


Algorithmic Complexity of EEG for Prognosis of Neurodegeneration in Idiopathic Rapid Eye Movement Behavior Disorder (RBD)

GIULIO RUFFINI ¹, DAVID IBAÑEZ,¹ ELENI KROUPI,¹ JEAN-FRANÇOIS GAGNON,² JACQUES MONTPLAISIR,² RONALD B. POSTUMA,² MARTA CASTELLANO,¹ and AURELI SORIA-FRISCH¹

¹Starlab Barcelona, Barcelona, Spain; and ²Centre for Advanced Research in Sleep Medicine, Hôpital du Sacré-Cœur de Montréal, Montreal, Canada

(Received 3 April 2018; accepted 6 August 2018; published online 30 August 2018)

Associate Editor Xiaoxiang Zheng oversaw the review of this article.

Abstract—Idiopathic rapid eye movement sleep behavior disorder (RBD) is a serious risk factor for neurodegenerative processes such as Parkinson’s disease (PD). We investigate the use of EEG algorithmic complexity derived metrics for its prognosis. We analyzed resting state EEG data collected from 114 idiopathic RBD patients and 83 healthy controls in a longitudinal study forming a cohort in which several RBD patients developed PD or dementia with Lewy bodies. Multichannel data from ~ 3 min recordings was converted to spectrograms and their algorithmic complexity estimated using Lempel–Ziv–Welch compression. Complexity measures and entropy rate displayed statistically significant differences between groups. Results are compared to those using the ratio of slow to fast frequency power, which they are seen to complement by displaying increased sensitivity even when using a few EEG channels. Poor prognosis in RBD appears to be associated with decreased complexity of EEG spectrograms stemming in part from frequency power imbalances and cross-frequency amplitude algorithmic coupling. Algorithmic complexity metrics provide a robust, powerful and complementary way to quantify the dynamics of EEG signals in RBD with links to emerging theories of brain function stemming from algorithmic information theory.

Keywords—Parkinson’s disease, Complexity, Algorithmic complexity, LZW, Lempel–Ziv–Welch compression, RBD, Time-frequency analysis, Dementia with Lewy bodies, DLB.

INTRODUCTION

Rapid eye movement (REM) behavior disorder (RBD) is a male-predominant sleep disorder characterized by vivid dreaming together with dream-acting behaviors. *Idiopathic* RBD occurs in the absence of any identified neurological disease or other cause

and its clinical course is generally chronic progressive.¹⁰ Several studies have shown that a majority of patients diagnosed with idiopathic RBD will eventually be diagnosed with a neurological disorder such as Parkinson disease (PD) and dementia with Lewy bodies (DLB).^{10,16,24} Because of this, idiopathic RBD has been suggested as a prodromal factor of the synucleinopathies PD, DLB and less frequently multiple system atrophy (MSA).¹⁶ Indeed, idiopathic RBD is now considered to be an early stage of α -synucleinopathy that can provide an early view of future brain health.¹⁴ On the other hand, RBD has an estimated prevalence of 15–60% in PD and has been proposed as characterizing a subtype of PD with relatively poor prognosis, reflecting a brainstem-dominant route of pathology progression (see Ref. 20 and references therein) with a higher risk for dementia or hallucinations. PD with RBD is characterized by more profound and extensive pathology—one not limited to the brainstem—with higher synuclein deposition in both cortical and sub-cortical regions.

The human brain can be modeled as a highly dimensional complex dynamical system which instantiates electrochemical communication and computation. Electroencephalographic (EEG) and magnetoencephalographic (MEG) signals are rich with information associated with these processes, and are accessible non-invasively. To a large extent, progress in the analysis of such signals has been driven by the study of classical temporal and spectral features in electrode space, and applied to the study the human brain in both health and disease. For example, the “slowing down” of EEG is known to characterize neurodegenerative diseases,^{9,26} and the slow to fast ratio (the ratio of power in delta and theta bands to

Address correspondence to Giulio Ruffini, Starlab Barcelona, Barcelona, Spain. Electronic mail: giulio.ruffini@starlab-int.com
Starlab Barcelona, <http://starlab.es>.

alpha and beta) has shown good discriminatory sensitivity.^{21,26} However, brain activity measurements exhibit non-linear dynamics and non-stationarity, limiting the usefulness of classical, linear approaches and calling for the use of novel methods capable of exploiting underlying spatiotemporal hierarchical structures. Deep learning techniques in particular and neural networks in general are bio-inspired by neural structure and function—the same biological systems generating the electric signals we aim to decode—and should be well suited for the task. In past work, for example, we studied a particular class of recurrent neural networks called Echo State Networks (ESNs) combining the power of networks for classification of temporal patterns and ease of training, reasoning that as they implement non-linear dynamics with memory they may be ideally poised for the classification of complex EEG time series data. Starting from a dataset of recordings of resting state EEG from idiopathic RBD patients who later converted to PD and from healthy controls (HC), we showed that using such recurrent networks using temporal series of EEG power led to a prognosis classification accuracy of 85% in the binary, balanced classification problem.²⁷

With the goal of developing metrics that capture non-linear dynamics from EEG signals, here we explore a different feature extracted from EEG data—upper bounds on algorithmic complexity. The notion of “algorithmic information” is mathematically formalized by the concept of algorithmic complexity or Kolmogorov complexity (\mathcal{K})³⁰: *the Kolmogorov complexity of a string is the length of the shortest program capable of generating it*. The precise length of this program depends on programming language only up to a string-independent constant. Algorithmic information is notoriously difficult to estimate. Here we rely on the notion of complexity as measured from Lempel–Ziv–Welch compression (LZW) or entropy rate.

As discussed in Ref. 30 and references therein, the healthy brain generates apparently complex (entropic) data. Complexity should be associated to cognitive health and conscious state, and complexity appears to decrease with age (see, e.g., Ref. 18 and references therein). As a universal computer, a brain may produce many different types of patterns—from simple to highly entropic. A healthy brain engaging in modeling, prediction and interaction with the world will more naturally produce complex-looking, highly entropic data which can presumably be measured from behavior or directly in brain activity. Entropy and LZW represent direct measures of apparent complexity and can be applied to, e.g., electrophysiological or metabolic brain data.^{3,5,15,31,32} We hypothesize here that global apparent EEG algorithmic complexity or entropy rate in our dataset will decrease with worsening

neurodegenerative disease prognosis or progression, in a similar manner to that already reported in the PD literature¹² and in the case of Alzheimer’s disease.^{1,8,11} As a starting point, we further assume that algorithmically relevant aspects in EEG data are present in compositional features in the time-frequency spectral amplitude representation.

METHODS

Participants

The used dataset has been previously described in Refs. 26 and 27, so an abridged description is provided here. Idiopathic RBD patients (called henceforth RBD for data analysis labeling) and healthy controls were recruited at the Center for Advanced Research in Sleep Medicine of the Hôpital du Sacré-Cœur de Montréal. All patients with the required full EEG montage for resting-state EEG recording at baseline and with at least one follow-up examination (after a mean of 4 years) after the baseline visit were included in the study (see Table 1 for baseline demographic and clinical data). The first valid EEG for each patient enrolled in the study was considered as baseline. Participants also underwent a complete neurological examination by a neurologist specialized in movement disorders and a cognitive assessment by a neuropsychologist. No controls reported abnormal motor activity during sleep or showed cognitive impairment on neuropsychological testing. The protocol was approved by the hospital’s ethics committee, and all participants gave their written informed consent to participate. For mode details, the reader is directed to Ref. 26.

EEG Dataset

The data was collected by the Hôpital du Sacré-Coeur, Montréal as described in Ref. 26, and consisted of resting-state EEG collected from recently awaken patients and healthy controls (within 30 min of waking up) using a subset of 14 scalp electrodes in the 10–20 EEG system (C3, C4, F3, F4, F7, F8, O1, O2, P3, P4, P7, P8, T7, T8). The recording protocol consisted of conditions with periods of “eyes open” of variable duration (approximately 2 min) followed by periods of “eyes closed” in which patients were not given any particular task. Resting EEG signals were digitized with 16-bit resolution at a sampling rate of 256 S/s. The amplification device used a hardware band pass filter between 0.3 and 100 Hz and a line-noise notch filter at 60 Hz. All recordings were referenced to linked ears.

After quality check with rejection of subjects with insufficient data or subjects with abnormal cognition at

the time of acquisition, the dataset (of initially 213 subjects) consisted of eyes-closed resting EEG data from a total of 114 patients who were diagnosed with idiopathic RBD at the time of acquisition and 83 healthy controls without sleep complaints in which RBD was excluded. EEG data was collected in every patient at baseline, i.e., when they were still diagnosed with idiopathic RBD. After 1–10 years of clinical follow-up, 19 patients developed Parkinson disease (PD), 12 Lewy body dementia (DLB) and the remaining 83 ones remained idiopathic RBD. Summary demographic data are provided in Table 1. Group averaged EEG spectral power density plots for each group are provided in Fig. 1.

Data Preparation

First, the EEG dataset from each subject was converted into a spectrogram “stack” of “frames” (time frequency bi-dimensional arrays) (see Fig. 2), with

TABLE 1. Baseline sociodemographic and clinical data for the four groups.

Group/data	HC	RBD	PD	DLB
N (F/M)	83 (30/53)	83 (18/65)	19 (5/14)	12 (2/10)
(age)	66 ± 9	67 ± 7	67 ± 7	71 ± 6
(follow-up)	1 ± 1	4 ± 3	4 ± 2	4 ± 2

Number of subjects per group with gender breakdown (female/male), mean and standard deviation for age and follow-up time (in years).

HC healthy controls, *RBD* idiopathic RBDs who remained idiopathic RBD on follow-up, *PD* idiopathic RBD who converted to PD on follow-up, *DLB* idiopathic RBD who converted to DLB on follow-up.

each stack representing about 3 min of eyes-closed resting EEG. Each multichannel spectrogram frame was generated from 20 s long artifact free sequences (with an artifact threshold of 100 μ V after detrending of each 20 s segment) from each of the 14 EEG channels using a sliding window of 1 s between frames. This length of time was chosen to capture dynamic aspects of the power envelopes of the signals, but using shorter windows (10 s) did not change the results very much. To generate the spectrogram stack frames, EEG data for each channel was processed using a Fast Fourier Transform (FFT) after detrending 1-s data blocks with a Hann window (50% overlap). The FFT with resolution was thus of 1Hz, and we kept the frequencies in the range 1–50 Hz.

The resulting data frames are “tensors” of the form [channels (14)] \times [FFT bins (50)] \times [Time bins (39)], and the full stack is four dimensional, with the fourth dimension indexing frame number or epoch. The exact temporal span of each stack varies, as we fixed the number of frames in each stack as a tradeoff between quantity and exclusion of subjects. Fixing the number of frames in each subject stack is important, since measures that estimate complexity are sensitive to data length.^{19,36}

Starting from the spectrogram frame stack dataset we carried out an analysis of the complexity for each subject, as we discuss next.

Lempel–Ziv–Welch Complexity and Entropy Rate

As discussed above, in order to provide an upper bound to description length or algorithmic complexity we use LZW compression. The LZW algorithm seeks increasingly long reappearing patterns in the data, and

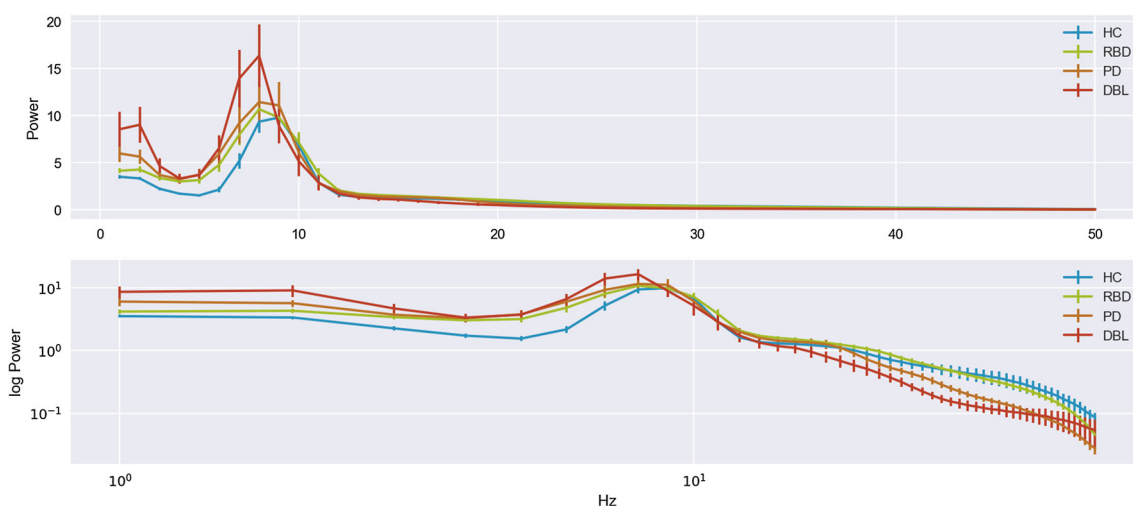


FIGURE 1. Group PSDs. Average power Spectral Density (all channels) for each group with standard error of the mean (SEM), displaying the characteristic “slowing” of EEG with more power at low frequencies.

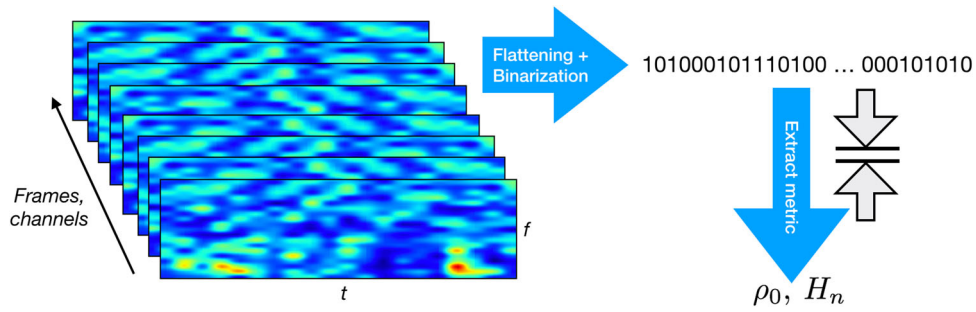


FIGURE 2. Computation of description length and entropy metrics from subject spectrogram frame stack. To compute a global measure of complexity for each subject, the frame stack $S(t, f, ch, F)$ is flattened and binarized prior compression or entropy computation, where t denotes time in each spectrogram, f frequency, ch channel, and F frame number.

compresses information by reusing them: instead of rewriting a previously seen sequence, it will refer to the identifier of the one seen last.^{22,35} After applying LZW to a string of length n , we are provided with a set of words (or phrases, as they are sometimes called) that form a dictionary of length $c(n)$, which can be thought of as a complexity counter.³⁶ The length of the compressed string will, in general, be $l_{LZW}n$. We note that LZW can be used to obtain an upper bound on algorithmic complexity, but, given its reduced programming repertoire (LZW is the Kolmogorov complexity computed with a limited set of programs that only allow copy and insertion in strings¹⁹), it will fail to compress random-looking data generated by simple, but highly recursive programs, e.g., the sequence of binary digits of π , with more sophisticated regularities than sequence repetition (deep programs²⁷). As an example of how LZW or entropy rate are limited tools for compression, and therefore coarse approximations to algorithmic complexity, consider a string and its bit-flipped version, or a “time-reversed” version in a file with the ordering of symbols temporally inverted, or a string and “time-dilated” string where each symbol is repeated, say twice. Such simple algorithmic manipulations will not be detected and exploited by LZW. Despite these limitations, such compression algorithms can be useful, as we shall see below.

Let $H_0(p) = -p \log p - (1 - p) \log(1 - p)$ denote the univariate or zero order entropy, with p the probability of a Bernoulli (binary) process (Markov chain of order zero). By the entropy rate of the stochastic process $\{X_i\}$, we will mean

$$\mathcal{H}(X) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, \dots, X_n),$$

when this limit exists, with H denoting the usual multivariate entropy of X , $H(X) = -E_X[\log(P(X))]$. We note that entropy rate of a stochastic processes is non-increasing as a function of order, that is, $0 \leq \mathcal{H} \leq \dots \leq H_q \leq \dots \leq H_0 \leq 1$. A fundamental relation is that description length computed from an

algorithm closely related to LZW (LZ77, see Ref. 7 for a brief review of the Lempel–Ziv variants) is related to entropy rate, that is,

$$l_{LZW} \approx c(n) \log_2 c(n) \rightarrow n\mathcal{H}.$$

(see Refs. 7 and 29). Thus, we expect that the description length of the sequence encoded by LZW can be approximated by the number of phrases times the number of bits needed to identify a seen phrase. In order to apply LZW we need first to digitize the data. Here we binarize the data to reduce the length of strings needed for LZW to stabilize. A reasonable strategy in analysis of algorithmic information of EEG data is to preserve as much information as possible in the resulting transformed string. In this sense, using methods that maximize the entropy of the resulting series are recommended, such as using the median for thresholding (this is guaranteed to yield $H_0 = 1$ and makes the result independent of overall scale).

Two associated metrics are commonly used in the field: $c(n)$ and l_{LZW} . Of the two, the latter is more closely related to Kolmogorov complexity or description length. Both contain similar information (in fact one is a monotonic function of the other). A natural manner to normalize this metric is to divide description length by the original string length n , $\rho_0 = l_{LZW}/n \rightarrow \mathcal{H}$, with units of bits per character. This is the LZW compression ratio metric we will use here. For details and code used to compute LZW complexity see Ref. 29.

Data Flattening and Binarization

LZW and entropy rate can be sensitive to inherent choices in data flattening (that is, how the multidimensional 4D arrays are “flattened” into a one dimensional string). Here we have focused on the search for temporal patterns. For this purpose, the stack data tensor for each subject was converted to a one-dimensional array maintaining temporal adja-

gency in each frame (with flattening in this order: time in each frame, epoch, channel, frequency). The spectrogram hypercube flattened arrays for each subject were then binarized using the median as a threshold. We note here that binarization will produce sequences with information related to burst or “bump” events in each band—including non-stationary aspects of the dynamics. Moreover, these sequences are independent of the overall scale (or units) of the data, since we used the median for binarization. Using the median also produces maximal entropy (H_0) strings (since by definition there are the same number of ones and zeros in the binarized string). We computed ρ_0 and the entropy rate for orders 0 (H_0)–5 (H_5 , there are $2^6 = 64$ cases or bins in the associated histogram, which is a reasonable number given the length of the strings analyzed). Since complexity estimators can be sensitive to data string length,^{19,36} for the computation of global complexity we chose a fixed stack string length for all subjects data prior compression. As a reasonable tradeoff between number of subjects and data quantity we generated the binarized strings from the first 169 frames in each subject stack, so each string was of length $n = 14 \times 50 \times 39 \times 169 = 4.6 \times 10^6$. In another analysis variant, we carried out the complexity analysis frame by frame, producing an estimate for each frame and then an average of frame complexity across epochs per subject. In this case, we could use all the available data per subject, since string length equality in LZW was then guaranteed regardless of number of epochs used for averaging.

Mutual Information and Derived Graph Theoretic Metrics

A measure such as ρ_0 applied to multi-channel, multi-epoch spectrograms reflects any detectable regu-

larities across time, channels and frequencies in the frame stacks, and exploit, in that sense, the integration of multichannel, multifrequency EEG data. We can apply it in a global sense by compressing the entire dataset, that is, all channels and frequencies and epochs forming a single input data string. Or, we can study different data subsets to further dissect the driving elements of integration (compression) of the global dataset. We can ask, for example, how much is to be gained in the compression of a spectrogram stack by feeding all the stack (channels and frequencies) into the compression engine, compared to what can be achieved compressing, for example, each frequency independently and then adding the results. The more “integrated” the data is, the more is to be gained by global compression. If there is mutual algorithmic information across frequency bands, for, example, LZW may be able to use it, making a better job than compressing one channel at the time. In this vein, we can define several compression “integration” measures, depending along what axis we want to segment spectrogram stack data, for example channel or frequency.

To do this we start from the *mutual algorithmic information (MAI)* between two strings,¹³ the algorithmic analog of Shannon mutual information. Let $\mathcal{K}(x)$ denote the description length of a string x , which we approximate by the LZW description length. The MAI between two strings is then

$$\begin{aligned} I_{\mathcal{K}}(x:y) &\approx \mathcal{K}(y) + \mathcal{K}(x) - \mathcal{K}(x,y) \\ &\approx l_{LZW}(x) + l_{LZW}(y) - l_{LZW}(x,y) \end{aligned}$$

where $\mathcal{K}(x, y)$ denotes the complexity of the concatenated strings, and where we again approximate Kolmogorov complexity by LZW description length (see Ref. 30, annotated version). We now define the mutual

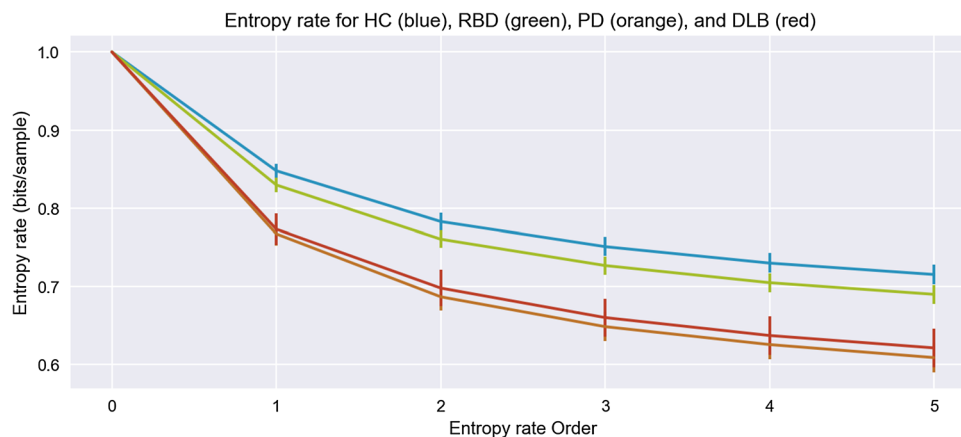


FIGURE 3. Entropy rate H_n as a function of order. Entropy rate plot with error bar (standard error of the mean) as a function of order.

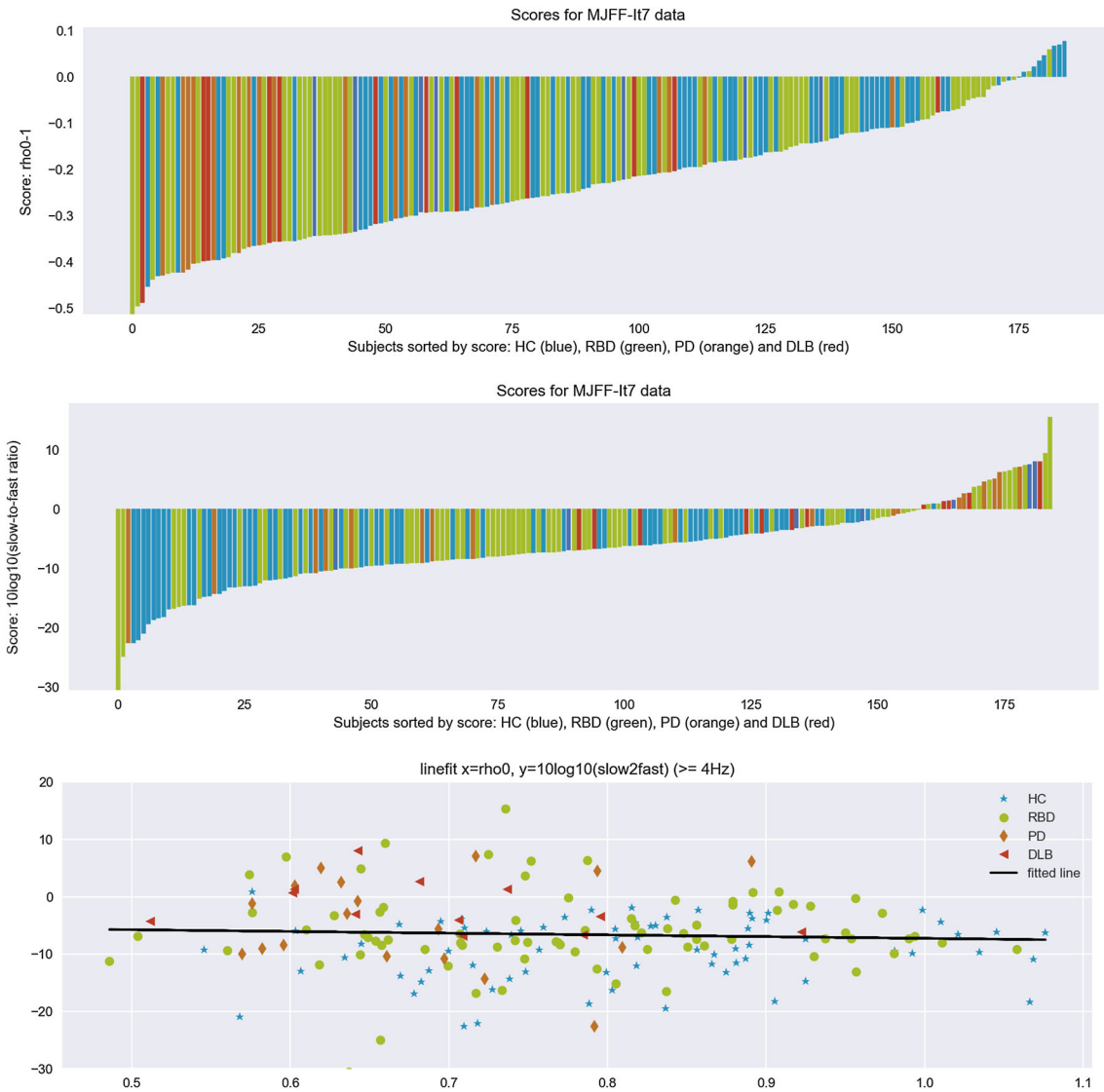


FIGURE 4. Complexity and (log) slow-to-fast scores per subject using full-band spectrograms. Top: Complexity metric ($\rho_0 - 1$ is shown for convenience) per subject for each class (HCs in blue, RBDs in green, PDs in orange and DLBs in red). Middle: log of slow-to-fast metric. Bottom: scatterplot of ρ_0 vs. log of the slow-to-fast (line fit: slope: slope: -3.03 ± 3.91 , $r = -0.06$, p value: $p < 0.44$), displaying no correlation.

algorithmic information coefficient between two strings to be

$$\mu_0(x, y) = \frac{I_{\mathcal{K}}(x: y)}{\mathcal{K}(y) + \mathcal{K}(x)} \approx 1 - \frac{l_{LZW}(x, y)}{l_{LZW}(x) + l_{LZW}(y)}$$

This metric is closely related to the Normalized Compression Distance.^{6,23}

Using $\mu_0(x, y)$, we can build the adjacency matrix of a graph defined by a set of nodes, where a node represents a given frequency band, and with a link across nodes if $\mu_0(x, y)$ for the stack frequency band binarized strings x and y is above a chosen threshold. In this way, we can associate a mutual information graph for each subject spectrogram stack. We can then analyze

these graphs using standard graph theory metrics such as the average degree (the average number of links per node) or average clustering index C (a measure of the extent to which nodes in a graph tend to cluster together),^{2,25} thus defining new metrics for each stack. This analysis can also help understand how compression of the data is actually taking place, as we discuss below.

Slow to Fast Ratio

For comparison with earlier work, we have also computed the *slow-to-fast* frequency power ratio, normally defined as $[(\delta + \theta)/(\alpha + \beta)]$, with δ [0.5, 4

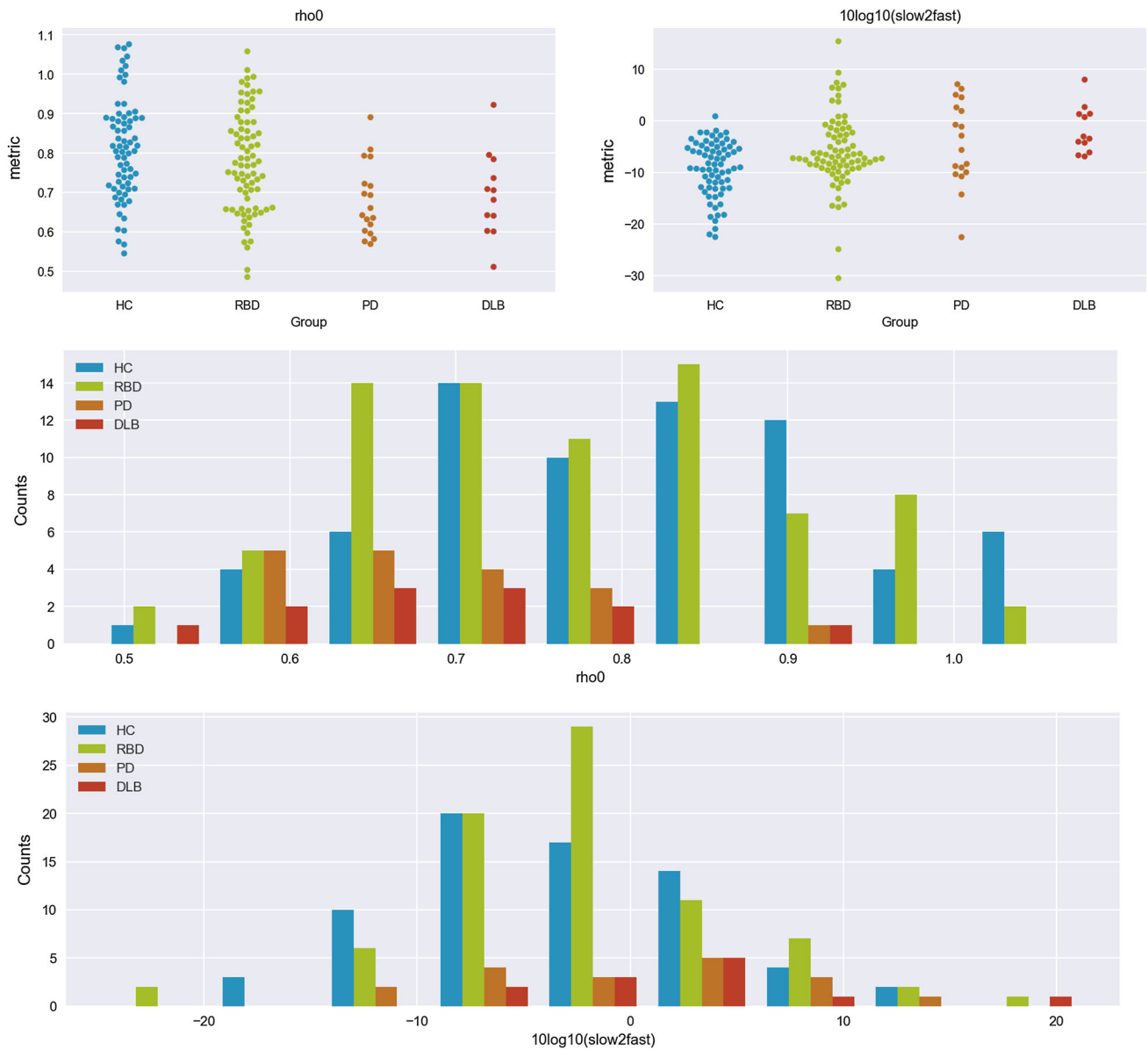


FIGURE 5. Histograms and scatter plots for ρ_0 complexity and slow-to-fast metrics from full-band spectrograms. Scatter plots (top), and histograms for ρ_0 and slow-to-fast with each class (HCs in blue, RBDs in green, PDs in orange and DLBs in red).

TABLE 2. Mean complexity (ρ_0), entropy rate (H_5), frequency mutual algorithmic information mean clustering coefficient ($\langle C_f \rangle$) (for a threshold of 0.0425) and slow-to-fast metrics for each group, with standard error of the mean (SEM).

Group (N)	ρ_0	H_5	$\langle C_f \rangle$	Slow-to-fast
HC (70)	0.81 ± 0.02	0.71 ± 0.01	0.13 ± 0.01	0.21 ± 0.02
RBD (78)	0.78 ± 0.03	0.69 ± 0.02	0.18 ± 0.03	1.16 ± 0.96
PD (18)	0.68 ± 0.02	0.61 ± 0.02	0.28 ± 0.02	1.21 ± 0.36
DLB (12)	0.69 ± 0.02	0.62 ± 0.02	0.23 ± 0.03	1.20 ± 0.38
Kruskal–Wallis p value	1×10^{-4}	3×10^{-4}	8×10^{-7}	1×10^{-4}

In the last row, the four-group Kruskal–Wallis p value for each metric is provided.

Hz), θ [4, 8 Hz), α [8, 13 Hz), and β [13, 32 Hz).²¹ Here we calculated it as the ratio of power in the interval [4,

8) Hz (θ) frequencies vs. [8, 32) Hz ($\alpha + \beta$), which we saw gave much a more robust performance than the

canonical one probably due to low frequency artifacts in the data.

RESULTS

Entropy rate estimates as a function of order for each group are shown in Fig. 3, displaying clear differences across two super-groups (HC + RBD vs. PD + DLB) and also a flattening out with increasing order. LZW complexity metrics per subject are shown, sorted, in Fig. 4 together with the slow-to-fast ratio (further discussed below), and summarized in Fig. 5. In terms of statistical performance, the related complexity related metrics we tested (ρ_0 , entropy rate and log-log entropy rate slope) produced similar results. Statistical results for the main metrics are summarized in Table 2 (mean with standard error of the mean and four-way Kruskal–Wallis test). Since H_5 entropy rate or power law slope led to very similar results to those using ρ_0 (which is much faster to compute), we will drop them in what follows. Table 3 provides the two-way Wilcoxon ranksum statistic test for several comparisons of the three main metrics.

For completeness, we explored complexity applied on an epoch by epoch basis, which allowed us to use all the data from every subject (string length prior compression not being dependent on number of epochs). Complexity per epoch was averaged, leading to an overall increase in complexity but slightly improved statistical performance ($p < 3 \times 10^{-5}$ four group Kruskal–Wallis).

The slow-to-fast frequency power ratio, computed as the ratio of power in the interval [4, 8) Hz (θ) frequencies vs. [8, 32) Hz ($\alpha + \beta$), also provided good discrimination (see Tables 2 and 3). This frequency range choice gave much better performance than the canonical one (which produced barely significant four-

FIGURE 6. Graph analysis using connectivity from algorithmic mutual information (MAI, μ_0) across frequency bands. Top: clustering coefficient as a function of threshold for the different groups. Middle: corresponding scatter plots. Bottom. Average graphs (threshold = 0.0425), with each node representing a frequency bin (label in Hz). Note the increasing MAI within low and within high bands in the second supergroup (PD + DLB).

group $p < 0.05$ Kruskal–Wallis statistical test results), probably due to low frequency noise in the data.

Finally, we studied the impact of using a few channels instead of the full set of 14. Complexity metrics remained robust, indicating that relevant information for statistical discrimination is available in each channel. For example, using ρ_0 with stacks generated from channels P4 and P7 led to a statistically significant results in the four-group Kruskal–Wallis analysis ($p < 2 \times 10^{-4}$). The same analysis using the slow-to-fast ratio metric was not significant.

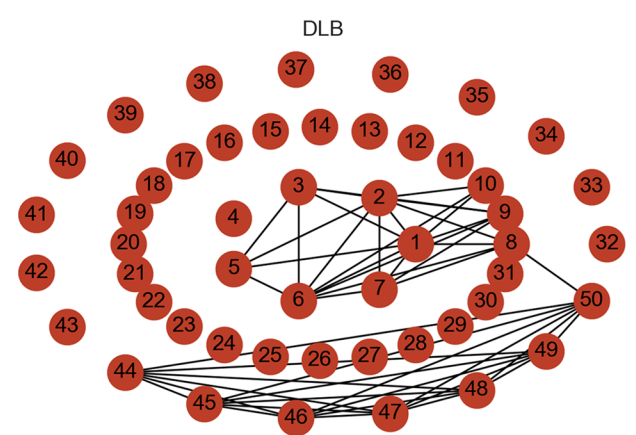
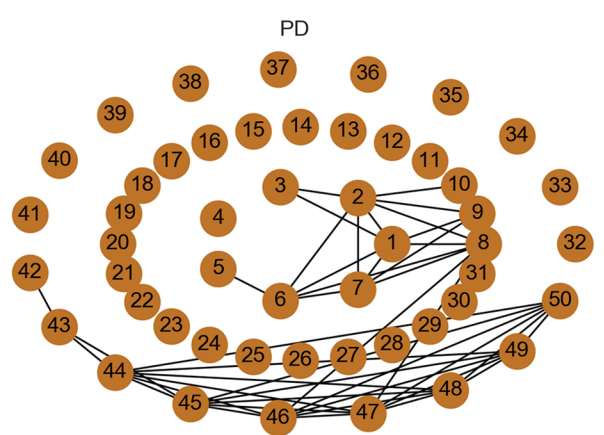
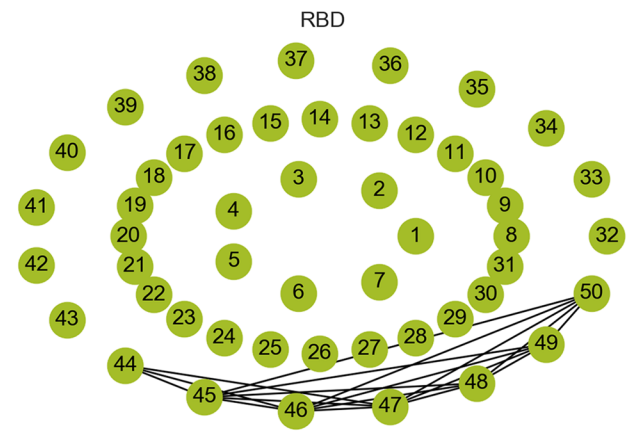
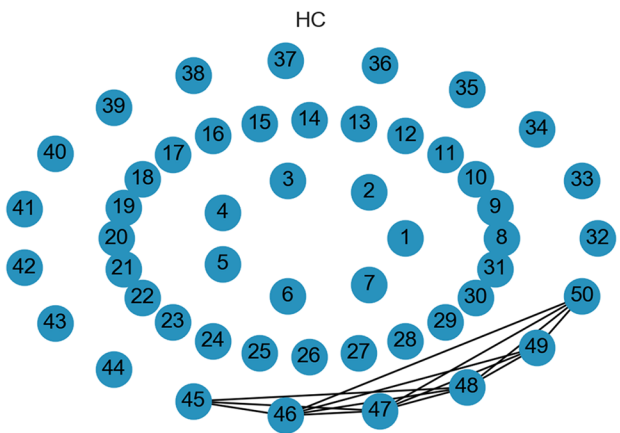
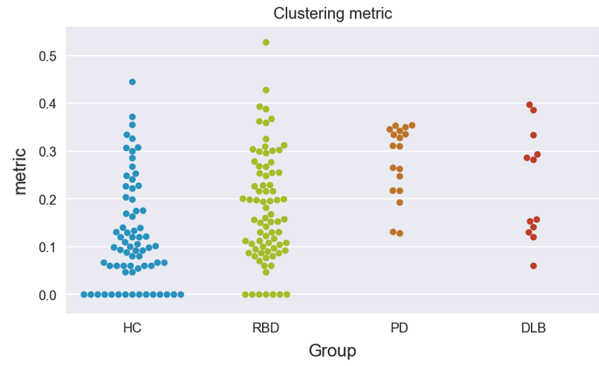
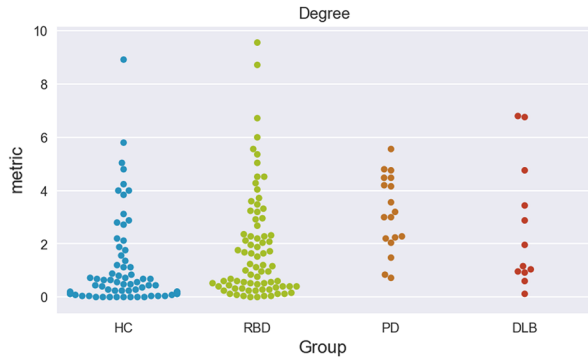
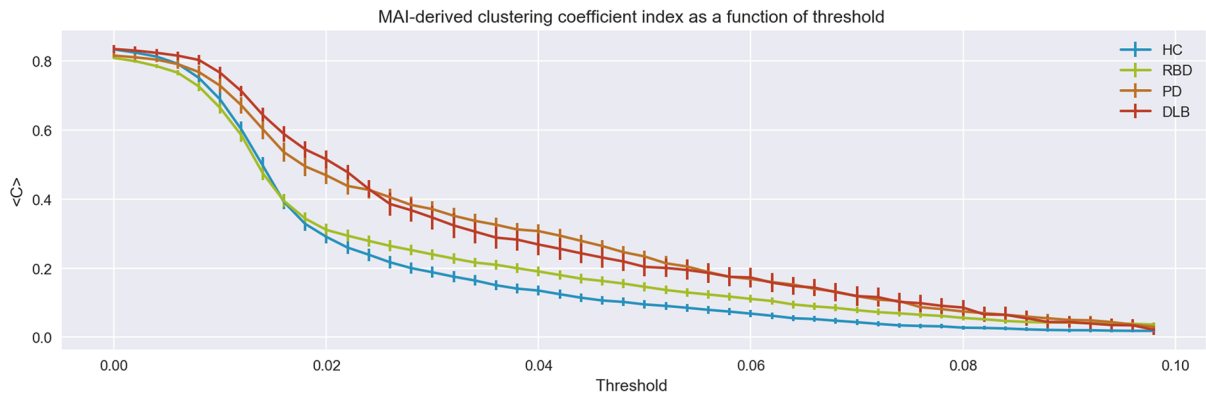
The Role of Frequency in Complexity and Compressibility

To better understand changes in compressibility in PD and DLB converter subject group data, we investigated the role of high vs. low frequency bands in complexity. Shortly, if only low frequency data ($f < 14$ Hz) is used to produce binarized strings, complexity metrics fail to discriminate the subject classes well. Similarly, restricting binarization and analysis to either the β or γ band alone eliminated statistical differences in ρ_0 across the groups. If only high frequency data is retained ($f > 12$ Hz), statistical performance (separability) was partially maintained ($p < 2 \times 10^{-2}$ four-group Kruskal–Wallis).

Next, using the concept of mutual algorithmic information described above, we studied differences across groups from graphs constructed using the MAI

TABLE 3. Statistics for comparisons between groups for selected metrics: Wilcoxon rank-sum statistic significance two sided p value and the Area Under the Curve (AUC, in parenthesis).

Comparison	ρ_0	$\langle C_f \rangle$	Slow-to-fast
HC vs. RBD	1×10^{-1} (57%)	2×10^{-3} (61%)	5×10^{-3} (63%)
HC vs. PD	1×10^{-4} (80%)	2×10^{-6} (87%)	2×10^{-2} (68%)
HC vs. DLB	3×10^{-3} (77%)	4×10^{-3} (75%)	9×10^{-5} (85%)
RBD vs. PD	2×10^{-3} (73%)	3×10^{-4} (77%)	5×10^{-1} (55%)
RBD vs. DLB	3×10^{-2} (69%)	2×10^{-1} (63%)	8×10^{-3} (74%)
PD vs. DLB	6×10^{-1} (56%)	2×10^{-1} (64%)	4×10^{-1} (59%)
HC vs. PD + DLB	7×10^{-6} (78%)	3×10^{-7} (82%)	6×10^{-5} (75%)
HC vs. rest	3×10^{-3} (63%)	1×10^{-5} (69%)	2×10^{-4} (67%)
HC+RBD vs. PD + DLB	2×10^{-5} (75%)	4×10^{-6} (77%)	1×10^{-3} (69%)
RBD vs. PD + DLB	5×10^{-4} (72%)	5×10^{-4} (72%)	5×10^{-2} (62%)



index across frequency bands (μ_0) using complex network theory,² reasoning that there may be an increase in mutual information *across* bands that partly explains changes in compressibility of spectrograms (as discussed further below). Table 2 provides the main statistics for the mean frequency clustering coefficient—the clustering coefficient $\langle C_f \rangle$ derived from the thresholded mutual algorithmic information adjacency matrix—and Fig. 6 provides the average degree and clustering coefficient as a function of threshold and the associated graph for a representative threshold.

DISCUSSION

Using a few minutes of eyes-closed resting EEG data to generate spectrogram stacks, we have seen here that RBD patients who later developed PD or DLB display diminished complexity compared to HCs or to RBD patients who remained disease-free. The results of our statistical analysis are compelling, displaying significant differences between two main super-groups: HCs and RBDs vs. PDs and DLBs (see Tables 2 and 3). Data from subjects in the latter super-group display lower complexity and lower entropy rates with increased cross-frequency algorithmic amplitude coupling. We have also seen that as expected, LZW and entropy rate provide very similar metrics (although LZW is much faster to compute). The clustering coefficient at an appropriately chosen threshold gave very good discrimination across all four groups, but in general it separated robustly the two supergroups across many thresholds (see Fig. 6).

These differences are remarkable as they reflect brain state at the moment of idiopathic RBD diagnosis, years before conversion to PD or DLB. The performance of these metrics is slightly superior and complementary to those from power analysis,^{21,26} with which they are well aligned. Unlike in previous studies in neurodegenerative diseases,^{4,8,11} power ratio and complexity metrics were not correlated here (see Fig. 4). We note that correlation of power metrics such as slow-to-fast ratio and LZW or entropic complexity would not be entirely surprising, since complexity analysis is based on compression of spectrograms and EEG power imbalances that can lead to more compressible spectrograms after binarization). In addition to improving statistical separability, the global algorithmic complexity metrics discussed here provide a rather general detection mechanism of regularities in the data. For example, imbalances in power ratio across arbitrary different bands can lead to lower complexity after data binarization. Complexity metrics can detect such regularities in the data without the need to define a priori which specific aspects to use

(e.g., which bands to use for a ratio computation). In this sense, they provide a rather assumption-free analysis of the data which can then be further dissected. Moreover, we saw they remained robust even if applied (and then averaged) epoch by epoch or using a few channels, which could be useful for screening in clinical practice.

In our approach, we have introduced a global complexity analysis of the data, where information redundancies across epochs, channels and frequency bands can be exploited for compression and studied to understand what features lead to simpler data streams. This limits also the multiple comparisons problem.

Analysis of Sources of Compressibility

As suggested above, even if complexity metrics detect differences across groups, it is interesting to investigate where specifically the relevant regularities lie in the data. They may lie in the temporal series within bands and across epochs, similarities across bands or across channels—that is, in frequency, time and space. Where is the loss of complexity taking place in patients with poor prognosis?

To find out in detail, we carried out tests reordering and randomizing the data in different ways to detect which lead to a loss of discrimination between the groups using complexity metrics (see Fig. 8 for a visual representation of the different data re-shuffling methods). In general, all such tests led to an overall increase in complexity, highlighting the fact that compression exploits regularities across channels, across time (within and across frames), across frequency and across epochs, and that by randomizing the data it becomes harder to compress (more complex). However, interestingly, not all the above reshuffling methods led to a loss of discrimination across groups.

We first randomized the frequency index at each time point within each frame independently across frames. As expected, this shuffling destroyed the discrimination performance of both complexity and power ratio metrics. However, if the frequency randomization was held constant within each spectrogram frame (but randomized across frames), only the performance of the power ratio metric (slow-to-fast) was severely affected. Similarly, applying the same randomization procedure across the temporal dimension (reshuffling time within each frame, which leaves power ratio metrics unaffected, as they rely on unordered time averages) increased overall complexity but did not affect the discriminability of the complexity metric. Thus, complexity metrics appear to rely on within-frame temporal regularities across some frequencies (similar patterns appearing in several fre-

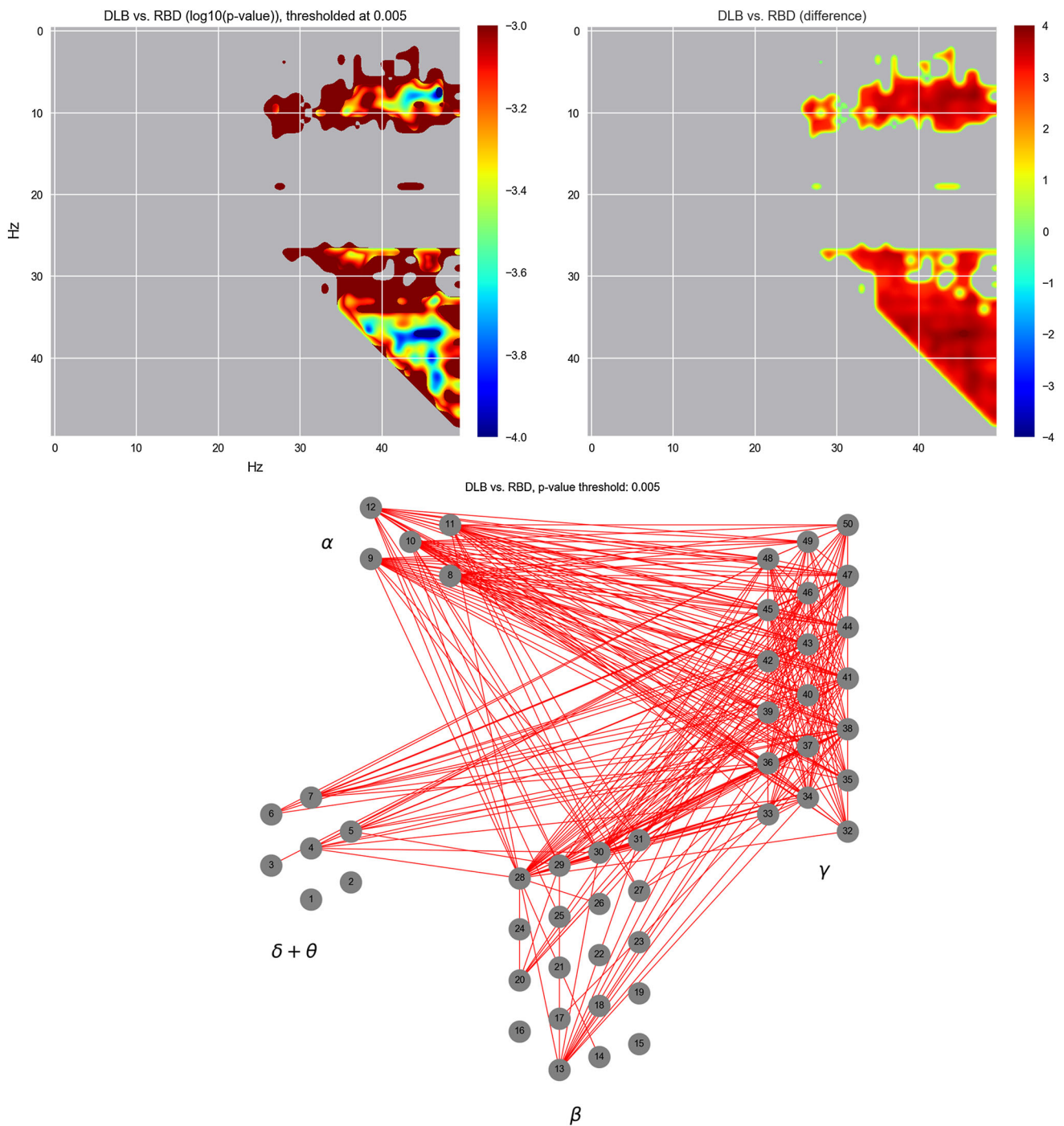


FIGURE 7. Statistics of MAI graphs (μ_0) across frequency bands. **Left:** statistical differences between two groups at the edge level using per band binarization to remove amplitude scale differences. **Right:** Associated graph displaying links with increased (red) or decreased (blue) scale-independent MAI connectivity in the first group with respect to the second. In this particular case, we observe an increase in connectivity, especially in the $\alpha + \theta$ to γ bands.

quencies or at several time points in a frame), with highest complexity for HC and RBD, then PD and finally DLB—irrespective of the actual shape of the spectrum. Poor prognosis in RBD appears to be associated with a decrease in cross-frequency, within-frame temporal regularities in spectrograms. A partial

explanation of the above rests on the observation that binarization by the median of the global stack can lead to simple sequences in some frequencies bands which naturally tend to be far away from the global median, i.e., low and high frequencies, which have power typically either above or below the median. Frequency

bands whose values lie far from the median will tend to look simpler (producing mostly chains of 0's or 1's). Thus, our complexity metric applied to global stacks in which either frequencies or channels lie far from the median produces lower complexity estimates. The complexity metric is insensitive as to where these regularities are (in which frequency or channel)—it detects events of similar symbol aggregation—chains of 0's or of 1's.

A further test we carried out was to binarize each band independently, which in essence removes differences in power scale across bands since binarization then uses each band's median. This resulted in a loss of discriminability using ρ_0 , indicating that imbalances in power across bands are an important element in ρ_0 complexity. However, as discussed above, we observed that cross-frequency MAI differences across groups remained significant and discriminative despite this manipulation (see, as an example, Fig. 7). This implies that there exist scale-independent cross-frequency

(amplitude) temporal regularities in the data (typically an increase in regularity with disease progression).

Our results with regard to the role of high frequencies in changes in complexity can be related to those in Ref. 4, where multiscale entropy was employed as complexity metric to study the differences in brain signal variability in PD patients who developed dementia at follow-up with respect who did not. The authors found significant changes, with lower signal variability at timescales sensitive to higher frequencies (i.e., more compressibility at higher frequencies) in patients that developed dementia than those who did not or in controls. We do note, however, that multi-scale entropy and LZW (or entropy rate) estimate complexity in different ways.

Finally, one can ask how much does stack compression benefit from having access to all frames from a subject as opposed to compressing each frame independently. The answer is that PD and DLB data benefit the most (with a ratio of per frame vs. per stack

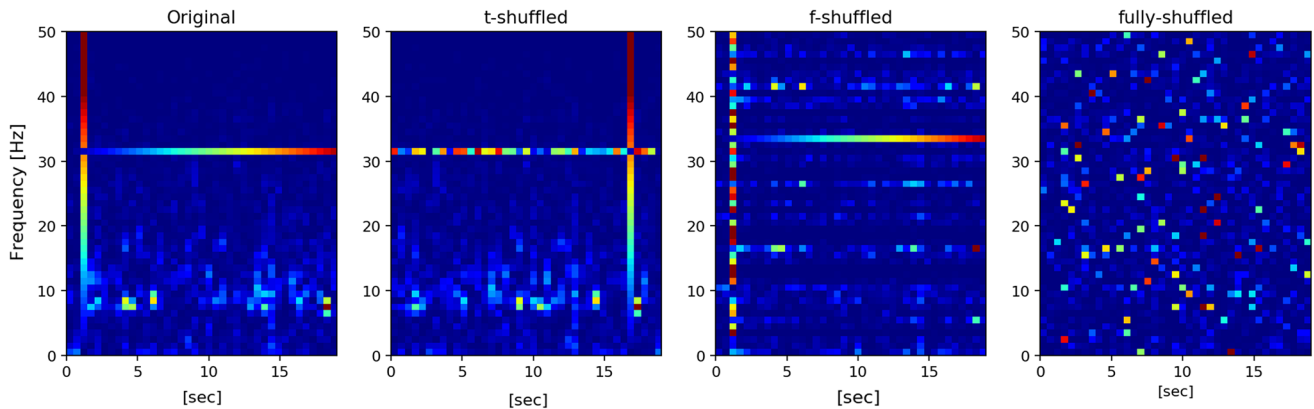


FIGURE 8. Frame (intra-epoch) spectrogram shuffling. From left to right: original, temporally shuffled, frequency and fully shuffled spectrograms. A horizontal and a vertical line has been added for reference.

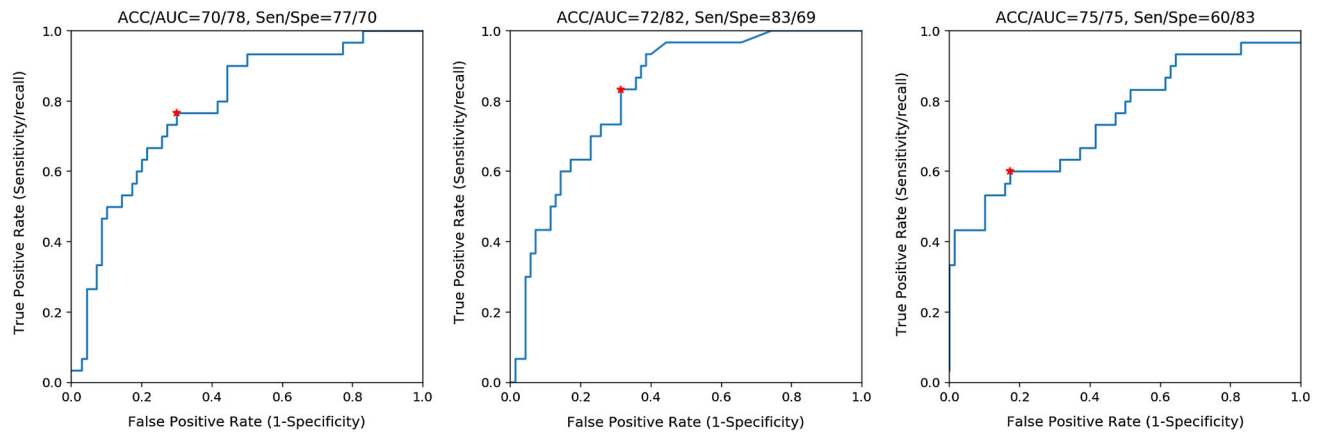


FIGURE 9. ROC analysis for the HC vs. PD + DLB problem. From left to right, ROC curves using ρ_0 , $\langle C_f \rangle$ and slow-to-fast metrics, accuracy (ACC), area under-the-curve (AUC), and with Sensitivity and Specificity (%) at the optimal point (shown in red).

ρ_0 of 1.26). That is, using the full stack data improves compression by 26% in these groups, compared to 22 and 23% in HC and RBDs ($\pm 1\%$, $p < 6 \times 10^{-5}$ four group Kruskal–Wallis). There are longer term persistent regularities in the PD and DLB groups than in HC or RBD.

Sensitivity, Specificity and Machine Learning Analysis

We have also explored the potential of these metrics for classification, focusing on the problem of differentiating HC ($n = 70$) vs. PD + DLB ($n = 30$).

First, we have computed the accuracy, sensitivity and specificity of the metrics at the optimal point. In Fig. 9 we provide receiver operating characteristic curves (ROC) for ρ_0 , $\langle C_f \rangle$ and slow-to-fast metrics.

Next, using the three metrics as a combined set of features, we have trained a random forest classifier and evaluated it using leave-pair-out (test-set) cross-fold validation. This results in an accuracy of 73% and an AUC of 84%, with a specificity of 75% and sensitivity of 72%. Using one feature at the time results in poor results, (AUC of 75, 79 and 70% respectively for ρ_0 , $\langle C_f \rangle$ and slow-to-fast metrics), highlighting the complementarity of these metrics.

CONCLUSION

In Ref. 30 it is hypothesized that while a system capable of universal computation may produce many different types of patterns, both simple (e.g., constant or repetitive) and complex, a healthy brain engaging in modeling and prediction of complex input-output strings will produce complex-looking, highly entropic data. Such apparent complexity is what we actually measure when using entropy rate and LZW compression of electrophysiological or metabolic brain data.^{3,5,15,31,32} First order entropy, entropy rate or LZW provide (poor) *upper bounds* on algorithmic complexity.

From our results with this dataset we conclude that LZW and, equivalently, entropy rate highlight losses of estimated algorithmic complexity in spectrograms of RBD patients likely to evolve into PD or DLB compared to those that remained in RBD or to healthy controls. Similar findings have been reported in earlier related work comparing healthy controls and PD patients.¹² The loss of complexity takes place both in low and high frequencies, and is partly due to increased mutual information within and across bands. These results indicate that information *differentiation* (global complexity) is a potentially relevant metric for the prognosis of RBD, and are in line with current views of the brain stemming from information theory

connecting theories of cognition and consciousness with the phenomenology of brain health, and in particular, neurodegeneration.^{30,34}

Our full band complexity scores are generally more discriminative than spectral ratios,²⁶ do not require a pre-selection of the spectral bands to study, and remain robust even when using a few channels. However, we believe it is more interesting to view complexity and spectral ratio metrics as complementary, as they display low correlation and hence can be combined using machine learning.

Although we have focused here on a particular biomarker, we note that the use of EEG for prognosis of idiopathic RBD represents a practical technology of clinical interest. It could be used, for example, as a pre-screening tool for more expensive and invasive technologies, such as positron emission tomography (PET) or single-photon emission computed tomography (SPECT), which require the injection of radio tracers. Indeed, the method studied here can be carried out with a relatively simple EEG system and very light protocol—we found it could be implemented with a few channels, requiring an acquisition of a few minutes of data that can easily be collected as an add-on to RBD polysomnographic studies (as in Ref. 26), or even during routine visits to the neurologist.

Finally, one of the limitations in this study is the limited availability of data associated to the problem at hand, including deeper longitudinal studies that could shed light on the power of EEG biomarkers for prediction of early conversion of RBD.¹⁴ With the creation of larger databases, the method proposed here—possibly in combination with other algorithms and machine learning^{17,27,28,33}—should become of increasing practical interest.

AUTHOR CONTRIBUTIONS

DI and MC pre-processed EEG data to produce artifact free spectrograms. EK and ASF contributed to code development and revised the manuscript. JFG, RP and JM collected the EEG data and revised the manuscript. GR provided the complexity analysis code, carried out the complexity and statistical analysis, and wrote the manuscript.

ACKNOWLEDGMENTS

This work has been partly supported by the Michael J. Fox Foundation project “Discovery of EEG biomarkers for Parkinson Disease and Lewy Body Dementia using advanced Machine Learning techniques” (Rapid Response Innovation Awards 2013), the FET Open Luminous project (H2020-FETOPEN-

2014- 2015-RIA under agreement No. 686764) as part of the European Union's Horizon 2020 research and training programme 2014-2018, the Canadian Institutes of Health Research (CIHR) and the W. Garfield Weston Foundation.

DISCLOSURE

Starlab and Neuroelectrics authors have an interest in developing commercial, translational applications from this research. GR is a shareholder of both companies.

REFERENCES

- ¹Abásolo, D., R. Horner, C. Gómez, M. García, and M. López. Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and central tendency measure. *Med. Eng. Phys.* 28(4):315–322, 2006.
- ²Albert, R., and A.-L. Barabasi. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47, 2002.
- ³Andrillon, T., A. T. Poulsen, L. K. Hansen, D. Leger, and S. Kouider. Neural markers of responsiveness to the environment in human sleep. *J. Neurosci.* 36(24):6583–6596, 2016.
- ⁴Bertrand, J.-A., A. McIntosh, R. Postuma, N. Kovacevic, V. Latreille, M. Panisset, S. Chouinard, and J. Gagnon. Brain connectivity alterations are associated with the development of dementia in Parkinson's disease. *Brain Connect.* 6(3):216–224, 2016.
- ⁵Casali, A. G., O. Gosseries, M. Rosanova, M. Boly, S. Sarasso, K. R. Casali, S. Casarotto, M.-A. Bruno, S. Laureys, G. Tononi, and M. Massimini. A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* 5(198):105, 2013.
- ⁶Cerra, D., and M. Datcu. Expanding the algorithmic information theory frame for applications to earth observation. *Entropy* 15:407–415, 2013.
- ⁷Cover, T. M., and J. A. Thomas. *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.
- ⁸Dauwels, J., K. Srinivasan, M. R. Reddy, T. Musha, F.-B. Vialatte, C. Latchoumane, J. Jeong, and A. Cichocki. Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin? *Int. J. Alzheimer's Dis.* 2011:539621, 2011.
- ⁹Fantini, L., J. Gagnon, D. Petit, S. Rompre, A. Decary, J. Carrier, and J. Montplaisir. Slowing of electroencephalogram in rapid eye movement sleep behavior disorder. *Ann. Neurol.* 53(6):774–780, 2003.
- ¹⁰Fulda, S. Idiopathic REM sleep behavior disorder as a long-term predictor of neurodegenerative disorders. *EPMA J.* 2(4):451–458, 2011.
- ¹¹Gomez, C., R. Hornero, D. Abasolo, M. Lopez, and A. Fernandez. Decreased Lempel-Ziv complexity in Alzheimer's disease patients' magnetoencephalograms. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 5:4514–4517, 2005.
- ¹²Gomez, C., K. T. E. O. Dubbelin, C. J. Stam, D. Abasolo, H. W. Berendse, and R. Hornero. Complexity analysis of resting-state MEG activity in early-stage Parkinson's disease patients. *Ann. Biomed. Eng.* 39(12):2935–2944, 2011.
- ¹³Grunwald, P., and P. Vitanyi. Shannon information and kolmogorov complexity. [arXiv:cs/0410002](https://arxiv.org/abs/cs/0410002), 2004.
- ¹⁴Högl, B., A. Stefani, and A. Videnovic. Idiopathic REM sleep behaviour disorder and neurodegeneration—an update. *Nat. Rev. Neurol.* 14:40–55, 2018.
- ¹⁵Hudetz, A. G., X. Liu, S. Pillay, M. Boly, and G. Tononi. Propofol anesthesia reduces Lempel-Ziv complexity of spontaneous brain activity in rats. *Neurosci. Lett.* 628:132–135, 2016.
- ¹⁶Iranzo, A., A. Fernández-Arcos, E. Tolosa, M. Serradell, J. L. Molinuevo, F. Valldeoriola, E. Gelpi, I. Vilaseca, R. Sánchez-Valle, A. Lladó, C. Gaig. Neurodegenerative disorder risk in idiopathic REM sleep behavior disorder: study in 174 patients. *PLoS ONE* 9(2):89741, 2014.
- ¹⁷Javier, J. M., A. S.-F. Aureli, D. I. Nez, S. Dunne, C. Grau, G. Ruffini, J. Rodrigues-Brazete, R. Postuma, c. G. Jean-Fran J. Montplaisir, and A. Pascual. Advanced machine learning for classification of EEG traits as Parkinson's biomarker. *Front. Neuroinform.* 00071, 2014.
- ¹⁸Jia, Y., H. Gu, Q. Luo. Sample entropy reveals an age-related reduction in the complexity of dynamic. *Brain. Sci. Rep.* 7(1):7990, 2017.
- ¹⁹Kaspar, F., and H. G. Schuster. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys. Rev. A* 36(2):842–848, 1987.
- ²⁰Kim, Y., Y. E. Kim, E. O. Park, C. W. Shin, H.-J. Kim, and B. Jeon. REM sleep behavior disorder portends poor prognosis in Parkinson's disease: a systematic review. *J. Clin. Neurosci.* 47:6–13, 2017.
- ²¹Latreille, V., J. Carrier, B. Gaudet-Fex, J. Rodrigues-Brazete, M. Panisset, S. Chouinard, R. B. Postuma, and J.-F. Gagnon. Electroencephalographic prodromal markers of dementia across conscious states in Parkinson's disease. *Brain* 138:1189–1199, 2016.
- ²²Lempel, A., and J. Ziv. On the complexity of finite sequences. *IEEE Trans. Inf. Theory* 22(1):75–81, 1976.
- ²³Li, M., X. Chen, X. Li, B. Ma, and P. M. B. Vitanyi. The similarity metric. *IEEE Trans. Inf. Theory* 50(12):113859, 2004.
- ²⁴Postuma, R., J. Gagnon, M. Vendette, M. Fantini, J. Massicotte-Marquez, *et al.* Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder. *Neurology* 72:1296–1300, 2009.
- ²⁵Ray, C., G. Ruffini, J. Marco-Pallarés, L. Fuentemilla, and C. Grau. Complex networks in brain electrical activity. *Eur. Phys. Lett.* 79:38004, 2007.
- ²⁶Rodrigues-Brazete, J., J. Gagnon, R. Postuma, J. Bertrand, and M. J. D Petit. Electroencephalogram slowing predicts neurodegeneration in rapid eye movement sleep behavior disorder. *Neurobiol. Aging* 37:74–81, 2016.
- ²⁷Ruffini, G., D. Ibañez, M. Castellano, S. Dunne, and A. Soria-Frisch. EEG-driven RNN classification for prognosis of neurodegeneration in at-risk patients. In: *ICANN 2016*. Cham: Springer, 2016.
- ²⁸Ruffini, G., D. Ibañez Soria, L. Dubreuil, M. Castellano, J.-F. Gagnon, J. Montplaisir, and A. Soria-Frisch. Deep learning with eeg spectrograms in rapid eye movement behavior disorder. *BioRxiv* 240267, 2018.
- ²⁹Ruffini, G. Lempel-Ziv complexity reference. [arXiv:1707.09848](https://arxiv.org/abs/1707.09848) [cs.IT], 2017.
- ³⁰Ruffini, G. An algorithmic information theory of consciousness. *Neurosci. Conscious.* 3(1):012 2017.

- ³¹Schartner, M., A. Seth, Q. Noirhomme, M. Boly, M.-A. Bruno, S. Laureys, *et al.* Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. *PLoS ONE* 10(8):0133532, 2015.
- ³²Schartner, M. M., R. L. Carhart-Harris, A. B. Barrett, A. K. Seth, and S. D. Muthukumaraswamy. Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, lsd and psilocybin. *Sci.Rep.* 7:46421, 2017.
- ³³Soria-Frisch, A., J. Marin, D. I. Ibañez, S. Dunne, C. Grau, G. Ruffini, J. Rodrigues-Brazète, R. Postuma, J.-F. Gagnon, J. Montplaisir, and A. Pascual-Leone. Machine learning for a Parkinson's prognosis and diagnosis system based on EEG. In: *Proceedings of the International Pharmacology-EEG Society Meeting PEG 2014*. Leipzig: Germany, 2014.
- ³⁴Tononi, G., M. Boly, M. Massimini, and C. Koch. Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17:450–461, 2016.
- ³⁵Welch, T. A technique for high-performance data compression. *Computer* 17(6):8–18, 1984.
- ³⁶Zhang, X.-S., R. J. Roy, and E. W. Jensen. EEG complexity as a measure of depth of anesthesia for patients. *IEEE Trans. Biomed. Eng.* 48(12):1424–1433, 2001.