of a scale-free background with canonical rhythmic frequencies (Haller et al., 2018; Kosciessa, Grandy, Garrett, & Werkle-Bergner, 2019) that are spatially specific and dynamically modulated during spontaneous cognition and evoked task states (e.g., Keitel & Gross, 2016; Vidaurre et al., 2018). In the face of such spectral complexity, signal variance may impact entropy estimates in complex ways depending on the frequency composition of the target signal. Note that if the signal is constrained to narrowband frequencies, its variance corresponds directly to spectral

power. This problem persists at coarser scales, where entropy results remain partially dependent on the similarity criterion, and thus the variance of the remaining frequencies. Hence, even when adapted thresholds are used, the variance used to normalize entropy estimates may introduce inter-individual, condition, and/or group differences that could invalidly be attributed as unique to entropy rather than simply being shared with (or determined by) spectral variance.

1.3 Are "fast" and "slow" entropy estimates valid estimates of fast and slow processes?

A multiscale entropy approach is primarily motivated by the goal to derive additional insight into the time scales at which complex neural dynamics occur. Hence, the aim is to characterize signal irregularity along a continuum of time scales varying from fast dynamics to slow fluctuations. In turn, observed scale-dependent effects are commonly interpreted with reference to dynamical systems theory (Breakspear, 2017) and structural connectomics (Sporns, 2010). Specifically, it is often assumed that events at fine time scales closely relate to fast dynamics and vice versa (McIntosh et al., 2014), with theoretical and empirical work indicating that the time scale of neural dynamics is related to intrinsic activity time constants that depend at least in part on structural properties of the underlying neural circuits (Buzsaki, Logothetis, & Singer, 2013; Fries, 2009; Mejias, Murray, Kennedy, & Wang, 2016; von Stein & Sarnthein, 2000; X. J. Wang, 2010). To align with such interpretations, entropy effects at fine scales should ideally reflect the pattern irregularity of fast dynamics, whereas those at coarse scales ought to mainly characterize slower dynamics. This expectation is sometimes made explicit in claims that "the structure of variability at short time scales, *or high frequencies*, has been linked to local neural population processing, whereas variability at longer time scales, *or lower frequencies*, has been linked to large-scale network processing" (Courtiol et al., 2016, p. 176; emphases added). Such expectations may however be violated by standard MSE estimation procedures. Notably, the dependence of coarse-scale estimates on high-frequency power when invariant similarity criteria are used (see section 1.2) challenges the fundamental assumption that estimates at coarser time scales exclusively reflect slow neural dynamics. This motivates Hypothesis B (see section 1.5). In addition, a time scale mismatch may also be present at finer time scales. Specifically, while entropy estimates at original sampling rates are often interpreted as indicating 'fast' events, they characterize and are (scale-dependently) normalized by broadband variance. Importantly, broadband variance represents the sum of power across individual frequency bands, with most neural signals exhibiting a scale-free (or 1/f) power distribution, for which variance is maximal at low frequencies (Buzsaki & Mizuseki, 2014; He, 2014). Hence, when broadband signals are analyzed, pattern similarity is traditionally referenced to signal variance dominated by slow fluctuations (see Figure 1B). In principle, this may reliably manifest as an association between spectral slopes and fine-scale entropy that has been observed both across subjects and wakefulness states (Bruce et al., 2009; Miskovic et al., 2019; Waschke, Wostmann, & Obleser, 2017). As sample entropy has been shown to be sensitive also to the autocorrelative properties of the signal (Courtiol et al., 2016; Kaffashi, Foglyano, Wilson, & Loparo, 2008), it is hence unlikely that fine-scale entropy is specific to the irregularity of high frequency activity. Taken together, this prior evidence motivates Hypothesis C (see section 1.5). In worst-case scenarios, a conjunction of the mechanisms described above may thus lead to a reflection of fast dynamics at coarse scales and a reflection

of slow dynamics at fine time scales, potentially *inverting* the interpretation of MSE time scales in general.

We argue that narrowband rhythms provide an optimal test case to assess a proper mapping of neural irregularity to specific time scales (see Figure 1C), given that they are a well-researched characteristic of brain function, and given their specific definition of the time scale of events (i.e., period = inverse of frequency). While previous work has assessed the relation between multiscale entropy estimates and autocorrelative features (Courtiol et al., 2016), little work has focused on the mapping of spectral frequencies and entropy time scales. Rather, existing simulations have produced puzzling results that have received little attention in the literature so far; while a linear mapping between simulated rhythmicity and its reflection in entropy timescales has been observed, added rhythmic regularity appeared to *increase* entropy above baseline (Park, Kim, Kim, Cichocki, & Kim, 2007; Takahashi et al., 2010; Vakorin & McIntosh, 2012). This notably contrasts with the intuition that added signal regularity should rather reduce observed entropy. Targeted simulations are thus necessary to assess the intuitive notion that rhythmicity should be anticorrelated with entropy, and to assess whether this phenomenon occurs at appropriate time scales.

1.4 Age differences in neural irregularity at fast and slow time scales

An unambiguous mapping between the spectral frequency of neural events and their reflection in entropy time scales is arguably crucial to accurately infer the potential mechanisms behind entropy modulations. One principal application of multiscale entropy is research into lifespan covariations between functional neural dynamics and structural network ontogeny (for a review see McIntosh, 2019). Within this line of inquiry, it has been proposed that structural brain alterations across the lifespan manifest as entropy differences at distinct time scales (McIntosh, Kovacevic, & Itier, 2008; McIntosh et al., 2014; H. Wang et al., 2016; Waschke et al., 2017). In particular, it has been suggested that coarse-scale entropy decreases and fine-scale entropy rises with increasing adult age as a reflection of senescent shifts from global to increasingly local information processing (McIntosh et al., 2014; H. Wang et al., 2016). Crucially, this suggestion mirrors observations based on spectral power, where age-related decreases in the magnitude of low-frequencies (Leirer et al., 2011; Vlahou, Thurm, Kolassa, & Schlee, 2014) are accompanied by increases in high-frequency activity, conceptualized also as a flattening of power spectral density (PSD) slopes (McIntosh et al., 2014; Voytek et al., 2015; H. Wang et al., 2016; Waschke et al., 2017). While these results seemingly converge towards a joint decrease of low-frequency power and slow scale entropy in older adults (and an increase for both regarding fast dynamics), this correspondence is surprising upon closer inspection given the presumed anticorrelation between the magnitude of stereotypic rhythm dynamics and their estimated entropy. Given uncertainty regarding the unique information offered by entropy modulations, as well as concerns regarding the valid interpretation of time scales of entropy effects, we attempted to reconcile these various issues by investigating the relation between cross-sectional age effects on both MSE and spectral power.

1.5 Hypotheses and current study

We used simulations and empirical EEG data to probe the relationship between spectral power and multiscale sample entropy (MSE), with a specific focus on the relation between rhythmic frequencies and entropy time scales. We formulated the following general hypotheses regarding the link between spectral variance and MSE:

A. The magnitude of the variance-based similarity criterion is negatively correlated with entropy estimates.
B. 'Original' scale-invariant similarity criteria produce increasingly biased thresholds for the detection of time series pattern similarity towards coarser time scales. The magnitude of this bias scales with the amount of excluded high frequency variance. This produces scale-to-frequency mismatches, wherein power differences at high frequencies manifest as differences in coarse-scale entropy.
C. When fine time scales characterize signals that include both fast and slow fluctuations, fine-scale entropy estimates (and age differences therein) will relate to PSD slopes. Such an association will be absent when slow fluctuations are removed.

Extending these hypotheses to the domain of age-related differences in EEG-based MSE and spectral power, we assessed the following hypotheses:

D. Using 'Original' MSE, older adults will exhibit higher entropy at finer time scales and *artificially* lower entropy at coarser time scales compared to younger adults (e.g., McIntosh et al., 2014). Concurrently, older adults will have shallower PSD slopes than younger adults, as represented by higher power at high frequencies and lower power at low frequencies (Voytek et al., 2015; Waschke et al., 2017). Based on Hypotheses B & C, a relation of these effects is hypothesized as follows:
    D1. Scale-invariant similarity criteria introduce coarse-scale entropy differences as a function of high frequency power (cf. Hypothesis B). Hence, coarse-scale age differences relate to group differences in high frequency power and disappear when scale-invariant threshold biases are removed.
    D2. Age differences at fine time scales relate to age differences in PSD slopes, with higher entropy in older adults relating to steeper PSD slopes. This association is dependent on the presence of slow fluctuations during the entropy calculation (cf. Hypothesis C). No fine-scale age differences will be indicated when slow fluctuations are removed from the signal.

In line with our expectations, we observed that 'Original' MSE leads to a strong dependence of fine time scales on low-frequency power and coarse time scales on high-frequency power. To highlight the neuroscientific relevance of these associations, we used novel resting state data to replicate two previous findings in the literature: (1) an age-related shift in entropy from dominantly coarse to fine-scale entropy and (2) a strong association of fine-scale entropy with the slope of power spectral density. By varying filter settings, we show how these entropy effects may be explained in the context of spectral variance differences, but at opposing time scales to those observed for entropy. Finally, we highlight that narrowband implementations of

entropy approximate frequency-specific signal irregularity as the inverse of the rate of stereotypic spectral events.

2 Methods

2.1 Simulations of relations between rhythmic frequency, amplitude, and MSE

To assess the influence of rhythmicity on entropy estimates, we simulated varying amplitudes (0 to 7 arbitrary units in steps of 0.5) of 10 Hz (alpha) rhythms on a fixed 1/f background. This range varies from the absence to the clear presence of rhythmicity (see Supplementary Figure 1 for an example). The background consisted of $\frac{1}{f^x}$-filtered Gaussian white noise (mean = 0; std = 1) with x = 1 that was generated using the function f_alpha_gaussian (Stoyanov, Gunzburger, & Burkardt, 2011). The background was additionally band-pass filtered between .5 and 70 Hz using 4th order Butterworth filters. Eight second segments (250 Hz sampling rate) were simulated for 100 artificial, background-varying trials, and phase-locked 10 Hz sinusoids were superimposed. The alpha rhythm was chosen as it constitutes the largest and most prevalent human rhythm in scalp EEG data (Kosciessa et al., 2019) and therefore is commonly present and modulated in data that is used for entropy analyses. To analyze the reflection of rhythmic frequency on time scales and to replicate a previously observed linear frequency-to-timescale mapping between the spectral and entropy domains (Park et al., 2007; Takahashi et al., 2010; Vakorin & McIntosh, 2012), we repeated our simulations with sinusoids of different frequencies (5 Hz, 10 Hz, 20 Hz, 40 Hz, 80 Hz), that covered the entire eight second-long segments.

2.2 Resting state data and preprocessing

To investigate the influence of similarity criteria and filter ranges in empirical data, we used resting-state EEG data collected in the context of a larger assessment prior to task performance and immediately following electrode preparation. Following exclusion of three subjects due to recording errors, the final sample contained 47 younger (mean age = 25.8 years, SD = 4.6, range 18 to 35 years; 25 women) and 52 older adults (mean age = 68.7 years, SD = 4.2, range 59 to 78 years; 28 women) 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. Participants reported to be in good health with no known history of neurological or psychiatric incidences, and were paid for their participation (10 € per hour). All older adults had Mini Mental State Examination (MMSE) (Folstein, Robins, & Helzer, 1983; Kessler, Markowitsch, & Denzler, 2000) scores above 25. All participants gave written informed consent according to the institutional guidelines of the Deutsche Gesellschaft für Psychologie (DGPS) ethics board, which approved the study.

Participants were seated at a distance of 80 cm in front of a 60 Hz LCD monitor in an acoustically and electrically shielded chamber. Following electrode placement, participants were instructed to rest for 3 minutes with their eyes open and closed, respectively. During the eyes open interval, subjects were instructed to fixate on a centrally presented fixation cross. An

auditory beep indicated to the subjects when to close their eyes. Only data from the eyes open resting state were analyzed here. EEG was continuously recorded from 64 active (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 (Oostenveld & Praamstra, 2001), 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. Online, signals were digitized at a sampling rate of 1 kHz.

Preprocessing and analysis of EEG data were conducted with the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 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 0.2 to 125 Hz. Subsequently, data were downsampled to 500 Hz and all channels were re-referenced to mathematically averaged mastoids. Blink, movement and heart-beat artifacts were identified using Independent Component Analysis (ICA; Bell & Sejnowski, 1995) and removed from the signal. Artifact-contaminated channels (determined across epochs) were automatically detected using (a) the FASTER algorithm (Nolan, Whelan, & Reilly, 2010), and by (b) 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 on recursive outlier detection. Finally, recordings were segmented to participant cues to open their eyes, and were epoched into non-overlapping 3 second pseudo-trials. To enhance spatial specificity, scalp current density estimates were derived via 4th order spherical splines (Perrin et al., 1989) using a standard 10-05 channel layout (conductivity: 0.33 S/m; regularization: 1^-05; 14th degree polynomials).

2.3 Calculation of standard and "modified" multiscale entropy

The calculation of standard MSE and the point averaging procedure followed (Costa et al., 2002, 2005). In short, sample entropy quantifies the irregularity of a time series of length *N* by assessing the conditional probability that two sequences of *m* consecutive data points will remain similar when another sample (*m+1*) is included in the sequence (for a visual example see Figure 1A). The embedding dimension *m* was set to 2 in our applications. Sample entropy is defined as the inverse natural logarithm of this conditional similarity: $\mathrm{SampEn}(m, r, N) = -\log\left(\frac{p^{m+1}(r)}{p^{m}(r)}\right)$. Crucially, the similarity criterion (*r*) defines the tolerance within which time points are considered similar and is traditionally defined relative to the standard deviation (i.e., square root of signal variance; here set to *r* = .5). Note that a larger, more liberal, similarity criterion increases the likelihood of finding matching patterns, hence reducing entropy estimates (see Figure 1A). Furthermore, in traditional applications (e.g., Costa et al., 2005; Courtiol et al., 2016), the *r* parameter is calculated once from the entire broadband signal (i.e., in a scale-invariant manner) based on original recommendations by Richman and Moorman (2000). With progressive reduction of signal variance during the coarse-graining procedure, this leads to disproportionally high, increasingly liberal, similarity thresholds; and thus

decreasing entropy estimates (see section 1.2). Hence, fixed thresholds introduce dependencies between the 1/f shape of the frequency spectrum and entropy estimates (Nikulin & Brismar, 2004). To remedy this problem, a scale-wise recalculation of the similarity criterion has been proposed (Nikulin & Brismar, 2004; Sleimen-Malkoun et al., 2015; Valencia et al., 2009). We compared the implementation of MSE with a fixed and a scale-dependent r parameter (.5*STD of scale-wise signal variance) and assessed the differences in resulting entropy estimates.

To assess entropy at coarser time scales, while the original MSE method coarse-grains the data by averaging time points within discrete time bins (i.e., 'point averaging'; equivalent to applying a finite-impulse response (FIR) filter to the original time series followed by down-sampling (Courtiol et al., 2016; Valencia et al., 2009), we employed dedicated filtering prior to point skipping to down-sample the data (Semmlow, 2008; Valencia et al., 2009). Specifically, a 6th order Butterworth filter was used for either high- or low-pass filtering the signal at the approximate time scales. At each scale (also referred to as the embedding dimension; here: 1 to 42), the low-pass frequency was defined as $\mathrm{LP_{freq}} = \frac{1}{scale} * nyquist$. Similarly, high-pass cut-offs were defined as $\mathrm{HP_{freq}} = \frac{1}{scale+1} * nyquist$ and band-pass frequencies represented narrowband estimates bounded by $\mathrm{LP_{freq}}$ and $\mathrm{LH_{freq}}$ This definition ensures that each scale captures information that is unique to that frequency band. The down-sampling procedure consisted of skipping points according to the time scale and was identical across filter settings, except in the 'Original' case. To avoid biases arising from different starting points of the skipping procedure, pattern sequences were assessed for all possible starting points and entropy estimates were computed based on their summed counts. As down-sampling represents a form of low-pass filter, it is not employed in the 'high-pass' case. Thus, estimates are based on the original sampling rate (i.e., embedding dimension of 1) with an exclusive modulation of the spectral content according to the high-pass filter. Hence, we dissociated the embedding dimension from the frequency content of the signal. As entropy (re-)calculation at the original sampling rate introduces higher computational demands, scales were sampled in step sizes of 3 for empirical data and later spline-interpolated. As the interpretation of time scales is bound to the sampling rate of the data (to assess scale-wise sampling rates) as well as the remaining spectral content, our figures indicate the Nyquist frequency at each scale, except for the high-pass case (see above). Note that the sampling rate of the simulated data was 250 Hz, whereas the empirical data had a sampling rate of 500 Hz, which renders consideration of the Nyquist frequency particularly important. We refer to a traditional implementation with scale-invariant similarity criterion and time point averaging as 'Original' in both the main text and Figures.

Further, an adapted version of MSE calculations was used for all settings (Grandy, Garrett, Schmiedek, & Werkle-Bergner, 2016), in which scale-wise entropy is estimated across discontinuous data segments. The estimation of scale-wise entropy across trials allows for reliable estimation of coarse-scale entropy without requiring long, continuous signals (Grandy et al., 2016).

For the code of the MSE algorithm and a tutorial see https://github.com/LNDG/mMSE.

2.4 Calculation of power spectral density (PSD)

Power spectral density estimates were computed by means of a Fast Fourier Transform (FFT) over 3 second pseudo-trials for 41 logarithmically spaced frequencies between 2 and 64

Hz (employing a Hanning-taper; segments zero-padded to 10 seconds) and subsequently averaged. Spectral power was $\log_{10}$-transformed to render power values more normally distributed across subjects. Power spectral density (PSD) slopes were derived by linearly regressing power values on log-transformed frequencies. The spectral range from 7-13 Hz was excluded from the background fit to exclude a bias by the narrowband alpha peak (Voytek et al., 2015; Waschke et al., 2017).

2.5 Detection of single-trial spectral events

Spectral power, even in the narrowband case, is unspecific to the occurrence of systematic rhythmic events as it also characterizes periods of absent rhythmicity (e.g., Jones, 2016). Dedicated rhythm detection alleviates this problem by specifically detecting rhythmic episodes in the ongoing signal. To investigate the potential relation between the occurrence of stereotypic spectral events and narrowband entropy, we detected single-trial spectral events using the extended BOSC method (Caplan, Madsen, Raghavachari, & Kahana, 2001; Kosciessa et al., 2019; Whitten, Hughes, Dickson, & Caplan, 2011) and probed their relation to individual entropy estimates. In short, this method identifies stereotypic 'rhythmic' events at the single-trial level, with the assumption that such events have significantly higher power than the 1/f background and occur for a minimum number of cycles at a particular frequency. This effectively dissociates narrowband spectral peaks from the arrhythmic background spectrum. Here, we used a one cycle threshold during detection, while defining the power threshold as the 95th percentile above the individual background power. A 5-cycle wavelet was used to provide the time-frequency transformations for 49 logarithmically-spaced center frequencies between 1 and 64 Hz. Rhythmic episodes were detected as described in Kosciessa et al. (2019). Following the detection of spectral events, the rate of spectral episodes longer than 3 cycles was computed by counting the number of episodes with a mean frequency that fell in a moving window of 3 adjacent center frequencies. This produced a channel-by-frequency representation of spectral event rates, which were the basis for subsequent significance testing. Event rates and statistical results were averaged within frequency bins from 8-12 Hz (alpha) and 14-20 Hz (beta) to assess relations to narrowband entropy and for the visualization of topographies. To visualize the stereotypic depiction of single-trial alpha and beta events, the original time series were time-locked to the trough of individual spectral episodes and averaged across events (c.f., Sherman et al., 2016). More specifically, the trough was chosen to be the local minimum during the spectral episode that was closest to the maximum power of the wavelet-transformed signal. To better estimate the local minimum, the signal was low-pass filtered at 25 Hz for alpha and bandpass-filtered between 10 and 25 Hz for beta using a 6th order Butterworth filter. A post-hoc duration threshold of one cycle was used for the visualization of beta events, whereas a three-cycle criterion was used to visualize alpha events. Alpha and beta events were visualized at channels POz and Cz, respectively.

2.6 Statistical analyses

Spectral power and entropy were compared across age groups within condition by means of independent samples t-tests; cluster-based permutation tests (Maris & Oostenveld, 2007) were performed to control for multiple comparisons. Initially, a clustering algorithm formed

clusters based on significant t-tests of individual data points (p <.05, two-sided; cluster entry threshold) with the spatial constraint of a cluster covering a minimum of three neighboring channels. Then, the significance of the observed cluster-level statistic, based on the summed t-values within the cluster, was assessed by comparison to the distribution of all permutation-based cluster-level statistics. The final cluster p-value that we report in all figures was assessed as the proportion of 1000 Monte Carlo iterations in which the cluster-level statistic was exceeded. Cluster significance was indicated by p-values below .025 (two-sided cluster significance threshold). Effect sizes for MSE age differences with different filter settings were computed on the basis of the cluster results in the 'Original' version. This was also the case for analyses of partial correlations. Raw MSE values were extracted from channels with indicated age differences at the initial three scales 1-3 (>65 Hz) for fine MSE and scales 39-41 (<6.5 Hz) for coarse MSE. R^2 was calculated based on the t-values of an unpaired t-test: $R^2 = \frac{t^2}{t^2+df}$ (Lakens, 2013). The measure describes the variance in the age difference explained by the measure of interest, with the square root being identical to Pearson's correlation coefficient between continuous individual values and binary age group. Effect sizes were compared using the r-to-z-transform and a successive comparison of the z-value difference against zero: $Z_{Diff} = \frac{z1-z2}{sqrt(\frac{1}{N1-3}+\frac{1}{N2-3})}$ (Brandner, 1933). Unmasked t-values are presented in support of the assessment of raw statistics in our data (Allen, Erhardt, & Calhoun, 2012).

3 Results

3.1 Simulations indicate nonlinear relations between rhythmic power and entropy

Traditional MSE algorithms assess signal entropy relative to the standard deviation of the broadband signal. Crucially, most neural time series are characterized by a scale-free 1/f frequency distribution, indicating that lower frequency fluctuations have the highest amplitudes and contribute most to the overall variance. Hence, the similarity criterion relevant for fine-scale patterns is predominantly based on the amplitude of low frequencies, leading to large similarity criteria (*r* values). Such a large threshold could bias most of the actual fine-scale patterns by the dominant fluctuations of slow signals, with fast time series patterns treated as highly similar regardless of actual pattern fluctuations (see Figure 1AB). Low entropy values could result at fast entropy scales simply for this reason. In principle, this problem could be alleviated by using spectral filters to constrain signals to the frequency range of interest. In particular, we expected that scale-dependent low-pass filters would lead to a low-frequency representation also at finer time scales, whereas slow fluctuations would exclusively modulate entropy at coarser time scales if high-pass filters were applied (Figure 1C).

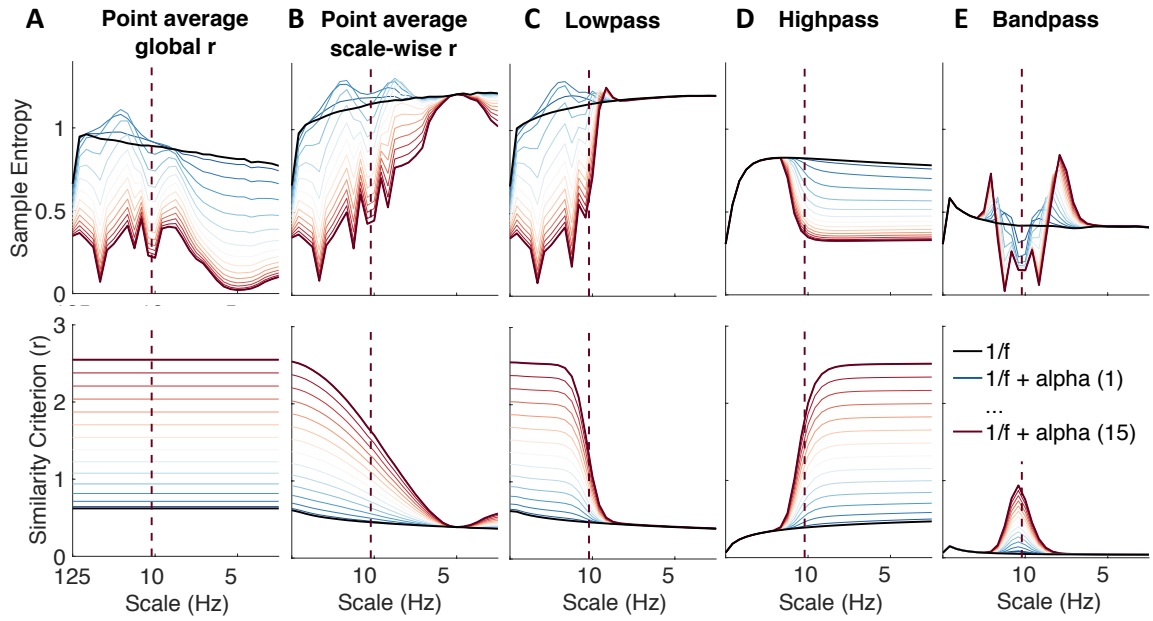

Figure 2: Rhythmic power manifests at different time scales depending on filter choice and similarity criterion. Simulations indicate at which time scales the addition of varying magnitudes of stereotypic narrowband 10 Hz rhythms (blue-to-red line gradient) modulate entropy compared to the baseline 1/f signal (black line). Simulations indicate that increases in rhythmicity strongly reduce entropy estimates alongside increases in the similarity criterion. The affected scales vary as a function of global vs. scale-dependent similarity criteria and the spectral filtering used to derive coarser time scales. Crucially, in 'Original' implementations, added narrowband rhythmicity decreased entropy with low scale-specificity, in line with global increases in the r parameter (A). In contrast, the use of scale-varying thresholds (B) and dedicated filtering (C-E) increased specify regarding the time scales at which rhythmicity was reflected. Note that timescales are presented in Hz to facilitate the visual assessment of rhythmic modulation. For all versions except high pass, the scale represents the upper Nyquist bound of the embedding dimension. For the high pass variant, the scale represents the high pass frequency (see methods). Time scales are log-scaled.

To probe the relationship between low-frequency rhythmic power and estimated multiscale sample entropy, we systematically varied the magnitude of simulated alpha power and assessed its influence on estimated MSE using different filter settings. Our first aim was to establish an inversion between similarity criteria and MSE estimates. In line with Hypothesis A, variations in the similarity criterion as a function of rhythmic power tightly covaried with entropy estimates; increased rhythmic power rendered the higher similarity criterion easier to surpass, in turn decreasing entropy estimates by increasing pattern matches (see Figure 1A, Figure 2). Importantly for scale-dependent inferences, with 'Original' settings, the effect of alpha power on *r* and MSE estimates was not specific to the time scale corresponding to the simulated frequency (Figure 2A). This can be attributed to the broadband similarity criterion, which by definition prohibits scale-specific allocations of the added signal variance. In contrast, when scale-dependent similarity criteria were used (Figure 2BC), strong alpha rhythmicity systematically decreased entropy at finer time scales than the simulated frequency (decreases from baseline to the left of the vertical line in Figure 2C). Hence, the presence of the low frequency rhythm diffusely affected fine-scale MSE estimates. This results from the low-pass filter (LPF) characteristics of the scale-wise estimation procedure for which the low-frequency rhythm is removed by LPFs < 10 Hz (see schematic in Figure 1C). As in previous work (Valencia et al., 2009), dedicated low-pass filtering provided a better spectral suppression compared with 'Original' point-averaging (Figure 2B), but with otherwise comparable results.

In contrast to low-pass filter results, when high-pass filters were used, rhythmicity reduced entropy at time scales below 10 Hz, hence leading to estimates of high frequency entropy that were independent of low frequency power (Figure 2D). Finally, when band-pass filters were used (Figure 2E), rhythmicity modulated entropy at the target frequency (although they also produced edge artifacts surrounding the time scale of rhythmicity). In sum, these analyses highlight that power increases of narrowband rhythms can diffusely modulate diverging temporal scales as a function of the MSE implementation. In addition, these analyses highlight that decreases in estimated entropy are often accompanied by comparable increases in the liberality of similarity criteria.

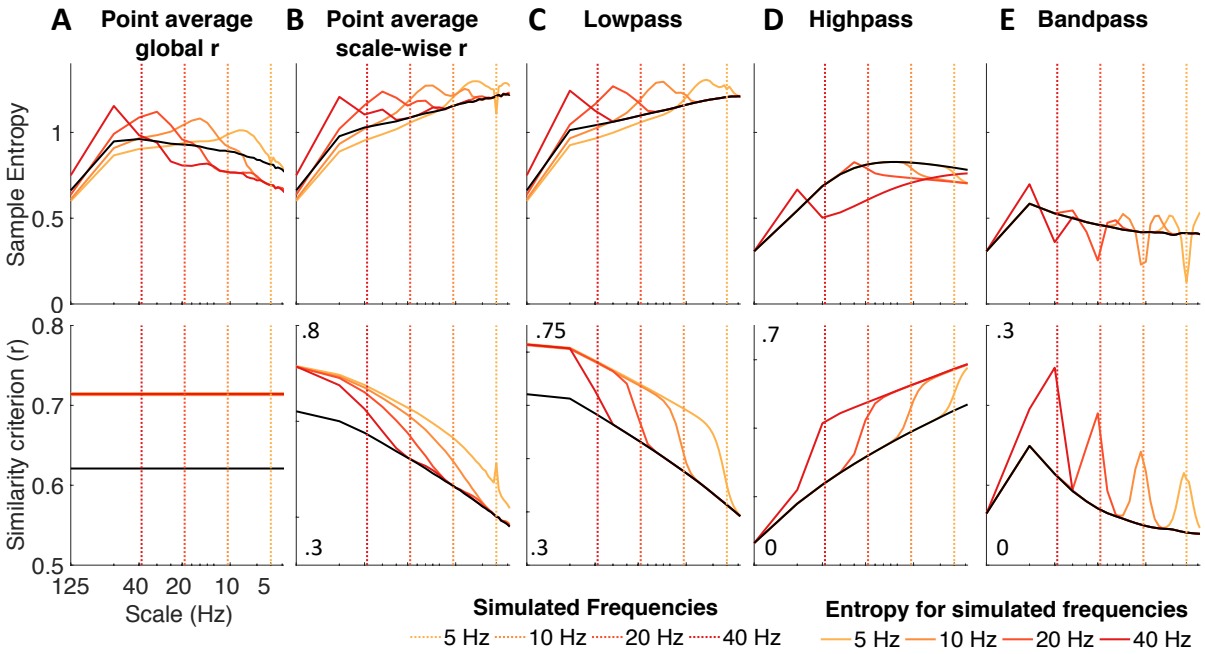

Figure 3: Influence of rhythmic frequency on MSE estimates and r parameters across different MSE variants. Simulations of different frequencies indicate a linear frequency-to-scale mapping of simulated sinusoids. Broken vertical lines indicate the simulated frequency. Low-pass MSE variants show increased entropy at time scales finer than the simulated frequency in combination with a global entropy decrease. Low-, high- and band-pass variants exhibit the properties observed in the alpha case, with a reduction above/below or at the simulated frequency. Time scales are log-scaled.

Whereas we observed a diffuse broadband decrease in 'Original' entropy under conditions of strong rhythmicity, previous simulations have presumed a rather constrained linear mapping between the frequency of simulated rhythms and their reflection in entropy time scales (Park et al., 2007; Takahashi et al., 2010; Vakorin & McIntosh, 2012). Furthermore, those studies indicated entropy *increases* with added rhythmicity, in contrast with the marked decreases in entropy observed here. How can these seemingly divergent results be reconciled? To answer this question, we simulated different frequencies superimposed on 1/f backgrounds and investigated their modulation of entropy timescales. Importantly, Figure 2A-C suggested that the amplitude of rhythmicity may be of crucial importance here, as transient entropy increases were indeed observed at low levels of rhythmicity. Hence, we focused on a comparatively low level of rhythmicity (amplitude level = 2; cf. exemplary alpha-band time series shown in Supplementary Figure 1). Similar to previous reports, we observed a linear association between simulated frequencies and peak entropy time scales (Figure 3) across implementations. Hence,

rhythms of higher frequency increased entropy at slightly finer time scales than the simulated frequency (see increases in entropy above baseline to the left of the dotted vertical lines in Figure 3A-C). Importantly, such sharp entropy increases were only observed with low-pass implementations (Figure 3A-C). Moreover, with scale-invariant *r* parameters (Figure 3A), these increases were paralleled by decreasing entropy at coarser time scales (i.e., to the right of the dotted lines in Figure 3A). This is in line with our observation of relatively broadband, amplitude-dependent, entropy decreases (cf., Figure 2A). Crucially, increased entropy relative to baseline is counterintuitive to the idea that the addition of a stereotypic pattern should decrease rather than increase pattern irregularity. Moreover, the results suggest that combinations of amplitude-varying contributions of spectral content can induce ambiguous scale-dependent results. In sum, our simulations highlight that the choice of similarity criterion and the signal's spectral content grossly affect the interpretation of entropy time scales. Furthermore, our frequency-resolved simulations suggest that a previously observed linear frequency-to-scale mapping does not provide sufficient evidence that entropy towards finer time scales dominantly represents the pattern irregularity of faster neural dynamics. Rather, such assumptions rely on puzzling entropy increases with the addition of faint rhythmic regularity that are counteracted by more dominant, and expected, decreases in entropy when the signal contains strong rhythmic predictability.

3.2 Probing the impact of spectral power on entropy in a cross-sectional age comparison

Our simulations suggest profound influences of the choice of similarity criterion and a signal's spectral content on scale-dependent MSE estimates. However, it remains uncertain if and how these factors alter inferences in traditional applications. Age-related entropy changes are an important area of application (Garrett et al., 2013), with previous applications suggesting scale-dependent differences across the lifespan (for a review see McIntosh, 2019). However, our theoretical considerations question whether such observations reflect veridical differences in the entropy of neural activity patterns or whether such effects can alternatively be accounted for by differences in spectral power (see Hypothesis D). To assess the relations between age differences in spectral power and multiscale entropy during eyes open rest, we used the following strategies: (1) we statistically compared spectral power and MSE between two age groups of younger and older adults; (2) we assessed the impact of scale-wise similarity criteria and different filtering procedures on age differences in MSE and (3) we probed the relationship between the *r* parameter and MSE.

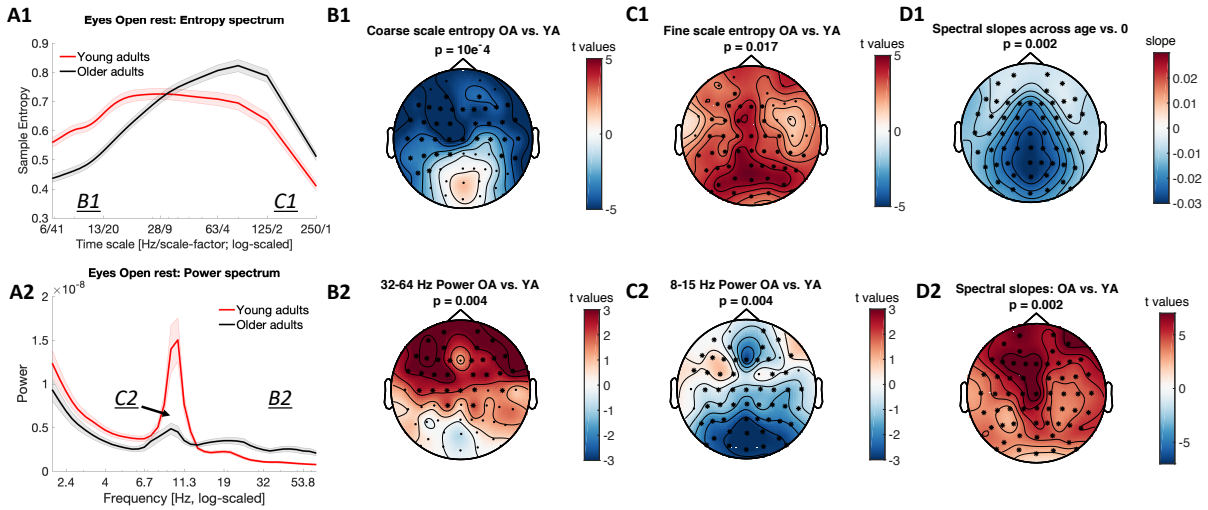

Figure 4: Timescale-dependent age differences in spectral power and entropy during eyes open rest. (A) MSE (A1) and power (A2) spectra for the two age groups. Error bars show standard errors of the mean. Note that in contrast to standard presentations of entropy, the log-scaled x-axis is sorted by decreasing scale/increasing frequency to enable a better visual comparison with the power spectra. T-values of power age contrast are shown in Supplementary Figure 2. (B, C) Topographies of age differences indicate mirrored age differences in fast entropy and low frequency power, as well as coarse entropy and high frequency power. Significant differences are indicated by asterisks. (D1) Spectral slopes across age groups. (D2) Age differences in spectral slopes.

Using traditional ('Original') settings, we replicated previous observations of scale-dependent entropy differences between younger and older adults (Figure 4A1, Figure 5A). Specifically, compared with younger adults, older adults exhibited lower entropy at coarse scales, while they showed higher entropy at fine scales (Hypothesis D; Figure 4A1). Mirroring these results in spectral power, older adults had lower parieto-occipital alpha power and increased frontal high frequency power (Figure 4A2) compared to younger adults. This was globally associated with a shift from steeper to shallower PSD slopes with increasing age (Figure 4D). At face value, this suggests joint shifts of both power and entropy, in the same direction and at matching time scales. Crucially, however, the spatial topography of differences in entropy inversely mirrored differences in power between fast and slow dynamics (Figure 4B & C; cf., upper and lower topographies), such that frontal high frequency power differences appeared inversely represented in coarse entropy scales (Figure 4B), while parieto-occipital age differences in slow frequency power more closely resembled fine-scale entropy differences (Figure 4D). This rather suggests scale-mismatched associations between entropy and power in line with our simulations and theoretical expectations (Hypothesis D1 & D2). We investigated their potential relationships more closely in the following sections regarding the potential mechanistic associations proposed in Hypotheses B and C.

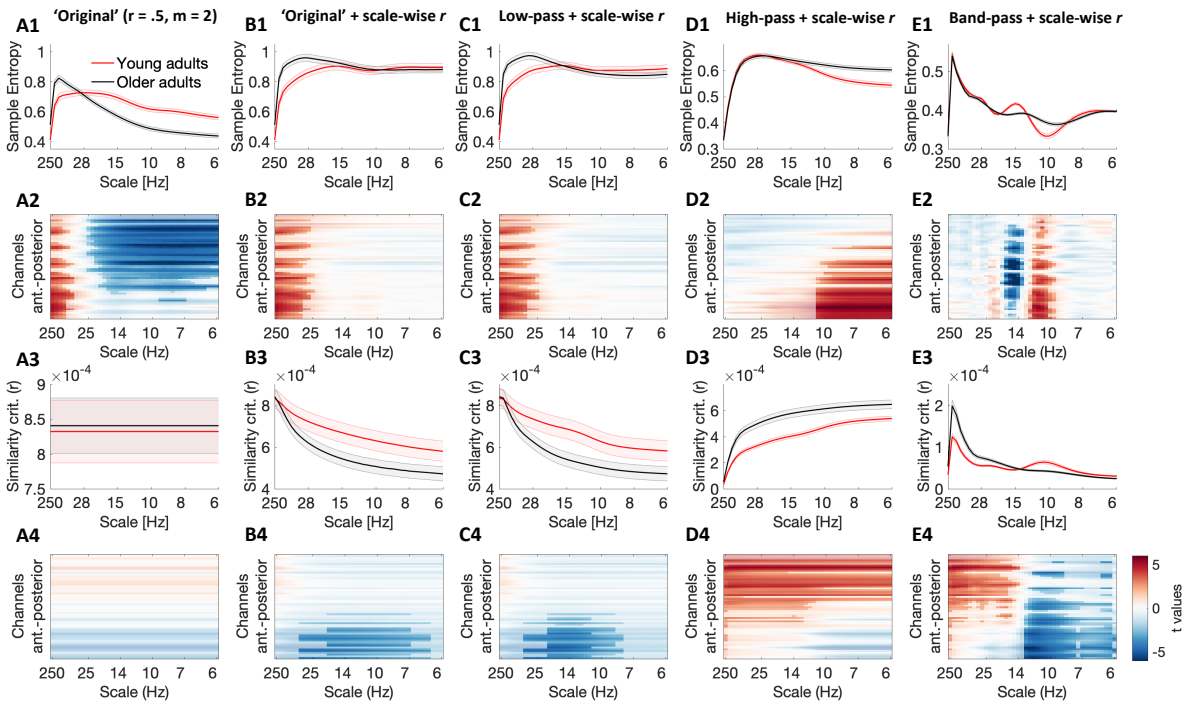

Figure 5: Average multiscale entropy and similarity criterion by age depend on the specifics of the estimation method. Grand average traces of entropy (1st row) and similarity criteria (3rd row) alongside t-maps from statistical contrasts of age differences (2nd + 4th row). Age differences were assessed by means of cluster-based permutation tests and are indicated via opacity. Original MSE (A) replicated reported scale-dependent age differences, with older adults exhibiting higher entropy at fine scales and lower entropy at coarse scales, compared with younger adults. The coarse-scale difference was exclusively observed when using invariant similarity criteria, whereas the fine-scale age difference was indicated with all low-pass versions (A, B, C), but not when signals were constrained to high-frequency or narrow-band ranges (D, E). In contrast, narrowband MSE indicated inverted age differences within the alpha and beta band (E).

Importantly, as suggested by our simulations, filter choice affected the estimation of age differences in entropy alongside differences in similarity thresholds (Figure 5). As described above, 'Original' settings indicated increased fine-scale and decreased coarse-scale entropy for older compared to younger adults, whereas no group differences in the global *r* parameter were indicated (Figure 5A). In contrast, scale-wise similarity criteria Figure 5B) abolished age differences in coarse-scale entropy (effect size was significantly reduced from r = .58 to r = .07; p=6.8*10^-5), while fine-scale entropy differences remained unchanged when low-pass filters were used (Figure 5B/C) (from r = .44 to r = .45; p=.934). However, when constraining the signal at fine scales to high frequency content (via high-pass filters; Figure 5D), no fine-scale age differences were observed and the age effect was significantly reduced (r = .09; p = .008. An age effect was only indicated once low-frequency dynamics contributed to the entropy estimation again at coarse scales. Both of these effects were in line with our Hypotheses D1 and D2 regarding the influence of spectral filtering on entropy estimates. Interestingly, we observed inverted age differences in the bandpass version (Figure 5E), with larger 'narrowband' entropy indicated in the alpha range and lower entropy in the beta range for older adults compared with younger adults. In the following sections, we investigate these results more closely with regard to the putative mechanisms linking spectral power and entropy.

605 3.3 Scale-invariant similarity criteria increasingly bias entropy towards coarser scales
606

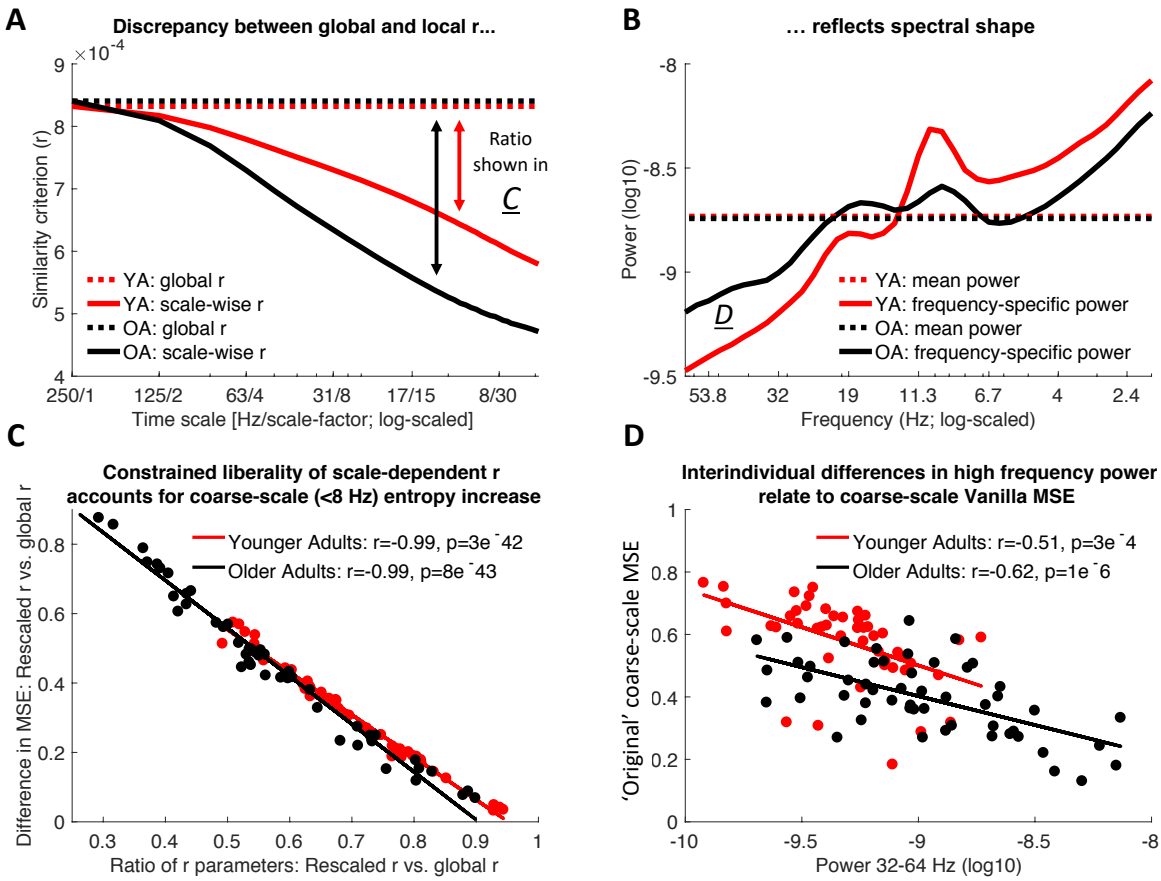

Figure 6: Mismatches between scale-specific signal variance and global similarity criteria (r parameters) can account for age differences in coarse-scale entropy. (A, B) A global similarity criterion does not reflect the spectral shape, thus leading to disproportionally liberal criteria at coarse scales following the successive removal of high-frequency variance. Scale-dependent variance (as reflected in *r*) is more quickly reduced in older compared to younger adults (A) due to their removal of more high-frequency variance (B). This leads to a differential bias, as reflected in the increasingly mismatched distance between the two invariant and scale-dependent similarity criteria towards coarser scales. This mismatch, in turn, should scale with the amount of variance removed up to a particular scale. Letter insets refer to the relevant subplots. (C) The r adjustment in the rescaled version is associated with a comparable increase in coarse-scale entropy. This shift is more pronounced in older adults. (D) With global similarity criteria, coarse-scale entropy strongly reflects high frequency power due to the proportionally more liberal similarity threshold associated. Data in A and B are global averages, data in C and D are averages from
607 frontal Original effect cluster (see Figure 4B) at time scales below 6 Hz.
608

609 Scale-dependent entropy effects in the face of scale-invariant similarity criteria (as observed
610 in the 'Original' implementation; Figure 5A) may intuitively suggest scale-wise variations in
611 pattern irregularity in the absence of variance differences. However, a fixed similarity criterion
612 is an artificial constraint that does not reflect the spectral shape of the broadband signal, leading
613 to potentially profound mismatches between the scale-dependent signal variance and the
614 invariant similarity criterion. That is, the total broadband variance may be similar while spectral
615 slopes and/or narrow-band frequency content differ. This is true for the case of aging as can be
616 appreciated by comparing the global *r* parameter with the age-specific frequency spectra
617 (Figure 6A & B). As this scale-invariant criterion thresholds a successively low-pass filtered
618 signal, this induces a relative mismatch between the scale-specific variance and the global

similarity criterion that successively increases towards coarser scales (Figure 6A). Importantly, the same broadband variance will pose a relatively higher (i.e. liberal) similarity threshold if low-pass filtering removes more high-frequency variance. In turn, the coarse-scale MSE estimate would be modulated as a function of high frequency power (i.e., Hypothesis B). To assess this hypothesis, we probed the link between the change in *r* and MSE between the use of a global and a scale-varying similarity criterion. As expected, we observed a strong anti-correlation between inter-individual differences in *r* and MSE (Figure 6C). That is, the more individual thresholds were re-adjusted to the lower scale-wise variance, the more entropy estimates increased. Crucially, this difference was more pronounced for the older adults (paired t-test; r: p = 5e-6; MSE: p = 3e-4). That is, due to their increased high frequency power, low-pass filtering decreased older adults' variance proportionally more than younger adults' variance. Thus, in 'Original' settings, older adults' global criterion presented a more liberal threshold at coarser scales than the threshold of younger adults, which can account for the 'lower' MSE estimates observed for older adults with 'Original' settings. In line with this assumption, individual high frequency power at frontal channels was inversely related to coarse-scale entropy estimates when a scale-invariant similarity criterion was applied (Figure 6C), but not when the similarity criterion was recomputed for each scale (YA: r = -0.15; p = .302; OA: r = .2, p = .146). This is further in line with the observation that coarse-scale age differences (Figure 5A) disappeared when a scale-wise similarity criterion was used (Figure 5B). Taken together, this indicates that the observed age difference at coarse entropy scales can be largely accounted for by high frequency power differences between young and old adults and provides an explanation for the inverse group differences between high frequency power and coarse-scale entropy (Hypothesis D1).

3.4 Low-frequency contributions render fine-scale entropy a proxy measure of PSD slope

A common observation in the MSE literature is a high sensitivity to task and behavioral differences even at the original sampling rates (i.e., fine scales), which are commonly assumed to reflect fast dynamics. This sensitivity is surprising given that little power generally exists in high-frequency ranges in humans or animals (Hipp & Siegel, 2013). Interestingly, multiple previous studies suggest that fine-scale entropy reflects the slope of power spectral density (e.g., Bruce et al., 2009; Waschke et al., 2017). Given that this slope can be approximated by the ratio of high to low-frequency power, and that 'Original' MSE implementations contain both components due to the assessment of a broadband signal, we probed to what extent fine-scale associations with PSD slopes depended on the presence of slow fluctuations (Hypothesis C) and whether such association may account for fine-scale entropy age differences (Hypothesis D2).

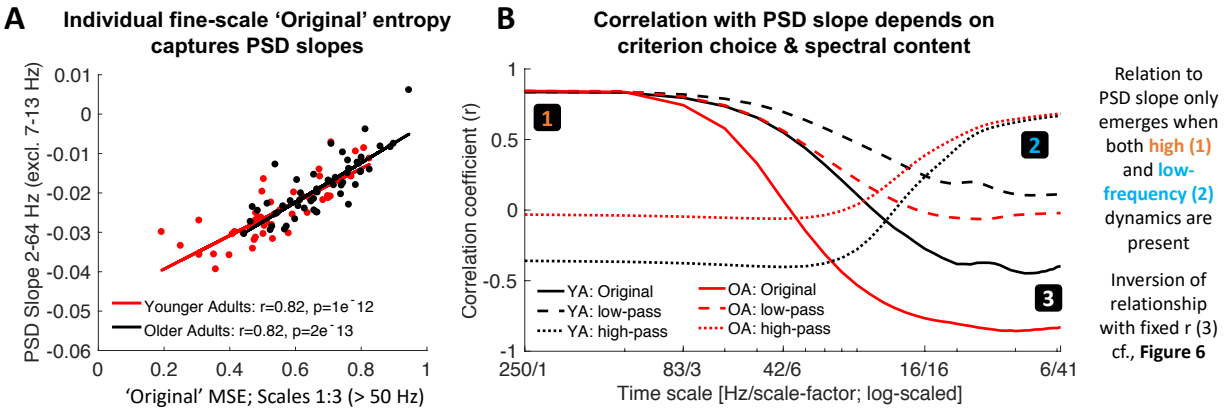

Figure 7: A) Sample entropy at fine time scales represents the slope of power spectral density. (B) The presence of both slow and fast dynamics is required for positive associations with PSD slopes to emerge. The direction and magnitude of correlations of scale-wise entropy with PSD slopes depends on the choice of fixed vs. rescaled r parameters as well as the choice of filtering. Original entropy inverts from a positive correlation with PSD slope at fine scales to a negative association at coarse scales. Rescaling of the r parameter abolishes the negative correlation of coarse-scale entropy with PSD slopes. Supplementary Figure 3 presents scatter plots of these relationships.

As expected (Hypothesis C), individual fine-scale entropy was strongly and positively related to the slope of power spectral density (Figure 7A) in both younger and older adults. This suggests that in low-pass scenarios, in which the target signal is dominated by low frequency fluctuations, fine-scale entropy is sensitive to the ratio of high-to-low frequency variance, as captured by PSD slopes. To highlight that fine-scale entropy does not exclusively relate to the signal irregularity of high-frequency activity, we observed that following a high-pass filter to the signal, the positive relation of fine-scale entropy to PSD slopes disappeared in both age groups (Figure 7B, dotted lines), and turned negative in older adults (see Supplementary Figure 3), alongside age differences in fine-scale entropy (Figure 5D). In turn, relations between PSD slopes and age differences re-emerged once low-frequency content was included in the entropy estimation (Figure 7C, dotted lines). Hence, the positive relation of fine-scale entropy to PSD slopes was conditional on the presence of both low- and high-frequency dynamics.

In line with the hypothesis that fine-scale age differences are dependent on the presence of slow fluctuations, we observed no age differences in fine-scale entropy when signals exclusively contained high-frequency content (see section 3.2). To assess whether age differences in PSD slope could account for fine-scale age differences in 'Original' entropy, we computed partial correlations between the measures. In line with fine-scale entropy primarily reflecting PSD slope variations, no significant prediction of age group status by fine-scale entropy was observed when controlling for the high collinearity with PSD slopes ($r = -.06$, $p = .59$). In contrast, PSD slopes significantly predicted age group status when controlling for MSE ($r = .38$, $p < .001$), suggesting that differences in PSD slopes primarily account for observed age differences in MSE, but not vice-versa (in line with Hypothesis D2).

On a side note, spectral slopes were anticorrelated with coarse-scale entropy when global similarity criteria were used (Figure 7C, continuous lines), but not when criteria were scale-wise re-estimated (Figure 7C, broken lines). This likely reflects the bias described in section 3.2. That is, subjects with shallower slopes (more high frequency power) had increasingly liberal-biased thresholds towards coarse scales, thereby resulting in decreased entropy estimates.

Jointly, these empirical examples indicate that the use of global similarity criteria, as well as the presence of large amplitude low frequency dynamics can severely bias scale-wise MSE. Hence, differences in the spectral power and the *r* parameter (typically neglected as measures of interest when estimating MSE) may actually account for a large proportion of reported MSE effects; in this scenario, the pattern irregularity of fast dynamics *per se* may do little to drive MSE estimates.

3.5 Narrowband MSE indicates age differences in signal irregularity in alpha and beta band

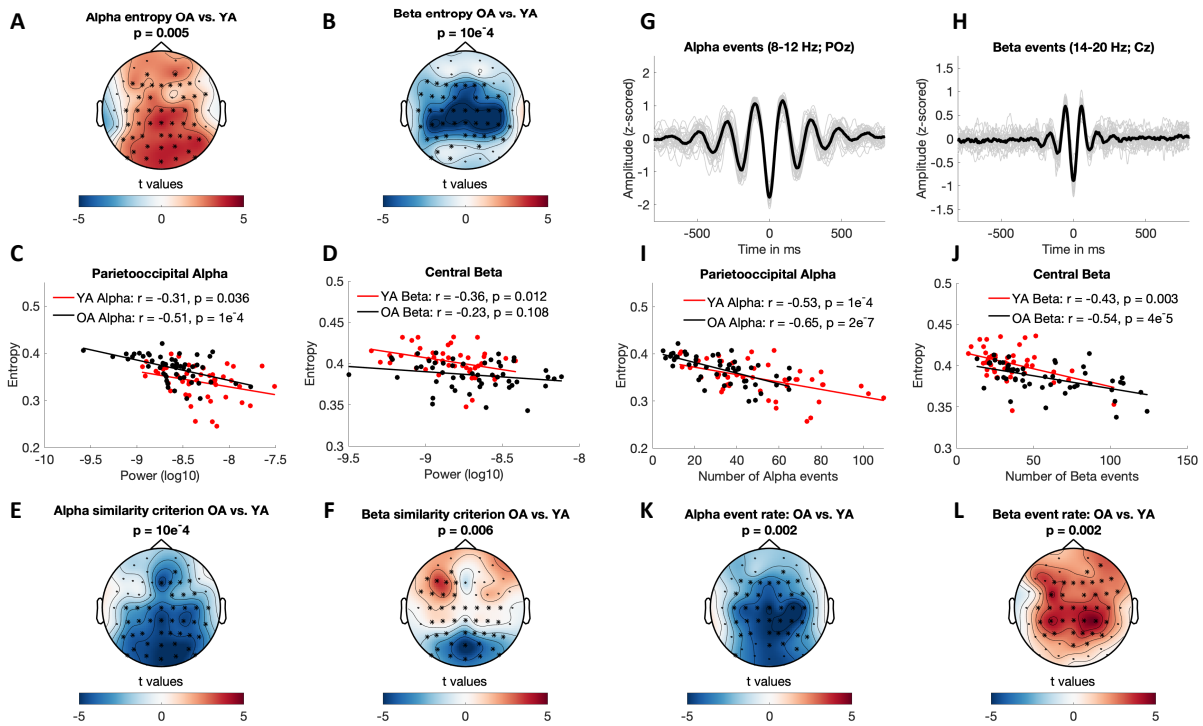

Figure 8: Narrowband MSE reflects age differences in alpha- and beta-specific event (ir)regularity. (A, B) Narrowband MSE indicates age differences in the pattern complexity at alpha (A) and beta (B) frequencies. (C, D) Alpha, but not beta power consistently correlates negatively with individual narrowband entropy within clusters of age differences. (E, F) Similarly, alpha but not beta similarity criteria show an inverted age effect with similar topography. (G, H) Single-trial rhythm detection highlights a more transient appearance of beta compared with alpha events. (I, J) The rate of stereotypical single-trial alpha and beta events is anticorrelated with individual narrowband entropy. (K, L) The rate of spectral events exhibits age differences that mirror those observed for entropy.

The previous analyses highlighted how the interpretation of scale-dependent results critically depends on the spectral content of the signal, in some cases giving rise to mismatching time scales. However, our simulations also suggest an accurate mapping between entropy and power when scale-wise bandpass filters are used (Figure 3A). Concurrently, the empirical band-pass results indicate a partial decoupling between entropy and variance age differences as reflected in the *r* parameter (Figure 5E). Specifically, older adults exhibited higher parieto-occipital entropy at alpha time scales (~8-12 Hz) and lower central entropy at beta time scales (~12-20 Hz) than in younger adults (Figure 5; Figure 8AB). Whereas alpha-band entropy was moderately and inversely correlated with alpha power (Figure 8C) and the age difference was inversely reflected in the similarity criterion in a topographically similar fashion (Figure 8E), the same was not observed for entropy in the beta range for both age groups (Figure 8DF). Promisingly, this indicates evidence for what many who employ MSE measures in cognitive

neuroscience presume; that power and entropy *can* be decoupled, providing complementary signatures of neural dynamics. This divergence of entropy and power in the beta band is particularly interesting as beta events have been observed to exhibit a more transient waveform shape (Sherman et al., 2016; Shin, Law, Tsutsui, Moore, & Jones, 2017), while occupying a lower total duration during rest than alpha rhythms (Kosciessa et al., 2019). This may explain a divergence of entropy estimates from spectral power as it should be the rate of stereotypic spectral events that reduces pattern irregularity rather than the overall power within a frequency band. To test this hypothesis, we applied single-trial rhythm detection to extract the individual rate of alpha (8-12 Hz) and beta (14-20 Hz) events. As predicted, individual alpha events had a more sustained appearance compared with beta events as shown in Figure 8G & H (events were time-locked to the trough of individual events; see section 2.6). Importantly, both individual alpha and beta event rate were inversely and moderately correlated with individual beta entropy estimates (Figure 8IJ) at matching time scales in the band-pass version. The relationships remained stable after controlling for individual event rate and entropy in the age cluster of the other frequency band (Alpha YA: r = -.63, p = 3e-6; Alpha OA: r = -.70, p = 1e-8; Beta YA: r = -.54, p = 1e-4; Beta OA: r = -.61, p = 2e-6), suggesting separable associations between event rate and entropy within the two frequencies bands. This is important, as our simulations suggest increased entropy estimates around narrow-band filtered rhythmicity (see Figure 2A). Furthermore, a permutation test indicated age differences in beta rate that were opposite in sign to the entropy age difference (see Figure 8L). In particular, older adults had a higher number of central beta events during the resting state compared with younger adults, thus rendering their beta-band dynamics more stereotypic. In sum, these results suggest that narrowband MSE estimates approximate the irregularity of spectral events at matching time scales.

4 Discussion

For entropy to be a practical and non-redundant measure in cognitive neuroscience, both its convergent and discriminant validity to known signal characteristics has to be established. Spectral features have a long history in cognitive electrophysiology and many procedures and theoretical work are available for their interpretation. In the face of this existing literature, it has been proposed that entropy is sensitive to non-linear time series characteristics that can complement linear spectral information. If and to what extent these measures are independent is however often not assessed, but tacitly inferred from applying a variance-based 'normalization' during the entropy calculation. Contrary to orthogonality assumptions, our analyses suggest that differences in the similarity criterion may account for a significant proportion of entropy effects in the literature, and thereby fundamentally affect the interpretation of observed effects. In traditional applications, these effects can be differentiated into separable effects of (a) biases arising from scale-invariant similarity criteria and (b) challenges in the presence of broadband, low-frequency dominated, signals (see Figure 9A for a schematic summary). In the following, we discuss these effects and how they can affect traditional inferences regarding signal irregularity.

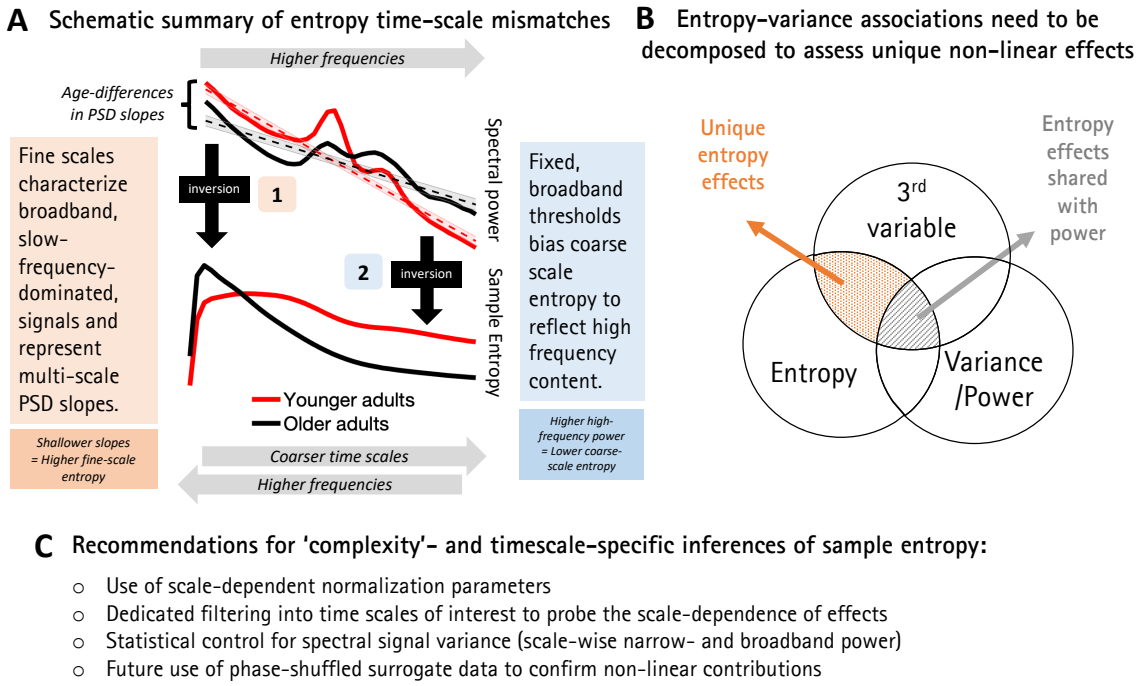

Figure 9: Summary of the identified time-scale mismatches and recommendations for future studies. (**A**) We highlight two scale-dependent mismatches that run counter to the intuition that entropy at fine scales primarily refers to fast dynamics, and vice-versa: (1) Fine-scale entropy characterizes scale-free 1/f slopes whenever broadband signals include slow frequency content. (2) Coarse-scale entropy is biased towards reflecting high-frequency content when increasingly signals of decreasing variance are compared to a fixed, and successively mismatched similarity criterion. (**B**) Beyond time-scale mismatches, entropy and variance are often strongly anticorrelated, in part due to their shared description of signal features, such as rhythmicity. To identify complementary and unique relations of pattern complexity compared to more established measures of variance, explicit statistical control is required for the latter. (**C**) We propose multiple strategies to safeguard against the highlighted issues problems in traditional applications.

4.1 Narrowband rhythmicity diffusely affects entropy scales

The use of MSE is often motivated by its sensitivity to non-linear properties of brain dynamics, that are assumed to reflect phenomena such as spontaneous network reconfigurations and brain state transitions (e.g., Deco, Jirsa, & McIntosh, 2011, 2013; Misic, Vakorin, Paus, & McIntosh, 2011). However, the variance-dependence of internal normalization parameters and the general dominance of slow fluctuations in broadband signals (from which sample entropy is typically calculated) suggest that traditional linear variance properties strongly contribute to entropy estimates (Hypothesis A). Hence, we argue that a consideration of spectral signal content is crucial to properly characterize entropy at distinct time scales of interest. Total signal variance can be dissociated into two components: broadband 'noise' and narrowband rhythmic peaks (Haller et al., 2018; Kosciessa et al., 2019) with the latter themselves being temporal averages of potentially non-stationary spectral events (Kosciessa et al., 2019). Notably, associations between pattern irregularity and the prevalence of these components are theoretically anticipated (Vakorin & McIntosh, 2012). In particular, as rhythmic events are defined by their periodic repetition, their occurrence should be associated with a decrease in signal irregularity. Due to this straightforward prediction, and their clear time scale definition, we simulated narrowband rhythms of different magnitude and frequency to assess their mapping onto MSE scales. As predicted, entropy decreased in the presence of strong

rhythmicity, however not exclusively at corresponding time scales. This was most apparent for 'Original' implementations, in which the scale-invariance of thresholds decreased estimates in a global fashion, in line with the constraints posed by the global similarity criterion that was increased in parallel. When scale-varying thresholds were used in conjunction with traditional low-pass filters, rhythms exclusively modulated entropy estimates across finer time scales. This highlights that low-pass filters render multiscale entropy especially sensitive to variance at low frequencies, while further suggesting that slow events (e.g. event-related potentials) will be reflected in a broad-scale manner. In contrast, we verified that the manipulation of spectral content via high- or band-pass filters controlled the reflection of rhythms in MSE time scales. The diffuse reflection of rhythms across many entropy time scales may initially seem at odds with previous simulations that suggested a linear mapping of increasing frequencies onto coarse-to-fine 'Original' MSE scales (Park et al., 2007; Takahashi et al., 2010; Vakorin & McIntosh, 2012). Curiously, such previous simulations indicated the frequency-to-scale mapping by considering the reflection of rhythms in positive entropy peaks. While we replicate such increases, we highlight their dependence on low rhythm strength. Specifically, whereas strong rhythmicity led to a sizeable reduction in entropy, fainter rhythmicity increased entropy at slightly finer time scales above baseline. However, increases in entropy contrast with our expectations that the addition of a more stereotypic pattern would decrease sample entropy and were quickly counteracted by more diffuse entropy decreases once rhythm magnitude increased. While the mechanistic origin of entropy increases with faint regularity remains unclear, previous conclusions may thus have overemphasized the scale-specificity of rhythmic influences. Hence, while rhythms of different frequencies modulate entropy at appropriate time scales, they also induce broadband effects, thereby leading to potential scale-to-frequency mismatches.

In addition to diffuse scale effects, we observed that rhythm-induced changes in sample entropy were strongly anti-correlated to changes in the *r* parameter, confirming Hypothesis A. However, we note that in the case of simulated rhythmicity, increases in variance (and *r*) are collinear with increases in signal regularity. Hence, entropy is not exclusively determined by the similarity criterion, but also by the reduction in pattern irregularity due to the addition of a predictable sinusoidal signal. This presents a challenge for dissociating valid differences in pattern irregularity that covary with spectral power from erroneous entropy decreases due to increased similarity criteria. To probe the main contributor to observed sample entropy effects, we replicated our analyses using permutation entropy, a measure that does not use an intrinsic similarity criterion (see Supplementary Materials). Crucially, we observed similar filter influences on the scale-wise representation of rhythmicity, suggesting that an explicit similarity criterion is not necessary to produce diffuse reflections of narrowband rhythms across multiple temporal scales. Rather, when entropy is applied to broadband signals, low-frequencies with high variance contribute in large part to fine-scale estimates (see also section 4.3).

4.2 Global similarity criteria bias coarse-scale entropy estimates

The global impact of frequency-specific events in 'Original' implementations is directly coupled to the use of global similarity criteria and challenges the notion of an accurate frequency-timescale mapping. The theoretical necessity of introducing scale-wise adaptations of similarity criteria has previously been noted (Nikulin & Brismar, 2004; Valencia et al.,

2009), and is highlighted here with a practical example. In particular, Nikulin and Brismar (2004) discussed the ambiguity between variance and pattern irregularity that arises from using scale-invariant criteria: "However, in the MSE approach the same *r* value is used for different scales. Therefore, the changes in MSE on each scale will depend on both the regularity and variation of the coarse-grained sequences. […] Therefore, the outcome of the MSE algorithm does not allow one to make a clear conclusion as to what extent this separation is based on the affected regularity or variation" (Nikulin & Brismar, 2004). In short, when the similarity criterion is fixed in the presence of scale-dependent spectral content, the liberality of thresholds systematically varies across scales. This introduces fundamental mismatches between the origin of group differences (pattern irregularity vs. variance), and the time scales at which differences manifest. These mismatches are independent of the values of the global similarity criterion – which did not differ across groups here – and rather depend on the slope of the power spectrum. The critical insight is thus that the bias relates to the relative amount of *removed* variance at the scale of interest. This leads to puzzling results, in that the entropy of white noise signals, which by definition are equally irregular at each time scale, decreases towards coarser scales, whereas pink noise signals, which have comparatively small contributions from high frequencies, receive relatively constant entropy estimates over the time scales typically examined (Nikulin & Brismar, 2004). While such reflection of PSD slopes across scales has been replicated, it has surprisingly been used to *validate* the method (Courtiol et al., 2016; Miskovic et al., 2016) rather than to indicate the presence of a systematic bias in estimation[1]. Importantly, the dependence of such biases on the spectral shape of the signal also indicates that they cannot be accounted for by choosing different constants of the similarity criterion. Importantly, this has practical implications for functional inferences. In the current resting state EEG data, we observed that an age-related increase in high frequency power manifested as a decrease in coarse-scale entropy due to group differences in the scale-wise mismatch between the (low-passed) signal variance and the global *r* parameter. Specifically, older adults' increased high frequency power strongly reduced variance with successive low-pass filtering towards coarser scales. As the similarity criterion was fixed across time scales relative to the total variance, this quickly invoked an increasingly liberal threshold. In comparison, less high-frequency variance was removed for younger adults at coarse scales. Given comparable global similarity criteria between groups, younger adults' criterion was thus more conservative, affording higher entropy estimates at coarser time scales (see Figure 9A). Crucially, coarse-scale group differences were not observed when scale-wise similarity criteria were applied, or when permutation entropy – a measure without a dedicated similarity threshold – was used (see Supplementary Materials),

[1] This appears to be mainly motivated by the questionable assumption that "changes of the variance due to the coarse-graining procedure are related to the temporal structure of the original time series, and should be accounted for by the entropy measure" (Costa et al., 2005, p. 5). However, as we show, such time-scale dependence can be explained by mismatched thresholds and hence, scale-dependent biases. Note that in previous simulations (see Figure 1 in Courtiol et al., 2016), MSE slopes varied from positive to negative as a function of spectral slopes, i.e., the ratio of high-to-low variance. In the most extreme case of blue noise signals with positive slopes, dominant high-frequency variance is quickly removed, leading to the highest rate of entropy decrease. With shallowing of slopes (and reduced high-frequency contributions), the rate of entropy reduction decreased, until eventually turning into entropy gains for signals with strong negative PSD slopes, for which biases were presumably minimal.

therefore highlighting the dependence of the group difference on mismatched thresholds. Note that we presume that this age difference arises from a relative bias. Pink noise signals, such as those observed here, have a relatively low contribution from high compared to low frequencies, rendering the absolute bias lower than for white noise signals with equal variance of these two components (and therefore a quicker 'bias rate' towards coarser scales as more high frequency variance is removed). However, variations in high-frequency variance (and thus the resulting bias) suffice, even at low levels, to systematically impact coarse-scale estimates. This may be independent from the main source of variance in course-scale entropy. Hence, the latter may be dominated by slower fluctuations, while even a relatively low contribution of high-frequency 'bias' could specifically explain variance in a third variable of interest (e.g., age; see Figure 9B). Thus, beyond bias controls noted above, we argue for rigorous statistical controls when evaluating the shared and unique predictive utility of power and multiscale entropy in neural time series data.

While difficulties with scale-invariant thresholds have been noted early on, scale-invariant similarity criteria remain prevalent in recent work (e.g., Carpentier et al., 2019; Grandy et al., 2016; Hadoush, Alafeef, & Abdulhay, 2019; Heisz, Shedden, & McIntosh, 2012; Jaworska et al., 2018; Kaur et al., 2019; Miskovic et al., 2016; Mizuno et al., 2010). We hope that our practical example of coarse-scale biases thus highlights the dangers of resulting mismatches and motivate the adoption of scale-varying parameters. We perceive little justification for invariant parameters unless one specifically aims to render the MSE spectrum sensitive to PSD slopes as a function of normalization bias. While this has been a desired property in previous validations, we highlight next that such slopes are already captured within fine scales when broadband signals are characterized.

4.3 Fine-scale entropy as an index of desynchronized cortical states

Fine-scale entropy has been proposed as a signature of desynchronized cortical states (Waschke, Tune, & Obleser, 2019; Waschke et al., 2017) that describe a suppression of low-frequency power with a concurrent increase in high frequency dynamics (Contreras & Steriade, 1997; Harris & Thiele, 2011; Marguet & Harris, 2011). This synergy is thought to benefit local information processing by regulating cortical gain and is under control of the local E/I balance. Spectral slopes, characterizing the scale-free 'background' or 'noise' component of the total variance, have been proposed as an index of such E/I balance (Gao, Peterson, & Voytek, 2017; Peterson, Rosen, Campbell, Belger, & Voytek, 2018; Voytek et al., 2015). By linking fine-scale entropy to measures of scale-free background slope (Hypothesis C), we replicate previous observations of increasing fine-scale entropy with shallower slopes (Bruce et al., 2009; Miskovic et al., 2019; Sheehan, Sreekumar, Inati, & Zaghloul, 2018; Waschke et al., 2017). This is further in line with the observation that linear autocorrelative properties of the global signal (as indicated by spectral slopes) are directly related to the entropy at fine time scales (Courtiol et al., 2016; Kaffashi et al., 2008; Vakorin & McIntosh, 2012). Similar effects have been observed for permutation entropy[2] (see Supplementary Materials; Waschke et al., 2017),

[2] The observation of this link in permutation entropy further suggests that the association between PSD slopes and fine entropy is not primarily dependent on the similarity criterion, but naturally arises from the characterization of a broadband signal.

in line with generally high correspondence between entropy variants (Gudmundsson, Runarsson, Sigurdsson, Eiriksdottir, & Johnsen, 2007; Kuntzelman, Jack Rhodes, Harrington, & Miskovic, 2018). The association between broadband signal entropy and spectral slopes coheres with the notion that shallower slopes (i.e., more high frequency content) have a more 'noisy' or irregular appearance in the time domain. Thus, the shallowness of spectral slopes of the broadband signal and its pattern irregularity can be conceptualized as different perspectives on the same signal characteristic. In line with this argument, a previous study has found a strong overlap in the predictive power of spectral slopes and fine entropy on memory performance (Sheehan et al., 2018).

Crucially, our analyses suggest that fine-scale entropy does not specifically reflect the pattern similarity of high frequency dynamics, but that the presence of both high- and low-frequency dynamics at fine time scales is necessary for a link between power spectral density slopes and fine signal entropy to emerge. If low frequency information is removed and entropy becomes specific to high-frequency content, the association with power spectral density fails to persist. In this case, entropy may however provide a sensitive index of high frequency activity (Werkle-Bergner et al., 2014). While there is a general relationship between the 1/f slope and fine-scale entropy for broadband signals, it is also worth noting that our simulations suggest an influence of band-limited power on fine entropy scales. This introduces ambiguities in the interpretation of fine scales, as they appear sensitive to both arrhythmic and rhythmic content. While similar problems are encountered in the frequency domain, overt rhythmic peaks are generally excluded prior to fitting spectral slopes to increase the specificity to arrhythmic variance (Haller et al., 2018; Kosciessa et al., 2019; Peterson et al., 2018; Voytek et al., 2015; Waschke et al., 2017). Without similar procedures, this is difficult to achieve in the case of sample entropy.

In sum, our analyses provide insights into the sensitivity of fine-scale entropy to fluctuations in the synchrony of cortical states and highlight the role of slow fluctuations for such associations. Crucially, our results suggest that fine-scale entropy modulations do not specifically relate to "patterns" of neural activity at high frequencies, but that it rather arises from the presence of broadband frequency signals in traditional entropy computations. Notably, this highlights that fine-scale entropy provides a multi-scale characterization, i.e., PSD slope, even without a scale-wise recalculation due to the broadband nature of the analyzed signals.

4.4 Relevance of identified time scale mismatches to previous work

Our results of time scale mismatches have high relevance for the interpretation of neural signal entropy by highlighting associations with spectral characteristics that have not been appreciated. While some studies have shown parallel group differences between MSE and spectral power (Carpentier et al., 2019; Heisz et al., 2012; Jaworska et al., 2018; Lippe, Kovacevic, & McIntosh, 2009; McIntosh et al., 2014; Mizuno et al., 2010; Raja Beharelle, Kovacevic, McIntosh, & Levine, 2012; Sleimen-Malkoun et al., 2015; Szostakiwskyj, Willatt, Cortese, & Protzner, 2017; Takahashi et al., 2009; H. Wang et al., 2016), others identified unique entropy effects (Catarino, Churches, Baron-Cohen, Andrade, & Ring, 2011; Misic et al., 2015; Takahashi et al., 2010; Ueno et al., 2015) within which the (mis)match between time-scales and frequencies is not always readily apparent. Some of these discrepancies likely stem from a combination of the reported effects: the global similarity criterion renders MSE sensitive

to the shape of the frequency spectrum across scales, whereas the low-pass procedure leads to a strong sensitivity to low-frequency content. While many papers perform control analyses with band-limited spectral power, such mechanisms may obscure key links between the two measures.

Our results are particularly relevant for understanding multiscale entropy differences across the lifespan, although our findings and suggestions presumably apply to any scenario in which MSE is a measure of interest, such as for the assessment of clinical outcomes (e.g., Takahashi et al., 2010) or prediction of cognitive performance (e.g., McIntosh et al., 2008), independent of modality (e.g., Shafiei et al., 2019). Previous applications in the domain of aging (Courtiol et al., 2016; McIntosh et al., 2014; H. Wang et al., 2016) have shown inversions of age differences in the entropy spectrum, with older adults exhibiting lower coarse-scale entropy and higher entropy at fine time scales compared with younger adults. In the power spectrum, these effects were inverted, with older subjects showing enhanced high-, and reduced low-frequency power. This was previously taken as evidence that older adults' high-frequency dynamics were not only enhanced in magnitude, but also more unpredictable compared with younger adults' dynamics. While we replicate those results with relatively minimal resting-state data here, our analyses question the validity of these intuitive previous interpretations. In particular, our results suggest that an apparent age-related increase of coarse-scale entropy is not due to valid group differences in pattern irregularity, but results from inadequate similarity criteria that render coarse-scale entropy sensitive to high frequency power (Hypothesis D1). No coarse-scale age differences were observed with scale-varying thresholds or permutation entropy (see Supplementary Materials), in line with previous work (Sleimen-Malkoun et al., 2015). Similarly, our analyses indicate that differences in fine-scale 'pattern irregularity' rely on variations in the magnitude of slow fluctuations, and describe age-related changes in PSD slopes (Hypothesis D2). Taken together, our results thus fundamentally challenge mechanistic inferences by suggesting that previously described age differences in entropy may be minimal beyond a misattribution of traditional age differences in the magnitude of fluctuations (i.e., signal variance). This is further in line with a previous application using surrogate data that highlighted that age group differences were mainly captured by linear auto-correlative properties (see appendix in Courtiol et al., 2016).

In contrast to existing 'broad-band' applications, our narrowband analyses suggested age-related entropy increases in the posterior-occipital alpha band and decreases in central beta entropy. Whereas alpha power and MSE were inversely related and the similarity criterion showed an inverted age effect, the situation was less clear for the beta band. One explanation for such divergence is that many Fourier-based methods assume stationary sinusoidal rhythms, whereas stereotypical spectral features, particularly in the beta band (Lundqvist, Herman, Warden, Brincat, & Miller, 2018; Lundqvist et al., 2016; Sherman et al., 2016; Shin et al., 2017), are transient in time, such that time-averaged spectral power is an imperfect index of the presence of stereotypical spectral events (Jones, 2016; Kosciessa et al., 2019). In contrast, entropy should closely relate to the extent of stereotypy that is indexed by the occurrence of such non-stationary events. In line with this prediction, entropy consistently decreased with more stereotypic spectral events, suggesting that narrowband entropy can indeed reflect the (ir-)regularity of rhythmic episodes. Posterior-occipital decreases in alpha power and frequency with age are considered fundamental features of age-comparative studies (Ishii et al., 2017) that may in part reflect structural shifts in the generating networks (Knyazeva, Barzegaran,

Vildayski, & Demonet, 2018). While age-related increases in beta power are not observed as consistently (see e.g., Ishii et al., 2017 for a review), age-related increases in the relative duration of their engagement has been observed during eyes open rest (Caplan, Bottomley, Kang, & Dixon, 2015). In addition, beta-band power increases over contralateral motor cortex during rest have been hypothesized to reflect greater GABAergic inhibition in healthy aging (Rossiter, Davis, Clark, Boudrias, & Ward, 2014). While our results are not hemisphere-specific, they may similarly reflect increased inhibition in older adults, potentially reflected in the number of stereotypical beta events (Shin et al., 2017). As our aims were methods-focused here, the functional interpretation of the observed changes still necessitates caution pending further research. Our results however highlight that modulation of the spectral signal content can reveal novel, scale-specific effects regarding frequency-specific event irregularity.

4.5 Recommendations for future applications

The problems raised in the present work suggest that additional steps need to be taken to validate the accurate interpretation of scale-dependent effects and to infer a unique contribution of non-linear signal characteristics to obtained entropy estimates. We advocate the following steps (see Figure 9C): (a) use of scale-wise similarity criteria to avoid mismatches between the scale-wise signal variance and its normalization, (b) dedicated filtering into time scales of interest to probe the time-scale specificity of effects and its dependence on the spectral signal content, (c) statistical control for signal variance and (d) the future use of phase-shuffled surrogate data to confirm non-linear contributions. In combination, such controls may go a long way towards establishing non-linear effects that can be validly attributed to signal entropy at matching time scales. We discuss these steps in more detail below.

a) As noted in section 4.2, we see little motivation for the use of scale-invariant similarity criteria (i.e., fixed *r* criteria) as they introduce additional challenges without providing apparent benefits. In particular, they bias coarse-scale entropy to the extent that variance has been removed, thereby rendering traditional spectral controls difficult. Furthermore, results obtained from multiscale permutation entropy more closely aligned with results from scale-varying criteria (see Supplement), highlighting higher reliability across entropy definitions. In sum, we therefore recommend to abandon scale-invariant *r* parameters.
b) We further recommend spectral filters to validate the scale-specificity of effects. For example, if effects are observed at coarse-temporal scales with a low-pass filter, more specific high-pass filters may inform about the spectral extent of the effect. Similarly, if effects are observed at fine scales, band-pass filtering may indicate whether effects are spectrally specific (e.g., due to rhythmicity) or broad-band. For entropy estimates of slow dynamics, traditional low-pass filter settings already apply this principle. In this regard, a major advantage of estimating entropy across discontinuous segments (Grandy et al., 2016) is the ability to estimate entropy at coarse timescales with sparse neuroimaging data. This may also allow for improved comparisons with established slow fluctuations, and for characterizations of the complex dynamics of their engagement. In extreme cases, if the signal is filtered into dedicated frequency ranges, inferences regarding pattern irregularity become narrowband-specific. While this enforces a more rhythmic appearance than the raw signal may convey (S. Cole & Voytek, 2018), it makes scale-wise entropy estimates specific to the local spectral content. We note also that while we highlight the importance of

appropriate filter ranges and spectral power for the interpretation of entropy results, we do not suggest that the chosen filter settings are optimal for any particular application, and should be used with caution given that any filter will alter the underlying signal characteristics (Widmann, Schroger, & Maess, 2015). Thus, we believe that parameters should be optimized based on the spectral features of interest.

c) Furthermore, we regard statistical control as necessary to establish entropy-specific effects that are not captured by traditional linear indices (such as spectral power or signal variance). This requires an identification of the features to control for. As shown here, this should include both rhythmic frequencies and the arrhythmic signal background. Importantly, as the scale-wise *r* parameter is a crucial normalization tool, it should at the very least be controlled for. Importantly, the choice of features may further be aided by comparing effect topographies of spectral power and entropy, as done here. An important point to note is the relevance of statistical controls for relations to third variables (see Figure 9B). While some studies highlight scale-dependent associations of entropy with power, a large amount of shared variance (e.g., of coarse-scale entropy with slow frequency power) does not guarantee that a smaller portion of residual variance (e.g., shared with high frequency biases; see section 4.2) relates to effects of interest. This is equally relevant for identifying unique non-linear contributions. For example, while we observed moderate associations between band-specific rhythm events and entropy here, this non-redundant association nevertheless leaves room for the two measures to diverge in relation to third variables. Hence, they are related but may not always be redundant. This is in line with prior work (Courtiol et al., 2016) showing that despite a dominant influence of linear characteristics on entropy estimates, non-linear contributions, uniquely explained a (smaller) portion of entropy variance. Hence, specific controls are necessary to indicate unique non-linear effects that may otherwise be obscured by potentially dominant linear signal characteristics.

d) Finally, a principled way to dissociate non-linear signal characteristics from linear signal variance is the use of phase-shuffled surrogate data (Garrett, Grandy, & Werkle-Bergner, 2014; Grandy, Garrett, Lindenberger, & Werkle-Bergner, 2013; Theiler, Eubank, Longtin, Galdrikian, & Farmer, 1992), as is common practice in connectivity analyses (Aru et al., 2015). Phase randomization effectively alters original time series patterns while preserving the original power spectrum of the data. While this has been done in select entropy applications (e.g., appendix of Courtiol et al., 2016; Vakorin & McIntosh, 2012), and is frequently used to highlight entropy's non-linear sensitivity (e.g., Miskovic et al., 2016; Shafiei et al., 2019), it has not become common practice, likely due to high computational demands. A two-tier analysis strategy may overcome such limitations by first reducing data dimensionality. Specifically, in an initial stage, MSE may be used to explore potentially non-linear effects in the data. Then, a more focused (and therefore lower-dimensional) confirmatory analysis could be conducted with a selective focus on the relevant time scales or channels, using surrogate data to verify the contribution of non-linear signal characteristics.

5 Conclusions

Many inferences regarding neural multiscale entropy rely on the assumption that estimates uniquely relate to pattern irregularity at specific temporal scales. Here we show that both

assumptions may be invalid depending on the consideration of signal normalization and spectral content. Using simulations and empirical examples, we highlight how power differences can introduce entropy effects that are inversely mapped in time scale (i.e., differences in the high frequency power may be reflected in coarse entropy and vice versa; see Figure 9A). As these results suggest fundamental challenges to traditional analysis procedures and inferences, we highlight the need to test for unique entropy effects (Figure 9B) and recommend best practices and sanity checks (Figure 9C) to increase confidence in the complementary value of pattern (ir)regularity for cognitive neuroscience. While the claim has been made that "it would be unreasonable simply to reduce sample entropy to autocorrelation, spectral power, non-stationarity or any of their combinations" (Vakorin & McIntosh, 2012), it is plausible that in any given application, one or more of these contributors could suffice to mechanistically explain entropy effects of interest. We thus propose that differences in sample entropy may be taken as a starting point to explore the linear and nonlinear features that may (alone or in conjunction) explain the entropy differences (e.g., Simpraga et al., 2017), thereby proceeding from sensitivity to mechanistic specificity. As neural signal entropy is often a behaviorally relevant marker, we believe that a convergence with extant measures and indication of unique non-linear predictive utility supports the quest for reliable mechanistic indicators of brain dynamics across the lifespan, and in relation to cognition, health, and disease.

6 Funding

This study was conducted within the 'Lifespan Neural Dynamics Group' within the Max Planck UCL Centre for Computational Psychiatry and Ageing Research in the Max Planck Institute for Human Development (MPIB) in Berlin, Germany. DDG and NAK were supported by an Emmy Noether Programme grant (to DDG) from the German Research Foundation, and by the Max Planck UCL Centre for Computational Psychiatry and Ageing Research. JQK is a pre-doctoral fellow supported by the International Max Planck Research School on Computational Methods in Psychiatry and Ageing Research (IMPRS COMP2PSYCH). The participating institutions are the Max Planck Institute for Human Development, Berlin, Germany, and University College London, London, UK. For more information, see https://www.mps-ucl-centre.mpg.de/en/comp2psych.

7 Acknowledgements

We thank our research assistants and participants for their contributions to the present work. The authors declare that there are no conflicts of interest.

8 Data availability statement

The raw data will be made available on osf.io, and code used to replicate the analyses in the paper will be published on github.com upon acceptance. The code implementing the MSE algorithm is available on https://github.com/LNDG/mMSE.

References

Al-Nashash, H., Sabesan, S., Krishnan, B., George, J., Tsakalis, K., Iasemidis, L., & Tong, S. (2009). Single-Channel EEG Analysis. In S. Tong & N. Thakor (Eds.), *Quantitative EEG Analysis Methods and Clinical Applications* (pp. 73-90): Artech House.

Allen, E. A., Erhardt, E. B., & Calhoun, V. D. (2012). Data Visualization in the Neurosciences: Overcoming the Curse of Dimensionality. *Neuron, 74*(4), 603-608. doi:10.1016/j.neuron.2012.05.001

Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., . . . Vicente, R. (2015). Untangling cross-frequency coupling in neuroscience. *Current Opinion in Neurobiology, 31*, 51-61. doi:10.1016/j.conb.2014.08.002

Bell, A. J., & Sejnowski, T. J. (1995). An Information Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation, 7*(6), 1129-1159. doi:10.1162/neco.1995.7.6.1129

Brandner, F. A. (1933). A test of the significance of the difference of the correlation coefficients in normal bivariate samples. *Biometrika, 25*, 102-109. doi:10.1093/biomet/25.1-2.102

Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience, 20*(3), 340-352. doi:10.1038/nn.4497

Bruce, E. N., Bruce, M. C., & Vennelaganti, S. (2009). Sample Entropy Tracks Changes in Electroencephalogram Power Spectrum With Sleep State and Aging. *Journal of Clinical Neurophysiology, 26*(4), 257-266. doi:10.1097/WNP.0b013e3181b2f1e3

Buzsaki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304*(5679), 1926-1929. doi:10.1126/science.1099745

Buzsaki, G., Logothetis, N., & Singer, W. (2013). Scaling Brain Size, Keeping Timing: Evolutionary Preservation of Brain Rhythms. *Neuron, 80*(3), 751-764. doi:10.1016/j.neuron.2013.10.002

Buzsaki, G., & Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. *Nature Reviews Neuroscience, 15*(4), 264-278. doi:10.1038/nrn3687

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. doi: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.

Carpentier, S. M., McCulloch, A. R., Brown, T. M., Ritter, P., Wang, Z., Salimpoor, V., . . . McIntosh, A. R. (2019). Complexity matching: brain signals mirror environment information patterns during music listening and reward. *bioRxiv*.

Catarino, A., Churches, O., Baron-Cohen, S., Andrade, A., & Ring, H. (2011). Atypical EEG complexity in autism spectrum conditions: A multiscale entropy analysis. *Clinical Neurophysiology, 122*(12), 2375-2383. doi:10.1016/j.clinph.2011.05.004

Cole, S., & Voytek, B. (2018). Cycle-by-cycle analysis of neural oscillations. *bioRxiv*.

Cole, S. R., & Voytek, B. (2017). Brain Oscillations and the Importance of Waveform Shape. *Trends in Cognitive Sciences, 21*(2), 137-149. doi:10.1016/j.tics.2016.12.008

Contreras, D., & Steriade, M. (1997). Synchronization of low-frequency rhythms in corticothalamic networks. *Neuroscience, 76*(1), 11-24.

Costa, M., Goldberger, A. L., & Peng, C. K. (2002). Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters, 89*(6). doi:10.1103/PhysRevLett.89.068102

Costa, M., Goldberger, A. L., & Peng, C. K. (2004). Comment on "Multiscale entropy analysis of complex physiologic time series" - Reply. *Physical Review Letters, 92*(8). doi:10.1103/PhysRevLett.92.089804

Costa, M., Goldberger, A. L., & Peng, C. K. (2005). Multiscale entropy analysis of biological signals. *Physical Review E, 71*(2). doi:10.1103/PhysRevE.71.021906

Courtiol, J., Perdikis, D., Petkoski, S., Muller, V., Huys, R., Sleimen-Malkoun, R., & Jirsa, V. K. (2016). The multiscale entropy: Guidelines for use and interpretation in brain signal analysis. *Journal of Neuroscience Methods, 273*, 175-190. doi:10.1016/j.jneumeth.2016.09.004

Deco, G., Jirsa, V. K., & McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience, 12*(1), 43-56. doi:10.1038/nrn2961

Deco, G., Jirsa, V. K., & McIntosh, A. R. (2013). Resting brains never rest: computational insights into potential cognitive architectures. *Trends Neurosci, 36*(5), 268-274. doi:10.1016/j.tins.2013.03.001

Diaz, J., Bassi, A., Coolen, A., Vivaldi, E. A., & Letelier, J. C. (2018). Envelope analysis links oscillatory and arrhythmic EEG activities to two types of neuronal synchronization. *Neuroimage, 172*, 575-585. doi:10.1016/j.neuroimage.2018.01.063

Folstein, M. F., Robins, L. N., & Helzer, J. E. (1983). The Mini-Mental State Examination. *Archives of General Psychiatry, 40*(7), 812-812.

Fries, P. (2009). Neuronal Gamma-Band Synchronization as a Fundamental Process in Cortical Computation. *Annual Review of Neuroscience, 32*, 209-224. doi:10.1146/annurev.neuro.051508.135603

Gao, R., Peterson, E. J., & Voytek, B. (2017). Inferring synaptic excitation/inhibition balance from field potentials. *Neuroimage, 158*, 70-78. doi:10.1016/j.neuroimage.2017.06.078

Garrett, D. D., Grandy, T. H., & Werkle-Bergner, M. (2014). *The neural forest and the trees: On distinguishing the variance of a brain signal from its information content*. Paper presented at the Annual Alpine Brain Imaging Meeting, Champéry, Switzerland.

Garrett, D. D., Samanez-Larkin, G. R., MacDonald, S. W., Lindenberger, U., McIntosh, A. R., & Grady, C. L. (2013). Moment-to-moment brain signal variability: a next frontier in human brain mapping? *Neurosci Biobehav Rev, 37*(4), 610-624. doi:10.1016/j.neubiorev.2013.02.015

Grandy, T. H., Garrett, D. D., Lindenberger, U., & Werkle-Bergner, M. (2013). *Exploring the limits of complexity measures for the analysis of age differences in neural signals*. Paper presented at the Dallas Aging and Cognition Conference, Dallas, TX, USA.

Grandy, T. H., Garrett, D. D., Schmiedek, F., & Werkle-Bergner, M. (2016). On the estimation of brain signal entropy from sparse neuroimaging data. *Scientific Reports, 6*. doi:10.1038/srep23073

Gudmundsson, S., Runarsson, T. P., Sigurdsson, S., Eiriksdottir, G., & Johnsen, K. (2007). Reliability of quantitative EEG features. *Clinical Neurophysiology, 118*(10), 2162-2171. doi:10.1016/j.clinph.2007.06.018

Hadoush, H., Alafeef, M., & Abdulhay, E. (2019). Brain Complexity in Children with Mild and Severe Autism Spectrum Disorders: Analysis of Multiscale Entropy in EEG. *Brain Topogr*. doi:10.1007/s10548-019-00711-1

Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., . . . Voytek, B. (2018). Parameterizing neural power spectra. *bioRxiv*.

Harris, K. D., & Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience, 12*(9), 509-523. doi:10.1038/nrn3084

He, B. Y. J. (2014). Scale-free brain activity: past, present, and future. *Trends in Cognitive Sciences, 18*(9), 480-487. doi:10.1016/j.tics.2014.04.003

Heisz, J. J., Shedden, J. M., & McIntosh, A. R. (2012). Relating brain signal variability to knowledge representation. *Neuroimage, 63*(3), 1384-1392. doi:10.1016/j.neuroimage.2012.08.018

Hipp, J. F., & Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00338

Ishii, R., Canuet, L., Aoki, Y., Hata, M., Iwase, M., Ikeda, S., . . . Ikeda, M. (2017). Healthy and Pathological Brain Aging: From the Perspective of Oscillations, Functional Connectivity, and Signal Complexity. *Neuropsychobiology, 75*(4), 151-161. doi:10.1159/000486870

Jaworska, N., Wang, H. Y., Smith, D. M., Blier, P., Knott, V., & Protzner, A. B. (2018). Pre-treatment EEG signal variability is associated with treatment success in depression. *Neuroimage-Clinical, 17*, 368-377. doi:10.1016/j.nicl.2017.10.035

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

Kaffashi, F., Foglyano, R., Wilson, C. G., & Loparo, K. A. (2008). The effect of time delay on Approximate & Sample Entropy calculations. *Physica D-Nonlinear Phenomena, 237*(23), 3069-3074. doi:10.1016/j.physd.2008.06.005

Kaur, Y., Ouyang, G., Junge, M., Sommer, W., Liu, M., Zhou, C., & Hildebrandt, A. (2019). The reliability and psychometric structure of Multi-Scale Entropy measured from EEG signals at rest and during face and object recognition tasks. *J Neurosci Methods, 326*, 108343. doi:10.1016/j.jneumeth.2019.108343

Keitel, A., & Gross, J. (2016). Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *Plos Biology, 14*(6). doi:10.1371/journal.pbio.1002498

Kessler, J., Markowitsch, H., & Denzler, P. (2000). *Mini-mental-status-test (MMST)*. Göttingen: Beltz Test GMBH.

Knyazeva, M. G., Barzegaran, E., Vildayski, V. Y., & Demonet, J. F. (2018). Aging of human alpha rhythm. *Neurobiology of Aging, 69*, 261-273. doi:10.1016/j.neurobiolaging.2018.05.018

Kosciessa, J. Q., Grandy, T. H., Garrett, D. D., & Werkle-Bergner, M. (2019). Single-trial characterization of neural rhythms: potentials and challenges. *bioRxiv*.

Kuntzelman, K., Jack Rhodes, L., Harrington, L. N., & Miskovic, V. (2018). A practical comparison of algorithms for the measurement of multiscale entropy in neural time series data. *Brain Cogn, 123*, 126-135. doi:10.1016/j.bandc.2018.03.010

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*. doi:10.3389/fpsyg.2013.00863

Leirer, V. M., Wienbruch, C., Kolassa, S., Schlee, W., Elbert, T., & Kolassa, I. T. (2011). Changes in cortical slow wave activity in healthy aging. *Brain Imaging and Behavior, 5*(3), 222-228. doi:10.1007/s11682-011-9126-3

Lippe, S., Kovacevic, N., & McIntosh, A. R. (2009). Differential maturation of brain signal complexity in the human auditory and visual system. *Frontiers in Human Neuroscience, 3*. doi:10.3389/neuro.09.048.2009

Lopes da Silva, F. (2013). EEG and MEG: relevance to neuroscience. *Neuron, 80*(5), 1112-1128. doi:10.1016/j.neuron.2013.10.017

Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L., & Miller, E. K. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nature Communications, 9*. doi:10.1038/s41467-017-02791-8

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and Beta Bursts Underlie Working Memory. *Neuron, 90*(1), 152-164. doi:10.1016/j.neuron.2016.02.028

Marguet, S. L., & Harris, K. D. (2011). State-Dependent Representation of Amplitude-Modulated Noise Stimuli in Rat Auditory Cortex. *Journal of Neuroscience, 31*(17), 6414-6420. doi:10.1523/Jneurosci.5773-10.2011

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190. doi:10.1016/j.jneumeth.2007.03.024

McIntosh, A. R. (2019). Neurocognitive Aging and Brain Signal Complexity: Oxford University Press.

McIntosh, A. R., Kovacevic, N., & Itier, R. J. (2008). Increased Brain Signal Variability Accompanies Lower Behavioral Variability in Development. *Plos Computational Biology, 4*(7). doi:10.1371/journal.pcbi.1000106

McIntosh, A. R., Vakorin, V., Kovacevic, N., Wang, H., Diaconescu, A., & Protzner, A. B. (2014). Spatiotemporal Dependency of Age-Related Changes in Brain Signal Variability. *Cerebral Cortex, 24*(7), 1806-1817. doi:10.1093/cercor/bht030

Mejias, J. F., Murray, J. D., Kennedy, H., & Wang, X. J. (2016). Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. *Science Advances, 2*(11). doi:10.1126/sciadv.1601335

Misic, B., Doesburg, S. M., Fatima, Z., Vidal, J., Vakorin, V. A., Taylor, M. J., & McIntosh, A. R. (2015). Coordinated Information Generation and Mental Flexibility: Large-Scale Network Disruption in Children with Autism. *Cerebral Cortex, 25*(9), 2815-2827. doi:10.1093/cercor/bhu082

Misic, B., Vakorin, V. A., Paus, T., & McIntosh, A. R. (2011). Functional embedding predicts the variability of neural activity. *Frontiers in Systems Neuroscience, 5*, 90. doi:10.3389/fnsys.2011.00090

Miskovic, V., MacDonald, K. J., Rhodes, L. J., & Cote, K. A. (2019). Changes in EEG multiscale entropy and power-law frequency scaling during the human sleep cycle. *Human Brain Mapping, 40*(2), 538-551. doi:10.1002/hbm.24393

Miskovic, V., Owens, M., Kuntzelman, K., & Gibb, B. E. (2016). Charting moment-to-moment brain signal variability from early to late childhood. *Cortex, 83*, 51-61. doi:10.1016/j.cortex.2016.07.006

Mizuno, T., Takahashi, T., Cho, R. Y., Kikuchi, M., Murata, T., Takahashi, K., & Wada, Y. (2010). Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. *Clinical Neurophysiology, 121*(9), 1438-1446. doi:10.1016/j.clinph.2010.03.025

Nikulin, V. V., & Brismar, T. (2004). Comment on "Multiscale entropy analysis of complex physiologic time series". *Physical Review Letters, 92*(8). doi:10.1103/PhysRevLett.92.089803

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. doi: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. doi: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*. doi:10.1155/2011/156869

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713-719. doi:10.1016/S1388-2457(00)00527-7

Park, J. H., Kim, S., Kim, C. H., Cichocki, A., & Kim, K. (2007). Multiscale entropy analysis of EEG from patients under different pathological conditions. *Fractals-Complex Geometry Patterns and Scaling in Nature and Society, 15*(4), 399-404. doi:10.1142/S0218348x07003691

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. doi:10.1016/0013-4694(89)90180-6

Peterson, E. J., Rosen, B. Q., Campbell, A. M., Belger, A., & Voytek, B. (2018). 1/f neural noise is a better predictor of schizophrenia than neural oscillations. *bioRxiv*.

Raja Beharelle, A., Kovacevic, N., McIntosh, A. R., & Levine, B. (2012). Brain signal variability relates to stability of behavior after recovery from diffuse brain injury. *Neuroimage, 60*(2), 1528-1537. doi:10.1016/j.neuroimage.2012.01.037

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology, 278*(6), H2039-H2049.

Rossiter, H. E., Davis, E. M., Clark, E. V., Boudrias, M. H., & Ward, N. S. (2014). Beta oscillations reflect changes in motor cortex inhibition in healthy ageing. *Neuroimage, 91*, 360-365. doi:10.1016/j.neuroimage.2014.01.012

Semmlow, J. L. (2008). *Biosignal and medical image processing*: CRC press.

Shafiei, G., Zeighami, Y., Clark, C. A., Coull, J. T., Nagano-Saito, A., Leyton, M., . . . Misic, B. (2019). Dopamine Signaling Modulates the Stability and Integration of Intrinsic Brain Networks. *Cerebral Cortex, 29*(1), 397-409. doi:10.1093/cercor/bhy264

Sheehan, T. C., Sreekumar, V., Inati, S. K., & Zaghloul, K. A. (2018). Signal Complexity of Human Intracranial EEG Tracks Successful Associative-Memory Formation across Individuals. *Journal of Neuroscience, 38*(7), 1744-1755. doi:10.1523/Jneurosci.2389-17.2017

Sherman, M. A., Lee, S., Law, R., Haegens, S., Thorn, C. A., Hamalainen, M. S., . . . Jones, S. R. (2016). Neural mechanisms of transient neocortical beta rhythms: Converging evidence from humans, computational modeling, monkeys, and mice. *Proceedings of the National Academy of Sciences of the United States of America, 113*(33), E4885-E4894. doi:10.1073/pnas.1604135113

Shin, H., Law, R., Tsutsui, S., Moore, C. I., & Jones, S. R. (2017). The rate of transient beta frequency events predicts behavior across tasks and species. *Elife, 6*. doi:10.7554/eLife.29086

Simpraga, S., Alvarez-Jimenez, R., Mansvelder, H. D., van Gerven, J. M. A., Groeneveld, G. J., Poil, S. S., & Linkenkaer-Hansen, K. (2017). EEG machine learning for accurate detection of cholinergic intervention and Alzheimer's disease. *Scientific Reports, 7*. doi:10.1038/s41598-017-06165-4

Sleimen-Malkoun, R., Perdikis, D., Muller, V., Blanc, J. L., Huys, R., Temprado, J. J., & Jirsa, V. K. (2015). Brain Dynamics of Aging: Multiscale Variability of EEG Signals at Rest and during an Auditory Oddball Task(1,2,3). *Eneuro, 2*(3). doi:10.1523/ENEURO.0067-14.2015

Sporns, O. (2010). *Networks of the Brain*: MIT Press.

Stam, C. J. (2005). Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical Neurophysiology, 116*(10), 2266-2301. doi:10.1016/j.clinph.2005.06.011

Stoyanov, M., Gunzburger, M., & Burkardt, J. (2011). Pink noise, 1/f (alpha) noise, and their effect on solutions of differential equations. *International Journal for Uncertainty Quantification, 1*(3), 257-278. doi:10.1615/Int.J.UncertaintyQuantification.2011003089

Szostakiwskyj, J. M. H., Willatt, S. E., Cortese, F., & Protzner, A. B. (2017). The modulation of EEG variability between internally-and externally-driven cognitive states varies with maturation and task performance. *Plos One, 12*(7). doi:10.1371/journal.pone.0181894

Takahashi, T., Cho, R. Y., Mizuno, T., Kikuchi, M., Murata, T., Takahashi, K., & Wada, Y. (2010). Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: A multiscale entropy analysis. *Neuroimage, 51*(1), 173-182. doi:10.1016/j.neuroimage.2010.02.009

Takahashi, T., Cho, R. Y., Murata, T., Mizuno, T., Kikuchi, M., Mizukami, K., . . . Wada, Y. (2009). Age-related variation in EEG complexity to photic stimulation: A multiscale entropy analysis. *Clinical Neurophysiology, 120*(3), 476-483. doi:10.1016/j.clinph.2008.12.043

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992). Testing for Nonlinearity in Time-Series - the Method of Surrogate Data. *Physica D-Nonlinear Phenomena, 58*(1-4), 77-94. doi:10.1016/0167-2789(92)90102-S

Ueno, K., Takahashi, T., Takahashi, K., Mizukami, K., Tanaka, Y., & Wada, Y. (2015). Neurophysiological basis of creativity in healthy elderly people: A multiscale entropy approach. *Clinical Neurophysiology, 126*(3), 524-531. doi:10.1016/j.clinph.2014.06.032

Vakorin, V. A., & McIntosh, A. R. (2012). Mapping the Multiscale Information Content of Complex Brain Signals. *Principles of Brain Dynamics: Global State Interactions*, 183-208.

Valencia, J. F., Porta, A., Vallverdu, M., Claria, F., Baranowski, R., Orlowska-Baranowska, E., & Caminal, P. (2009). Refined Multiscale Entropy: Application to 24-h Holter Recordings of Heart Period Variability in Healthy and Aortic Stenosis Subjects. *Ieee Transactions on Biomedical Engineering, 56*(9), 2202-2213. doi:10.1109/Tbme.2009.2021986

Vidaurre, D., Hunt, L. T., Quinn, A. J., Hunt, B. A. E., Brookes, M. J., Nobre, A. C., & Woolrich, M. W. (2018). Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. *Nature Communications, 9*. doi:10.1038/s41467-018-05316-z

Vlahou, E. L., Thurm, F., Kolassa, I. T., & Schlee, W. (2014). Resting-state slow wave power, healthy aging and cognitive performance. *Scientific Reports, 4*. doi:10.1038/srep05101

von Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology, 38*(3), 301-313. doi:10.1016/S0167-8760(00)00172-0

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-Related Changes in 1/f Neural Electrophysiological Noise. *Journal of Neuroscience, 35*(38), 13257-13265. doi:10.1523/Jneurosci.2332-14.2015

Wang, H., McIntosh, A. R., Kovacevic, N., Karachalios, M., & Protzner, A. B. (2016). Age-related Multiscale Changes in Brain Signal Variability in Pre-task versus Post-task Resting-state EEG. *J Cogn Neurosci, 28*(7), 971-984. doi:10.1162/jocn_a_00947

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev, 90*(3), 1195-1268. doi:10.1152/physrev.00035.2008

Waschke, L., Tune, S., & Obleser, J. (2019). Neural desynchronization and arousal differentially shape brain states for optimal sensory performance. *bioRxiv*.

Waschke, L., Wostmann, M., & Obleser, J. (2017). States and traits of neural irregularity in the age-varying human brain. *Scientific Reports, 7*. doi:10.1038/s41598-017-17766-4

Werkle-Bergner, M., Grandy, T. H., Chicherio, C., Schmiedek, F., Lovden, M., & Lindenberger, U. (2014). Coordinated within-Trial Dynamics of Low-Frequency

Neural Rhythms Controls Evidence Accumulation. *Journal of Neuroscience, 34*(25), 8519-8528. doi:10.1523/Jneurosci.3801-13.2014

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. doi:10.1016/j.neuroimage.2010.08.064

Widmann, A., Schroger, E., & Maess, B. (2015). Digital filter design for electrophysiological data - a practical approach. *Journal of Neuroscience Methods, 250*, 34-46. doi:10.1016/j.jneumeth.2014.08.002

Yang, A. C., Wang, S. J., Lai, K. L., Tsai, C. F., Yang, C. H., Hwang, J. P., . . . Fuh, J. L. (2013). Cognitive and neuropsychiatric correlates of EEG dynamic complexity in patients with Alzheimer's disease. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 47*, 52-61. doi:10.1016/j.pnpbp.2013.07.022

Supplementary Information for

Standard multiscale entropy reflects spectral power at mismatched temporal scales

Julian Q. Kosciessa, Niels A. Kloosterman, and Douglas D. Garrett

Email: kosciessa@mpib-berlin.mpg.de; garrett@mpib-berlin.mpg.de

This PDF file includes:

Supplementary text
Figures S1 to S4
Supplementary references

SI Methods

Calculation of multiscale permutation entropy. Sample entropy's similarity criterion makes it difficult to differentiate between rhythmic modulations of MSE via added pattern regularity or the influence on similarity criteria. For this purpose, we extended our analyses to multiscale permutation entropy, a measure that assesses pattern irregularity independent of a similarity criterion. In particular, permutation entropy describes the randomness in the occurrence of symbolic sequences (rank-order permutations) (Bandt & Pompe, 2002; Riedl, Muller, & Wessel, 2013). To investigate the correspondence between sample entropy and permutation entropy, we repeated our analyses with identical settings as described for the MSE analyses. The calculation of permutation entropy followed previous implementations (e.g., Ouyang, Li, Liu, & Li, 2013). Specifically. for a given template length m (i.e., embedding dimension, here m = 4), all m! rank-order permutations π were assessed with regard to their relative occurrence: $p(\pi) = C(\pi)/(N-(m-1)l)$, where N is the number of samples and l is the time delay/lag (here l = 1). The permutation entropy of a signal was defined as $PE = -\sum_{m=1}^{m!} p(\pi) \ln p(\pi)$. We calculated a normalized version of permutation entropy with bounds between zero and one. Specifically, complete randomness of permutation occurrence would result in values of one, whereas increasing regularity results in lower values. To assess the convergence between sample and permutation entropy, we repeated the simulations noted in the main text, and probed age differences in the traditional (i.e., low-pass) implementation.

SI Results

Dissociating between similarity criterion and spectral regularity using multiscale permutation entropy (MPE). In our MSE analyses, the intrinsic, variance-bound, similarity criterion makes it difficult to distinguish whether spectral events (e.g., narrowband rhythms) decrease entropy as a result of increasing the *r* parameter or via their contribution of added (sinusoidal) signal regularity. To probe this issue, we used multiscale permutation entropy (MPE) as a measure of signal complexity that does not use a variance-based threshold. In particular, permutation entropy assesses pattern complexity as the relative (im-)balance in the occurrence of symbolic patterns.

In simulations, rhythmicity modulated MPE in a similar fashion as MSE (Figure S4A, B). Notably, MPE did not indicate rhythm-dependent increases in entropy, although it should be noted that permutation entropy was at ceiling even at baseline. Crucially, we observed a similar decrease of entropy at fine scales in the absence of variance normalization, suggesting that added rhythmicity decreased broadband 'fine-scale' estimates due to the added rhythmic regularity. We further assessed age effects in the traditional low-pass scenario. Most notably, permutation entropy in the low-pass implementation did not exhibit an age difference at coarse (Figure S4C), in line with our suggestions that this MSE difference is exclusively induced by fixed similarity criteria. However, a fine-scale age difference was also observed in low-pass MPE (Figure S4C), suggesting that this effect is not exclusively related to the similarity criterion. As in the MSE analysis, fine-scale estimates characterized individual PSD slopes, underlining the broadband origin of the effect.

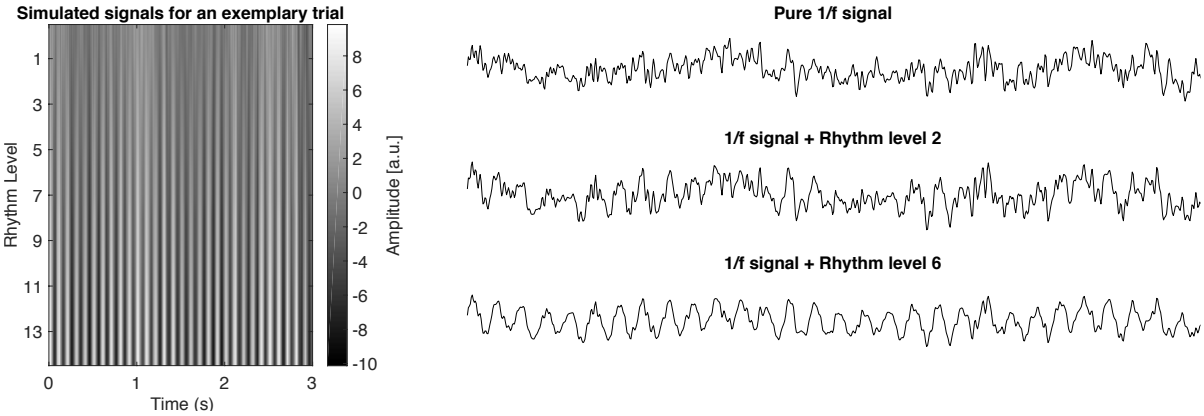

Figure S1: Examples of simulated data. Time series from an exemplary simulated trial for a pure 1/f signal pink noise signal and at different magnitudes of added alpha rhythmicity. The left presentation provides a top-down view of time-series amplitudes.

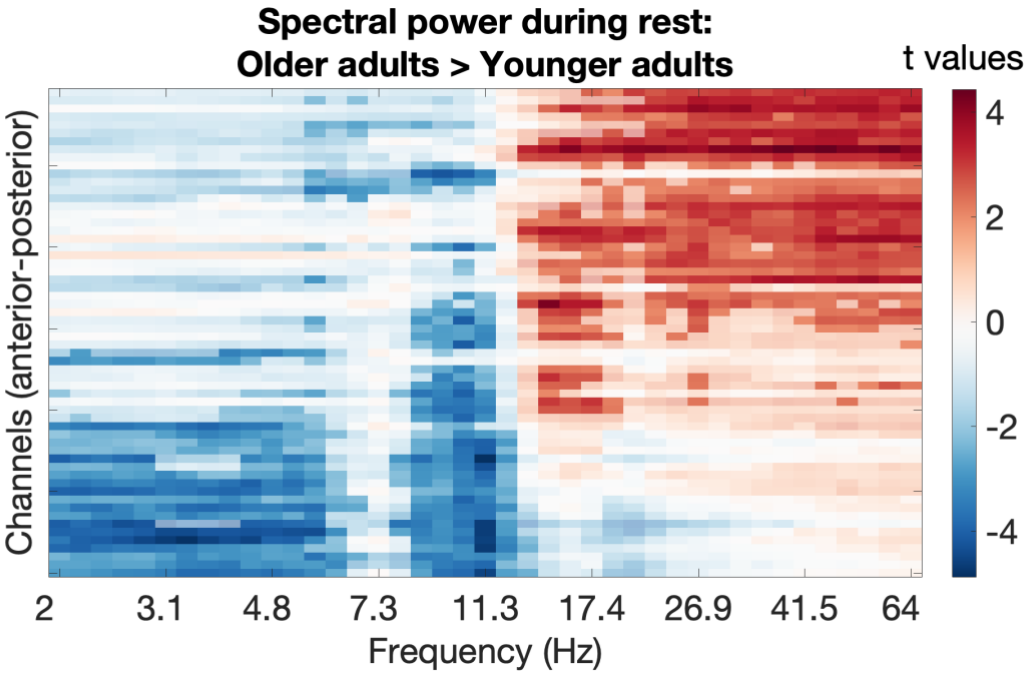

Figure S2: T-values for age group differences in spectral power (OA > YA). Statistical significance ($p < .05$) was assessed by means of cluster-based permutation tests and is indicated via opacity.

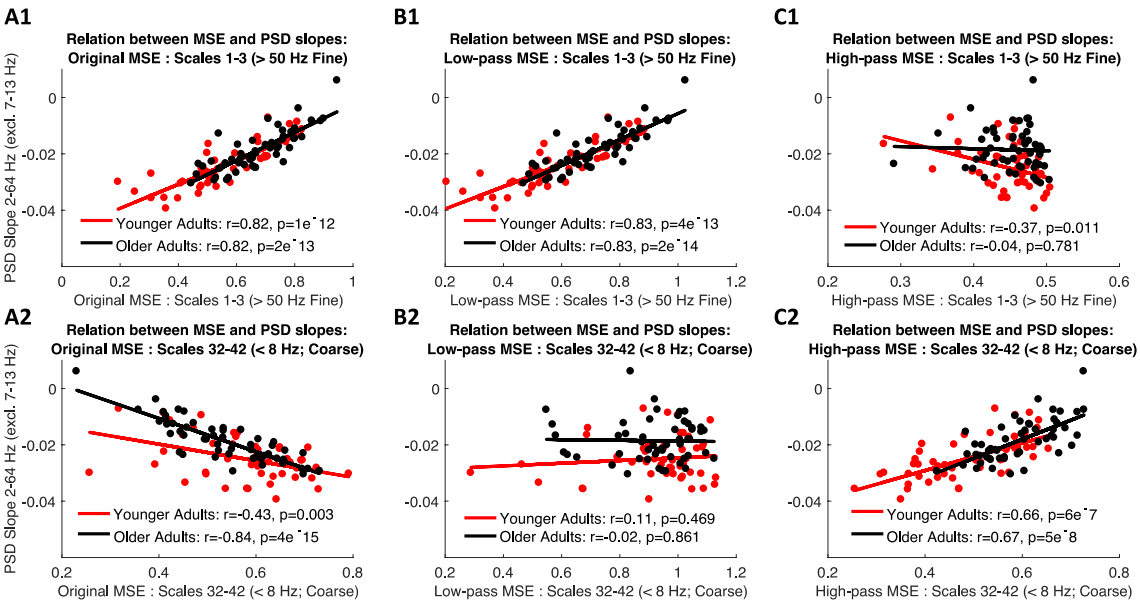

Figure S3: Methods- and scale-dependent associations between sample entropy and PSD slopes. 'Original' settings indicate a strong positive association at fine scales (A1) that turns negative at coarse scales (A2), likely due to coarse-scale biases by the scale-invariant similarity criterion. In line with this notion, scale-wise adaptation of thresholds retains the fine-scale effect (B1), while abolishing the coarse-scale inversion (B2). Crucially, the entropy of exclusively high-frequency signals does not positively relate to PSD slopes (C1), whereas the association reemerges once slow fluctuations are added into the signal (C2).

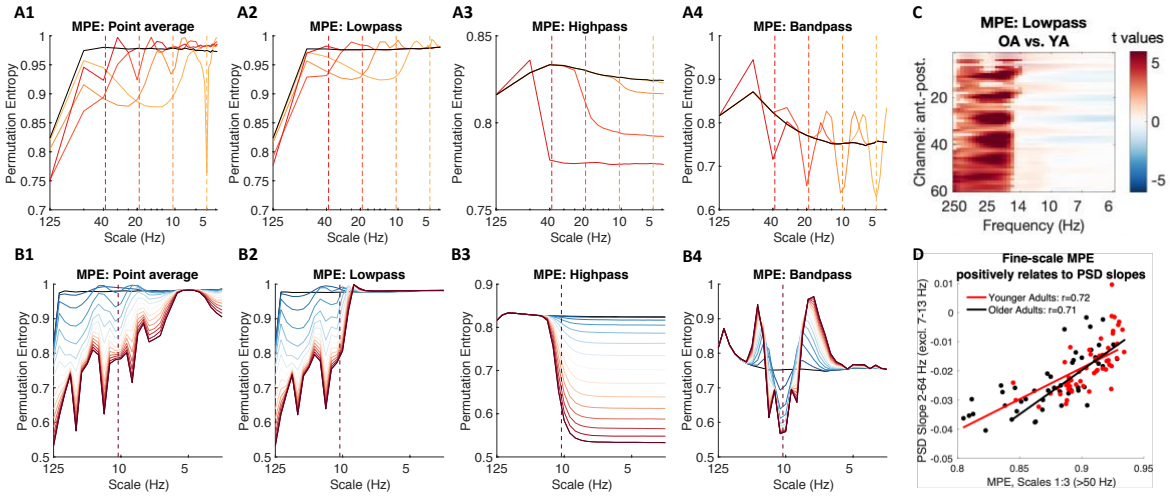

Figure S4. Permutation entropy reproduces dominant effects from sample entropy analysis. (A) The rhythmic representation on multiscale permutation entropy (MPE) is similar to that observed in MSE. (A1) Frequency-wise rhythm simulations indicate frequency-dependent decreases in permutation entropy. (A2) Low-pass filtering results in decreased entropy at frequencies above the simulated frequency, whereas the opposite effect is observed when using high-pass filters (A3). A difference to low-pass MSE is the absence of entropy increases above baseline. (B) Amplitude simulations of alpha rhythms indicate similar parametric effects as for sample entropy. The narrow bandpass filter introduces spurious entropy increases around filter boundaries (B4). (C) Lowpass MPE indicates higher fine-scale entropy, but no decreased coarse-scale entropy, for older compared to younger adults, in line with MSE results with scale-varying similarity criteria. (D) Fine-scale low-pass permutation entropy relates to individual PSD slopes.

Supplementary References

Bandt, C., & Pompe, B. (2002). Permutation entropy: A natural complexity measure for time series. *Physical Review Letters, 88*(17). doi:10.1103/PhysRevLett.88.174102

Ouyang, G. X., Li, J., Liu, X. Z., & Li, X. L. (2013). Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis. *Epilepsy Research, 104*(3), 246-252. doi:10.1016/j.eplepsyres.2012.11.003

Riedl, M., Muller, A., & Wessel, N. (2013). Practical considerations of permutation entropy. *European Physical Journal-Special Topics, 222*(2), 249-262. doi:10.1140/epjst/e2013-01862-7