Algorithmic complexity of EEG for prognosis of neurodegeneration in idiopathic rapid eye movement behavior disorder (RBD)

Giulio Ruffini[a,b,*], David Ibañez[a], Eleni Kroupi[a], Jean-François Gagnon[c,e], Jacques Montplaisir[c,f], Ronald B. Postuma[c,d], Marta Castellano[a], Aureli Soria-Frisch[a]

[a]*Starlab Barcelona, Avda. Tibidabo, 47bis, 08035 Barcelona, Spain*
[b]*Neuroelectrics Corporation, 210 Broadway, Cambridge, MA 02139, USA*
[c]*Centre for Advanced Research in Sleep Medicine, Hôpital du Sacré-Cœur de Montréal, Montreal, Canada*
[d]*Department of Neurology, Montreal General Hospital, Montreal, Canada*
[e]*Department of Psychology, Université du Québec à Montréal, Montreal, Canada*
[f]*Department of Psychiatry, Université de Montréal*

Abstract

Objectives: Idiopathic REM 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.

Methods: 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 ∼5 minute recordings was converted to spectrograms and their algorithmic complexity estimated using Lempel-Ziv-Welch compression (LZW).

Results: 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.

Conclusions: Poor prognosis in RBD appears to be associated with decreased

*Corresponding author, giulio.ruffini@neuroelectrics.com, +1 617-899 6911

complexity of EEG spectrograms stemming in part from frequency power imbalances and cross-frequency amplitude coupling.

Significance: 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: Biomarkers, EEG, LZW, PD, LBD

---

1. Introduction

REM Behavior Disorder (RBD) is a serious risk factor for neurodegenerative diseases such as Parkinson's disease (PD). RBD is a parasomnia characterized by vivid dreaming and dream-enacting behaviors associated with REM sleep without muscle atonia (Iranzo et al., 2014). Idiopathic RBD occurs in the absence of any neurological disease or other identified cause, is male-predominant and its clinical course is generally chronic progressive (Fulda, 2011). Several longitudinal studies conducted in sleep centers have shown that most patients diagnosed with the idiopathic form of RBD will eventually be diagnosed with a neurological disorder such as Parkinson disease (PD) and dementia with Lewy bodies (DLB) (Postuma et al., 2009; Fulda, 2011; Iranzo et al., 2014). In essence, idiopathic RBD has been suggested as a prodromal factor of the synucleinopathies PD, DLB and less frequently multiple system atrophy (MSA) (Iranzo et al., 2014).

RBD has an estimated prevalence of 15-60% in PD and has been proposed to define a subtype of PD with relatively poor prognosis, reflecting a brainstem-dominant route of pathology progression (v. Kim et al. (2017) and references therein) with a higher risk for dementia or hallucinations. PD with RBD is characterized by more profound and extensive pathology—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 dynami-

cal system which instantiates electrochemical communication and computation. Electroencephalographic (EEG) and magnetoencephalographic (MEG) signals contain rich information associated with these processes, and are partially 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 (Fantini et al., 2003; Rodrigues-Brazète et al., 2016), and the slow to fast ratio (the ratio of power in delta and theta bands to alpha and beta) has shown good discriminatory power (Rodrigues-Brazète et al., 2016; Latreille et al., 2016). 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 (Ruffini et al., 2016).

With the goal of developing metrics that capture non-linear dynamics from non-stationary signals, here we carry statistical analysis to explore a different feature extracted from EEG data—algorithmic complexity. As discussed in Ruffini (2017a), the healthy brain generates apparently complex (entropic) data. Complexity should be associated to cognitive health and conscious state.

In order to do so, we work with time-frequency analysis of EEG data to analyze amplitude spectral dynamics. We hypothesize here that global apparent EEG algorithmic complexity or entropy rate decrease with neurodegenerative disease prognosis or progression, in a similar manner to that already reported in the literature PD Gomez et al. (2011) and also in the case of Alzheimer's disease (Gomez et al., 2005; Abásolo et al., 2006; Dauwels et al., 2011). We further assume that algorithmically relevant aspects in EEG data are present in compositional features in the time-frequency spectral amplitude representation.

Algorithmic information is formalized by the mathematical concept of algorithmic complexity or Kolmogorov complexity ($\mathcal{K}$), co-discovered during the second half of the 20th century by Solomonoff, Kolmogorov and Chaitin. We recall its definition: *the Kolmogorov complexity of a string is the length of the shortest program capable of generating it.* More precisely, let $\mathcal{U}$ be a universal computer (a Turing Machine), and let p be a program. Then the Kolmogorov or algorithmic complexity of a string x with respect to $\mathcal{U}$ is given by $\mathcal{K}_{\mathcal{U}(x)} = \min_{p:\mathcal{U}(p)=x} l(p)$, i.e., the length $l(p)$ of the shortest program that prints the string x and then halts (see e.g., Cover and Thomas (2006); Li and Vitanyi (2008)). Crucially, although the precise length of this program depends on the programming language used, it does so only up to a string-independent constant. Algorithmic information is difficult to estimate. Here we analyze our hypothesis using the notion of apparent complexity as measured from Lempel-Ziv-Welch compression (LZW) or entropy rate, two closely related measures that provide an upper bound to algorithmic complexity.

2. Methods

2.1. Participants

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 Sacrè-Cœur de Montral. All patients with full EEG montage for resting-state EEG recording at baseline and with at least one follow-

Table 1: **Baseline sociodemographic and clinical data for the four groups.**

Group / data (yr)	**HC**	**RBD**	**PD**	**DLB**
N (F/M)	83 (30/53)	83 (18/65)	19 (5/14)	12 (2/10)
Mean age	66 ± 9	67 ± 7	67 ± 7	71 ± 6
Mean 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 (units are years). Acronyms: 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 LBD on follow-up.

up examination after the baseline visit were included in the study. The first valid EEG for each patient enrolled in the study was considered 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, see Rodrigues-Brazète et al. (2016).

2.2. EEG dataset

As with previous work (Ruffini, 2016), the first level data in this study consisted of resting-state EEG collected from awake patients and healthy controls using a subset of 14 scalp electrodes (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 two minutes) followed by periods of "eyes closed" in which patients were not asked to perform any particular task. Resting EEG signals were digitized with 16-bit resolution at a sampling rate of 256 S/s. The amplification device implemented hardware band pass filtering between 0.3 and 100 Hz and notch filtering at 60 Hz to minimize

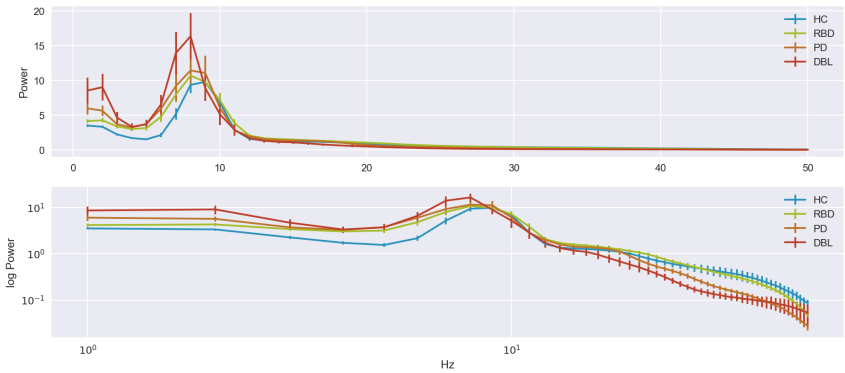

Figure 1: **Group PSDs.** Average power Spectral Density (all channels) for each group with SEM, displaying characteristic "slowing" of EEG.

the influence of power line noise. All recordings were referenced to linked ears.

After quality check and rejection of subjects with insufficient data, or rejection healthy subjects with abnormal cognition at the time of acquisition, the dataset (of initially 213 subjects) includes 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 PSDs are provided in Figure 1.

2.3. Lempel-Ziv-Welch complexity

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 compresses information by reusing them: instead of rewriting a previously seen sequence, it will refer to the identifier of the one seen last (Lempel and Ziv, 1976). After applying LZW to a string, we are provided with a set of words (or phrases, as they are sometimes called) $c(n)$ that form a dictionary. The length of the compressed string will,

in general, be $l_{LZW} \lesssim n$. 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 strings (Kaspar and Schuster, 1987)), 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 programs (Ruffini et al., 2016)). 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

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

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, $0 \leq \mathcal{H} \leq .. \leq H_q \leq ... \leq H_0 \leq 1$. A fundamental relation is that description length computed from LZW is closely related to entropy rate,

$$l_{LZW} = c(n) \log_2 [c(n) + \log_2 A] \approx c(n) \log_2 c(n) \longrightarrow n\mathcal{H}.$$

(Cover and Thomas, 2006; Ruffini, 2017b). 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, or

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

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 result in $H_0 = 1$).

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,

$$\rho_0 = l_{LZW}/n = \frac{c(n)\log_2[c(n) + A]}{n} \longrightarrow \mathcal{H}.$$

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 Ruffini (2017b).

A measure such as ρ_0 applied to multi-channel, multi-epoch spectrograms reflects any detectable regularities 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 (Grunwald and Vitanyi, 2004), 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

$$I_{\mathcal{K}}(x\!:\!y) \approx \mathcal{K}(y) + \mathcal{K}(x) - \mathcal{K}(x,y) \approx l_{LZW}(x) + l_{LZW}(y) - l_{LZW}(x,y)$$

where $K(x,y)$ denotes the complexity of the concatenated strings, and where we again approximate Kolmogorov complexity by LZW description length (see Ruffini (2017a), annotated version). We now define the mutual algorithmic information coefficient between to 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 (Li et al., 2004; Cerra and Datcu, 2013). Using it we can assess the MAI between frequency bands, for example, and build an adjacency matrix or graph defined by a set of nodes (e.g., frequencies) and edges or links (MAI between nodes using a threshold). In this way, we associate a mutual information graph with each subject frame stack. We can then analyze these graphs using standard metrics such as average degree or clustering index C (Albert and Barabasi, 2002; Ray et al., 2007). Such analysis can also help analyze how compression of the data is actually taking place.

2.4. Data analysis

First, using the EEG dataset we have generated a spectrogram stack of "frames" (time frequency bi-dimensional arrays) for each subject (see Figure 2), with each stack representing about 3 minutes of eyes-closed resting EEG. To create it, EEG data for each channel was processed using Fourier analysis (FFT)

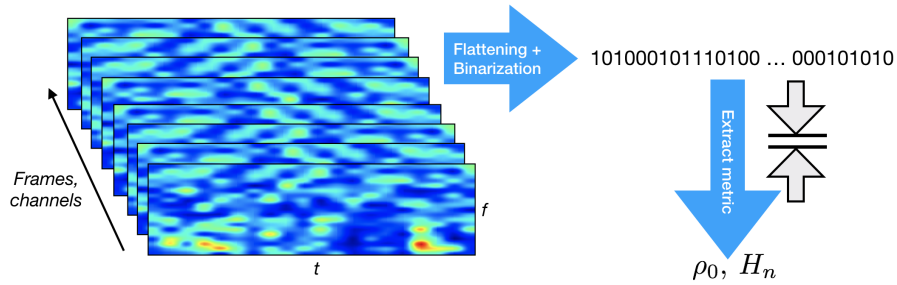

Figure 2: **Computation of description length and entropy metrics from subject spectrogram frame stack.** To compute a global measure of complexity for each subject, the frame stack $S(t, f, ch, F)$ is flattened and binarized prior compression or entropy computation, where t denotes time in each spectrogram,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

after detrending one-second data blocks with a Hann window (50% overlap FFT with resolution is 1 Hz in the range 1–50Hz). Each spectrogram frame was generated from 20 second long artifact free sequences for each of the 14 EEG channels using a sliding window of one second between frames. The resulting data frames are thus tensors of the form [channels (14)] x [FFT bins (50)] x [Time bins (39)], and the full stack is four dimensional, with the fourth dimension indexing epoch.

Using the spectrogram frame stack dataset we carried out an analysis of the complexity for each subject. 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 adjacency 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. We computed ρ_0 and the entropy rate for orders 0 (H_0) to 5 (H_5, there are 2^6 =64 cases or bins in the associated histogram), and computed the

log-log line fit slope and clustering coefficient of MAI derived graphs using μ_0 as an additional metric. Since these metrics can be sensitive to data string length (Kaspar and Schuster, 1987), for the computational of global complexity we chose a reasonable string length and fixed it for all subjects data prior compression (using the first N bits of the flattened spectrogram frame stacks, with $N = 4.6 \times 10^6$, or about 169 twenty-second epochs, as a reasonable tradeoff between number of subjects and data quantity per subject). In an 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.

3. Results

Entropy rates as a function of order are shown in Figure 3. Complexity metrics per subject are shown, sorted, in Figure 4 together with the slow-to-fast ratio (further discussed below), and summarized in Figure 5. In terms of statistical performance, using median binarization results in similar significance results for the relevant metrics (ρ_0, entropy rate and log-log entropy rate slope). Statistical results for the main metrics are summarized in Table 2 (mean with standard error of the mean with four-way Kruskal-Wallis test). Since H_5 entropy rate or power law slope led to very similar results to those from ρ_0 (which is much faster to compute), we will drop them in what follows. Table 3 provides the two-way Wilkoxon ranksum statistic test for several comparisons of the three main metrics.

For completeness, we also explored complexity 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).

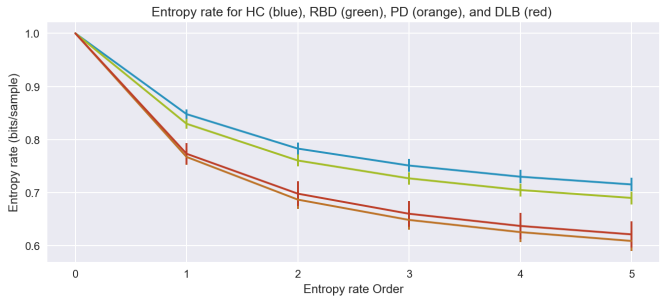

Figure 3: **Entropy rate H_n as a function of order.** Entropy rate plot with error bar (standard error of the mean) as a function of order.

For comparison, we have also computed the *slow-to-fast* frequency power ratio, normally defined as $[(\delta+\theta)/(\alpha+\beta)]$, with δ [0.5,4 Hz), θ [4,8 Hz), α [8,13Hz), and β [13,32Hz) (Latreille et al., 2016). Here we computed it as the ratio of power in the interval [4, 8) Hz (θ) frequencies versus [8, 32) Hz ($\alpha+\beta$), which we saw gave much better performance than the canonical one (which resulted a barely significant four group p<0.05 Kruskal-Wallis).

Finally, we explored the use of a few channels instead of the full set of 14. Complexity metrics remained robust. E.g., using ρ_0 with channels P4 and P7 led to $p < 2\times10^{-4}$ four group Kruskal-Wallis, unlike the slow-to-fast ratio (not significant).

3.1. The role of frequency in compressibility

To better understand changes in compressibility in PD and DLB converter subject data, we investigated the role of high versus 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 beta band or the gamma 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 ($p < 2\times10^{-2}$ four group Kruskal-Wallis).

Next, using the concept of mutual algorithmic information described above,

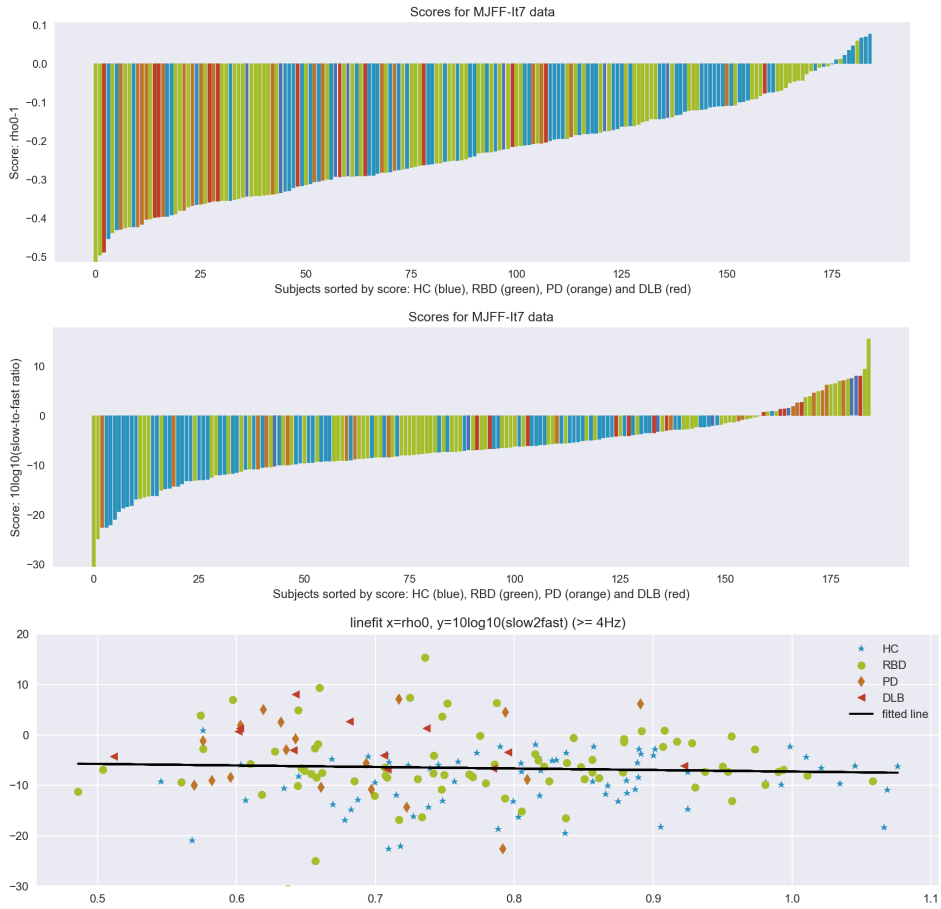

Figure 4: **Complexity and (log) slow-to-fast scores per subject using full-band spectrograms.** Complexity metric $\rho_0 - 1$ per subject for each class (HCs in blue, RBDs in green, PDs in orange and DLBs in red). The second one displays the log of slow-to-fast metric and the last displays ρ_0 vs log of slow-to-fast (line fit: slope: slope:-3.03 $\pm$ 0.06, r: -0.07, pval: $p < 0.5$). Using these two metrics leads to a training classification rate AUC of 78% using, e.g., Naive Bayes or logistic regression.

we have studied differences across groups from graphs constructed using MAI across frequency bands (μ_0), reasoning that there may be an increase in mutual information across bands that partly explains changes in compressibility of spectrograms. Table 2 provides the main statistics for the mean frequency clustering coefficient—the clustering coefficient derived from the mutual algorithmic

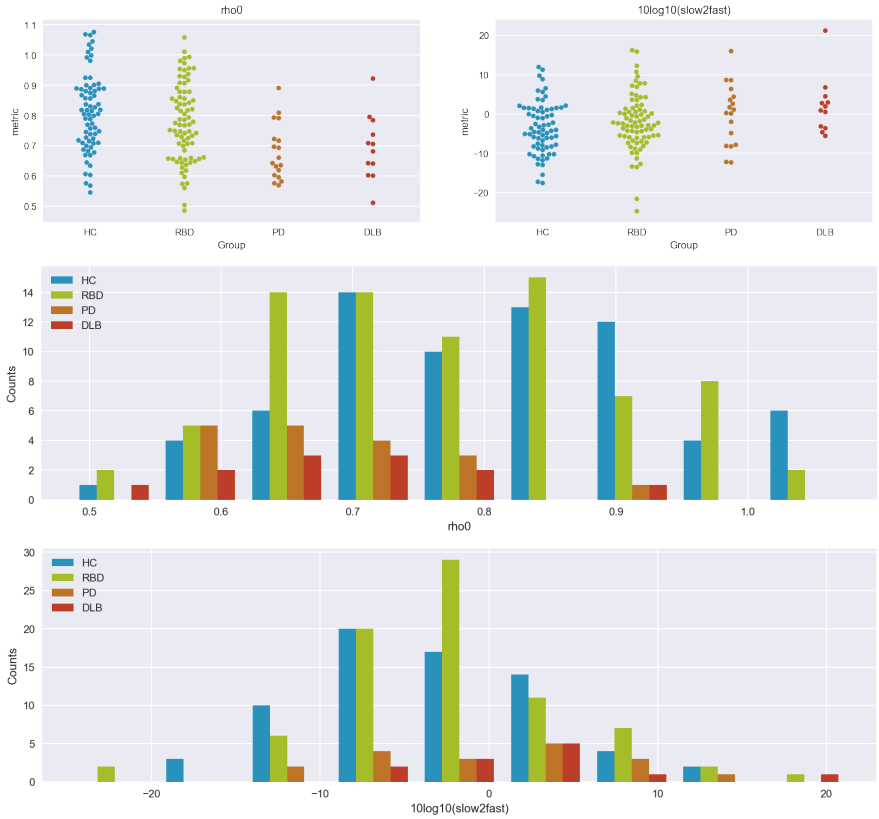

Figure 5: **Histograms and scatter plots for ρ_0 complexity and slow-to-fast metrics from full-band spectrograms.** Scatter plots (top), and histograms for ρ_0 and slow-to-fast with each class (HCs in blue, RBDs in green, PDs in orange and DLBs in red).

coefficient $\langle C_f \rangle$—, and Figure 6 provides the degree and clustering coefficient as a function of threshold and the associated graph for a representative threshold. We see again the increased compressibility of global stack data in PD and DLB groups can partly be explained by increased MAI within low and also within high frequency bands.

4. 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

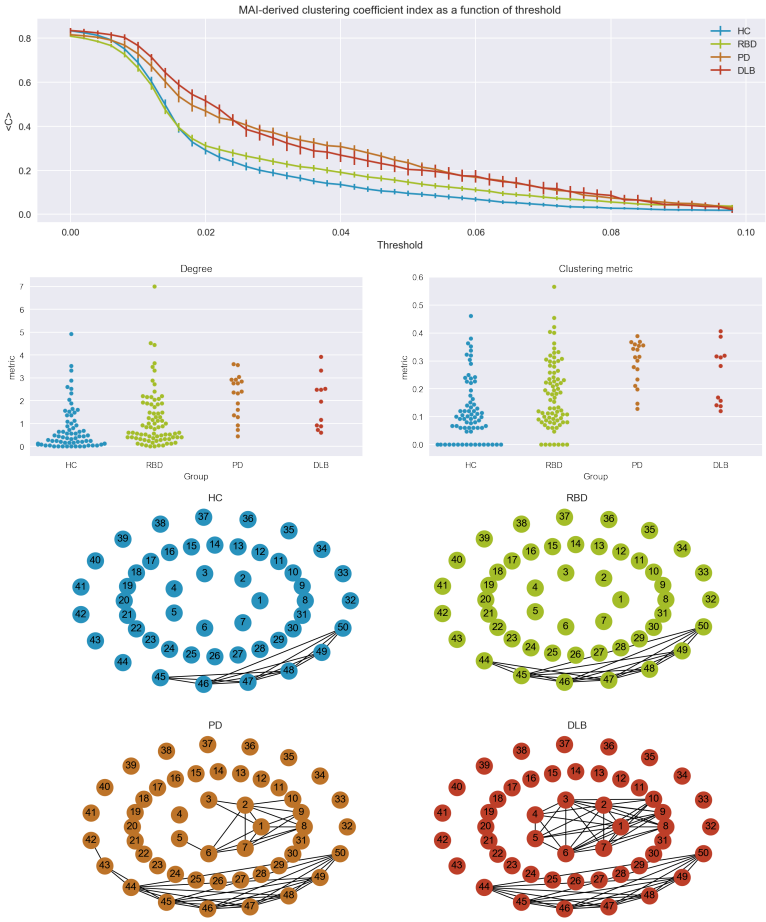

Figure 6: **Graph analysis using connectivity from algorithmic mutual information (μ_0) across frequency bands**. Top: clustering coefficient as a function of threshold for the different groups. Middle: corresponding scatter plots. Bottom. Average graphs (threshold = 0.0425), with each node representing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

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. Data from the latter group display lower complexity, entropy rates and increased cross-frequency amplitude coupling. We have also seen

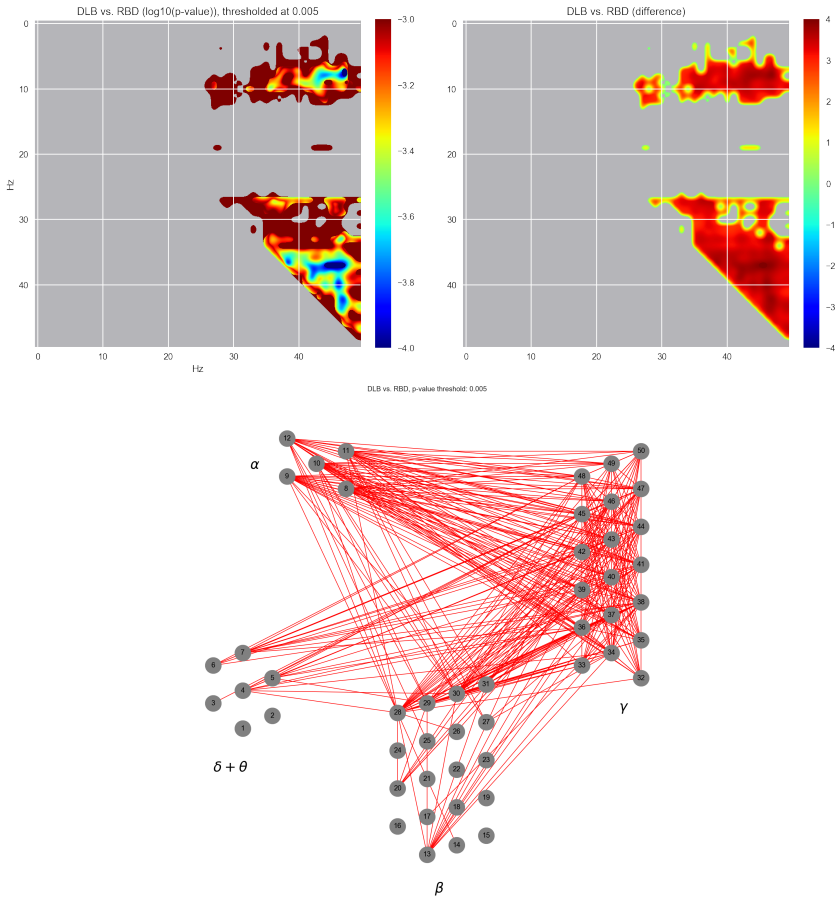

Figure 7: **Statistics of MAI graphs (μ_0) across frequency bands**. Top: statistical differences between two groups at the edge level using per band binarization to remove amplitude scale differences. Bottom. Associated graph displaying links with increased (red) or decreased (blue) scale-independent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

that as expected, LZW and entropy rate provide very similar metrics (although LZW is much faster to compute).

These differences are remarkable as they reflect brain state at the moment of RBD diagnosis, years before conversion to PD or DLB. The performance of these metrics is slightly superior and complementary to those from power anal-

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

Group (N)	ρ_0	H_5	$\langle C_f \rangle$	slow-to-fast
HC (70)	0.81 ± 0.02	0.72 ± 0.01	0.12 ± 0.01	0.21± 0.02
RBD (78)	0.78 ± 0.03	0.69 ± 0.03	0.18 ± 0.03	1.16 ± 0.96
PD (18)	0.68 ± 0.02	0.61 ± 0.02	0.29 ± 0.02	1.21 ± 0.36
DLB (12)	0.69 ± 0.02	0.62 ± 0.02	0.25 ± 0.02	1.20 ± 0.38
Kruskal-Wallis p-value	1×10^{-4}	2×10^{-4}	8×10^{-8}	1×10^{-4}

Table 3: Statistics for comparisons between groups for selected metrics: Wilkoxon 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%)	3×10^{-3} (63%)	5×10^{-3} (63%)
HC vs. PD	1×10^{-4} **(80%)**	7×10^{-7} **(88%)**	2×10^{-2} (69%)
HC vs. DLB	3×10^{-3} **(77%)**	7×10^{-4} **(82%)**	9×10^{-5} **(85%)**
RBD vs. PD	2×10^{-2} (73%)	1×10^{-4} **(80%)**	5×10^{-1} (55%)
RBD vs. DLB	3×10^{-2} (69%)	4×10^{-2} (70%)	8×10^{-3} (74%)
PD vs. DLB	6×10^{-1} (56%)	3×10^{-1} (62%)	4×10^{-1} (59%)
HC vs. PD+DLB	7×10^{-6} **(78%)**	2×10^{-8} **(86%)**	6×10^{-5} **(75%)**
HC vs. rest	3×10^{-3} (63%)	6×10^{-6} (69%)	2×10^{-4} (67%)
HC+RBD vs. PD+DLB	2×10^{-5} **(75%)**	2×10^{-7} **(80%)**	1×10^{-3} (69%)
RBD vs. PD+DLB	5×10^{-4} **(72%)**	4×10^{-5} **(76%)**	5×10^{-2} (62%)

ysis (Rodrigues-Brazète et al., 2016; Latreille et al., 2016), with which they are well aligned. Unlike in previous studies in neurodegenerative diseases (Gomez et al., 2005; Dauwels et al., 2011; Bertrand et al., 2016), power ratio and com-

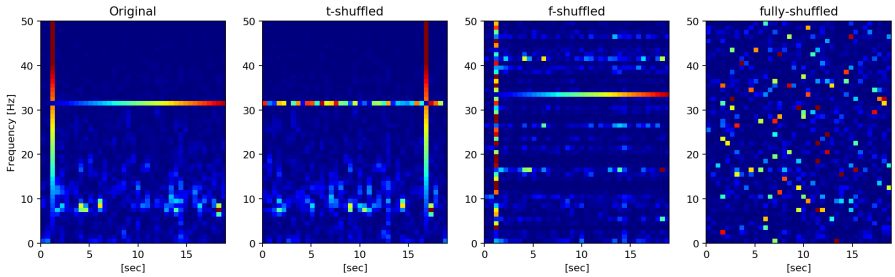

Figure 8: **Frame (intra-epoch) spectrogram shuffling.** From left to right: original, temporally shuffled, frequency and fully shuffled spectrograms. A horizontal and a vertical line has been added for reference.

plexity metrics were not correlated here (a correlation of power metrics such as slow-to-fast ratio and LZW or entropic complexity would not be entirely surprising, since our 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, we note that the global complexity metric provides 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, 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.

4.1. Analysis of sources of compressibility

As we discussed before, once 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.

We carried out tests to find out, reordering and randomizing the data in different ways to see which led to loss of discrimination (see Figure 8). All tests led to an overall increase in complexity, highlighting the fact that compression exploited regularities across channels, across time (within and across frames), across frequency and across epochs. However, 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 does not affect power ratio metrics) increased overall complexity but did not affect complexity metric discriminability. Thus, complexity metrics appear to rely on within-frame temporal regularities in some frequencies (similar patterns appearing in several frequencies 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 power 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 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, highlighting that imbalances in power across bands are an important element in ρ_0 complexity. However, we observed that cross-frequeny MAI differences across groups remained significant and discriminative (see, as an example, Figure 7). This implies that scale-independent cross-frequency (amplitude) temporal regularities remain in the data (typically an increase in connectivity with disease progression).

Our results with regard to the role of high frequencies in changes in complexity can be related to those in Bertrand et al. (2016), 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. They found significant changes, with lower signal variability at timescales sensitive to higher frequencies in patients that developed dementia than those who did not or controls (i.e., more compressibility at higher frequencies). We do note, however, that multiscale entropy and LZW (or entropy rate) measure complexity in different if related ways.

Finally, one can ask how much does stack compression benefit from having access to all frames as opposed to compressing each independently. The answer is that PD and DLB data benefit the most (with a ratio of per frame vs. per stack 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 $<6\times10^{-5}$ four group Kruskal-Wallis). There are longer term persistent regularities in PD and DLB groups.

5. Conclusion

In Ruffini (2017a) we 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 measure using entropy rate and LZW compression of electrophysiological or metabolic brain data (Casali et al., 2013; Schartner et al., 2015; Andrillon et al., 2016; Hudetz et al., 2016; Schartner et al., 2017). First order entropy, entropy rate or LZW provide *upper bounds* on algorithmic complexity.

From our results 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. Full band complexity scores are generally more discriminative than spectral ratios, do not require a pre-selection of the spectral bands to study, and remain robust even if using a few channels. The loss of complexity occurs both in low and high frequencies, and is partly due to increased mutual information within and across bands. These results indicate that differentiation (global complexity) is a good metric for 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 (Tononi et al., 2016; Ruffini, 2017a).

Disclosure

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

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 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.

References

Abásolo D, Horner R, Gómez C, García M, López M. Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and central tendency measure. Medical Engineering & Physics 2006;28(4):315–22.

Albert R, Barabasi AL. Statistical mechanics of complex networksks. Rev Mod Phys 2002;74(47).

Andrillon T, Poulsen AT, Hansen LK, Leger D, Kouider S. Neural markers of responsiveness to the environment in human sleep. The Journal of Neuroscience 2016;36(24):6583–96.

Bertrand JA, McIntosh A, Postuma R, Kovacevic N, Latreille V, Panisset M, Chouinard S, Gagnon J. Brain connectivity alterations are associated with the development of dementia in Parkinson's disease. Brain Connectivitry 2016;6(3):216–24.

Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, Casarotto S, Bruno MA, Laureys S, Tononi G, Massimini M. A theoretically based

index of consciousness independent of sensory processing and behavior. Sci Transl Med 2013;5(198).

Cerra D, Datcu M. Expanding the algorithmic information theory frame for applications to earth observation. Entropy 2013;15(407-415).

Cover TM, Thomas JA. Elements of information theory. 2nd ed. John Wiley & sons, 2006.

Dauwels J, Srinivasan K, Reddy MR, Musha T, Vialatte FB, Latchoumane C, Jeong J, Cichocki A. Slowing and loss of complexity in Alzheimer's EEG: Two sides of the same coin? International Journal of Alzheimer's Disease 2011;2011(Article ID 539621).

Fantini L, Gagnon J, Petit D, Rompre S, Decary A, Carrier J, Montplaisir J. Slowing of electroencephalogram in rapid eye movement sleep behavior disorder. Annals of neurology 2003;53(6):774–80.

Fulda S. Idiopathic REM sleep behavior disorder as a long-term predictor of neurodegenerative disorders. EPMA J 2011;2(4):451–8.

Gomez C, Dubbelin KTEO, Stam CJ, Abasolo D, Berendse HW, Hornero R. Complexity analysis of resting-state MEG activity in early-stage Parkinson's disease patients. Annals of Biomedical Engineering 2011;39(12):2935–44.

Gomez C, Hornero R, Abasolo D, Lopez M, Fernandez A. Decreased Lempel-Ziv complexity in Alzheimer's disease patients' magnetoencephalograms. Conf Proc IEEE Eng Med Biol Soc 2005;5:4514–7.

Grunwald P, Vitanyi P. Shannon information and kolmogorov complexity. arXiv:cs/0410002 2004;.

Hudetz AG, Liu X, Pillay S, Boly M, Tononi G. Propofol anesthesia reduces Lempel-Ziv complexity of spontaneous brain activity in rats. Neurosci Lett 2016;628:132–5.

Iranzo A, Fernández-Arcos A, Tolosa E, Serradell M, Molinuevo JL, Valldeoriola F, Gelpi E, Vilaseca I, Sánchez-Valle R, Lladó A, Gaig C, a JS. Neurodegenerative disorder risk in idiopathic REM sleep behavior disorder: Study in 174 patients. PLoS One 2014;9(2).

Kaspar F, Schuster HG. Easily calculable measure for the complexity of spatiotemporal patterns. Phys Rev A 1987;36(2):842–8.

Kim Y, Kim YE, Park EO, Shin CW, Kim HJ, Jeon B. REM sleep behavior disorder portends poor prognosis in Parkinson's disease: A systematic review. J Clin Neurosci 2017;.

Latreille V, Carrier J, Gaudet-Fex B, Rodrigues-Brazete J, Panisset M, Chouinard S, Postuma RB, Gagnon JF. Electroencephalographic prodromal markers of dementia across conscious states in Parkinson's disease. Brain 2016;138:1189–99.

Lempel A, Ziv J. On the complexity of finite sequences. IEEE Transactions on Information Theory 1976;IT-22(1):75–81.

Li M, Chen X, Li X, Ma B, Vitànyi PMB. The similarity metric. IEEE TRANSACTIONS ON INFORMATION THEORY 2004;50(12).

Li M, Vitanyi P. An Introduction to Kolmogorov Complexity and Its Applications. Spriger Verlag, 2008.

Postuma R, Gagnon J, Vendette M, Fantini M, Massicotte-Marquez J, et al. . Quantifying the risk of neurodegenerative disease in idiopathic REM sleep behavior disorder. Neurology 2009;72:1296–300.

Ray C, Ruffini G, Marco-Pallarés J, Fuentemilla L, Grau C. Complex networks in brain electrical activity. European Phys Lett 2007;79(38004).

Rodrigues-Brazète J, Gagnon J, Postuma R, Bertrand J, D Petit MJ. Electroencephalogram slowing predicts neurodegeneration in rapid eye movement sleep behavior disorder. Neurobiol Aging 2016;37(74-81).

Ruffini G. Models, networks and algorithmic complexity. Starlab Technical Note - arXiv:161205627 2016;TN00339(DOI: 10.13140/RG.2.2.19510.50249).

Ruffini G. An algorithmic information theory of consciousness. Neuroscience of Consciousness 2017a;3(1).

Ruffini G. Lempel-Ziv complexity reference. arXiv:170709848 [csIT] 2017b;.

Ruffini G, Ibañez D, Castellano M, Dunne S, Soria-Frisch A. EEG-driven RNN classification for prognosis of neurodegeneration in at-risk patients. ICANN 2016 2016;.

Schartner M, Seth A, Noirhomme Q, Boly M, Bruno MA, Laureys S, et al. . Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. PLoS ONE 2015;10(8).

Schartner MM, Carhart-Harris RL, Barrett AB, Seth AK, Muthukumaraswamy SD. Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, lsd and psilocybin. Scientific Reports 2017;7(46421).

Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: from consciousness to its physical substrate. Nature Reviews Neuroscience 2016;17:450–61.