

¹ Consciousness & Brain Functional Complexity
² in Propofol Anaesthesia

³ TF Varley^{1,2,5,6,+,*}, A Luppi^{1,2,7,#,*}, I Pappas^{1,2,8}, L Naci^{3,4}, R Adapa¹,
A Owen⁴, DK Menon^{1,9}, EA Stamatakis^{1,2}

⁴ June 23, 2019

⁵ ¹ Division of Anaesthesia, School of Clinical Medicine, University of Cambridge, UK

⁶ ² Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge,
⁷ UK

⁸ ³ Institute for Neuroscience, Trinity College, Dublin, Ireland

⁹ ⁴ The Brain and Mind Institute, Department of Psychology, The University of Western Ontario,
¹⁰ London, Ontario, CA

¹¹ ⁵ Centre for Complex Networks and Systems Research, Indiana University, Bloomington, IN,
¹² USA

¹³ ⁶ Department of Psychological & Brain Sciences, Indiana University, Bloomington, IN, USA

¹⁴ ⁷ Department of Psychology, Harvard University, Cambridge, MA, USA

¹⁵ ⁸ Department of Psychology, University of California Berkeley, Berkeley, CA, USA

¹⁶ ⁹ Wolfson Brain Imaging Centre, University of Cambridge, Cambridge, UK

¹⁷* These authors contributed equally to this research

¹⁸

¹⁹ Corresponding authors:

²⁰ + Thomas Varley: tvarley@iu.edu

²¹ # Andrea Luppi: al857@cam.ac.uk

22

Abstract

23

The brain is possibly the most complex system known to mankind, and its complexity has
24 been called upon to explain the emergence of consciousness. However, complexity can take
25 many forms: here, we investigate measures of algorithmic and process complexity in both the
26 temporal and topological dimension, testing them on functional MRI data obtained from indi-
27 viduals undergoing various levels of sedation with the anaesthetic agent propofol, in two separate
28 datasets. We demonstrate that the various measures are differently able to discriminate between
29 levels of sedation, with temporal measures showing higher sensitivity. Further, we show that
30 all measures are strongly related to a single underlying construct explaining most of the vari-
31 ance, as assessed by Principal Component Analysis, which we interpret as a measure of overall
32 complexity of our data. This overall complexity was also able to discriminate between levels of
33 sedation, supporting the hypothesis that consciousness is related to complexity - independent
34 of how the latter is measured.

35

1 Introduction

36

The science of complex systems has gained increasing prominence in the 21st century. It combines
37 the reductionist ideal of science, with the notion of emergence, whereby high-level phenomena can
38 result from the interactions of simple constituent parts, confirming Aristotle's saying that the whole
39 is more than the sum of its parts [1]. However, complexity science is also a discipline still in its
40 infancy. In particular, due to its appealing and apparently intuitive nature, the notion of complex-
41 ity has remained relatively ill-defined. The interdisciplinary nature of this science has resulted in
42 different fields applying the term complexity to multiple quantities, variously measured. Complex-
43 ity is perhaps best understood as the negation of simplicity. A system exhibits complex behaviour
44 when it is not uniform, stereotyped, or predictable. However, there is a key assumption that this
45 is not sufficient: complexity must emerge from the underlying orderly interactions of a system's
46 components, about which its behaviour must provide information in other words, its unpredictabil-
47 ity must be more than mere randomness, but rather the result of interesting behaviours emerging.
48 Thus, a complex system lies between complete order such as the perfectly predictable regularity of
49 a crystal and complete disorder, as exhibited for instance by the random motion of molecules of

50 a gas. Complexity can be identified in more than one dimension of the same system, too. It may
51 be due to the structure of the interactions between components, such as the connections in a social
52 or biological network. Or it may only become apparent over time, as when it is applied to signals
53 and temporal patterns. Furthermore, there are different ways in which something can be said to be
54 complex, reflected in the different ways that have been developed to estimate complexity. On the
55 one hand, methods from algorithmic information theory such as Shannon entropy and Lempel-Ziv
56 compressibility [2, 3] emphasise unpredictability as the key property for complexity. One downside of
57 such approach, however, is that they would treat a purely random sequence as maximally complex.
58 Alternatively, methods from the physics of dynamical systems focus on the aspect of interactions
59 in the process whether between the system's elements (e.g. synchronisability; [4]), between its
60 present and past states (e.g. Hurst exponent; [5]), or between different scales [6]. In this work,
61 we aim to explore the relation between algorithmic and process measures of complexity, in both
62 the topological and temporal dimensions. We choose to test these measures on a paradigmatically
63 complex system: the human brain. Not only is the brain the source of humans' widely diverse
64 range of behaviours and accomplishments, which is itself suggestive of a highly complex underlying
65 organisation; its structure is also that of a complex network of subnetworks, in turn made of multiple
66 kinds of neurons obeying nontrivial plasticity rules for their interactions. For these reasons, it has
67 been proposed that the brain's complexity may explain another unique property it possesses: con-
68 sciousness. Recent scientific theories of consciousness have emphasised, in one way or another, the
69 brain's complexity as a crucial requirement for consciousness [7, 8, 9, 10]. Anaesthetic drugs such as
70 the GABA-ergic agonist propofol provide a way to controllably and reversibly modulate the brain's
71 state of consciousness. Its complexity, in various aspects, may then be assessed based on signals
72 from noninvasive neuroimaging techniques. In particular, functional MRI (fMRI) has the advantage
73 of providing high spatial resolution, thus allowing for estimation of the brain's network properties in
74 greater detail than afforded by other methods such as EEG. Here, we chose to evaluate measures of
75 algorithmic and process complexity applied to the temporal and topological (network) dimensions,
76 derived from fMRI blood-oxygen-level-dependent (BOLD) signals of volunteers undergoing sedation
77 with propofol, in order to investigate the relationship between the different measures of complexity,
78 as well as determining whether they can be related to different levels of consciousness. We also

79 replicated our results in an independent dataset of propofol anaesthesia, in order to demonstrate
80 their robustness.

81 2 Methods

82 2.1 Data Acquisition & Preprocessing

83 2.1.1 Dataset A

84 Twenty-five healthy volunteer subjects were recruited for scanning. The acquisition procedures are
85 described in detail by Stamatakis et al, [11]: MRI data were acquired on a Siemens Trio 3T scanner
86 (WBIC, Cambridge). Each functional BOLD volume consisted of 32 interleaved, descending, oblique
87 axial slices, 3 mm thick with interslice gap of 0.75 mm and in-plane resolution of 3 mm, field of view
88 = 192x192 mm, repetition time = 2 s, acquisition time = 2 s, time echo = 30 ms, and flip angle 78.
89 We also acquired T1-weighted structural images at 1 mm isotropic resolution in the sagittal plane,
90 using an MPRAGE sequence with TR = 2250 ms, TI = 900 ms, TE = 2.99 ms and flip angle =
91 9°, for localization purposes. Of the 25 healthy subjects, 14 were ultimately retained: the rest were
92 excluded, either because of missing scans (n=2), or due to excessive motion in the scanner (n=9,
93 5mm maximum motion threshold).

94

95 Propofol Sedation

96 Propofol was administered intravenously as a target controlled infusion (plasma concentration mode),
97 using an Alaris PK infusion pump (Carefusion, Basingstoke, UK). Three target plasma levels were
98 used - no drug (baseline), 0.6 mg/ml (mild sedation) and 1.2 mg/ml (moderate sedation). A period
99 of 10 min was allowed for equilibration of plasma and effect-site propofol concentrations. Blood sam-
100 ples were drawn towards the end of each titration period and before the plasma target was altered,
101 to assess plasma propofol levels. In total, 6 blood samples were drawn during the study. The mean
102 (SD) measured plasma propofol concentration was 304.8 (141.1) ng/ml during light sedation, 723.3
103 (320.5) ng/ml during moderate sedation and 275.8 (75.42) ng/ml during recovery. Mean (SD) total
104 mass of propofol administered was 210.15 (33.17) mg, equivalent to 3.0 (0.47) mg/kg. The level of

105 sedation was assessed verbally immediately before and after each of the scanning runs. The three
106 conditions from this dataset are referred to as Awake, Mild and Moderate sedation respectively.

107 **2.1.2 Dataset B**

108 These data were generously provided by the Brain and Mind Institute, Department of Psychology,
109 The University of Western Ontario. Sixteen healthy volunteer subjects were recruited for scanning.
110 Scanning was performed using a 3 Tesla Siemens Tim Trio system with a 32-channel head coil, at
111 the Robarts Research Institute in London, Ontario, Canada. Participants lay supine in the scanner.
112 Function echo-planar images (EPI) were acquired (33 slices, voxel size: 3 x 3 x 3mm; inter-slice gap
113 of 25%, TR=2000ms, TE=30ms, matrix size=64x64, FA=75 degrees). An anatomical volume was
114 obtained using a T1-weighted 3D MPRAGE sequence (32 channel coil, voxel size: 1x 1 x 1mm, TA=
115 5 min, TE=4.25ms, matrix size=240x256, FA=9 degrees).

116

117 **Propofol Sedation**

118 Intravenous propofol was administered with a Baxter AS 50 (Singapore). The infusion pump was
119 manually adjusted using step-wise increases to achieve desired levels of sedation of propofol (Ramsay
120 level 5). Concentrations of intra-venous propofol were estimated using the TIVA Trainer (the
121 European Society for Intravenous Anaesthesia, eurosiva.eu) pharmacokinetic simulation program.
122 If Ramsay level was lower than 5, the concentration was slowly increased by increments of 0.3 $\mu\text{g}/\text{ml}$
123 with repeated assessments of responsiveness between increments to obtain a Ramsay score of 5.
124 Ramsay level 5 was determined as being unresponsive to verbal commands and rousable only by
125 physical stimulus. In contrast to Propofol Dataset A, the two conditions from this dataset are re-
126 ferred to by Awake and Deep sedation respectively, reflecting the substantial increase in sedation
127 depth present in this dataset.

128 **2.1.3 Image Pre-Processing**

129 All of the collected images were preprocessed using the CONN functional connectivity toolbox
130 [12]¹, using the default pre-processing pipeline, which includes realignment and unwarping (motion

¹<http://www.nitrc.org/projects/conn>

131 estimation and correction), slice-timing correction, outlier detection, structural coregistration and
132 spatial normalisation using standard grey and white matter masks, normalization to the Montreal
133 Neurological Institute space (MNI), and finally spatial smoothing with a 6mm full width at half-
134 maximum (FWHM) Gaussian kernel.

135 Temporal preprocessing included nuisance regression using anatomical CompCor to remove noise
136 attributable to white matter and CSF components from the BOLD signal, as well as subject-specific
137 realignment parameters (three rotations and three translations) and their first-order temporal deriva-
138 tives [13]. Linear detrending was also applied, as well as band-pass filtering in the default range of
139 [0.008, 0.09] Hz [14]. For a more detailed discussion of the details of the CONN default preprocessing
140 pipeline, see Whitefield-Gabrieli and Nieto-Castanon, 2012.

141 2.2 Complexity of BOLD Signals

142 To explore the space of different formalizations of complexity, we used algorithms from algorithmic
143 information theory (Lempel-Ziv compressibility, sample entropy, and principal component analysis),
144 as well as from dynamical systems physics (Higuchi fractal dimension, Hurst exponent). Before
145 analysis, the BOLD time-series were transformed by applying the Hilbert transform. The absolute
146 value of the transformed signal was then taken, to remove negative frequencies and ensure that all
147 series were positive. The Hilbert transform was also used to maintain consistency with earlier studies
148 exploring the complexity of brain activity as it relates to consciousness [15, 16].

149 2.2.1 Lempel-Ziv Complexity

150 The Lempel-Ziv algorithm is a computationally tractable method for quantifying the complexity of
151 a data-series by calculating the number of distinct patterns present in the data. For sufficiently large
152 datasets, it is a useful approximation of Kolmogorov complexity, which is famously uncomputable
153 for most strings [2]. The method used here is described in Shartner et al., (2015). Briefly: for every
154 ROI in our parcellated brain, a time-series $F(t)$ is binarized according to the following procedure:

$$F_B(t_i) = \begin{cases} 1, & \text{if } F(t_i) \geq \text{mean}(F(t)) \\ 0, & \text{otherwise} \end{cases}$$

155 The resulting time-series are stacked into a binary matrix $M(X, T)$, where every row corresponds
156 to the time-series $F_B(t)$ for every ROI $x \in X$ and every column is a time-point $t \in T$. The matrix is
157 then flattened orthogonally to T , resulting in a vector V of length $X \times T$, on which the Lempel-Ziv
158 analysis was performed.

159 The Lempel-Ziv algorithm creates a dictionary D , which is the set of binary patterns that make
160 up V and returns a value $LZ_C \propto |D|$. For every time-series $F_B(t) \in X$, a random time-series was
161 created, by shuffling all the entries in $F(t)$. These were stacked into a binary matrix M_{rand} , with the
162 same dimensions as M , however containing only noise. This random matrix was flattened and its
163 LZ_C value calculated. As the randomness of a string increases, $LZ_C \rightarrow 1$, so this value was used to
164 normalize the "true" value of LC_C , which was divided by $LZ_{C_{Rand}}$ to ensure all values were within
165 a range $(0, 1)$.

166 2.2.2 Sample Entropy

167 Sample Entropy (SampEn) quantifies how unpredictable a signal is [3] by estimating the probability
168 that similar sequences of observations in a timeseries will remain similar over time. To compute
169 SampEn, each time-series $X(t)$ of length N is divided into subsections S of length m and the
170 Chebychev distance between two sections S_i, S_j is calculated. Two sections are "similar" if their
171 distance is less than some tolerance r . The procedure is repeated for sections of length $m + 1$. We
172 then calculate the probability that, if two data sequences of length m have distance less than r , then
173 the same two sequences of length $m + 1$ also have distance less than r .

$$SampEn = -\log \frac{A}{B}$$

174 Where A is the number of chunks of length $m + 1$ that are similar (have Chebyshev distance less
175 than r), and B is the number of chunks of length m that are similar. Low values of SampEn would
176 indicate that the signal is highly stereotyped - with a perfectly predictable series, such as [1,1,1, ...]
177 having a SampEn of zero, and SampEn increasing as the series becomes more disordered.

178 SampEn depends on the choice of parameters m and r . Here, we used $m = 2$ and $r = 0.3 \times$
179 $\sigma(X(t))$, where $\sigma()$ is the standard deviation function.

180 SampEn has been used to test the level of sedation induced by propofol and remifentanil in
181 electrophysiological studies (Ferenets et al., 2007), and been shown to be associated with the degree
182 of sedation much like Lempel-Ziv complexity has.

183 **2.2.3 Hurst Exponent**

184 The Hurst Exponent returns an estimate of how predictable a time-series is by quantifying its
185 'memory,' or how dependent the value at time t is on the value at time $t - 1$ [5]. There are a number
186 of algorithms for estimating the Hurst Exponent; here we report results calculated using a rescaled
187 range approach. In it, a time-series $X(t)$ of length N is segmented into non-overlapping sections of
188 length n , $X_i(t)$. For each segment, the cumulative departure from the signal mean is calculated:

$$X'_i(t) = \sum_{t=0}^n x_t - \bar{x}$$

189 where \bar{x} is the mean of $X_i(t)$. The rescaled range of deviations (R/S) is then defined as:

$$\frac{R}{S} = \frac{\max(X'_i(t)) - \min(X'_i(t))}{\sigma(X_i(t))}$$

190 where $\sigma()$ is the standard deviation function. We then compute R/S for all $X_i(t)$ and average
191 them, generating $(R(n)/S(n))$, which is the average scaled range for all the subsections of $X(t)$ with
192 length n . We are left with a power relation, where:

$$\frac{R(n)}{S(n)} \propto n^{-H}$$

193 Where H is the Hurst exponent, and can be extracted by regression.

194 **2.2.4 Higuchi Fractal Dimension**

195 To calculate the temporal fractal dimension, we used the Higuchi method for calculating the self-
196 similarity of a one-dimensional time-series [6], an algorithm widely used in EEG and MEG analysis
197 [17]. From each time-series $X(t)$, we create a new time-series $X(t)_k^m$, defined as follows:

$$X(t)_k^m = x_m, x_{m+k}, x_{m+2k}, \dots, x_{m+\lfloor \frac{N-m}{k} \rfloor k}$$

198 where $m = 1, 2, \dots, k$.

199 For each time-series $X(t)_k^m$ in k_1, k_2, \dots, k_{max} , the length of that series, $L_m(k)$, is given by:

$$L_m(k) = \frac{\left(\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |x_{im+k} - x_{(i-1)k}| \right) \frac{N-1}{\lfloor \frac{N-m}{k} \rfloor k}}{k}$$

200 We then define the average length of the series $\langle L(k) \rangle$, on the interval $[k, L_m(k)]$ as:

$$\langle L(k) \rangle = \sum_{m=1}^k \frac{L_m(k)}{k}$$

201 If our initial time-series $X(t)$ has fractal character, then:

$$\langle L(k) \rangle \propto k^{-D}$$

202 Where D is our desired fractal dimension. The Higuchi algorithm requires a pre-defined k_{max}
203 value as an input, along with the target time-series. This value is usually determined by sampling
204 the results returned by different values of k_{max} and selecting a value based on the range of k_{max}
205 where the fractal dimension is stable. For both datasets, we sampled over a range of powers of
206 two ($2, \dots, 128$). Due to the comparably small size of BOLD time-series, the range of k_{max} values
207 that our algorithm could process without returning an error was limited. We ultimately decided on
208 $k_{max} = 32$ for Dataset A and $k_{max} = 64$ for the Dataset B.

209 2.2.5 PCA of BOLD Signals

210 Principal component analysis (PCA) is commonly used to compress data by finding the dimensions
211 that encode the maximal variance in a high-dimensional dataset. Here, we use PCA in a matter
212 similar to Lempel-Ziv complexity, to relate the complexity of sets of BOLD signals to their compress-
213 ability. The more algorithmically random the dataset, the more orthogonal dimensions are required
214 to describe the dataset, which we took advantage of to attempt to quantify the complexity of our
215 BOLD time-series data. We constructed a large array of un-binariized BOLD signals, $M(X, T)$ to

216 which we applied a standard feature scaler from Scikit-Learn [18] to ensure all values had a mean
217 of zero and unit variance, and then a PCA function, recording recorded how many dimensions were
218 required to cumulatively describe 95% of the variance in the original dataset. We used this value as
219 our measure of data complexity.

220 2.3 Complexity of Functional Connectivity Graphs

221 Networks are a common example of complex system, and perhaps none more so than the human
222 brain, which can be considered as a network at multiple scales. A network, or graph, is represented
223 mathematically as an object comprised of nodes (in this case, cortical regions) and the connections
224 between them, or edges (in this case, functional connectivity given by statistical association of
225 the regions' BOLD time-series). Investigating how the complexity of brain functional networks is
226 affected by the anaesthetic drug propofol is therefore a clear way of testing our hypothesis that loss
227 of consciousness should reduce the brain's level of complexity.

228 2.3.1 Formation of Functional Connectivity Networks

229 To construct brain functional connectivity networks, the preprocessed BOLD time-series data were
230 extracted from each brain in CONN and the cerebral cortex was segmented into distinct ROIs,
231 using the 234-ROI parcellation of the Lausanne atlas [19]. Each time-series $F(t)$ was transformed
232 by taking the norm of the Hilbert transform, to maintain consistency with the time-series analysis.

$$H(t) = |Hilbert(F(t))|$$

233 Every time-series $H(t)$ was then correlated against every other time-series, using the Pearson
234 Correlation, forming a matrix M such that:

$$M_{ij} = \rho(H_i(t), H_j(t))$$

235 The matrices were then filtered to remove self-loops, ensuring simple graphs, and all negative
236 correlations were removed:

$$M_{ij} = \begin{cases} 0, & \text{if } i = j \\ 0, & \text{if } M_{ij} < 0 \\ M_{ij}, & \text{otherwise} \end{cases}$$

²³⁷ Finally, the matrices were binarized with a $k\%$ threshold, such that:

$$M'_{ij} = \begin{cases} 1, & \text{if } M_{ij} \geq P_k \\ 0, & \text{otherwise} \end{cases}$$

²³⁸ The results could then be treated as adjacency matrices defining functional connectivity graphs,
²³⁹ where each row M_i and column M_j corresponds to an ROI in the initial cortical parcellation, and
²⁴⁰ their connection being represented by the corresponding cell in the matrix. For each graph theoretical
²⁴¹ analysis, a range of percentage thresholds ($k\%$) were tested to ensure that any observed effects were
²⁴² not an artefact of one particular threshold, and are consistent over different graph topologies.

²⁴³ 2.3.2 Algebraic Connectivity

²⁴⁴ Algebraic connectivity (AC) is a measure of graph connectivity derived from spectral graph theory
²⁴⁵ [4], which gives an upper bound on the classical connectivity of a graph. As such, it is often used as a
²⁴⁶ measure of how well-integrated a graph is and how robust it is to damage, in the sense of the number
²⁴⁷ of connections that must be removed before it is rendered disconnected. Unlike classical connectivity,
²⁴⁸ which must be calculated by computationally intensive brute-force methods, AC is quite easy to
²⁴⁹ find for even quite large graphs. AC is also a measure of graph synchronizability and emerges
²⁵⁰ from analysis of the Kuramoto model of coupled oscillators [20]. For a simple example, imagine
²⁵¹ placing identical metronomes at every vertex of a graph and allowing the vibrations to propagate
²⁵² along the edges. The synchronizability describes the limit behaviour of how long it will take all the
²⁵³ metronomes to synchronize. Here we use AC as a proxy measure of synchronisability to capture the
²⁵⁴ possible temporal dynamics of the brain networks modelled by our functional connectivity graphs.

²⁵⁵ The AC of a graph G is formally defined as the first non-zero eigenvalue of the Laplacian matrix
²⁵⁶ L_G associated with G . L_G is derived by subtracting the adjacency matrix A_G from the degree matrix

257 D_G :

$$L_G = D_G - A_G$$

258 As every row and column of L_G sum to zero, and it is symmetric about the diagonal, the
259 imaginary part of every eigenvalue in the spectrum of L_G is zero, and if G is a fully-connected
260 graph, then:

$$0 = \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_{max}$$

261 To ensure that we were capturing the full topology of the graph, we calculated λ_2 for each graph
262 at multiple thresholds [10, 20, 30, ... 90], creating a curve $\Lambda = [\lambda_{2_{10}}, \lambda_{2_{20}}, \lambda_{2_{30}} \dots \lambda_{2_{90}}]$. We then
263 integrated Λ using the trapezoid method to arrive at our final value $AC = \int \Lambda dx$.

264 **2.3.3 Graph Compressibility**

265 In contrast to AC, which we use to explore the limit behaviour of possible brain temporal dynamics,
266 our measure of graph compressibility is purely algorithmic, and estimates the Kolmogorov complexity
267 of a graph: that is, the size of a computer program necessary to fully recreate a given graph G . To
268 do this, we re-employ the Lempel-Ziv algorithm originally used to calculate the LZ_C score of BOLD
269 signals. Here we use it to calculate a related measure, LZ_G , which is the length of a dictionary
270 required to describe the adjacency matrix A_G of a given graph.

271 To calculate LZ_G , we take a binary adjacency matrix and flatten it into a single vector V , and
272 then run the Lempel-Ziv algorithm on that vector. As a binary vector of length l can be used to
273 perfectly reconstruct an adjacency matrix defining a graph with \sqrt{l} vertices (so long as l is a square
274 number, of course), we take V to be equivalent to a program defining G . As with AC, to ensure that
275 we were capturing the full topology of G , we calculated the Lempel-Ziv complexity of the binary A_G
276 at the same nine thresholds [10...90], and then defined LZ_G as the integral of the resulting curve of
277 complexity values.

278 2.4 Higher-Order Measures

279 Once we had calculated individual measures of complexity, we tested how they related to each-other,
280 and (for Dataset A) serum concentrations of propofol. We correlated each one against all others to
281 construct a correlation matrix which describes, how different metrics cluster.

282 We also did a principal component analysis on the set of all results. We hypothesized that,
283 despite variability in the effectiveness of the individual measures, there should be a single, underlying
284 component reflecting a shared factor of "complexity". We further hypothesised that this underlying
285 factor should be predictive of both the level of consciousness, and (in Dataset A), of the individual
286 serum concentration of propofol.

287 2.5 Statistical Analysis

288 All analysis was carried out using the Python 3.6 programming language in the Spyder IDE ², using
289 the packages provided by the Anaconda distribution ³. All packages were in the most up-to-date
290 version. Packages used include SciKit-Learn [18], NumPy [21], SciPy [22], and NetworkX [23]. Unless
291 otherwise specified, all the significance tests are non-parametric: given the small sample sizes and
292 heterogeneous populations, normal distributions were not assumed. Wilcoxon Signed Rank test was
293 used to compare drug conditions against their respective control conditions.

294 3 Results

295 3.1 Temporal Algorithmic Complexity

296 3.1.1 Lempel-Ziv Compressibility

297 The first measure of algorithmic complexity we used was normalized Lempel-Ziv compressibility
298 [15, 16] of BOLD signals. We found significant differences between conditions in both Dataset
299 A and Dataset B. In Dataset A, Kruskal-Wallis Analysis of Variance found significant differences
300 between all three conditions ($H(10.57)$, $p=0.005$), and post-hoc analysis with Wilcoxon Signed-Rank
301 test found significant differences between the Awake and Mild conditions ($W(21)$, $p=0.05$), Awake

²<https://github.com/spyder-ide/spyder>

³<https://www.anaconda.com/download>

and Moderate conditions ($W(9)$, $p=0.006$), and Mild and Moderate conditions ($W(4)$, $p=0.002$). In Dataset B, the Wilcoxon test found significant difference between the Awake and Deep conditions ($H(19)$, $p=0.011$). In both datasets, the Awake condition had the highest complexity, and as the depth of sedation increased, the associated LZC decreased. In Dataset A, the transition from Awake to Moderate showed $\Delta = -0.029 \pm 0.027$, and in Dataset B, the transition from Awake to Deep showed $\Delta = -0.029 \pm 0.039$. We note that these two results are remarkably similar, although this is likely a coincidence. For full results from Dataset A, see Table 1, and for Dataset B, Table 2. In the propofol sedation conditions of Dataset A (Mild and Moderate), we found significant negative correlations between LZC and serum concentrations of propofol ($r=-0.55$, $p=0.002$), see Figure 3A.

These results are consistent with previous findings that Lempel-Ziv compressibility of spontaneous brain activity is discriminative of level of consciousness in humans [15, 16] and animals [24].

Of all the time-series measures described, the LZC algorithm described here is distinct in that it communicates information about the spatial complexity as well as the temporal complexity. This is because, unlike other measures like Sample Entropy or Higuchi Fractal Dimension which are calculated on 234 individual time-series and then averaged, LZC is calculated on an entire dataset, which has been flattened column-wise, as was done in [15, 16, 25], by "stacking" each column on top of the next, resulting in a one-dimensional vector where the first 234 elements are the first column, the second 234 elements are the second column, etc. This means that the vector V (see Methods section) can be divided into 234 segments where every entry corresponds to the coarse activation of a distinct brain region at the same time. The result is that each entry in the dictionary D created by the Lempel-Ziv algorithm corresponds, not to a series of samples from a single ROI, but rather a distribution of cortical regions that are "on" or "off."

3.1.2 Sample Entropy

We found significant decreases in the Sample Entropy of BOLD signals under anaesthesia in both Datasets A and B. In Dataset A, Kruskal-Wallis Analysis of Variance found a significant difference between all three conditions ($H(12.94)$, $p=0.002$) and post-hoc analysis with the Wilcoxon Signed-Rank test found significant differences between the Awake and Moderate conditions ($W(6)$, $p=0.004$), and Mild versus Moderate conditions ($W(8)$, $p=0.005$), but not the Awake versus Mild conditions.

330 In Dataset B there was a significant difference between the Awake and Deep conditions ($W(21)$,
331 $p=0.015$). As with the LZC analysis, the Awake condition had the highest Sample Entropy in
332 both Datasets A and B, with the mean value decreasing with increasing sedation. In Dataset A,
333 we observed $\Delta = 0.036 \pm 0.035$ from Awake to Moderate, and in Dataset B we observed $\Delta =$
334 -0.023 ± 0.031 . In the Mild and Moderate conditions of Dataset A, we found a significant negative
335 correlation between serum concentration of propofol and Sample Entropy of BOLD signals ($r=-0.53$,
336 $p=0.003$).

337 These results are consistent with both the LZC results reported above and the findings of Ferentes
338 et al., (2007), who found that Sample Entropy decreased with increasing sedation in much the same
339 way that LZC does.

340 3.1.3 PCA of BOLD Signals

341 As with LZC, the PCA-based measure of BOLD signal complexity returns a measure of how com-
342 pressible the set of data are as a proxy for complexity, by identifying the number of components
343 required to explain a fixed proportion of the variance in the data. A larger number of components
344 to explain the same amount of variance would indicate less compressibility of the data. Thus, we
345 hypothesized that as level of sedation increased, so would the compressibility of BOLD signals,
346 as measured by the number of components required to explain 95% of the variance. In Dataset
347 A, Kruskal-Wallis analysis of variance found significant differences between all three conditions
348 ($H(8.13)$, $p=0.017$), and post-hoc testing found significant differences between all three sets of con-
349 ditions: Awake vs. Mild ($W(9)$, $p=0.03$), Awake vs. Moderate ($W(6)$, $p=0.016$), and Mild vs.
350 Moderate ($W(11)$, $p=0.048$). In Dataset B, we found a significant difference between Awake and
351 Deep ($W(4)$, $p=0.002$). As before, there was a consistent pattern of increasing mean compressibility
352 (and a consequent decreasing number of required components) as sedation increased. In Dataset A,
353 $\Delta = -2.214 \pm 2.73$ from Awake to Moderate, and in Dataset B, $\Delta = -5.188 \pm 4.68$. Here, the Δ is
354 negative because the number of components decreased between the Awake and sedated conditions,
355 and is non-integer because it is the average over all subjects in the datasets. Of all the measures
356 of BOLD signal compressibility, this was the only measure that did not significantly correlate with
357 serum propofol concentration in the Mild and Moderate conditions in Dataset A.

358 As with LZC and the SampEn, these results indicate that as propofol-induced sedation increases,
359 the algorithmic complexity of BOLD signals decreases. All measures of complexity discussed so far
360 support each-other, despite being a variety of linear and non-linear algorithms.

361 3.2 Temporal Process Complexity

362 3.2.1 Hurst Exponent

363 The Hurst Exponent was the only measure that we hypothesized would increase as consciousness
364 was lost, rather than decrease, since as a signal becomes more predictable, its Hurst Exponent tends
365 towards unity [5]. In both Datasets A and B we found significant differences between conditions. In
366 Dataset A, Kruskal-Wallis Analysis of Variance found an omnibus difference ($H(9.11)$, $p=0.01$), and
367 post-hoc testing found significant differences between Awake and Mild ($W(16)$, $p=0.022$), Awake
368 and Moderate ($W(8)$, $p=0.005$), and Mild and Moderate ($W(17)$, $p=0.026$). In Dataset B, we found
369 significant differences between the Awake and Deep conditions ($W(26)$, $p=0.02$). Unlike the previous
370 two metrics, and as we expected, we found a relative increase in the Hurst Exponent as sedation
371 increased: in Dataset A, we found $\Delta = 0.014 \pm 0.014$ from Awake to Moderate sedation, and in
372 Dataset B we found $\Delta = 0.01 \pm 0.016$ from Awake to Deep sedation. In Dataset A, we found a
373 significant correlation between serum concentration of propofol and Hurst Exponent in the Mild and
374 Moderate sedation conditions ($r=0.393$, $p=0.039$).

375 This is consistent with our initial hypothesis that as sedation increased and consciousness was lost,
376 the BOLD signals would become more predictable, as measured by an increasing Hurst Exponent.

377 3.2.2 Higuchi Fractal Dimension

378 The Higuchi Fractal Dimension was one of the least sensitive measures of BOLD signal complex-
379 ity sampled here. In Dataset A, Kruskal-Wallis Analysis of Variance found a significant difference
380 between all three conditions ($H(8.27)$, $p=0.016$), and post-hoc analysis found significant differences
381 between the Awake and Moderate conditions ($W(15)$, $p=0.019$) and the Mild and Moderate condi-
382 tions ($W(17)$, $p=0.026$), but not the Awake and Mild conditions. In Dataset B we found a significant
383 difference between the Awake and Deep conditions ($W(23)$, $p=0.02$). As with LZC and Sample En-

384 tropy, the Awake condition had the highest mean fractal dimension in both samples, which went
385 down as sedation increased: in Dataset A $\Delta = -0.024 \pm 0.032$ from Awake to Moderate and in
386 Dataset B, $\Delta = -0.018 \pm 0.027$ from Awake to Deep. Surprisingly, the Higuchi Fractal dimen-
387 sion showed a very strong negative correlation with serum propofol concentration in the Mild and
388 Moderate conditions of Dataset A ($r=-0.614$, $p=0.0005$).

389 The finding that Higuchi Fractal dimension was relatively less able to discriminate between level
390 of consciousness than LZC or Sample Entropy but more predictive of serum propofol concentration
391 is interesting. While it is hard to come up with a definitive interpretation, it may suggest that there
392 is some variable factor in individuals that makes their level of consciousness more or less resistant
393 to the changes in brain activity (as measured by Higuchi Fractal Dimension) induced by propofol,
394 or that the plasma concentration data offer more resolution, extending beyond the three artificially
395 imposed bins of Awake, Mild and Moderate sedation.

396 3.3 Topological Algorithmic Complexity

397 3.3.1 Algebraic Connectivity

398 Our first of two measures of functional network complexity is algebraic connectivity, which returns
399 information about the robustness of the network to removal of elements [26]. In Dataset A, Kruskal-
400 Wallis analysis found a significant difference in algebraic connectivity between all three conditions
401 ($H(9.654)$, $p=0.008$). Post-hoc analysis found significant differences between the Awake and Mod-
402 erate conditions ($W(12)$, $p=0.011$) and the Mild and Moderate conditions ($W(15)$, $p=0.019$), but
403 not the Awake versus Mild conditions. In Dataset B, we found a significant difference between the
404 Awake and Deep conditions ($W(23)$, $p=0.02$). As before, in Datasets A and B the Awake condition
405 had the highest mean algebraic connectivity, with mean values dropping as sedation increased. In
406 Dataset A, $\Delta = -107.67 \pm 116.98$ from Awake to Moderate, while in Dataset B, -906.09 ± 1235.07 .
407 Despite the ability of algebraic connectivity to discriminate between conditions, there was no signif-
408 icant correlation with serum propofol concentration in the Mild and Moderate conditions of Dataset
409 A.

410 These results suggest that, while graph theoretical measures may be predictive of level of con-

411 consciousness in propofol anaesthesia, algebraic connectivity in particular seems to lack the discrimi-
412 native power of direct analysis on BOLD signals. Nevertheless, these results are promising as they
413 show that the topological complexity of functional brain networks can communicate information
414 relevant to the level of consciousness of an individual.

415 **3.4 Topological Process Complexity**

416 **3.4.1 Graph Lempel-Ziv Compressibility**

417 The final metric we tested, and the second measure of topological complexity, was the compressibility
418 of functional connectivity adjacency matrices using the Lempel-Ziv algorithm. This was the weakest
419 of all the measures explored: the only significant difference was in Dataset A, between the Awake
420 and Moderate conditions ($W(11)$, $p=0.009$), although in Dataset B there was a similar trend that
421 approached, but did not reach significance ($W(32)$, 0.06). The general trend of Awake having the
422 highest value which decreased under increasing sedation was conserved (although note the large
423 standard deviations): in Dataset A $\Delta = -2885.71 \pm 4937.34$ and in Dataset B $\Delta = -4923.44 \pm$
424 8967.29. There was no significant correlation between graph compressibility and serum propofol
425 concentrations in Dataset A.

426 While this is clearly the weakest result, in the context of the others, we still find its success at
427 discriminating between the Awake and Moderate conditions of Dataset A intriguing, and suspect
428 that in a larger set of data it may have more discriminative power. The relationship between
429 consciousness and network compressibility may not be as direct as when performing analysis such
430 as LZC on BOLD signals, but these results suggest this is an area worth exploring.

431 **3.5 Higher Order Analysis of Overall Complexity**

432 Every metric, when correlated against every other metric, showed a highly significant correlation
433 (see Figure 2), all of which were significant with the sole exception of the correlation between the
434 number of PCA components required to explain the majority of the variance and the Hurst expo-
435 nent in Dataset A. We had hypothesized that, if the different kinds of complexity explored here
436 (algorithmic and process-based, in both the temporal and topological dimensions) all were ways to

437 quantify an underlying construct of overall complexity, then there should be a single component
438 that explains the majority of the variance of the results. In Dataset A, we found that the principal
439 component explained 67.07% of the variance in the set of results and in Dataset B the principal
440 component explained 71.05% of the variance of the results. In both datasets, this component corre-
441 lated extremely highly with each metric: in Dataset A it correlated most highly with LZC ($r=-0.947$,
442 $p \leq 1 \times 10^{-5}$), followed by Sample Entropy ($r=-0.929$, $p \leq 1 \times 10^{-5}$). In Dataset B, these two were
443 also the most highly correlated with the principal component, although the order was flipped, with
444 Sample Entropy having the highest correlation ($r=-0.95$, $p \leq 1 \times 10^{-5}$), followed by LZC ($r=-0.932$,
445 $p \leq 1 \times 10^{-5}$). When broken down by condition, in both Datasets, the principal component was
446 able to discriminate between states of consciousness: in Dataset A the Kruskal-Wallis test found a
447 significant difference between all three conditions ($H(12.048$, $p=0.002$), and post-hoc testing found
448 significant differences between the Awake and Moderate conditions ($W(8)$, $p=0.005$), and the Mild
449 and Moderate conditions ($W(3)$, $p=0.001$) but not the Awake and Mild conditions. In Dataset B
450 we found a significant difference between the Awake and Deep conditions ($W(8)$, $p=0.002$). In the
451 Mild and Moderate conditions of Dataset A, the principal component significantly correlated with
452 serum concentrations of propofol ($r=0.531$, $p=0.004$). Thus, the principal component derived from
453 multiple specific measures of complexity can be related to states of consciousness in the human
454 brain, and may be identified with the overall complexity of the dataset.

455 4 Discussion

456 In the present work, we have investigated measures of complexity from algorithmic information
457 theory and the physics of dynamical systems, as they apply to the temporal and topological (network)
458 dimensions of functional MRI brain data from individuals under different levels of propofol sedation.
459 Two main insights can be derived from our results. The first is that, at least in the context of
460 the human brain, different measures purporting to quantify complexity are indeed related to some
461 underlying common construct, regardless of the dimension along which they measure complexity,
462 or the aspect of complexity that they measure. This provides much-needed validation to the idea
463 that a dataset - and the system from which it derives - can be considered complex *tout court*,

464 rather than just being complex in a specific dimension, and according to a specific way of assessing
465 complexity. We term this the overall complexity of the system or dataset. In turn, this suggests
466 that it is appropriate to use the term complexity for the various specific measures, because there
467 does seem to exist a common underlying property of the data that they tap into. In particular, we
468 have demonstrated that the complexity of the human brain activity, as inferred from fMRI BOLD
469 signals, is modulated by one's state of consciousness. This was observed both with the individual
470 measures - validating and extending previous results - and, most importantly, with the underlying
471 construct of overall complexity, which demonstrates its validity as a construct. The latter is also
472 reinforced by the fact that we were able to replicate this finding with a separate dataset.

473 Secondly, it is important to observe that different complexity measures, though correlated to each
474 other and related to the same underlying construct of overall complexity, are nevertheless sensitive
475 to different aspects of the data. In particular, measures operating along the temporal dimension
476 appeared especially sensitive at discriminating between levels of sedation; conversely, topological
477 measures failed to discriminate between Awake and Mild conditions in Dataset A, and also did
478 not correlate with propofol serum levels. This suggests that the temporal dimension of the human
479 brain's complexity, as derived from BOLD signal timeseries (despite their limited temporal resolution
480 compared to EEG), may be especially vulnerable to loss of consciousness, at least as it is induced
481 by the GABAergic agent propofol. Further work may seek to identify whether this effect is uniform
482 across cortical regions, or whether specific areas' timeseries are more largely affected by propofol
483 than others. This represents a novel insight regarding the ways in which anaesthetic drugs such as
484 propofol intervene on the brain to cause unconsciousness. Additionally, it would be worth exploring
485 whether this observation of different sensitivity of temporal and topological measures of complexity
486 is drug-specific, or if instead it is a generalisable feature of how the brain loses consciousness. Thus,
487 one future direction of research is to apply these same metrics to states of consciousness induced by
488 different anaesthetic agents, whose molecular mechanisms of action can vary widely. Disorders of
489 consciousness (DOC) due to severe brain injury may also represent a crucial future step for research:
490 unlike anaesthetics, DOC involve changes in the physical structure of the brain, which is bound to
491 impact the topology of brain networks. Investigating how this impacts the relation between different
492 measures and dimensions of complexity will provide further understanding into the relation between

493 complexity and consciousness in the brain. Additionally, as MRI is already a routine part of care for
494 DOC patients, algorithms such as those explored here might be helpful in determining the presence
495 or absence of consciousness in ambiguous states such as minimally conscious state.

496 Importantly, our results also show that, despite the relative temporal paucity of information in
497 BOLD signals, these signals carry sufficient information to discriminate between states of conscious-
498 ness. While preliminary, these findings suggest that the process complexity of individual BOLD
499 signals is at least partially re-encoded as topological complexity when forming functional connec-
500 tivity networks. One possible avenue of future work is to explore the parameters under which
501 this conservation of complexity is maximized (different similarity functions, different thresholding
502 procedures, etc), in order to increase the sensitivity of these measures. Crucially, even higher dis-
503 criminative power may be achieved by applying the same analyses to measures with higher temporal
504 information, such as EEG, which may then improve anaesthetists' ability to detect unwanted resid-
505 ual consciousness in patients, thereby avoiding the rare but extremely distressing condition known
506 as intraoperative awareness [27].

507 Nevertheless, our work also presents a number of limitations, and these should be borne in
508 mind when evaluating the present results. Firstly, as already mentioned the temporal information
509 available in the BOLD signal is limited, and it is also not a direct measure of neural activity.
510 Additionally, our analysis pipeline involved removing the negative correlations between brain regions,
511 before the network analysis. While negative correlations are unclear in origin and interpretation, and
512 removing them is the most common approach, it is known that they are altered during anaesthesia
513 and other states of unconsciousness [28, 29, 30]; thus, ignoring them may have different effects
514 on conscious versus unconscious brain networks, which could explain the reduced sensitivity of
515 topological measures. Thirdly, in Dataset A the state of consciousness was determined based on the
516 estimated propofol concentration, rather than behaviour, so that different individuals' susceptibility
517 to the drug may have led to different levels of sedation, despite the same level of propofol. However,
518 this concern is mitigated by the replication of our results in Dataset B, where sedation was deeper and
519 it was assessed behaviourally, so that all individuals met the same criteria. Finally, the measures
520 of complexity explored here are but a subset of those that have been proposed over the years in
521 the literature. Future research could benefit from expanding this repertoire, for instance including

522 estimates of Phi, a measure of integrated information derived from neural complexity [10], which
523 has been proposed to quantify a system's consciousness [9, 31]

524 **5 Conclusion**

525 We have investigated measures of algorithmic and process complexity of fMRI BOLD signal in both
526 the temporal and topological dimensions, at various levels of consciousness induced by propofol
527 sedation. Our results demonstrate that complexity measures are differently able to discriminate
528 between levels of sedation, with temporal measures showing higher sensitivity. Additionally, all
529 measures were strongly correlated, and most of the variance could be explained by a single underlying
530 construct, which may be interpreted as a more general quantification of complexity, and which
531 also proved capable of discriminating between levels of sedation, demonstrating a relation between
532 consciousness and complexity.

533 **Acknowledgements**

534 This work was supported by a grant from the Wellcome Trust: Clinical Research Training Fel-
535 lowship to Ram Adapa (Contract grant number: 083660/Z/07/Z); from the Canada Excellence
536 Research Chairs program (215063) and the Canadian Institute for Advanced research (CIFAR)[to
537 AMO]; Cambridge Biomedical Research Centre and NIHR Senior Investigator Awards [to DKM], the
538 Stephen Erskine Fellowship at Queens College, Cambridge [to EAS], the LOral-Unesco for Women
539 in Science Excellence Research Fellowship [to LN]; the British Oxygen Professorship of the Royal
540 College of Anaesthetists [to DKM] and the Gates Cambridge Trust [to AIL]. I Pappas received
541 funding from the Oon Khye Beng Ch'Hia Tsio Studentship for Research in Preventive Medicine, ad-
542 ministered via Downing College, University of Cambridge; The research was also supported by the
543 NIHR Brain Injury Healthcare Technology Co-operative based at Cambridge University Hospitals
544 NHS Foundation Trust and University of Cambridge. We would like to thank Victoria Lupson and
545 the staff in the Wolfson Brain Imaging Centre (WBIC) at Addenbrookes Hospital for their assistance
546 in scanning. We would like to thank Dian Lu and Olaf Sporns for useful discussions, and all the

547 participants for their contribution to this study.

548 References

- 549 [1] James P Crutchfield. Between order and chaos. *Nature Physics*, 8(1):17, 2012.
- 550 [2] Gregory J. Chaitin. On the Simplicity and Speed of Programs for Computing Infinite Sets of
551 Natural Numbers. *J. ACM*, 16(3):407–422, July 1969.
- 552 [3] J. S. Richman and J. R. Moorman. Physiological time-series analysis using approximate en-
553 tropy and sample entropy. *American Journal of Physiology. Heart and Circulatory Physiology*,
554 278(6):H2039–2049, June 2000.
- 555 [4] Ljupco Kocarev and Gbor Vattay. Synchronization in Complex Networks. In Ljupco Kocarev
556 and Gbor Vattay, editors, *Complex Dynamics in Communication Networks*, Understanding
557 Complex Systems, pages 309–328. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.
- 558 [5] H. E. Hurst. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society*
559 *of Civil Engineers*, 116(1):770–799, 1951.
- 560 [6] T. Higuchi. Approach to an irregular time series on the basis of the fractal theory. *Physica D:*
561 *Nonlinear Phenomena*, 31(2):277–283, June 1988.
- 562 [7] Robin Lester Carhart-Harris, Robert Leech, Peter John Hellyer, Murray Shanahan, Amanda
563 Feilding, Enzo Tagliazucchi, Dante R Chialvo, and David Nutt. The entropic brain: a theory of
564 conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in human*
565 *neuroscience*, 8:20, 2014.
- 566 [8] Robin L Carhart-Harris. The entropic brain-revisited. *Neuropharmacology*, 2018.
- 567 [9] Giulio Tononi. An information integration theory of consciousness. *BMC neuroscience*, 5(1):42,
568 2004.

- 569 [10] Giulio Tononi, Olaf Sporns, and Gerald M Edelman. A measure for brain complexity: relating
570 functional segregation and integration in the nervous system. *Proceedings of the National
571 Academy of Sciences*, 91(11):5033–5037, 1994.
- 572 [11] Emmanuel A. Stamatakis, Ram M. Adapa, Anthony R. Absalom, and David K. Menon. Changes
573 in Resting Neural Connectivity during Propofol Sedation. *PLOS ONE*, 5(12):e14224, December
574 2010.
- 575 [12] Susan Whitfield-Gabrieli and Alfonso Nieto-Castanon. Conn: a functional connectivity toolbox
576 for correlated and anticorrelated brain networks. *Brain Connectivity*, 2(3):125–141, 2012.
- 577 [13] Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T. Liu. A component based noise
578 correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1):90–
579 101, August 2007.
- 580 [14] Michael D. Fox, Abraham Z. Snyder, Justin L. Vincent, Maurizio Corbetta, David C. Van Essen,
581 and Marcus E. Raichle. The human brain is intrinsically organized into dynamic, anticorrelated
582 functional networks. *Proceedings of the National Academy of Sciences of the United States of
583 America*, 102(27):9673–9678, July 2005.
- 584 [15] Michael Schartner, Anil Seth, Quentin Noirhomme, Melanie Boly, Marie-Aurelie Bruno, Steven
585 Laureys, and Adam Barrett. Complexity of Multi-Dimensional Spontaneous EEG Decreases
586 during Propofol Induced General Anaesthesia. *PLOS ONE*, 10(8):e0133532, August 2015.
- 587 [16] Michael M. Schartner, Andrea Pigorini, Steve A. Gibbs, Gabriele Arnulfo, Simone Sarasso,
588 Lionel Barnett, Lino Nobili, Marcello Massimini, Anil K. Seth, and Adam B. Barrett. Global
589 and local complexity of intracranial EEG decreases during NREM sleep. *Neuroscience of Con-
590 sciousness*, 3(1), January 2017.
- 591 [17] Srdjan Kesi and Sladjana Z. Spasi. Application of Higuchi's fractal dimension from basic to
592 clinical neurophysiology: A review. *Computer Methods and Programs in Biomedicine*, 133:55–
593 70, September 2016.

- 594 [18] Fabian Pedregosa, Gal Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion,
595 Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Van-
596 derplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and douard
597 Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*,
598 12:28252830, October 2011.
- 599 [19] Alessandro Daducci, Stephan Gerhard, Alessandra Griffa, Alia Lemkadem, Leila Cammoun,
600 Xavier Gigandet, Reto Meuli, Patric Hagmann, and Jean-Philippe Thiran. The Connectome
601 Mapper: An Open-Source Processing Pipeline to Map Connectomes with MRI. *PLOS ONE*,
602 7(12):e48121, December 2012.
- 603 [20] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer Series in Synergetics.
604 Springer-Verlag, Berlin Heidelberg, 1984.
- 605 [21] Stfan van der Walt, S. Chris Colbert, and Gal Varoquaux. The NumPy Array: A Structure for
606 Efficient Numerical Computation. *Computing in Science & Engineering*, 13(2):22–30, March
607 2011.
- 608 [22] Eric Jones, Travis Oliphant, and Pearu Peterson. SciPy: Open Source Scientific Tools for
609 Python. January 2001.
- 610 [23] Aric Hagberg, Daniel Schult, and Pieter Swart. Exploring Network Structure, Dynamics, and
611 Function using NetworkX. 2008.
- 612 [24] Anthony G. Hudetz, Xiping Liu, Siveshigan Pillay, Melanie Boly, and Giulio Tononi. Propofol
613 anesthesia reduces Lempel-Ziv complexity of spontaneous brain activity in rats. *Neuroscience
Letters*, 628:132–135, August 2016.
- 615 [25] Michael M. Schartner, Robin L. Carhart-Harris, Adam B. Barrett, Anil K. Seth, and Suresh D.
616 Muthukumaraswamy. Increased spontaneous MEG signal diversity for psychoactive doses of
617 ketamine, LSD and psilocybin. *Scientific Reports*, 7:46421, April 2017.
- 618 [26] Reinhard Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer-Verlag, Berlin
619 Heidelberg, 5 edition, 2017.

- 620 [27] Beverley A Orser. Depth-of-anesthesia monitor and the frequency of intraoperative awareness,
621 2008.
- 622 [28] Altered local coherence in the default mode network due to sevoflurane anesthesia. *Brain*
623 *research*, 1318:110–21, mar 2010.
- 624 [29] Pierre Boveroux, Audrey Vanhaudenhuyse, and Christophe Phillips. Breakdown of within- and
625 between-network Resting State during Propofol-induced Loss of Consciousness. *Anesthesiology*,
626 113(5):1038–1053, 2010.
- 627 [30] Carol Di Perri, Mohamed Ali Bahri, Enrico Amico, Aurore Thibaut, Lizette Heine, Georgios
628 Antonopoulos, Vanessa Charland-Verville, Sarah Wannez, Francisco Gomez, Roland Hustinx,
629 Luaba Tshibanda, Athena Demertzi, Andrea Soddu, and Steven Laureys. Neural correlates of
630 consciousness in patients who have emerged from a minimally conscious state: a cross-sectional
631 multimodal imaging study. *The Lancet. Neurology*, 15(8):830–842, 2016.
- 632 [31] Max Tegmark. Improved Measures of Integrated Information. *PLoS Computational Biology*,
633 12(11), 2016.

	LZ_C	SampEnt	PCA	Hurst	Higuchi	AlgConn	LZ_{Graph}	Serum Protein
Awake	0.967 ± 0.013	0.662 ± 0.013	28.5 ± 7.771	0.764 ± 0.009	0.867 ± 0.015	470.433 ± 73.893	375316.786 ± 4256.981	N/A
Mild	0.962 ± 0.012	0.654 ± 0.01	27.429 ± 7.5	0.769 ± 0.008	0.864 ± 0.013	436.068 ± 70.707	375353.214 ± 3931.755	286.025 ± 4.0
Moderate	0.939 ± 0.028	0.626 ± 0.035	26.286 ± 7.314	0.778 ± 0.013	0.842 ± 0.025	362.762 ± 85.815	372431.071 ± 4277.4	626.126 ± 1.0

Table 1: The values for all the complexity measures, temporal and spatial, for Dataset A.

	LZ _C	SampEnt	PCA	Hurst	Higuchi	AlgConn	LZ _{Graph}
Awake	0.967 ± 0.018	0.659 ± 0.017	45.812 ± 2.744	0.738 ± 0.008	0.981 ± 0.011	4693.704 ± 826.809	372236.562 ± 3794.728
Deep	0.938 ± 0.05	0.636 ± 0.038	40.625 ± 4.702	0.749 ± 0.016	0.963 ± 0.03	3787.616 ± 1338.374	367313.125 ± 8565.381

Table 2: The values for all the complexity measures, temporal and spatial, for Dataset B.

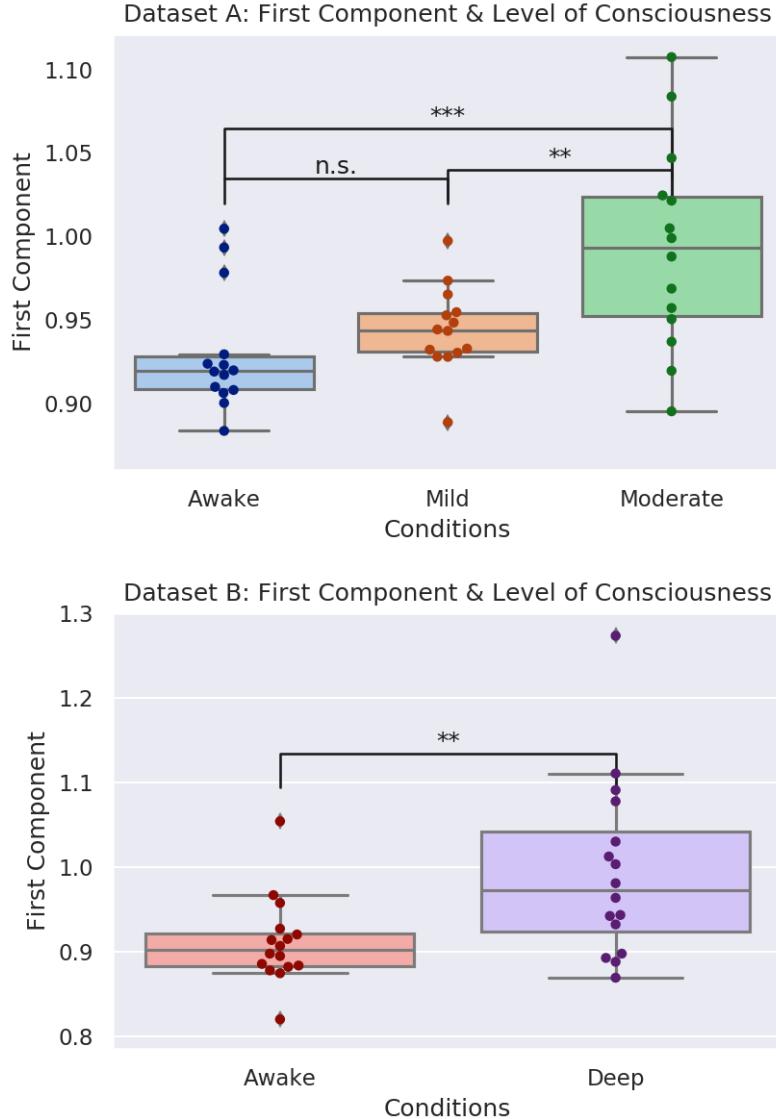


Figure 1: Here are the differences in the first principal component generated from all the measures from Datasets A and B. Interestingly, in Dataset A, there was no significant difference between the Awake and Mild condition, while there were differences between both of those states and the Moderate condition. While this may be a reflection of lack of sensitivity, it is worth noting that, between the Awake and Mild conditions, consciousness was not actually lost: volunteers experienced conscious sedation, while the difference in level of consciousness between the Awake and Moderate conditions was much more dramatic. In Dataset B, where consciousness was fully lost in the Deep condition, a significant difference appeared. Note that, despite the measures of complexity generally dropping as consciousness was lost (with the notable exception of the Hurst exponent analysis), the PCA analysis returned a Hurst-like pattern, with the values in the component increasing as consciousness is lost. This does not indicate an increase in complexity in any sense, but rather, is an artefact of how the dimensionality reduction transforms values. To ensure that this was not being driven by the Hurst exponent in any way, we ran the analysis after multiplying each Hurst exponent by -1 (so that the value decreased with loss of consciousness), and found no difference in the result.²⁹

Correlation Matrices

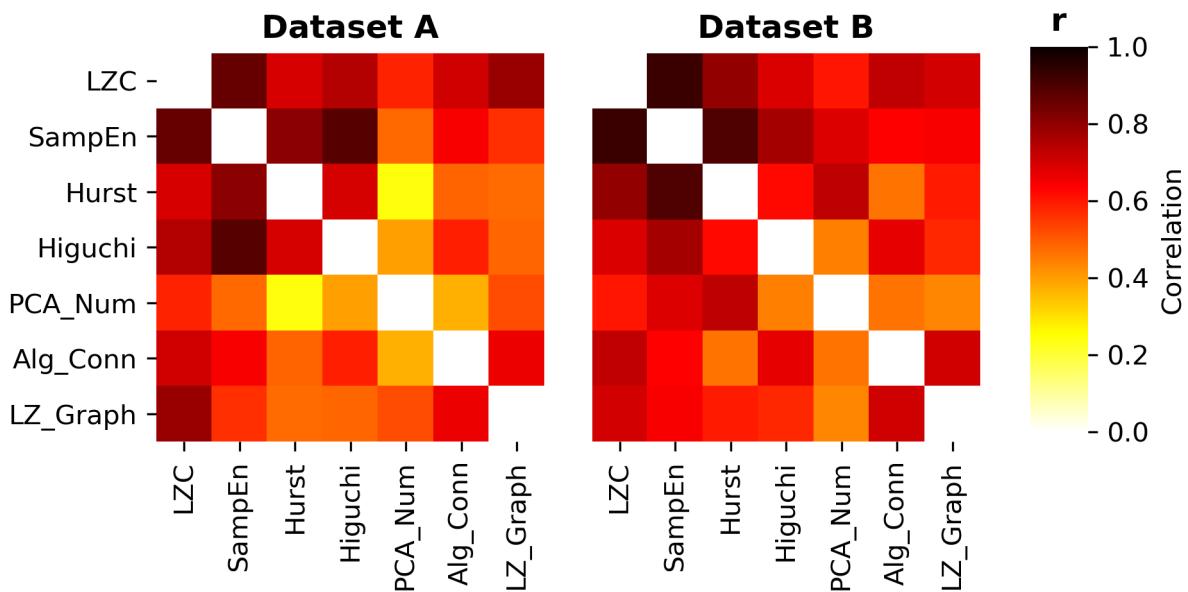


Figure 2: The correlation matrices between all the different metrics from Datasets A and B. All entries along the diagonal have been removed. There are some typical patterns: the graph measures (LZ_Graph and Algebraic Connectivity) are both generally more highly correlated, as are LZC, Sampen and Hurst). With the exception of a single correlation between the PCA Number and the Hurst Exponent in Dataset A. The p-values ranged over many orders of magnitude from 10^{-2} to 10^{-20}

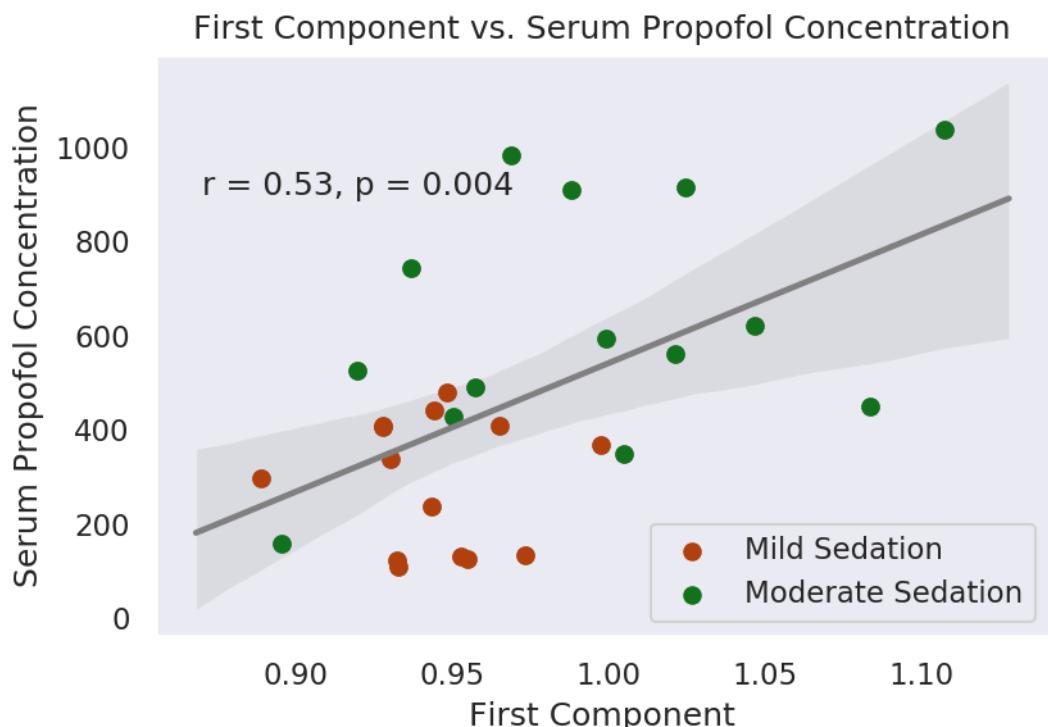


Figure 3: There was a significant correlation between the first component and serum concentration of propofol, with patients in the Mild condition ($r = 0.53$, p -value = 0.004) clustering together with low concentrations, and increasing, with larger variances, as the propofol concentration climbs. As with the plots shown above, the incongruous increase in the values of the component does not reflect a relative increase in complexity in this case, but is an artefact of the PCA algorithm. No Awake volunteers were included in this analysis, as all would have had a blood propofol concentration of exactly zero.