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 J. Lau[1], & S.H. Annabel Chen[1]

[1] Nanyang Technological University

Author Note

                                        Abstract

Complexity quantification, such as entropy, information and fractal dimension indices, are gaining a renewed traction in psychopsyiology, 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 indices, in turn hindering reproducibility, replicability, and clarity in the field. In this study, we systematically compared 125 indices of complexity by their computational weight, their representativeness of a multidimensional space of latent dimensions, and empirical proximity with other indices. We propose that the selection of a few indices, such as ..., might offer a complimentary choice in regards to the quantification of the complexity of a time series.

    Keywords: chaos, complexity, fractal, physiology, noise

    Word count: X

28   **The Structure of Chaos: An Empirical Comparison of Fractal Physiology**

29   **Complexity Indices using NeuroKit2**

30   **Introduction**

31   Complexity is an umbrella term for concepts derived from information theory, chaos

32   theory, and fractal mathematics, used to quantify unpredictability, entropy, and/or

33   randomness. Using these tools to characterize signals (a subfield commonly referred to as

34   "fractal physiology," Bassingthwaighte et al., 2013) has shown promising results in

35   physiology in the assessment and diagnostic of the state and health of living systems Ehlers

36   (1995).

37   There has been a large and accelerating increase in the number of complexity indices

38   in the past few decades. These new procedures are usually mathematically well-defined and

39   theoretically promising. However, few empirical evidence exist to understand their

40   differences and similarities. Moreover, some can be very expensive in terms of computation

41   power and thus, time, which can become an issue in some applications such as high

42   sampling-rate techniques (e.g., M/EEG) or real-time settings (brain-computer interface).

43   As such, having a general view depicting the relationship between the indices with

44   information about their computation time would be useful, for instance to guide the indices

45   selection in settings where time or computational power is limited.

46   One of the contributing factor of this lack of empirical comparison is the lack of free,

47   open-source, unified, and easy to use software for computing various complexity indices.

48   Indeed, most of them are described mathematically in journal articles, and reusable code is

49   seldom made available, which limits their further application and validation. *NeuroKit2*

50   (Makowski et al., 2021) is a Python package for physiological signal processing that aims at

51   providing the most comprehensive, accurate and fast pure Python implementations of

52   complexity indices.

⁵³ Leveraging this tool, the goal of this study is to empirically compare a vast number of

⁵⁴ complexity indices, inspect how they relate to one another, and extract some

⁵⁵ recommendations for indices selection, based on their added-value and computational

⁵⁶ efficiency. Using *NeuroKit2*, we will compute more than a hundred complexity indices on

⁵⁷ various types of signals, with varying degrees of noise. We will then project the results on a

⁵⁸ latent space through factor analysis, and report the most interesting indices in regards to

⁵⁹ their representation of the latent dimensions.

⁶⁰ **Methods**

⁶¹ The script to generate the data can be found at

⁶² **github.com/neuropsychology/NeuroKit/studies/complexity_benchmark**

⁶³ We started by generating 5 types of signals, one random-walk, two oscillatory signals

⁶⁴ made (one made of harmonic frequencies that results in a self-repeating - fractal-like -

⁶⁵ signal), and two complex signals derived from Lorenz systems (with parameters

⁶⁶ ($\sigma = 10, \beta = 2.5, \rho = 28$); and ($\sigma = 20, \beta = 2, \rho = 30$), respectively). Each of this signal was

⁶⁷ iteratively generated at . . . different lengths (). The resulting vectors were standardized

⁶⁸ and each were added 5 types of $(1/f)^\beta$ noise (namely violet $\beta = -2$, blue $\beta = -1$, white

⁶⁹ $\beta = 0$, pink $\beta = 1$, and brown $\beta = 2$ noise). Each noise type was added at 48 different

⁷⁰ intensities (linearly ranging from 0.1 to 4). Examples of generated signals are presented in

⁷¹ **Figure 1**.

⁷² The combination of these parameters resulted in a total of 6000 signal iterations. For

⁷³ each of them, we computed 128 complexity indices, and additionally basic metric such as

⁷⁴ the standard deviation (*SD*), the *length* of the signal and its dominant *frequency*. The

⁷⁵ parameters used (such as the time-delay τ or the embedding dimension) are documented in

⁷⁶ the data generation script. For a complete description of the various indices included,

⁷⁷ please refer to NeuroKit's documentation (https://neuropsychology.github.io/NeuroKit).

Results

⁷⁸

⁷⁹   The data analysis script, the data and the code for the figures is fully available at

⁸⁰ **ADD LINK**. The analysis was performed in R using the *easystats* collection of packages

⁸¹ (Lüdecke et al., 2021; Lüdecke et al., 2020; Makowski et al., 2020/2022, 2020).

⁸² **Computation Time.**   Despite the relative shortness of the signals considered (a

⁸³ few thousand points at most), the fully-parallelized data generation script took 12h to run

⁸⁴ on a 48-cores machine. After summarizing and sorting the indices by computation time,

⁸⁵ the most striking feature are the orders of magnitude of difference between the fastest and

⁸⁶ slowest indices. Some of them are also particularly sensitive to the data length, a property

⁸⁷ which combined with computational expensiveness leads to indices being 100,000 slower to

⁸⁸ compute than other basic metrics.

⁸⁹   Multiscale indices are among the slowest, due to their iterative nature (a given index

⁹⁰ is computed multiple times on coarse-grained subseries of the signal). Indices related to

⁹¹ Recurrence Quantification Analysis (RQA) are also relatively slow and don't scale well

⁹² with signal length.

⁹³   For the subsequent analyses, we removed statistically redundant indices, such as

⁹⁴ *PowEn* (identical to *SD*), *CREn (100)* (identical to *CREn (10)*), and *FuzzyRCMSEn*

⁹⁵ (identical to *RCMSEn*).

⁹⁶ **Correlation.**   The Pearson correlation analysis revealed that complexity indices,

⁹⁷ despite their multitude, their unicities and specificities, do indeed share similarities. They

⁹⁸ form clusters, with two major ones easily appearing to the naked eye (the blue and the red

⁹⁹ groups). These two anti-correlated groups are driven by the fact that some indices, by

¹⁰⁰ design, index the "predictability", whereas other the "randomness", and thus are negatively

¹⁰¹ related to one-another (see **Figure 2**).

¹⁰² **Factor Analysis.**   The agreement procedure for the optimal number of factors

¹⁰³ suggested that the 125 indices can be mapped on a multidimensional space of 14

104 orthogonal latent factors, that we extracted using a *varimax* rotation. We then took

105 interest in the loading profile of each indices, and in particular the latent dimension that

106 maximally related to each index (see **Figure 3**).

107   The first factor is the closest to the largest amount of indices. Many indices with

108 positive and strong loadings are particularly sensitive to the deviation of consecutive

109 differences (e.g., *ShanEn - D*, *NLDFD*, *PFD - D*). It was negatively loaded by indices

110 related to Detrended Fluctuation Analysis (DFA), which tend to index the presence of

111 long-term correlations. This latent factor might encapsulate the predominance of

112 short-term vs. long-term unpredictability. Indices that are the most representative

113 (positively and negatively) and have a relatively low computational cost include *ShanEn -*

114 *D*, *NLDFD*, *PFD - D*, and *AttEn*, *PSDFD*, *FuzzyMSEn*. The second factor was loaded

115 maximally by signal *length* and *SD*, and thus might not capture features of complexity *per*

116 *se.* Indices the most related to it were indices known to be sensitive to signal length, such

117 as *ApEn.* The third factor included multiscale indices, such as *MSWPEn.* The fourth

118 factor included indices that quantified the diversity of the tendency of a signal to revisit a

119 past state (within a certain tolerance threshold). It was positively loaded by *ShanEn - r*

120 and negatively by *RQA - Reccurence Rate.* The fifth factor was loaded by permutation

121 entropy indices, such as *WPEn.* The sixth factor was driven by indices that were based on

122 converting the signal into a number of bins. The seventh factor was loaded positively by

123 the amount of noise, and negatively by multifractal indices such as *MFDFA - Increment*,

124 suggesting a sensitivity to regularity. The last notable result is that indices based on a

125 symbolization (discretization) of the time series do tend to create factors alongside the

126 symbolization method. Finally, as a manipulation check of our factorization method, the

127 random vector does indeed form its own factor, and doesn't load unto anything else.

128   **Correlation Network.** For illustration purposes, we also represented the

129 correlation matrix as a connectivity graph (see **Figure 4**).

130 **Hierarchical Clustering.** We ran hierarchical clustering (with a Ward D2

131 distance) to provide additional information or confirmation about the groups discussed

132 above. This allowed us to fine-grain our recommendations of complimentary complexity

133 indices (see **Figure 5**).

Discussion

135 As complexity science grows in size and application, a systematic approach to

136 compare their "performance" becomes necessary to increase the clarity and structure of the

137 field. The word *performance* is here to be understood in a relative sense, as any such

138 endeavor faces the "hard problem" of complexity science. The indices are sensitive to

139 specific objective properties of a signal that we consider part of over-arching concepts such

140 as "complex" and "chaotic". But it is unclear how these high-level concepts transfer back,

141 in a top-down fashion, into a combination of lower-level features, such as short-term

142 vs. long-term variability, auto-correlation, information, randomness, and so on. As such, it

143 is conceptually complicated to benchmark complexity measures against "objectively"

144 complex vs. non-complex signals. In other words, we know that different characteristics can

145 contribute to the "complexity" of a signal, but there is not a one-to-one correspondence

146 between the latter and the former.

147 Hence the choice of the current paradigm, in which we generated different types of

148 signals to which we systematically added different types and amount of perturbations.

149 However, we did not seek at measuring how complexity indices can discriminate between

150 these features or systems, nor did we attempt at mimicking real-life signals or scenarios.

151 The goal was instead to generate enough variability to reliably map the relationships

152 between the indices.

153 The plurality of underlying components of empirical complexity (what is measured by

154 complexity indices) seems to be somewhat confirmed by our results, showing that

155 complexity indices are more or less sensitive to various orthogonal latent dimensions. One

156 of the limitation of the current study has to do with the limited possibilities of

157 interpretation of these underlying dimensions, and future studies are needed to discuss

158 them.

159   Indices that were highlighted as encapsulating information about different underlying

160 dimensions at a relatively low computational cost include *ShanEn (D)*, *NLDFD*, *SVDEn*,

161 *AttEn*, *PSDFD*, *MFDFA (Mean)*, *FuzzyMSEn*, *MSWPEn*, *MFDFA (Increment)*, *ShanEn*

162 *(r)*, *Hjorth* and *WPEn*. These indices might be complimentary in offering a comprehensive

163 profile of the complexity of a time series. Future studies are needed to analyze the nature

164 of the dominant sensitivities of group of indices, so that results can be more easily

165 interpreted and integrated into new studies and novel theories.

₁₆₆   **References**

₁₆₇ Bassingthwaighte, J. B., Liebovitch, L. S., & West, B. J. (2013). *Fractal physiology.*

₁₆₈   Springer.

₁₆₉ Ehlers, C. L. (1995). Chaos and complexity: Can it help us to understand mood and

₁₇₀   behavior? *Archives of General Psychiatry, 52*(11), 960–964.

₁₇₁ Lau, Z. J., Pham, T., Annabel, S., & Makowski, D. (2021). *Brain entropy, fractal*

₁₇₂   *dimensions and predictability: A review of complexity measures for EEG in healthy and*

₁₇₃   *neuropsychiatric populations.*

₁₇₄ Lüdecke, D., Ben-Shachar, M., Patil, I., & Makowski, D. (2020). Extracting, computing

₁₇₅   and exploring the parameters of statistical models using R. *Journal of Open Source*

₁₇₆   *Software, 5*(53), 2445. https://doi.org/10.21105/joss.02445

₁₇₇ Lüdecke, D., Patil, I., Ben-Shachar, M. S., Wiernik, B. M., Waggoner, P., & Makowski, D.

₁₇₈   (2021). see: An R package for visualizing statistical models. *Journal of Open Source*

₁₇₉   *Software, 6*(64), 3393. https://doi.org/10.21105/joss.03393

₁₈₀ Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Methods and algorithms

₁₈₁   for correlation analysis in R. *Journal of Open Source Software, 5*(51), 2306.

₁₈₂   https://doi.org/10.21105/joss.02306

₁₈₃ Makowski, D., Lüdecke, D., Ben-Shachar, M. S., & Patil, I. (2022). *modelbased: Estimation*

₁₈₄   *of model-based predictions, contrasts and means* (Version 0.7.2.1) [Computer software].

₁₈₅   https://CRAN.R-project.org/package=modelbased (Original work published 2020)

₁₈₆ Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel,

₁₈₇   C., & Chen, S. H. A. (2021). NeuroKit2: A python toolbox for neurophysiological

₁₈₈   signal processing. *Behavior Research Methods, 53*(4), 1689–1696.

₁₈₉   https://doi.org/10.3758/s13428-020-01516-y
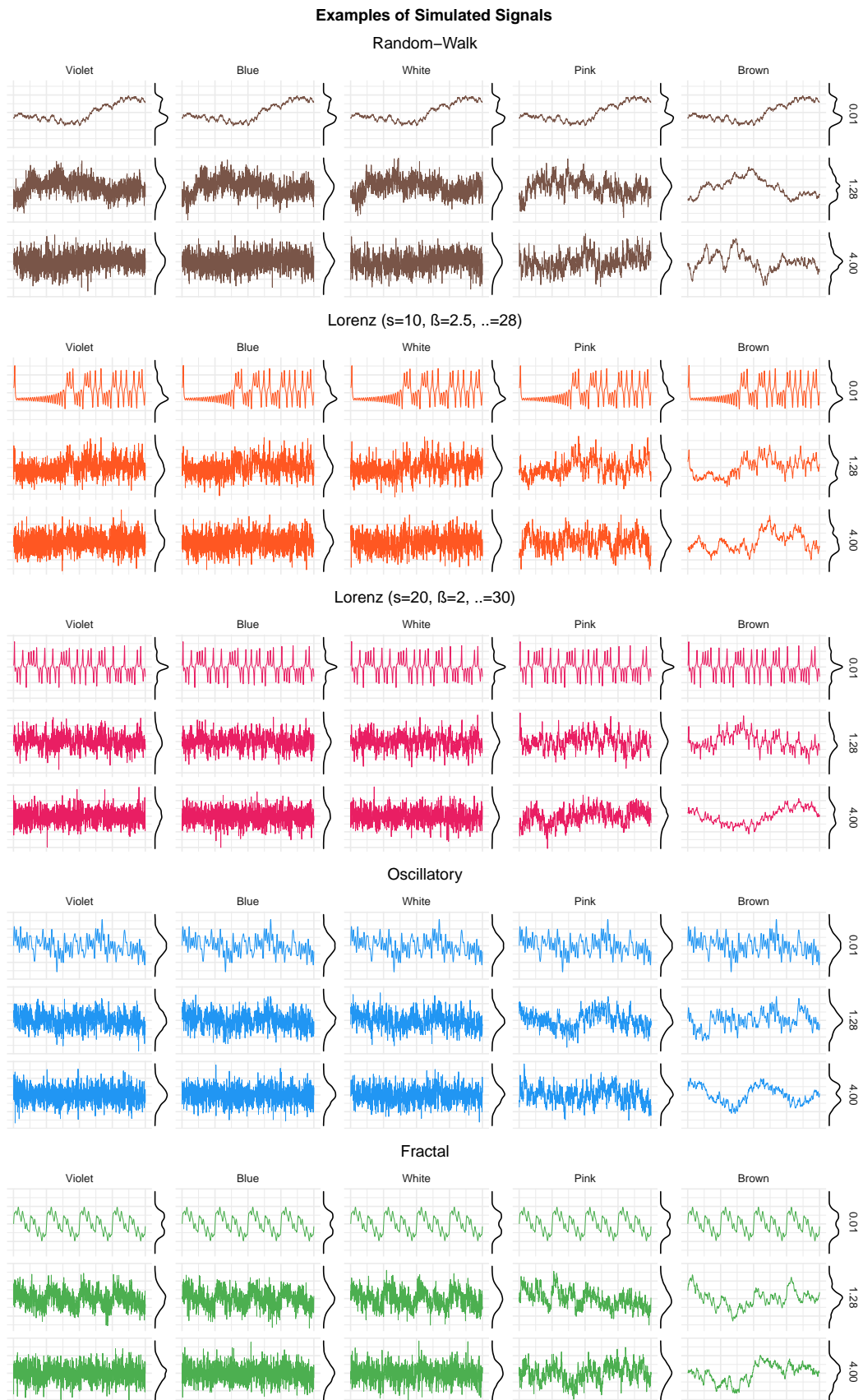
Figure 1. Different types of simulated signals, to which was added 5 types of noise (violet, blue, white, pink, and brown) with different intensities. For each signal type, the first row shows the signal with a minimal amount of noise, and the last with a maximal amount of noise. We can see

Figure 2. Correlation Matrix.
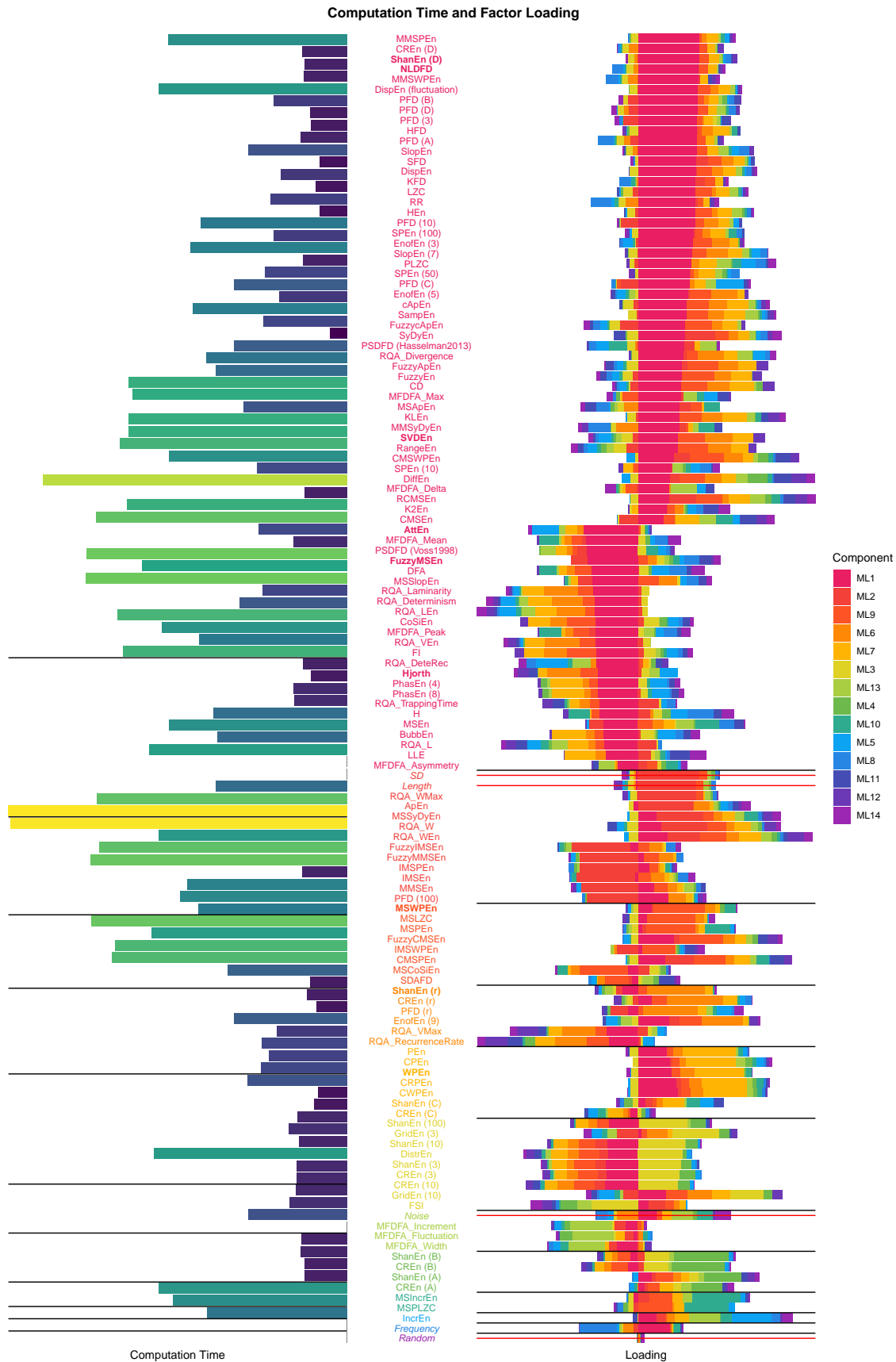
Figure 3. Factor loadings and computation times of the complexity indices, colored by the factor they represent the most.

Figure 4. Correlation network of the complexity indices. Only the links where |r| > 0.6 are displayed.
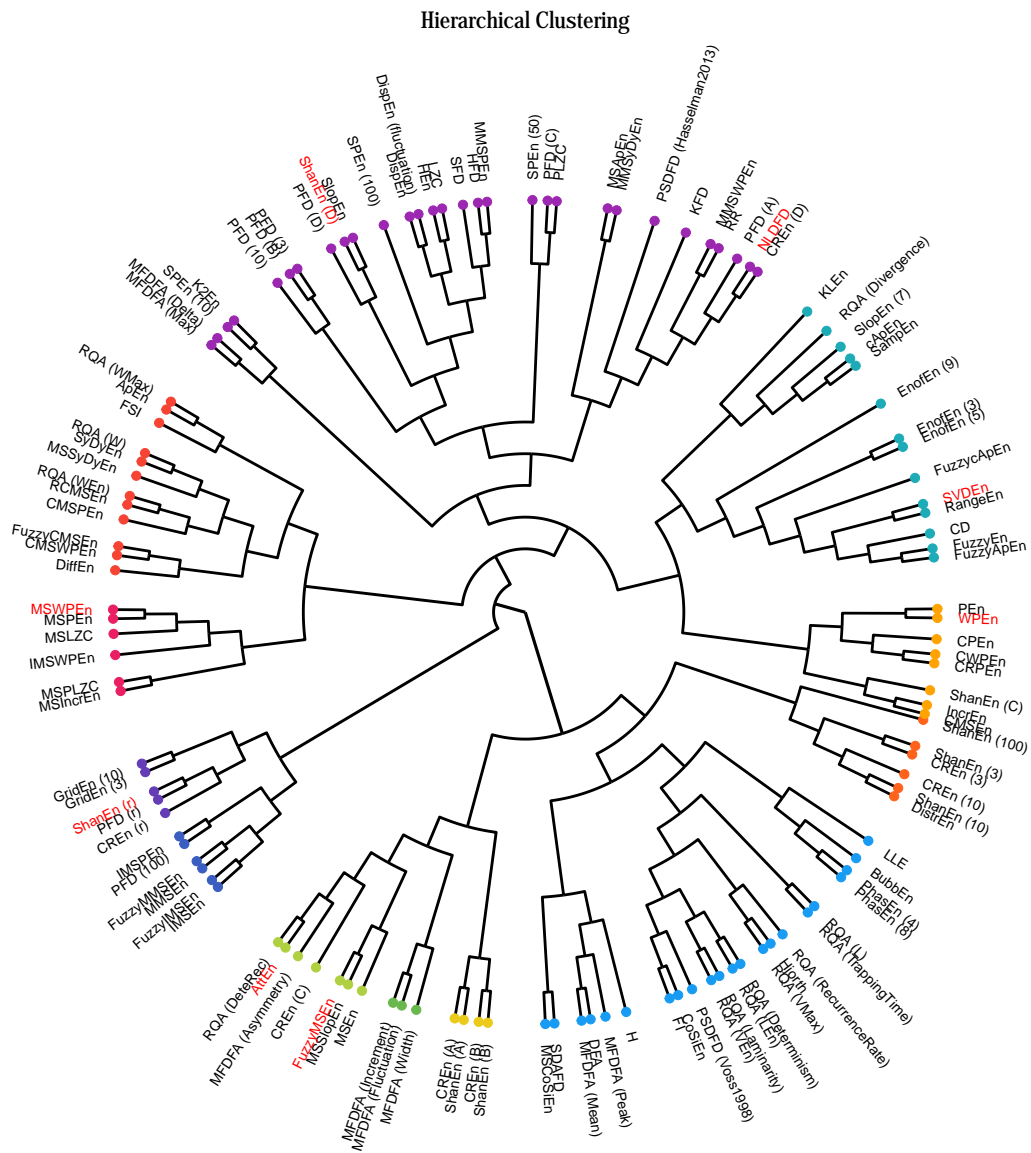
Figure 5. Dendrogram representing the hierarchical clustering of the complexity indices.