

Article

The Structure of Chaos: An Empirical Comparison of Fractal Physiology Complexity Indices Using NeuroKit2

Dominique Makowski ^{1,*} , An Shu Te ¹ , Tam Pham ¹ , Zen Juen Lau ¹  and S. H. Annabel Chen ^{1,2,3,4,*} 

¹ School of Social Sciences, Nanyang Technological University, Singapore 639818, Singapore; teanshu97@gmail.com (A.S.T.); phamtam17@gmail.com (T.P.); lauzenjuen@gmail.com (Z.J.L.)

² LKC Medicine, Nanyang Technological University, Singapore 639818, Singapore

³ National Institute of Education, Nanyang Technological University, Singapore 637616, Singapore

⁴ Centre for Research and Development in Learning, Nanyang Technological University, Singapore 639818, Singapore

* Correspondence: dom.makowski@gmail.com (D.M.); annabelchen@ntu.edu.sg (S.H.A.C.)

Abstract: Complexity quantification, through entropy, information theory and fractal dimension indices, is gaining a renewed traction in psychophysiology, as new measures with promising qualities emerge from the computational and mathematical advances. Unfortunately, few studies compare the relationship and objective performance of the plethora of existing metrics, in turn hindering reproducibility, replicability, consistency, and clarity in the field. Using the NeuroKit2 Python software, we computed a list of 112 (predominantly used) complexity indices on signals varying in their characteristics (noise, length and frequency spectrum). We then systematically compared the indices by their computational weight, their representativeness of a multidimensional space of latent dimensions, and empirical proximity with other indices. Based on these considerations, we propose that a selection of 12 indices, together representing 85.97% of the total variance of all indices, might offer a parsimonious and complimentary choice in regards to the quantification of the complexity of time series. Our selection includes *CWPE_n*, *Line Length (LL)*, *BubbEn*, *MSWPE_n*, *MFDEFA (Max)*, *Hjorth Complexity*, *SVDE_n*, *MFDEFA (Width)*, *MFDEFA (Mean)*, *MFDEFA (Peak)*, *MFDEFA (Fluctuation)*, *AttEn*. Elements of consideration for alternative subsets are discussed, and data, analysis scripts and code for the figures are open-source.

Keywords: chaos; complexity; fractal; physiology; noise



Citation: Makowski, D.; Te, A.S.; Pham, T.; Lau, Z.J.; Chen, S.H.A. The Structure of Chaos: An Empirical Comparison of Fractal Physiology Complexity Indices Using NeuroKit2. *Entropy* **2022**, *24*, 1036. <https://doi.org/10.3390/e24081036>

Academic Editor: Paolo Castiglioni

Received: 1 June 2022

Accepted: 22 July 2022

Published: 27 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Complexity is an umbrella term for concepts derived from information theory, chaos theory, and fractal mathematics, used to quantify unpredictability, entropy, and/or randomness. Using these approaches to characterize physiological signals (a subfield commonly referred to as “fractal physiology” [1]) has shown promising results in the assessment and diagnostic of the state and health of living systems [2–4].

Over the past few decades, there has been an important increase in the number of complexity indices [5]. Although these new procedures are usually mathematically well-defined and theoretically promising, limited empirical evidence is available to understand their similarities and differences [2,5]. Moreover, some of these methods are resource-intensive and require long computation times. This complicates their application with techniques that utilise high sampling-rates (e.g., M/EEG) and makes them impractical to implement in real-time settings—such as brain-computer interfaces [6,7]. As such, having empirical data about the computation time of various complexity indices would prove useful, for instance to guide their selection, especially in contexts where time or computational resources are constrained.

Additionally, the lack of a comprehensive open-source and user-friendly software for computing various complexity indices likely contributes to the scarcity of empirical

comparisons [8]. Indeed, many complexity indices are only described mathematically in journal articles, with reusable code seldom made available, therefore limiting their further application and validation [5,8]. Even when available and open-source, the code implementations of complexity measures are typically found scattered across different packages or scripts, or embedded within a larger goal-directed framework (e.g., *HCTSA*, a time-series comparison tool [9]). To address this lack of unified accessibility, we added a comprehensive set of complexity-related features to *NeuroKit2*, a Python package for physiological signal processing [10]. In doing so, we hope to provide users with an easy-to-use software capable of computing a wide range of complexity measures. The code is designed to run quickly and efficiently, while still being written in pure Python (with the help of standard dependencies such as *NumPy* or *Pandas* [11,12]) to maximize its reusability, transparency, and correctness.

Leveraging this tool, the goal of this study is to empirically compare a large number of complexity indices, inspect how they relate to one another, and derive recommendations for indices selection. More specifically, we will quantify the complexity of various types of signals with varying degrees of noise using 112 of the predominantly used indices that are available for computation with *NeuroKit2*. We will then project the results on a latent space through factor analysis, and review the indices that we find the most relevant and interesting in regards to their representation of the latent dimensions. This analysis will be complemented by hierarchical clustering. It should be noted that, even though this is one of the largest empirical comparison of complexity measures to date to our knowledge, the list of indices used is by no means exhaustive, with new indices constantly being developed, such as for instance *symmetry* [13].

2. Methods

The Python script to generate the data can be found at <https://github.com/DominiqueMakowski/ComplexityStructure> (accessed on 27 July 2022).

We started by generating 6 types of signals, one random-walk, two oscillatory signals (with one made of harmonic frequencies that results in a self-repeating - fractal-like - signal), two complex signals derived from Lorenz systems (with parameters $\sigma = 10, \beta = 2.5, \rho = 28$; and $\sigma = 20, \beta = 2, \rho = 30$, respectively) and one EEG-like simulated signal. Each of these signals was iteratively generated at 6 different lengths (ranging from 500 to 3000 by 500 samples). The resulting vectors were standardized and each were added 5 types of $(1/f)^\beta$ noise (namely violet $\beta = -2$, blue $\beta = -1$, white $\beta = 0$, pink $\beta = 1$, and brown $\beta = 2$ noise). Each noise type was added at 128 different intensities (linearly ranging from 0.001 to 3). Examples of generated signals are presented in Figure 1.

The combination of these parameters resulted in a total of 23040 signal iterations. For each of them, we computed 112 complexity indices, as well as additional basic metrics such as the *length* of the signal and its dominant *frequency*. We also included a *random* number at each iteration to ensure that our dimensionality analyses accurately discriminate this unrelated feature (as a manipulation check). The parameters used (such as the time-delay τ or the embedding dimension) are documented in the data generation script. For a complete description of the various indices included, please refer to *NeuroKit*'s documentation at <https://neuropsychology.github.io/NeuroKit> (accessed on 27 July 2022), in addition to the data generation script.

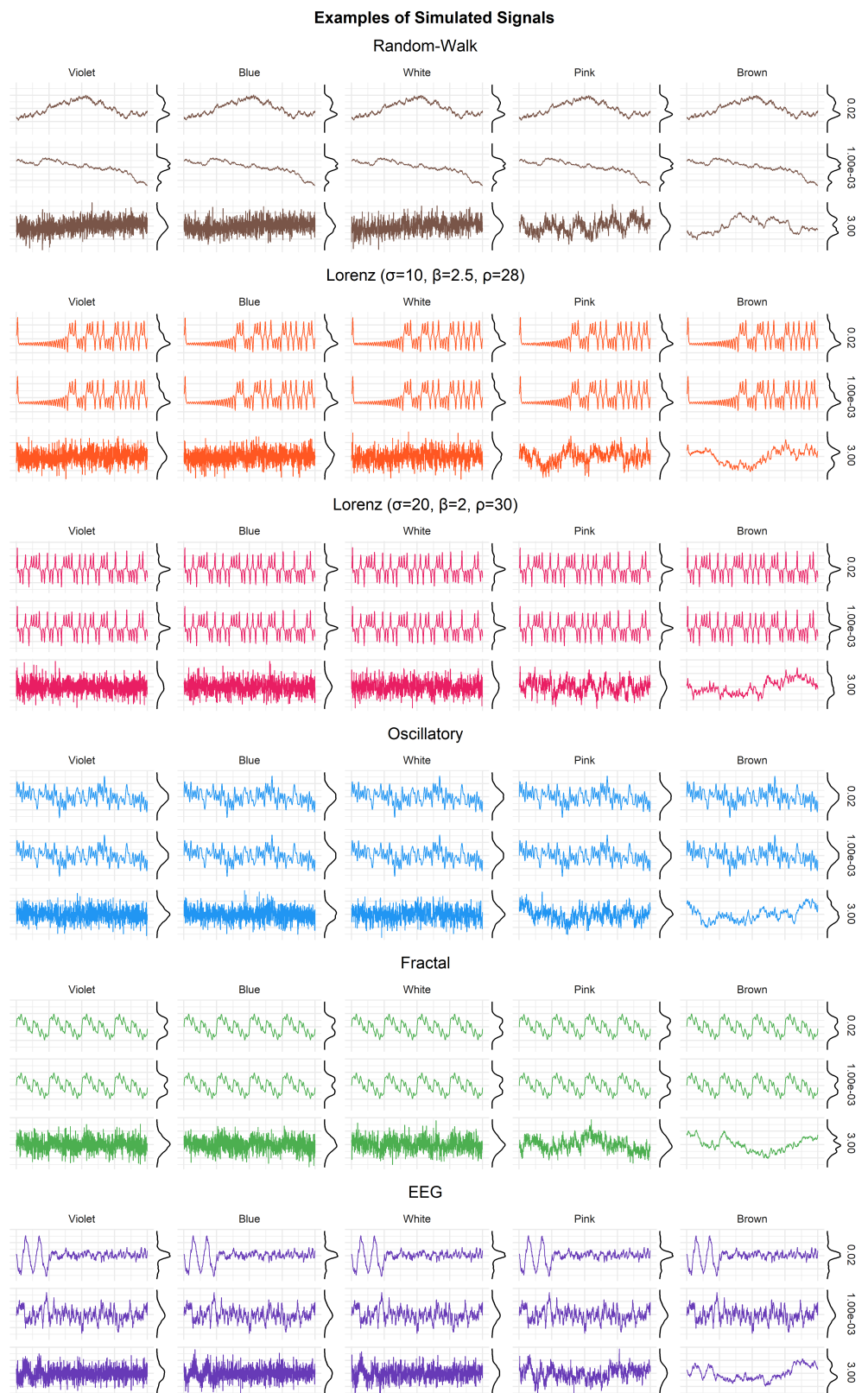


Figure 1. Different types of simulated signals, to which was added 5 types of noise (violet, blue, white, pink, and brown) with different intensities. For each signal type, the top row shows the signal with a minimal amount of noise, and the bottom row with a maximal amount of noise.

3. Results

The data analysis script and the data are fully available at <https://github.com/DominiqueMakowski/ComplexityStructure> (accessed on 27 July 2022). The analysis was performed in R using the *easystats* collection of packages [14–17]. As the results are primarily presented graphically via the figures, the code to fully reproduce them is also included in the analysis script.

3.1. Computation Time

Firstly, one should note that the computation times presented in Figure 2 are relative (in arbitrary units) and do not correspond to real times, as these would highly depend on the system specifications. Rather, the goal here was to convey some intuition on the differences between different classes of indices (using the same machine and the same language of implementation, i.e., Python). While it is possible that computational advances or improvements in the code efficiency might change some of these values, we believe that the “big picture” should remain fairly stable, as it is to a large extent driven by the inherent nature of the algorithms under consideration.

Despite the relative shortness of the signals considered (a few thousand points at most), the fully-parallelized data generation script took around 24 h to run on a 48-cores machine. After summarizing and sorting the indices by computation time, the most striking result is the order of magnitude of difference between the fastest and slowest indices. Additionally, some indices are particularly sensitive to the signal length, a property which, in combination with their computational cost, led to indices being 100,000 times slower to compute than others.

In particular, multiscale indices were among the slowest to compute due to their iterative nature (a given index being computed multiple times on coarse-grained sub-series of the signal). Indices related to Recurrence Quantification Analysis (RQA) were also relatively slow and did not scale well with signal length.

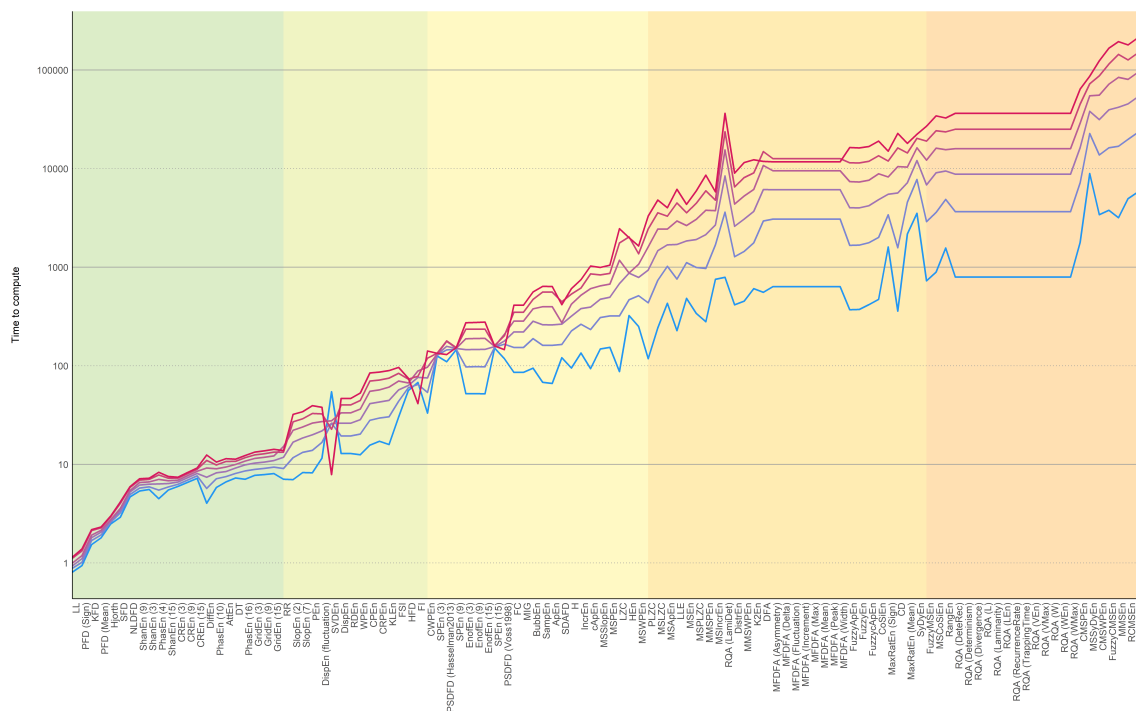


Figure 2. Median computation time difference between the different complexity indices algorithms, as well as variability as a function of signal lengths (represented by different line colors). The indices are grouped in sections (background color) according to their median computation time. Note that the time is expressed in arbitrary units as it is intended to convey differences, since the actual time would depend on the system specifications.

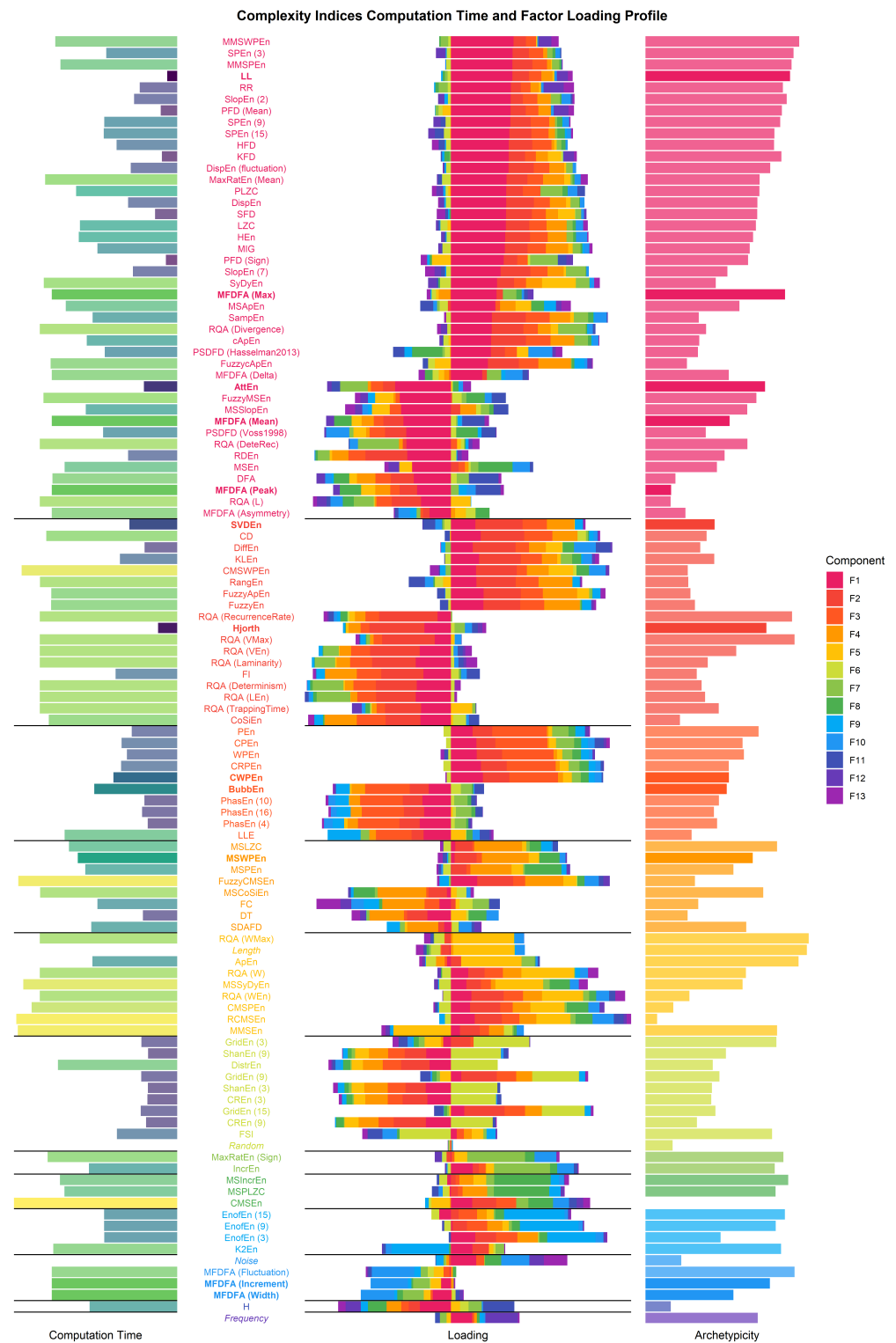


Figure 4. Factor loadings of the complexity indices, colored by the factor they represent the most (center). On the left, the median computation times and on the right, the archetypicity—the inverse of factor profile complexity (i.e., the extent to which each index is a pure representative of its dominant factor, which is low for indices that equally load on different factors).

The first extracted factor is the closest to the largest amount of indices, and is positively loaded by indices that are sensitive to the deviation of consecutive differences (e.g., *LL*, *PFDA (Mean)*) as well as indices that capture the amplitude of fluctuations (*DispEn (fluctuation)*, *MFDEFA (Max)*). In line with this, this factor was negatively loaded by indices related to

- *MSWPEn*: The Multiscale Weighted Permutation Entropy is the entropy of weighted ordinal descriptors of the time-embedded signal computed at different scales obtained by a coarse-graining procedure [22].
- *MFDEFA (Max)*: The value of singularity spectrum D corresponding to the maximum value of singularity exponent H .
- *Hjorth*: Hjorth's Complexity is defined as the ratio of the mean frequency of the first derivative of the signal to the mean frequency of the signal [23].
- *SVDEn*: The Singular Value Decomposition (SVD) Entropy quantifies the amount of eigenvectors needed for an adequate representation of the system [24].
- *MFDEFA (Width)*: The width of the multifractal singularity spectrum [25] obtained via Detrended Fluctuation Analysis (DFA).
- *MFDEFA (Mean)*: The mean of the maximum and minimum values of singularity exponent H , which quantifies the average fluctuations of the signal.
- *MFDEFA (Peak)*: The value of the singularity exponent H corresponding to peak of singularity dimension D . It is a measure of the self-affinity of the signal, and a high value is an indicator of high degree of correlation between the data points.
- *MFDEFA (Increment)*: The cumulative function of the squared increments of the generalized Hurst's exponents between consecutive moment orders [26].
- *AttEn*: The Attention Entropy is based on the frequency distribution of the intervals between the local maxima and minima of the time series [27].

Finally, we visualized the expected value of our selection of indices for different types of signals under different conditions of noise (see Figure 8). This confirmed that *LL* was primarily driven by the noise intensity (which is expected, as they capture the variability of successive differences). The other indices appear capable of discriminating between the various types of signals (when the signal is not dominated by noise).

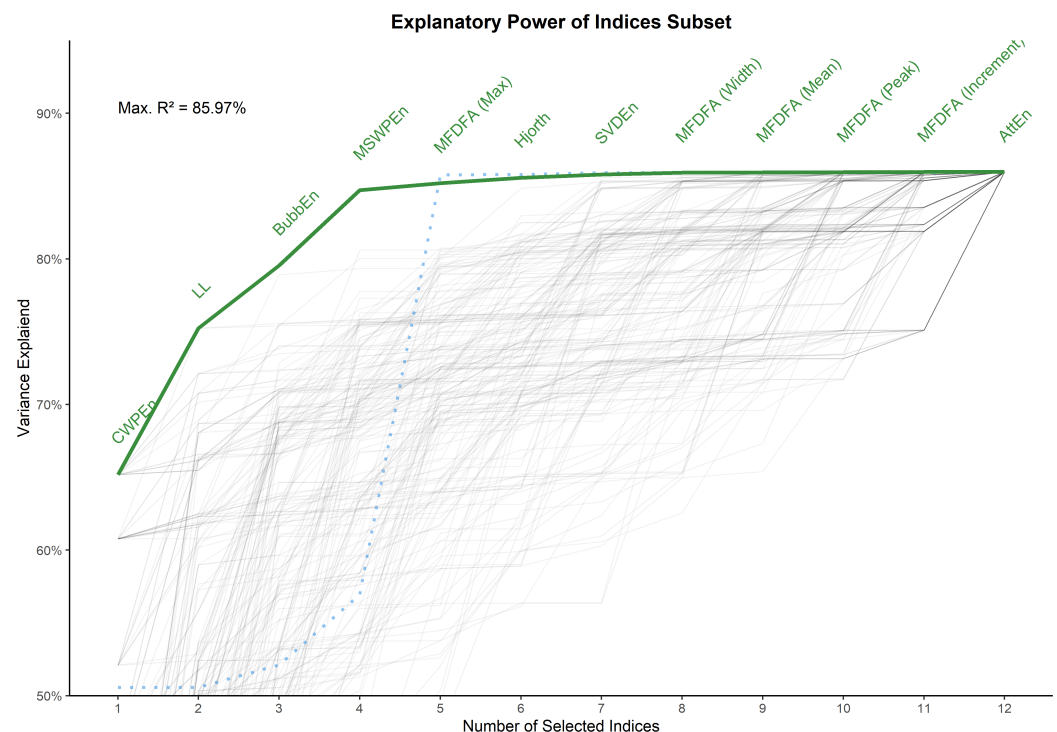


Figure 7. Variance of the whole dataset of indices explained by the subselection. Each line represents a random number of selected variables. The green line represents the optimal order (i.e., the relative importance) that maximizes the variance explained. The dotted blue line represents the cumulative relative median computation time of the selected indices, and shows that MFDEFA and multiscale indices are the most resource-costly algorithms.

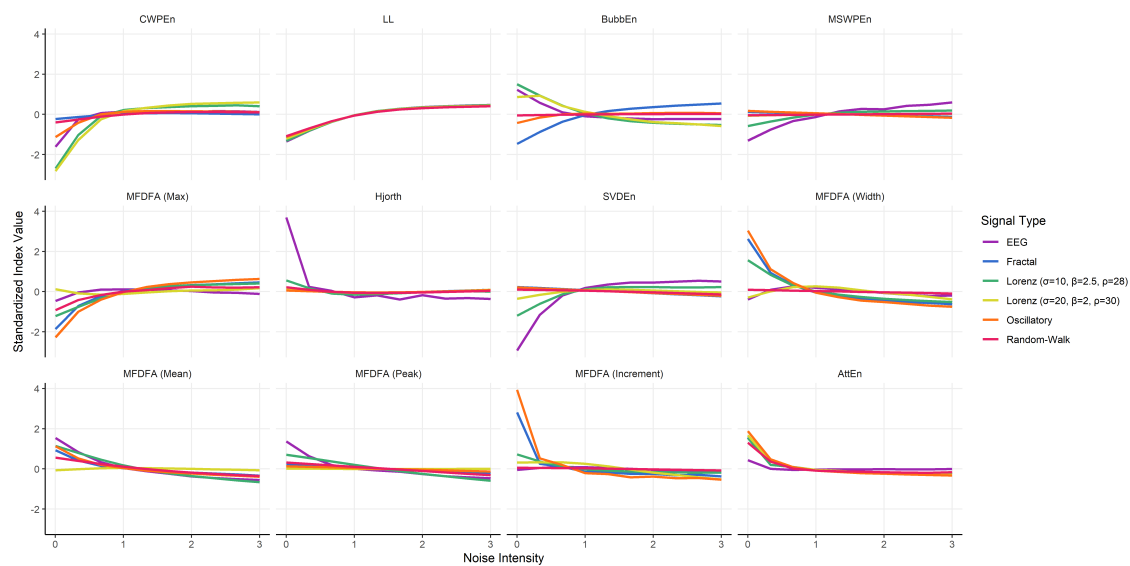


Figure 8. Visualization of the expected value of a selection of indices depending on the signal type and of the amount of noise.

4. Discussion

As the span and application of complexity science grows, a systematic approach to compare their “performance” becomes necessary to reinforce the clarity and structure of the field. The term *performance* used here is to be understood in a relative sense, as any such endeavor faces the “hard problem” of complexity science: various objective properties of signals (e.g., short-term vs. long-term variability, auto-correlation, information, randomness [28,29]) participate in forming overarching concepts such as “complex” and “chaotic”. Indices that are sensitive to some of these objective properties are thus conceptually linked through such overarching frameworks. However, it remains unclear how these high-level concepts transfer back, in a top-down fashion, into a combination of lower-level features. As such, it is conceptually complicated to benchmark complexity measures against “objectively” complex vs. non-complex signals. In other words, while we are aware that different objective signal characteristics can contribute to the “complexity” of a signal, there is not a one-to-one correspondence between the latter and the former.

To circumvent the aforementioned consideration, we adopted a paradigm where we generated different types of signals to which we systematically added distinct types—and amount—of perturbations. It should be noted that we did not seek to measure how complexity indices can discriminate between these signal types, nor did we attempt at mimicking real-life signals or scenarios. The goal was instead to generate enough variability to reliably map the relationships between the indices.

Our results empirically confirm the plurality of underlying components of complexity (although it is here defined somewhat circularly as what is measured by complexity indices), and more importantly show that complexity indices vary in their sensitivity to various orthogonal latent dimensions. However, the mostly descriptive interpretation of these dimensions is a limitation of the present investigation, and future studies are needed to investigate and discuss them in greater depth (for instance, by modulating specific properties of signals and measuring their impact on these latent dimensions).

Given the increasing role of complexity science as a field and the sheer number of complexity indices already published, our study aimed to empirically map the relationship between various indices and provide useful information to guide future researchers in their selection of complexity metrics. An example of indices subselection that encapsulates information about different underlying dimensions at a relatively low computational cost include *CWPEn*, *Line Length (LL)*, *BubbEn*, *MSWPEn*, *MFDEFA (Max)*, *Hjorth Complexity*, *SVDEn*, *MFDEFA (Width)*, *MFDEFA (Mean)*, *MFDEFA (Peak)*, *MFDEFA (Fluctuation)*, *AttEn*. These

indices might be complimentary in offering a parsimonious, yet comprehensive profile of the complexity of a time series. Moving forward, future studies are needed to validate, analyze and interpret the nature of the dominant sensitivities of various indices groups identified in the present work. In doing so, complexity findings in prospective studies can be more easily interpreted and integrated into new research and novel theories.

Author Contributions: The authors made the following contributions. D.M.: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft; A.S.T.: Software, Project administration, Writing—review and editing; T.P.: Software, Writing—review and editing; Z.J.L.: Software, Writing—review and editing; S.H.A.C.: Supervision, Project administration, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The study was funded partly by the Presidential Post-Doctoral Award to D.M. and Ministry of Education Academic Research Fund Tier 2 Grant (Project No.: MOE2019-T2-1-019) to S.H.A.C. The authors declare no conflict of interest, and the funding sponsors had no role in the design, execution, interpretation or writing of the study.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: <https://github.com/DominiqueMakowski/ComplexityStructure> (accessed on 27 July 2022).

Acknowledgments: We would like to thank the contributors of NeuroKit2, as well as the people that developed and shared open-source code which helped implementing the complexity algorithms in NeuroKit2. In particular, the contributors and maintainers of packages such as *nolds* [30], *AntroPy* [31], *pyEntropy*, and *EntropyHub* [8].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bassingthwaite, J.B.; Liebovitch, L.S.; West, B.J. *Fractal Physiology*; Springer: Berlin, Germany, 2013.
2. Lau, Z.J.; Pham, T.; Annabel, S.; Makowski, D. Brain Entropy, Fractal Dimensions and Predictability: A Review of Complexity Measures for EEG in Healthy and Neuropsychiatric Populations. 2021. Available online: <https://doi.org/10.31234/osf.io/f8k3x> (accessed on 27 July 2022)
3. Ehlers, C.L. Chaos and Complexity: Can It Help Us to Understand Mood and Behavior? *Arch. Gen. Psychiatry* **1995**, *52*, 960–964.
4. Goetz, S.J. Spatial Dynamics, Networks and Modelling? Edited by Aura Reggiani and Peter Nijkamp. *Pap. Reg. Sci.* **2007**, *86*, 523–524. <https://doi.org/10.1111/j.1435-5957.2007.00128.x>.
5. Yang, A.C.; Tsai, S.-J. Is Mental Illness Complex? From Behavior to Brain. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2013**, *45*, 253–257. <https://doi.org/10.1016/j.pnpbp.2012.09.015>.
6. Azami, H.; Rostaghi, M.; Abásolo, D.; Escudero, J. Refined Composite Multiscale Dispersion Entropy and Its Application to Biomedical Signals. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2872–2879.
7. Manis, G.; Aktaruzzaman, M.; Sassi, R. Low Computational Cost for Sample Entropy. *Entropy* **2018**, *20*, 61. <https://doi.org/10.3390/e20010061>.
8. Flood, M.W.; Grimm, B. EntropyHub: An Open-Source Toolkit for Entropic Time Series Analysis. *PLoS ONE* **2021**, *16*, e0259448. <https://doi.org/10.1371/journal.pone.0259448>.
9. Fulcher, B.D.; Jones, N.S. Hctsa: A Computational Framework for Automated Time-Series Phenotyping Using Massive Feature Extraction. *Cell Syst.* **2017**, *5*, 527–531. <https://doi.org/10.1016/j.cels.2017.10.001>.
10. Makowski, D.; Pham, T.; Lau, Z.J.; Brammer, J.C.; Lespinasse, F.; Pham, H.; Schölzel, C.; Chen, S.H.A. NeuroKit2: A Python Toolbox for Neurophysiological Signal Processing. *Behav. Res. Methods* **2021**, *53*, 1689–1696. <https://doi.org/10.3758/s13428-020-01516-y>.
11. Harris, C.R.; Millman, K.J.; Walt, S.J. van der; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array Programming with NumPy. *Nature* **2020**, *585*, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>.
12. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; Volume 445, pp. 51–56.
13. Girault, J.-M.; Menigot, S. Palindromic Vectors, Symmetry and Symmetry as Symmetry Descriptors of Binary Data. *Entropy* **2022**, *24*, 82. <https://doi.org/10.3390/e24010082>.
14. Makowski, D.; Ben-Shachar, M.; Patil, I.; Lüdecke, D. Methods and Algorithms for Correlation Analysis in R. *J. Open Source Softw.* **2020**, *5*, 2306. <https://doi.org/10.21105/joss.02306>.

15. Lüdecke, D.; Patil, I.; Ben-Shachar, M.S.; Wiernik, B.M.; Waggoner, P.; Makowski, D. see: An R Package for Visualizing Statistical Models. *J. Open Source Softw.* **2021**, *6*, 3393. <https://doi.org/10.21105/joss.03393>.
16. Lüdecke, D.; Ben-Shachar, M.; Patil, I.; Makowski, D. Extracting, Computing and Exploring the Parameters of Statistical Models Using R. *J. Open Source Softw.* **2020**, *5*, 2445. <https://doi.org/10.21105/joss.02445>.
17. Makowski, D.; Lüdecke, D.; Ben-Shachar, M.S.; Patil, I. modelbased: Estimation of Model-Based Predictions, Contrasts and Means. 2022. Available online: <https://CRAN.R-project.org/package=modelbased> (accessed on 27 July 2022)
18. Unakafov, A.M.; Keller, K. Conditional Entropy of Ordinal Patterns. *Phys. D Nonlinear Phenom.* **2014**, *269*, 94–102.
19. Esteller, R.; Echazuz, J.; Tchong, T.; Litt, B.; Pless, B. Line Length: An Efficient Feature for Seizure Onset Detection. In Proceedings of the 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Istanbul, Turkey, 25–28 October 2001; IEEE: Piscataway, NJ, USA, 2001; Vol. 2, pp. 1707–1710.
20. Kalauzi, A.; Bojić, T.; Rakić, L. Extracting Complexity Waveforms from One-Dimensional Signals. *Nonlinear Biomed. Phys.* **2009**, *3*, 8.
21. Manis, G.; Aktaruzzaman, M.; Sassi, R. Bubble Entropy: An Entropy Almost Free of Parameters. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2711–2718. <https://doi.org/10.1109/tbme.2017.2664105>.
22. Fadlallah, B.; Chen, B.; Keil, A.; Principe, J. Weighted-Permutation Entropy: A Complexity Measure for Time Series Incorporating Amplitude Information. *Phys. Rev. E* **2013**, *87*, 022911.
23. Hjorth, B. EEG Analysis Based on Time Domain Properties. *Electroencephalogr. Clin. Neurophysiol.* **1970**, *29*, 306–310.
24. Roberts, S.J.; Penny, W.; Rezek, I. Temporal and Spatial Complexity Measures for Electroencephalogram Based Brain-Computer Interfacing. *Med. Biol. Eng. Comput.* **1999**, *37*, 93–98.
25. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Havlin, S.; Bunde, A.; Stanley, H.E. Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series. *Phys. A Stat. Mech. Its Appl.* **2002**, *316*, 87–114.
26. Faini, A.; Parati, G.; Castiglioni, P. Multiscale Assessment of the Degree of Multifractality for Physiological Time Series. *Philos. Trans. R. Soc. A* **2021**, *379*, 20200254.
27. Yang, J.; Choudhary, G.I.; Rahardja, S.; Franti, P. Classification of Interbeat Interval Time-Series Using Attention Entropy. *IEEE Trans. Affect. Comput.* **2020**. doi:10.1109/TAFFC.2020.3031004.
28. Namdari, A.; Li, Z. A Review of Entropy Measures for Uncertainty Quantification of Stochastic Processes. *Adv. Mech. Eng.* **2019**, *11*, 168781401985735. <https://doi.org/10.1177/1687814019857350>.
29. Xiong, W.; Faes, L.; Ivanov, P.Ch. Entropy Measures, Entropy Estimators, and Their Performance in Quantifying Complex Dynamics: Effects of Artifacts, Nonstationarity, and Long-Range Correlations. *Phys. Rev. E* **2017**, *95*, 062114. <https://doi.org/10.1103/physreve.95.062114>.
30. Schölzel, C. *Nonlinear Measures for Dynamical Systems*; Zenodo: Geneva, Switzerland, 2019. <https://doi.org/10.5281/zenodo.3814723>.
31. Vallat, R. *Antropy: Entropy and Complexity of (EEG) Time-Series in Python*; GitHub: San Francisco, CA, USA, 2022. Available online: <https://github.com/raphaelvallat/antropy> (accessed on 27 July 2022)