bioengineering

MDPI

Article

A Multiscale Structural Analysis Approach Based on Horizontal Visibility Graphs for Electrogastrographic Signals

Luis Gabriel Gómez Acosta * and Max Chacón Pacheco

Departamento de Ingeniería Informática, Facultad de Ingeniería, Universidad de Santiago de Chile (USACH), Av. Víctor Jara 3659, Estación Central, Santiago 9170124, Chile; max.chacon@usach.cl

* Correspondence: luis.gomez.ac@usach.cl

Abstract

Analysis of electrogastrographic (EGG) signals requires methods capable of characterizing structural organization across scales while retaining temporal-order information. We present a multiscale approach based on horizontal visibility graphs (HVGs), consecutive outerplanar motifs, grouped probability distributions (GPDs), and information-theoretic quantifiers. The methodology is illustrated using two representative 28-sample EGG segments and subsequently evaluated using 19 complete recordings across motif sizes $m = 3, \ldots, 50$. Finite-sample effects were assessed by progressively varying signal length, the grouped representation was compared with ungrouped HVG topology counts, and recording-level grouped-HVG quantifiers were compared with conventional EGG spectral descriptors. Finite-sample deviations progressively decreased with increasing signal length, whereas the ungrouped representation exhibited rapid growth in the number of observed topological states at larger motif sizes. GHVGSC exhibited exploratory negative correlations with the percentage of normogastria at $m = 42 - 50$, which were significant under within-family BH-FDR correction but did not survive global correction across all 384 tests. Overall, the proposed framework provides a reproducible multiscale characterization of EGG structural organization, while the expanded evaluation establishes its finite-sample behavior and its relationships with ungrouped HVG representations and conventional spectral descriptors.

Keywords: horizontal visibility graphs (HVGs); outerplanar motifs; grouped probability distributions (GPDs); multiscale entropy and complexity

Received: 1 October 2026
Revised: 25 October 2026
Accepted: 29 October 2026
Published: 3 November 2026

Copyright: © 2026 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.

1. Introduction

This work focuses on the analysis of EGG signals using visibility graphs and ITQs. The study aims to describe the structural organization of EGG signals through a multiscale analysis of outerplanar motifs, grouped probability distributions (GPDs), and entropy and complexity quantifiers derived from horizontal visibility graphs (HVGs), thereby characterizing scale-dependent variations in their grouped motif distributions. Accordingly, the following research question is formulated: How can the structural organization of EGG signals be characterized using grouped ITQs derived from the HVG? The proposed descriptive hypothesis states that the structural organization of EGG signals can be represented through grouped visibility-motif distributions and corresponding entropy and complexity profiles across different motif sizes, while the methodological

behavior of these representations can be examined in terms of finite-sample effects, connectivity-based grouping, and their relationships with conventional EGG spectral descriptors.

EGG records gastric bioelectrical activity and can present analytical challenges associated with low signal-to-noise ratio, temporal variability, nonstationarity, and nonlinear characteristics. From a physiological perspective, the EGG signal reflects gastric myoelectrical activity associated with slow-wave propagation generated and coordinated by the interstitial cells of Cajal. Although conventional spectral and statistical approaches provide valuable information about EGG activity, some analyses may depend on assumptions or representations that do not explicitly describe the structural organization of observations across different scales.

Information-theoretic measures, including entropy and statistical complexity, have also been applied to bioelectrical signals to characterize their organization and variability [1]. In this context, visibility graphs—and HVGs in particular—provide a complementary framework for mapping time series into complex-network representations while retaining their temporal ordering and encoding visibility relationships among observations, without requiring a predefined probability distribution for the signal values. Their application across different types of time series supports the exploration of HVG-based representations for physiological signals such as EGG.

In the present research, previously established grouped-HVG concepts [2] are adapted and organized into a structured analytical workflow for the characterization of scale-dependent variations in EGG signals. The procedure comprises HVG construction, extraction of consecutive overlapping induced subgraphs across multiple motif sizes, grouping according to structural connectivity, construction and retention of recording- and motif-size-specific GPDs, and their subsequent information-theoretic characterization. In this context, the methodological gap addressed by the present study concerns the need for a systematic evaluation of how grouped-HVG representations behave when applied to EGG recordings, particularly with respect to finite-record length, motif size, connectivity-based grouping, and their relationships with conventional spectral descriptors.

In the implementation considered here, the GPDs are quantified using Shannon and Tsallis entropy formulations and their normalized forms, together with the corresponding Hellinger-based statistical complexity measures (GHVGSCSh and GHVGSC). Synthetic time series exhibiting different dynamical characteristics are incorporated as descriptive references for contextualizing the EGG representations in the complexity–entropy plane, rather than as theoretical bounds or classification regimes.

Thus, the methodological contribution of the present study lies in the development of a structured, systematic multiscale workflow that integrates previously established grouped-HVG concepts with explicit construction and retention of signal- and motif-size-specific GPDs under observed and fixed theoretical probability supports, including zero-probability classes in the latter. This framework enables the consistent computation of normalized information-theoretic quantifiers and complexity–entropy representations from the retained distributions, together with a systematic evaluation of finite-sample behavior, grouped versus ungrouped representations, and exploratory relationships with conventional EGG spectral descriptors. The contribution therefore concerns the integration, computational implementation, and methodological assessment of these components, rather than the original introduction of grouped-HVG or a claim of superior analytical performance.

The proposed workflow encompasses HVG construction, extraction of consecutive overlapping induced subgraphs, connectivity-based motif grouping, and the construction and explicit retention of signal- and motif-size-specific GPDs. These distributions are subsequently characterized using normalized Shannon and Tsallis entropies, Hellinger-based

disequilibrium, and the corresponding statistical complexity measures, followed by their representation in the complexity–entropy plane.

Beyond the illustrative examples, the methodological evaluation is extended to 19 complete EGG recordings to examine finite-sample effects, the representational differences between grouped and ungrouped HVGs, and the relationships between grouped-HVG quantifiers and conventional EGG spectral descriptors. This evaluation is methodological in scope and does not establish physiological or diagnostic discrimination between clinical populations.

2. Materials and Methods

The methodology of this study is structured into two complementary phases. The first consisted of a documentary review aimed at consolidating the theoretical foundations of the HVG framework through the search, classification, and critical analysis of the literature, enabling the identification of its main properties, applications, entropy-based measures, and methodological gaps. The second phase involved methodological and computational development, in which a quantitative and reproducible procedure was implemented to transform bioelectrical time series into HVGs, extract consecutive overlapping induced subgraphs, organize their structural patterns into recording- and motif-size-specific GPDs, and compute ITQs for characterizing their structural organization across different motif sizes.

2.1. Background on Visibility-Graph Analysis

Visibility-graph methods provide a geometric framework for mapping time series into networks while preserving relationships among observations. Research in this field has progressively examined visibility criteria, stochastic and deterministic dynamics, local fluctuations, temporal asymmetry and irreversibility, as well as structural descriptors based on motifs, degree distributions, and local graph organization. Natural visibility graphs (NVGs), HVGs, and related weighted or modified formulations have consequently been applied to characterize different structural and dynamical properties of time series.

Within this framework, HVG-based analyses have increasingly incorporated information-theoretic measures to quantify the organization of graph-derived probability distributions. Motif-based representations are particularly relevant because they provide local structural descriptions that can subsequently be aggregated and examined across different motif sizes. Previous grouped-HVG approaches have further organized these structures according to connectivity characteristics, providing the methodological foundation for the GPD-based quantifiers considered in the present study [2].

Despite these developments, methodological challenges remain regarding the interpretation of graph-derived quantifiers across structural scales and their validation using real-world signals. These considerations motivate the present application to EGG signals, in which previously established grouped-HVG concepts are organized into a reproducible multiscale workflow for constructing motif-size-specific GPDs and subsequently characterizing them using ITQs.

2.2. Data and Computational Methods

Ethics Statement. The EGG recordings analyzed in this work were obtained from a previous human-subject study conducted at the Clinical Hospital of the University of Chile. The study protocol was reviewed and approved by the Scientific Ethics Committee of the Clinical Hospital of the University of Chile (Approval Act No. 184; approval granted on 25 June 2008). The study was conducted in accordance with the applicable national and international ethical principles governing biomedical research involving human participants, and all participants provided written informed consent. The present study uses

previously acquired EGG recordings for the methodological evaluation of the proposed HVG-based analytical framework; no new EGG recordings, interventions, or additional data collection involving human participants were performed.

Study Population and Sample Characteristics. The expanded methodological evaluation included 19 complete EGG recordings obtained from 19 control participants at the Clinical Hospital of the University of Chile. The control group comprised 11 women (57.9%) and 8 men (42.1%). Seventeen participants belonged to a historical cohort recorded between 2000 and 2002, including 10 women (mean age: 34.1 years) and 7 men (mean age: 39.5 years). Two additional controls, one woman and one man, both aged 47 years, were recorded in May 2008. Each participant contributed one complete EGG recording, which constituted the independent observational unit for the expanded methodological evaluation. The dataset was selected on the basis of the availability of complete EGG recordings from control participants in the historical cohort used for this research. These 19 recordings constituted the available control dataset for the expanded methodological evaluation, rather than a sample selected to represent a broader clinical population.

EGG Acquisition Protocol. The EGG recordings analyzed in this study were previously acquired at the Clinical Hospital of the University of Chile using the Polygram v5.0 system (Gastrosoft Inc., Irving, TX, USA). The acquisition system incorporated analog filtering for slow gastric activity (0.03–0.15 Hz) and small-intestinal activity (0.15–0.30 Hz), together with an 8-Hz low-pass anti-aliasing filter before digitization. Recordings were obtained by trained clinical personnel following the standardized acquisition and electrode-placement procedures routinely implemented at the institution.

According to subsequent confirmation from the clinical personnel responsible for the historical recordings, the acquired signals were visually reviewed and identified artifacts were removed. However, the specific artifact-rejection criteria and the number of removed segments were unavailable for retrospective verification. No additional digital filtering, interpolation, or baseline-drift correction was identified in the computational procedures used for the present analysis. An independent numerical audit of the 19 complete control recordings (1,179,686 samples) found no empty, invalid, or nonfinite entries. Nevertheless, these checks cannot establish the absence of acquisition-time sample loss, residual artifacts, or baseline drift.

The EGG signals were digitally sampled at 8 Hz; consequently, each 28-sample worked example represents 3.5 s of recorded signal. These short segments were intentionally selected as representative control-signal examples for the step-by-step illustration and reproducibility of the grouped-HVG calculations and were not used for statistical inference.

Given the methodological objective of this study, EGG-1 and EGG-2 are used throughout each stage of the proposed procedure as worked examples, from HVG construction and grouped motif-distribution estimation to the computation of entropy and statistical-complexity quantifiers. Each segment is transformed into an HVG that encodes structural relationships among its time-ordered observations.

For the two 28-sample worked examples, each HVG was subsequently traversed using consecutive overlapping induced subgraphs of motif sizes m = 3, 4, 5, and 8. These motif sizes were selected to examine the multiscale behavior of the framework while accounting for the short length ($n = 28$) of the EGG segments. For a segment of length n, the number of available overlapping subgraphs decreases according to $K_m = n - m + 1$, yielding 26, 25, 24, and 21 subgraphs for $m = 3, 4, 5,$ and 8, respectively. Because increasing m simultaneously enlarges the space of admissible structural configurations while reducing the number of subgraphs available to estimate their occurrence probabilities, the worked examples were restricted to m = 3, 4, 5, and 8 to maintain an interpretable illustration of the GPD construction within the limited 28-sample segments; larger motif sizes

were instead examined in the expanded evaluation using complete EGG recordings. The resulting structural patterns were grouped according to their connectivity characteristics to construct the corresponding GPD at each motif size. These distributions were then used to compute normalized Shannon and Tsallis entropies (GSEHVG and GTEHVG) and their corresponding Hellinger-based statistical complexity measures (GHVGSCSh and GHVGSC), which were subsequently represented in the complexity–entropy plane.

For each signal segment and motif size m, the consecutive overlapping induced subgraphs are used collectively to construct a single empirical GPD, from which the corresponding entropy and statistical complexity quantifiers are calculated. The individual subgraphs are therefore not treated as statistically independent observations, and their number should not be interpreted as an increase in effective sample size. Rather, they constitute overlapping structural observations used to estimate the motif distribution of the same underlying signal segment at a given motif size. Accordingly, for the 28-sample worked examples, the resulting quantifiers represent within-segment, scale-dependent structural characterizations and should not be interpreted as inferential evidence from independent observations.

Although the present study remains methodological in scope, the proposed analytical pipeline is conceptually extensible to comparative studies involving independent groups or experimental conditions. Such applications could include control and clinical populations, age-stratified analyses, or other physiological conditions, provided that the sampling design, signal characteristics, and statistical analysis are appropriately adapted to the specific research question.

No between-group inferential analysis was performed, because the expanded evaluation was designed to characterize the methodological behavior of the proposed framework using 19 complete EGG recordings rather than to test differences between clinical populations. For the expanded methodological evaluation, each complete EGG recording was treated as an independent recording-level unit, whereas the overlapping subgraphs extracted at each motif size were used only to estimate the corresponding GPD and were not treated as independent observations. These analyses were designed to evaluate the finite-sample behavior of the grouped-HVG quantifiers and to support their methodological comparison with the corresponding ungrouped HVG representation and with conventional EGG spectral descriptors.

2.3 Methodological Evaluation Design and Analytical Procedures

To extend the methodological evaluation beyond the two 28-sample worked examples, 19 complete EGG recordings from control participants were analyzed, with one recording per participant. The complete recordings were processed using the same grouped-HVG framework described for the worked examples, but the motif-size range was extended to $m = 3, \ldots, 50$. The upper limit of m = 50 was adopted as an operational choice to examine a broad range of structural motif sizes while retaining a sufficient number of overlapping subgraphs for empirical GPD estimation within the analyzed recordings. This choice was not based on a presumed physiological timescale or a theoretically optimal motif size. The influence of signal length and motif size on the resulting quantifiers was evaluated through the finite-sample analysis described below.

For each recording and motif size, consecutive overlapping induced subgraphs were extracted and used collectively to construct a single recording- and scale-specific GPD. The corresponding normalized Tsallis entropy and Hellinger-based statistical complexity were then calculated from each GPD. Thus, each complete recording, rather than the individual overlapping subgraphs, constituted the independent recording-level unit for the expanded methodological evaluation.

Finite-sample effects were evaluated by progressively truncating each complete recording to common prefix lengths of N = 100, 250, 500, 1000, 2500, 5000, 10,000, and 16,000 samples. The grouped-HVG analysis was independently repeated at each signal length for all motif sizes $m = 3, \ldots, 50$. Because all 19 recordings contained at least 16,000 samples, the results obtained at $N_{ref} = 16{,}000$ were used as a common operational reference. For each shorter signal length, finite-sample deviations were quantified as the absolute differences in normalized Tsallis entropy and Hellinger-based statistical complexity relative to their corresponding values at N_{ref}, separately for each recording and motif size. The reference length was used solely to provide a common within-recording benchmark and should not be interpreted as an asymptotic limit or as a theoretically sufficient sample size.

To evaluate the effect of connectivity-based grouping, the complete recordings were additionally analyzed using an ungrouped HVG representation. For each recording and motif size $m = 3, \ldots, 50$, the same consecutive overlapping induced subgraphs used in the grouped analysis were retained as distinct topological states rather than being aggregated according to their connectivity group. The grouped and ungrouped representations were compared in terms of the number of observed states and, for the ungrouped representation, the proportion of available consecutive subgraphs exhibiting a distinct observed topology. This proportion was calculated as the number of observed ungrouped states divided by the total number of consecutive m-node subgraphs available in the corresponding recording. The comparison was intended to quantify the effect of grouping on the empirical state-space representation and not to establish superiority of either representation.

For comparison with conventional EGG spectral descriptors, each complete recording was analyzed using consecutive 256-s windows (2048 samples at 8 Hz), with a 60-s shift (480 samples). The 256-s window duration provides a nominal Fourier frequency-bin spacing of approximately 0.2344 cycles/min, supporting the estimation of dominant frequency within the conventional gastric frequency range. The 60-s shift corresponds to approximately 76.6% overlap between consecutive windows, allowing repeated spectral estimates throughout each recording while retaining the same window duration. This windowed spectral-analysis approach is consistent with established EGG methodology, although the 60-s shift represents the implementation used in the present study rather than the exact shift reported in earlier protocols. The frequency-band boundaries were retained as predefined operational criteria, recognizing that conventional EGG frequency ranges vary across published methodologies [3,4].

For each window, the dominant frequency was identified as the frequency corresponding to the maximum spectral power within $0 < f \leq 10$ cycles/min. Spectral windows were classified as bradygastric (0–2.2 cycles/min), normogastric (2.3–3.7 cycles/min), or tachygastric (3.8–10 cycles/min) according to their dominant frequency. Dominant frequencies falling within the gaps between these predefined intervals were retained as unclassified rather than reassigned to an adjacent frequency band. Because dominant frequencies were estimated at discrete Fourier frequency bins, the spectral resolution and predefined band boundaries determine the classification assigned to each window. Unclassified windows were retained in the total number of analyzed windows used to calculate the percentages, so the three reported band percentages were not required to sum to 100%. Recording-level descriptors comprised the mean dominant frequency and the percentages of analyzed windows classified within each of the three conventional frequency ranges. These measures were used as conventional spectral reference descriptors for the methodological comparison and not as diagnostic outcomes.

Associations between the grouped-HVG quantifiers and conventional EGG spectral descriptors were evaluated across the 19 complete recordings using Spearman's rank

correlation coefficient, with each recording contributing one pair of recording-level values to each correlation. For each motif size $m = 3, \ldots, 50$, GTEHVG and GHVGSC, calculated using the fixed theoretical support, were separately correlated with mean dominant frequency and the percentages of windows classified as bradygastric, normogastric, and tachygastric. The primary multiple-testing correction was performed using the Benjamini–Hochberg false discovery rate (BH-FDR) procedure separately across the 48 motif sizes for each quantifier–spectral descriptor pair, resulting in eight families of 48 tests. Each family was defined by a specific quantifier–spectral descriptor pair because the 48 motif sizes represent repeated evaluations of the same association across structural scales. The global correction was included to assess the sensitivity of the findings when all tested associations were considered as a single family.

As an additional sensitivity analysis, BH-FDR correction was applied jointly across all 384 correlations. Statistical significance was assessed at an FDR-adjusted threshold of 0.05 for each correction procedure. To characterize uncertainty in the observed GHVGSC–normogastria correlations at motif sizes m = 42–50, two-sided 95% percentile bootstrap confidence intervals were estimated using 10,000 resamples of the 19 recording-level paired observations, sampled with replacement. A fixed random seed (20261020) was used for reproducibility. These confidence intervals were calculated individually for each motif size and were not adjusted for multiple comparisons.

Collectively, these analyses were designed to assess the finite-sample behavior, representational characteristics, and relationship with conventional spectral descriptors of the grouped-HVG framework, while preserving the methodological scope of the study and without evaluating diagnostic or between-group discrimination.

3. Graph-Based Signal Processing and Structural Entropy Measures for Characterizing Complex Phenomena in Time Series

This section aims to provide the conceptual framework necessary to contextualize the proposed methodology, without intending to constitute an exhaustive review of the state of the art. The following section presents the results of the theoretical analysis of the background literature, which describes how the object of study, referred to as the HVG, has been addressed. On the basis of this mapping, the proposal is organized into seven dimensions: (i) the theoretical foundations of the phenomenon are established, showing how visibility graphs emerge as a natural extension of spectral theory and graph theory applied to signals; (ii) the main application domains in physical, economic, engineering, and biomedical contexts are synthesized, thereby justifying their transfer to the problem under study; (iii) the structural and dynamical properties that make them robust tools for characterizing the complexity and irreversibility of time series are analyzed; (iv) the formal construction of the HVG is detailed, with nodes, edges, and visibility criteria as the operational basis of the algorithm; (v) variants and extensions of visibility graphs that allow adapting the network architecture to the characteristics of the data are reviewed; (vi) graph-based measures are integrated to quantify irreversibility, temporal dependence, and dynamical organization; and (vii) graph entropies are incorporated as quantifiers of information and structural complexity, upon which the multiscale structural characterization and methodological evaluation developed in this study are based.

Although HVG construction, connectivity-based grouping of HVG motifs, GPDs, and the associated grouped ITQs have been established previously [2], the contribution of the present study lies in their organization and implementation within an EGG-specific multiscale analytical workflow. In particular, the study applies these established grouped-HVG constructs to EGG signal, examines their behavior across multiple motif sizes m, explicitly retains signal- and motif-size-specific GPDs throughout the analysis, and incorporates synthetic time series with different dynamical characteristics as descriptive

references in the complexity–entropy plane. Thus, the contribution is the structured integration and EGG-specific multiscale implementation of these elements rather than the original introduction of the grouped-HVG concept, the GPD formulation, or the associated entropy and statistical-complexity quantifiers.

3.1. Foundations of Signal Processing Using Visibility Graphs

Graph signal processing (GSP) extends conventional signal-processing concepts to signals defined on graphs and includes operations such as shifting, filtering, convolution, Fourier transformation, spectral decomposition, and impulse response [5–8]. Graph-based representations have been increasingly adopted for analyzing complex and nonlinear systems because they enable signal information to be examined through network topology while retaining structural relationships derived from the original time series [9–16]. Such representations have been applied to pattern characterization in univariate and multivariate signals and to the analysis of properties including temporal irreversibility, system complexity, and self-similarity [9,12,13,17–19].

Within this broader graph-based framework, visibility-graph methods transform time-series observations into nodes whose connections are determined by geometric visibility criteria [6,16,20–22]. The visibility graph algorithm (VGA), introduced by Lacasa and colleagues, provides a computationally efficient transformation that preserves temporal ordering and does not require additional tuning parameters [7,9,13,18,23,24], thereby enabling a representation of the original time series in the visibility-graph domain [5,14].

The horizontal visibility graph (HVG) constitutes a simplified form of the natural visibility graph (NVG) [19,20,25]. In an HVG, each observation is represented by a node, and two nodes are connected when a horizontal line joining their corresponding values does not intersect any intermediate data value [19,26]. This horizontal visibility criterion provides the graph construction subsequently used in the grouped-HVG analysis developed in the following sections.

3.2. Application Domains of Visibility Graphs and Complex

Visibility-graph methods, including HVG, have been applied across a broad range of complex systems and time-series analysis problems [2,8,9,12]. Their applications include classification and characterization, as well as the analysis of asymmetry, fluctuations, temporal irreversibility, nonlinear features, and hidden dynamical information [5,10,14–16,19,22,23]. They have also been investigated using synthetic and theoretical signals, including periodic, random, fractal, and chaotic dynamics, providing reference settings for examining stochastic, deterministic, and complex behavior [2,5,25].

Applications extend across economics and finance [2,7,16,20,22], spatial planning and transportation [13,26], Earth sciences and climate-related phenomena [11,18], process and fluid engineering [15,21], and mechanical engineering, robotics, and electronics [8,23]. These studies illustrate the broad applicability of visibility-based representations to signals and processes exhibiting different structural and dynamical characteristics.

Of particular relevance to the present study, visibility-graph methods have been applied to biomedical, physiological, and neurological signals for characterizing data acquired under different physiological or pathological conditions [9,10,13,16]. Reported applications include cardiovascular and heart-rate dynamics [7,18], neurological and neurodiagnostic problems [17,20,26], cognitive and meditative states [18,25], and electrophysiological signals such as EEG, ECG, and ECoG [25,26]. This biomedical use provides a relevant context for examining HVG-based representations of EGG signals.

3.3. Structural and Dynamic Properties of Visibility Graphs

HVGs derived from time series exhibit several structural properties relevant to their use in signal analysis [16,25]. They are connected and undirected graphs in which each node is connected to its nearest neighbors, preventing isolated nodes and providing a topology whose connectivity reflects aspects of the local structure of the original time series [10,14,15,18,26]. Although the resulting graph is undirected, the visibility-based representation retains structural information associated with the relationships among observations and their temporal ordering [2,10,15].

Local fluctuations in the time series generate recurrent network motifs and outerplanar subgraphs that characterize local topological organization [15,19,21,26]. These graph structures preserve information derived from the original signal after its transformation into a visibility representation [2,20,24]. Another relevant property is affine invariance: translations and linear rescaling of the data do not alter the visibility criterion and therefore preserve the resulting graph structure [5,10,11,15,16,20,25].

Visibility-graph topology can also reflect differences among dynamical regimes. Previous studies have reported characteristic graph structures and degree distributions for periodic, random, fractal, stochastic, and chaotic processes, supporting their use in examining nonlinear and nonstationary dynamics [13,18,19,22,25,26]. Consequently, graph-derived information has been used for structural characterization, pattern recognition, feature extraction, and analysis of subtle variations in complex signals [13,19,25,26].

From a practical perspective, visibility-graph methods offer computational efficiency, straightforward implementation, compact graphical representation, and construction without additional hyperparameter tuning [6,15,20]. They retain structural properties derived from the original signal in the associated graph representation [2,14]; however, the transformation is not information-preserving in an absolute sense, because some information from the original time series is inevitably lost when it is mapped to the resulting network structure [15].

3.4. Construction of the HVG from Time Series and Visibility Criteria

The construction of a visibility graph maps a time series $x(t_j)$, $j = 1, 2, \ldots, N$, onto a graph $G = (V, E)$, where each observation is represented by a vertex $v_j \in V$, and the edges E are established according to a visibility criterion [6,8,12,14,17]. In the HVG, this mapping preserves the temporal ordering of the observations, with each data point of the original time series corresponding to a unique node [5,9,12,13,18].

Two vertices i and j are connected when the corresponding observations satisfy the horizontal visibility criterion [9,13,17,26]. Specifically, an edge exists if every intermediate observation has an amplitude lower than both endpoint values [5,6,11,12,24]. Thus, for $i < j$, horizontal visibility is established when

$$x(t_k) < min\left(x(t_i), x(t_j)\right), \qquad \forall\, t_i < t_k < t_j \tag{1}$$

The visibility criterion uses a strict inequality; therefore, intermediate observations with amplitudes equal to or greater than either endpoint block horizontal visibility. This treatment of equal values is consistent with the implementation using the ts2vg.HorizontalVG library. This binary visibility criterion determines the edge structure of the HVG and provides the geometric representation from which its subsequent topological analysis is performed [7,8,13,16,21].

3.5. Variants of Visibility Graphs

Several methods have been proposed for mapping time series into complex networks, with different construction principles and connectivity rules [6,8,14,20,23]. Within the visibility-graph family, the two principal formulations are the natural visibility graph (NVG) [15,16,18] and the horizontal visibility graph (HVG) [5,13]. Extensions of these

formulations include weighted, directed, multilayer, and penetrable visibility graphs, among other variants designed to incorporate additional characteristics of temporal data [9,10,17,20,21].

Weighted visibility approaches, for example, incorporate information into the edges that is not explicitly represented by the binary connectivity of conventional VG or HVG formulations, whereas multilayer and penetrable variants modify the representation to address multivariate structures or alternative visibility conditions [9,10,12,17,21]. Other network-based transformations, including phase-space reconstruction and recurrence-based approaches, have also been proposed for the characterization of complex and chaotic time series [18,26].

The present study specifically employs the HVG formulation. Accordingly, the subsequent methodology is based on its horizontal visibility criterion and on the structural organization of the resulting graph, rather than on weighted, penetrable, multilayer, or other visibility-graph extensions.

3.6. Utility of the Graphs

Visibility-graph representations provide structural descriptors that have been used to investigate different properties of time series, including scaling behavior, multifractality, and temporal irreversibility [10,16,19]. In particular, irreversibility can be examined by comparing graph-derived distributions or sequential visibility patterns in the forward and backward temporal directions, using divergence measures such as the Jensen–Shannon or Kullback–Leibler divergence [16,19,21].

More generally, the connectivity structure of visibility graphs reflects variations in the organization of the underlying signal: observations with greater visibility tend to generate more highly connected nodes, whereas restricted visibility produces lower connectivity [7,14]. These structural differences provide the basis for extracting graph-derived probability representations and information-theoretic descriptors of time-series organization. In the present study, this utility is specifically exploited through the grouping of HVG subgraphs according to their connectivity characteristics and the subsequent construction of GPDs.

3.7. Graph-Based Entropies as Information Quantifiers

Information measures derived from graph representations are commonly referred to as graph entropies and provide quantitative descriptions of the information content associated with network structure [9,10,15]. In time-series analysis, these measures enable structural information encoded by visibility graphs to be represented through entropy-based quantifiers.

Different formulations have been proposed, including weighted visibility graph entropy (WVGE) [9,12], multiplex motif entropy (MME) [21], normalized HVG-based Shannon entropy [23], discrete generalized past entropy derived from HVG amplitude distributions [14,24], multiscale permutation entropy based on NVG representations [8], and relative-entropy approaches for comparing local network topology [10].

Of particular relevance to the present study, grouped horizontal visibility entropy (GHVE) characterizes structural organization through probability distributions obtained from grouped HVG patterns [18,19]. This formulation provides the conceptual basis for describing the information content of grouped connectivity structures and has been investigated in relation to dynamic fluctuations and physiological or pathological variations [2]. Building on this grouped-HVG representation, the worked examples subsequently evaluate Shannon- and Tsallis-based entropy formulations and their corresponding statistical-complexity measures from the empirical GPDs, whereas the expanded

methodological evaluation focuses on the normalized Tsallis entropy and its corresponding Hellinger-based statistical complexity.

4. Construction and Topological Analysis of HVG in EGG Signals: From Local Motifs to Multiscale Global Metrics

This section presents a five-phase methodological framework for the structural characterization of EGG signals using horizontal visibility graphs (HVGs) and grouped motifs. Phase 1 (Section 4.1) describes the transformation of EGG signals into HVGs, using EGG-1 and EGG-2 as illustrative examples. Phase 2 (Section 4.2) addresses the extraction of consecutive overlapping induced subgraphs and the characterization of their structural configurations according to additional visibility edges. Phase 3 (Section 4.3) defines the search strategies used to identify and enumerate admissible motif configurations according to motif size and connectivity constraints.

Phase 4 (Section 4.4) organizes the extracted motifs into connectivity-based structural classes and constructs the corresponding grouped probability distributions (GPDs). The computational implementation of the overall framework is detailed in Section 4.5, while Section 4.6 presents the GPD results and a worked numerical example demonstrating their traceability. Finally, Phase 5 (Section 4.7) introduces the multiscale information-theoretic quantification of the resulting GPDs through entropy, disequilibrium, and statistical complexity measures. The corresponding results and complexity–entropy representations are presented in Section 4.8, including descriptive contextualization using synthetic reference series with different dynamical characteristics. These references are not intended for statistically matched comparisons or classification.

4.1. Phase 1: Application of Grouped Horizontal Visibility Motifs (GHVM)

As a starting point—see the subsection Construction of the HVG from time series and visibility criteria—consider two intentionally selected 28-sample EGG segments from control participants as worked examples, denoted as EGG-1 = [319.79, 327.18, 330.13, 327.92, 324.96, 322.75, 321.27, 319.79, 316.84, 319.79, 322.01, 318.32, 314.62, 313.15, 312.41, 313.15, 311.67, 317.58, 324.96, 325.7, 326.44, 327.18, 328.66, 333.09, 330.87, 335.3, 341.21, 339.0] and EGG-2 = [258.04, 256.64, 259.44, 260.14, 257.34, 253.85, 253.15, 252.45, 255.24, 255.94, 253.15, 250.35, 246.85, 248.95, 249.65, 252.45, 250.35, 246.15, 242.66, 244.76, 244.06, 247.55, 247.55, 246.15, 245.45, 250.35, 251.75, 258.04].

Both worked examples consist of 28 consecutive EGG samples, corresponding to 3.5 s of signal at the 8-Hz sampling frequency. Given their intentionally limited length, the results derived from EGG-1 and EGG-2 are interpreted exclusively as within-segment structural characterizations for illustrating and reproducing the proposed analytical procedure, rather than as statistically robust estimates or inferential evidence for comparisons between experimental groups. Furthermore, given their short duration, these segments are not intended to characterize gastric slow-wave activity or support physiological interpretations; such analyses require substantially longer recordings.

For each segment, the corresponding HVG was constructed. Figure 1a,c shows the two original temporal sequences together with their non-adjacent horizontal-visibility connections, whereas Figure 1b,d shows the corresponding HVG representations. The additional visibility edges reveal structural relationships that are not represented by temporal adjacency alone. As illustrated in Figure 1, the two segments produce different connectivity patterns according to the relative amplitudes and temporal ordering of their observations. Some nodes establish visibility relationships with observations located several positions away in the original sequence, whereas others retain predominantly local connections. These differences in connectivity constitute the structural information

subsequently examined through consecutive induced subgraphs and grouped motif representations.

Each HVG node represents an observation of the time series, and an edge between two nodes i and j, with $i < j$, is established when every intermediate observation satisfies $x_k < min(x_i, x_j)$ for all $i < k < j$. Consecutive observations are therefore always connected, whereas additional edges may arise between temporally nonadjacent observations when the horizontal-visibility criterion is satisfied. This transformation encodes the ordering and relative-amplitude relationships of the original time series into the connectivity structure of the graph.

The resulting HVGs are shown in Figure 1b,d, with nodes arranged in temporal order and edges depicted as curved lines to facilitate visualization of graph connectivity. These representations illustrate the topological organization induced by the horizontal-visibility criterion [15] and provide the graph structures from which consecutive overlapping induced subgraphs are subsequently extracted and grouped according to their connectivity characteristics.

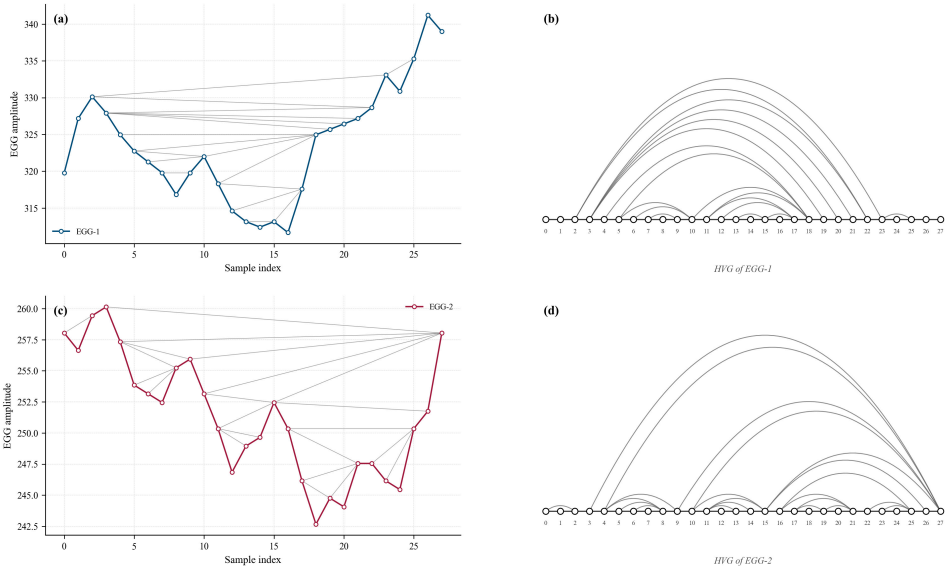
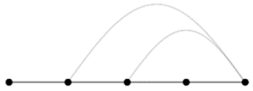
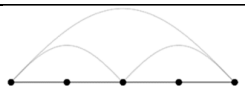


Figure 1. Representative EGG segments and their corresponding horizontal visibility graph (HVG) representations. (**a**) Representative 28-point EGG-1 segment with its non-adjacent horizontal visibility relationships; (**b**) HVG constructed from EGG-1; (**c**) representative 28-point EGG-2 segment with its non-adjacent horizontal visibility relationships; and (**d**) HVG constructed from EGG-2. In panels (**b**,**d**), nodes correspond to time-ordered signal observations and edges represent horizontal visibility relationships.

4.2. Phase 2: Motifs in the HVG

For the EGG-1 and EGG-2 worked examples, motif sizes $m = 3, 4, 5$, and 8 are considered, whereas the expanded evaluation using complete EGG recordings covers the full range $m = 3, \ldots, 50$, as described in Section 2.3. Table 1 illustrates the particular case $m = 5$, for which 22 distinct outerplanar connectivity configurations can be organized according to the number of additional visibility edges beyond the mandatory connections between consecutive nodes [2].

Table 1. HVG-realizable motif configurations for $m = 5$, organized according to the number of additional visibility edges beyond the mandatory connections between consecutive nodes.

Number of Additional Visibility Edges	Motifs	Number of Motifs	Illustration
0	()	1	
1	(0,2); (0,3); (0,4) (1,3); (1,4); (2,4)	6	 (2,4)
2	(0, 2), (0, 3); (0, 2), (0, 4) (0, 2), (2, 4); (0, 3), (0, 4) (0, 3), (1, 3); (0, 4), (1, 3) (0, 4), (1, 4); (0, 4), (2, 4) (1, 3), (1, 4); (1, 4), (2, 4)	10	 (0, 3), (1, 3)  (1, 4), (2, 4)
3	(0, 2), (0, 3), (0, 4) (0, 2), (0, 4), (2, 4) (0, 3), (0, 4), (1, 3) (0, 4), (1, 3), (1, 4) (0, 4), (1, 4), (2, 4)	5	 (0, 2), (0, 4), (2, 4)  (0, 2), (0, 3), (0, 4)

Although Table 1 shows the individual connectivity configurations for $m = 5$, the grouped-HVG approach does not treat each configuration as a separate category. Instead, motifs are grouped according to the number of additional visibility edges beyond the mandatory edges connecting consecutive nodes. Accordingly, the 22 configurations for $m = 5$ are organized into four structural classes containing 0, 1, 2, or 3 additional visibility edges, with 1, 6, 10, and 5 configurations, respectively.

Grouping motifs by the number of additional visibility edges is sufficient for the specific purpose of characterizing the distribution of additional connectivity across motif sizes, but it does not provide a complete topological description of individual motifs. As illustrated in Table 1, different connectivity configurations may be assigned to the same structural class, including configurations corresponding to non-isomorphic graphs. Consequently, the grouping operation discards information about the specific arrangement of visibility edges and the individual topological identities of the motifs. The resulting GPD therefore characterizes the frequency distribution of additional connectivity rather than the full diversity of HVG motif topologies.

Importantly, the admissibility of these configurations is not established by outerplanarity alone. HVGs are characterized as outerplanar graphs containing a Hamiltonian path consistent with the temporal ordering of the observations. In the present construction, this path is guaranteed by the mandatory visibility edges between consecutive nodes, $(i, i + 1)$. The configurations reported in Table 1 were therefore re-verified by retaining these mandatory edges and enumerating the additional visibility-edge configurations compatible with the HVG structural constraints. Figure 2 shows, for comparison, a configuration that violates the outerplanarity condition and is therefore excluded from the HVG motif space.

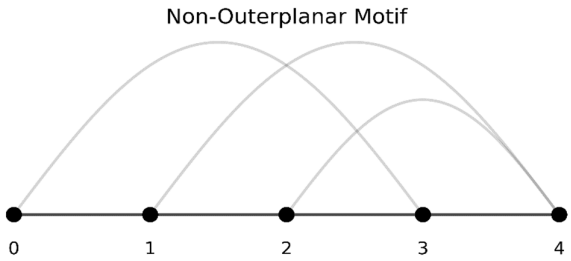


Figure 2. Schematic example of a nonouterplanar connectivity configuration, shown for comparison with the admissible outerplanar HVG motifs considered in the grouped structural analysis.

4.3. Phase 3: Search Strategy

By grouping motifs according to the number of additional visibility edges beyond the $m-1$ mandatory connections between consecutive nodes—as illustrated in Table 1—the set of individual motif configurations is reduced to a smaller number of structural connectivity classes. This grouping can be performed using different search strategies: (i) one option is to fix the subgraph size m and the maximum number of additional visibility edges k_{max} beyond the $m-1$ mandatory connections between consecutive nodes, thereby defining a finite set of admissible motif configurations to be evaluated; (ii) alternatively, targeted searches can be performed on specific regions or topological configurations of interest, previously identified based on differences in the distribution of connectivity groups across the signals or experimental conditions under investigation, thereby enabling a more detailed examination of the motifs comprising those groups; or (iii) a global exploration of the entire graph can be carried out to identify recurrent patterns without predefined constraints.

For the first search strategy, both parameters determine the set of possible motifs and, consequently, the granularity with which the structural organization of the signal is represented. A larger m value includes more nodes within each sliding window, thereby capturing broader interaction structures within the HVG. Similarly, increasing k_{max} expands the admissible motif space by allowing a larger number of additional visibility edges beyond the $m-1$ mandatory consecutive-node connections, thereby incorporating progressively more connected structural configurations.

The motif search space can be either restricted or complete, depending on the selected maximum number of additional visibility edges, k_{max}. A restricted search considers configurations satisfying $0 \le k \le k_{max} \le m-2$, whereas a complete search is obtained by setting $k_{max} = m-2$, thereby including all theoretically admissible additional-edge classes. For example, at $(m = 7)$, a complete search includes the 42 configurations belonging to class g_5, yielding 394 configurations across classes g_0 through g_5. Accordingly, the counts reported in Table 2 refer to the specified pairs (m, k_{max}) and do not necessarily represent the complete configuration space for each motif size.

For a window size m = 6 and a maximum of $k_{max} = 3$ additional visibility edges, 76 admissible outerplanar motif configurations are obtained and organized into four structural classes according to their number of additional edges: g_0, g_1, g_2, and g_3 (Table 2).

When the maximum number of additional visibility edges is increased from $k_{max} = 3$ to $k_{max} = 4$ for $m = 6$, the number of admissible outerplanar motif configurations increases from 76 to 90. For $m = 7$ and $k_{max} = 4$, the admissible motif space comprises 352 outerplanar configurations distributed across the five structural classes g_0, through g_4. The combinatorial growth becomes more pronounced as m and k_{max} increase; for example, for $m = 9$ and $k_{max} = 5$, the admissible motif space comprises 6413 outerplanar configurations distributed across the six structural classes g_0 through g_5. Consequently,

increasing m and k_{max} expands the admissible motif space and the range of structural configurations considered.

At the same time, the available signal length should be considered in relation to the motif-size range of interest, because larger motif sizes reduce the number of sequential subgraphs available for estimating the occurrence probabilities of the corresponding grouped structural classes and can increase finite-sample sensitivity. Conversely, using smaller values of m or restricting the maximum number of additional visibility edges reduces the range of structural configurations considered and simplifies the resulting motif space; the appropriate parameterization therefore depends on the objectives of the analysis and the amount of data available.

The motif configuration counts were independently validated using a Python implementation of the strict horizontal visibility criterion. For each pair (m, k_{max}) reported in Tables 1 and 2, candidate configurations were generated by combining the mandatory edges between consecutive vertices with up to k_{max} additional noncrossing edges. The resulting configurations were grouped according to their number of additional edges, k. Their realizability was then verified by enumerating weak-order amplitude patterns, including equal-valued observations, and checking whether each configuration could be produced by the strict horizontal visibility criterion. All reported configurations were reproduced, with exact agreement in the number of configurations per structural class.

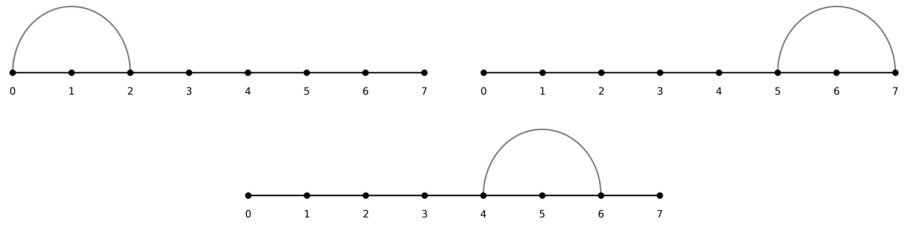
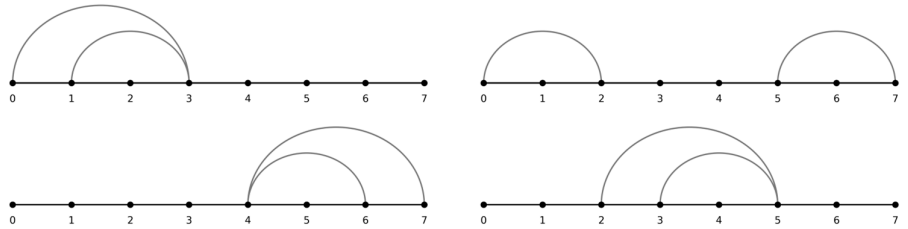
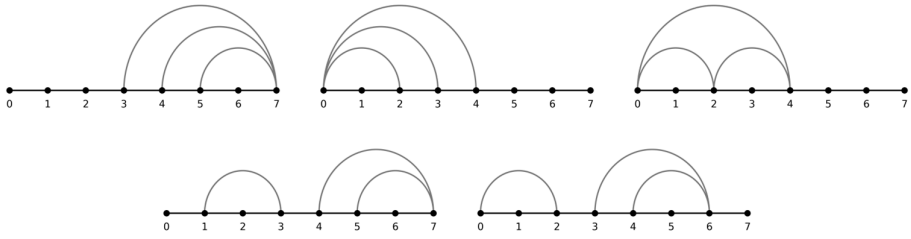
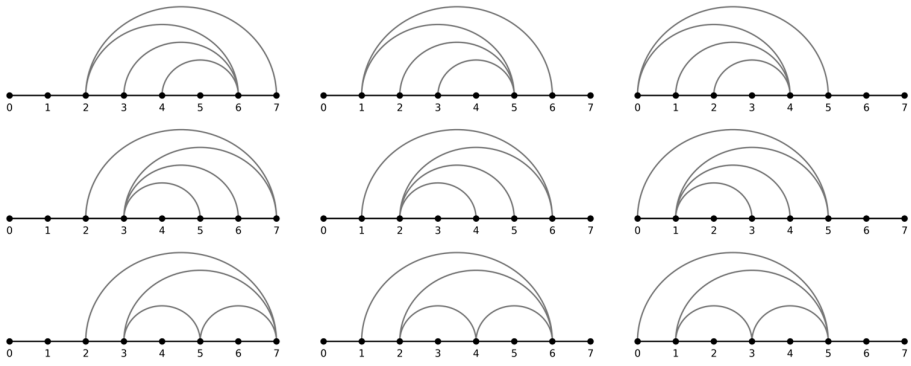
The motif positions and grouped frequencies presented in Tables 3 and 4 were also independently verified using the two illustrative EGG segments. The complete validation procedure, including the illustrative EGG input sequences and the expected table counts, is provided as Supplementary Code S1 (validate_hvg_motif_tables.py) to enable independent reproduction of the reported results.

Table 2. Number of HVG-realizable motif configurations according to motif size m and the maximum number of additional visibility edges.

Motif Size (m)	Max. Additional Visibility Edges (k_{max})	Motif Configurations per Structural Class
6	3	{g_0: 1, g_1: 10, g_2: 30, g_3: 35}
6	4	{g_0: 1, g_1: 10, g_2: 30, g_3: 35, g_4: 14}
7	4	{g_0: 1, g_1: 15, g_2: 70, g_3: 140, g_4: 126}
9	5	{g_0: 1, g_1: 28, g_2: 252, g_3: 1050, g_4: 2310, g_5: 2772}

Applying the third search strategy to EGG-2 with motif size $m = 8$ yields 21 consecutive overlapping induced subgraphs from the original HVG, corresponding to 28 − 8 + 1 = 21 window positions. These subgraph occurrences are assigned to four observed structural classes according to their number of additional visibility edges: g_1, g_2, g_3, and g_4. Specifically, the 21 subgraph occurrences are distributed as follows: g_1 contains the occurrences at window positions 0, 12, and 13; g_2 those at positions 5, 6, 7, and 20; g_3 those at positions 1, 11, 17, 18, and 19; and g_4 those at positions 2, 3, 4, 8, 9, 10, 14, 15, and 16 (Table 3).

Table 3. Visual grouping of the 21 overlapping induced subgraphs of EGG-2 at $m = 8$ according to the number of additional visibility edges. Window positions are zero-based, ranging from 0 to $N - m = 20$, and the vertices within each subgraph are locally indexed from 0 to $m - 1 = 7$.

Window Positions	Grouped Structural Class
0-12-13	
5-6-7-20	
1-11-17-18-19	
2-3-4-8-9-10-14-15-16	

For motif size $m = 5$, 24 consecutive overlapping induced subgraphs are obtained from each EGG signal segment, corresponding to 28 − 5 + 1 = 24 window positions. For EGG-1, these occurrences are distributed as 13 in g_0, 9 in g_1, and 2 in g_2, with none in g_3. For EGG-2, 9 occurrences are assigned to g_0, 8 to g_1, 4 to g_2, and 3 to g_3 (Table 4).

Table 4. Number of overlapping induced-subgraph occurrences assigned to each structural class g_k, defined by the number of additional visibility edges, for EGG-1 and EGG-2 across motif sizes $m = 3, 4, 5$, and 8. For each motif size, the theoretical support comprises the $N_m = m - 1$ classes $g_0, \ldots, g_{m-2}$. A value of zero indicates an admissible class that was not observed in the corresponding signal, whereas a dash indicates a class outside the theoretical support for that motif size.

Motif Size (m)	Signal	g_0	g_1	g_2	g_3	g_4	g_5	g_6	Total
3	EGG-1	22	4	—	—	—	—	—	26
	EGG-2	20	6	—	—	—	—	—	26

4	EGG-1	17	8	0	—	—	—	—	25
	EGG-2	14	8	3	—	—	—	—	25
5	EGG-1	13	9	2	0	—	—	—	24
	EGG-2	9	8	4	3	—	—	—	24
8	EGG-1	4	9	2	4	1	1	0	21
	EGG-2	0	3	4	5	9	0	0	21

In this study, the third search strategy was adopted to systematically examine all consecutive overlapping induced subgraphs observed along each HVG and assign their occurrences to the corresponding structural classes g_k according to the number of additional visibility edges. This strategy provides a direct characterization of the grouped structural organization observed within each signal across the motif sizes considered.

For the implementation adopted in this study, a sliding window of fixed motif size m is displaced one sample at a time along the HVG. Each window defines an induced subgraph composed of m consecutive vertices while preserving all visibility edges among those vertices. Consequently, adjacent windows overlap by $m-1$ vertices, and a signal segment of length n generates $K_m = n - m + 1$ overlapping induced-subgraph occurrences. Each occurrence is then assigned to its corresponding structural class g_k according to the number k of additional visibility edges. The occurrence frequencies of these classes, including zero occurrences for admissible but unobserved classes, are normalized by K_m to construct the grouped probability distribution (GPD) over the fixed theoretical support $S_m = \{g_0, \ldots, g_{m-2}\}$, with $N_m = m - 1$ states.

A limitation of this strategy is that increasing the motif size m expands the space of admissible structural configurations [2] while, for a fixed signal length n, simultaneously reducing the number of available overlapping subgraphs according to $K_m = n - m + 1$. Consequently, the empirical occurrence probabilities may be estimated from progressively fewer subgraph observations at larger m, which can affect the finite-sample stability of the resulting GPDs. This limitation is particularly relevant for the short EGG-1 and EGG-2 worked examples and motivates the finite-sample evaluation performed using the complete EGG recordings over progressively increasing signal lengths, as described in Section 2.3.

Therefore, exploring different motif sizes provides a multiscale description of the structural organization encoded by the HVG, allowing connectivity patterns to be examined across subgraphs of increasing size. Larger motif sizes incorporate structural relationships among a greater number of consecutive observations, although their interpretation must account for the decreasing number of available subgraphs in finite-length signals. In this context, the complete graph can be viewed in terms of the local motif occurrences that contribute to its overall topological organization [15].

4.4. Phase 4: GPDs

The distribution of subgraph occurrences across the structural classes g_k provides the basis for constructing the grouped probability distribution (GPD) at each motif size m. Each overlapping induced subgraph is assigned to a class g_k according to its number k of additional visibility edges and contributes one occurrence to that class. If $n_k^{(m)}$ denotes the number of occurrences assigned to g_k and $K_m = n - m + 1$ is the total number of overlapping subgraphs extracted from an EGG time series of length n, the empirical probability associated with class g_k is defined as $p_k^{(m)} = {n_k^{(m)}}/{K_m}$. The GPD at motif size m is therefore formed by the probabilities associated with the $N_m = m - 1$ structural classes in the fixed theoretical support = $\{g_0, \ldots, g_{m-2}\}$, including zero-probability states

when an admissible class is not observed, and constitutes the probability distribution used in the subsequent entropy and statistical-complexity calculation.

The cardinality of the theoretical grouped support can be established formally. For an HVG motif with $m \geq 3$ vertices, the $(m-1)$ mandatory edges between consecutive vertices and the outerplanar upper bound of $(2m-3)$ edges imply that the number of additional visibility edges satisfies $0 \leq k \leq m-2$. To establish that every integer in this interval is realizable, consider, for any $k \in \{0, \ldots, m-2\}$, the time series $x_0 = k + \frac{1}{2}$ for $x_i = 1, i = 1, \ldots, m-1$. Under the strict horizontal visibility criterion, the first observation has exactly k additional visibility connections, while the strictly increasing remaining observations introduce no other nonconsecutive connections. Thus, every class g_k is realizable by at least one HVG motif. Consequently, the complete theoretical grouped support contains exactly $N_m = m-1$ admissible classes, independently of which classes are observed in a particular recording.

4.5. Algorithmic Implementation of the Proposed Framework

Algorithm 1 summarizes the general computational architecture of the proposed grouped-HVG framework. For each EGG time series and motif size m, the HVG is constructed and consecutive induced subgraphs are extracted and assigned to connectivity groups according to $k = |E(W)| - |V(W)| + 1$. The resulting group occurrences are normalized to obtain the signal- and motif-size-specific grouped probability distribution $P^{(s,m)}$. Importantly, each GPD is explicitly retained before the subsequent calculation of ITQs, thereby preserving the probabilistic representation generated at each motif size for subsequent analyses.

The GPD constitutes the central probabilistic representation connecting the structural and information-theoretic stages of the framework. Once $P_r^{(s,m)}$ has been constructed and stored, it can be evaluated using an entropy formulation $\mathrm{e} \in \mathfrak{E}$ and a disequilibrium or distance measure $\mathrm{d} \in \mathfrak{D}$. For each GPD, P_e denotes the corresponding uniform reference distribution defined over the probability-vector support used in the calculation, whereas Θ contains any formulation-specific parameters. The normalized entropy H_e and disequilibrium Q_d are subsequently combined to obtain the statistical complexity $C_{e,d} = H_e Q_d$, and the pair $(H_e, C_{e,d})$ defines the corresponding coordinate in the complexity–entropy plane.

For the worked-example implementation presented in this section, the general framework described in Algorithm 1 was instantiated using the two signal segments EGG-1 and EGG-2, with motif sizes $m = \{3, 4, 5, 8\}$; the expanded methodological evaluation subsequently applied the framework to 19 complete EGG recordings over $m = 3, \ldots, 50$, as described in Section 2.3. The entropy formulations considered were Shannon entropy and Tsallis entropy, with $q = 0.25$ for the latter, while the Hellinger-based measure was used for the disequilibrium component of statistical complexity. Accordingly, the generalized sets of entropy formulations $\mathfrak{E}$ and disequilibrium/distance measures $\mathfrak{D}$ in Algorithm 1 represent the extensible architecture of the framework rather than a collection of formulations evaluated in the present analysis. Alternative entropy and distance formulations were therefore not comparatively assessed in the present study.

Algorithm 1. General HVG-Based Grouped-Motif Framework for EGG Signal Analysis

Require: S = {$s^{(1)}$, …, $s^{(J)}$}, where $s^{(j)} = (x_0^{(j)}, \ldots, x_{(n_j-1)}^{(j)})$ is a real-valued sequence of N_j samples; M = {m_1, …, m_l}; entropy formulations $\mathfrak{E}$; disequilibrium/distance measures $\mathfrak{D}$; support definitions $\mathfrak{K}$ = {observed, theoretical}; parameter set Θ required by the selected entropy and disequilibrium formulations.

Ensure: GPDStore, MetricsStore, CECoordinates

1: **Initialize GPDStore ← {}**

Initialize MetricsStore ← {}
Initialize CECoordinates ← {}
for each signal s ∈ S do
$N_s \leftarrow |s|$
$G \leftarrow$ ConstructHVG(s)
▷ *Connect $i < j$ iff $x_t < \min(x_i, x_j)$ for all $i < t < j$*
for each motif size m ∈ M do
if m < 3 or m > N_s then
continue
end if
$W_m \leftarrow$ ExtractConsecutiveInducedSubgraphs(G, m)
▷ *Zero-based indexing: $i = 0, \ldots, |V(G)| - m$; window vertices $i, \ldots, i + m - 1$*
Initialize $n_k^{(s,m)} \leftarrow 0$ for k = 0, ..., m − 2
for each induced subgraph W ∈ W_m do
$k \leftarrow |E(W)| - |V(W)| + 1$
Assign W to connectivity group g_k
$n_k^{(s,m)} \leftarrow n_k^{(s,m)} + 1$
end for
for each support definition r ∈ $\mathscr{R}$ do
if r = observed then
$S_{m,r} \leftarrow \{g_k : n_k^{(s,m)} > 0\}$
else
$S_{m,r} \leftarrow \{g_0, g_1, \ldots, g_{m-2}\}$
▷ *Complete admissible support with $N_m = m - 1$*
end if
$P_r^{(s,m)} \leftarrow$ NormalizeGroupOccurrences($\{n_k^{(s,m)}\}$, $S_{m,r}$)
▷ *$p_k^{(s,m)} = n_k^{(s,m)} / (N_s - m + 1)$; retain zero-probability classes*
GPDStore[s, m, r] ← $P_r^{(s,m)}$
$P_{e,r}^{(m)} \leftarrow$ UniformDistribution($S_{m,r}$)
▷ *Uniform reference over selected support; reported results use theoretical support*
for each entropy formulation e ∈ $\mathscr{E}$ do
$S_{e,r} \leftarrow$ Entropy($P_r^{(s,m)}$, e, Θ)
$H_{e,r} \leftarrow$ NormalizedEntropy($P_r^{(s,m)}$, e, Θ)
▷ *Normalize by maximum entropy for formulation e and support $S_{m,r}$*
for each measure d ∈ $\mathcal{D}$ do
$Q\mathcal{d}_{,r} \leftarrow$ Disequilibrium($P_r^{(s,m)}$, $P_{e,r}^{(m)}$, d, Θ)
$C_{e,}\mathcal{d}_{,r} \leftarrow H_{e,r}\, Q\mathcal{d}_{,r}$
MetricsStore[s, m, r, e, d] ← $\{S_{e,r}, H_{e,r}, Q\mathcal{d}_{,r}, C_{e,}\mathcal{d}_{,r}\}$
CECoordinates[s, m, r, e, d] ← $(H_{e,r}, C_{e,}\mathcal{d}_{,r})$
end for
end for
end for
end for
end for
return GPDStore, MetricsStore, CECoordinates

The proposed framework processes each signal through overlapping sliding windows of fixed motif size m. For each window starting at vertex (i), the induced subgraph is defined as $G[V_i] = (V_i, E_i)$, where $V_i = \{i, i+1, \ldots, i+m-1\}$ and $E_i = \{(u,v) \in E, v \in V_i\}$, with (E) denoting the edge set of the original HVG. Thus, every HVG edge connecting two vertices within the window is retained, including edges between temporally non-adjacent vertices. Each window contributes one induced-subgraph occurrence to the construction of the corresponding empirical GPD; these overlapping occurrences are computational units used to estimate group frequencies and are not treated as statistically

independent observations. As the number of analyzed windows and motif sizes increases, the computational workload also increases, while the general sequence of HVG construction, subgraph extraction, connectivity grouping, GPD storage, and information-theoretic quantification remains unchanged. Because the GPDs are explicitly retained, alternative information-theoretic formulations can subsequently be evaluated from the stored distributions without repeating the preceding grouped-HVG construction stage.

Regarding computational complexity, the HVG is constructed once for each signal and its edge list is reused across motif sizes. For grouped-HVG analysis, motif enumeration is implemented by determining the range of sliding windows containing each HVG edge. Edge contributions are accumulated using a difference array followed by a cumulative sum, avoiding explicit reconstruction of individual subgraphs. For a signal with (N) samples and (E) HVG edges, this counting procedure requires $O(E+N)$ time per motif size m. The subsequent construction of the grouped probability distribution and calculation of entropy and Hellinger disequilibrium require at most $O(m)$ additional time, since the theoretical grouped support contains $m-1$ states, including unobserved classes with zero probability. In contrast, ungrouped-HVG analysis explicitly extracts the internal-edge signatures of the overlapping subgraphs to distinguish their topologies, with a conservative time bound of $O(N+E\,log\,E+(N-m+1)m)$ per motif size. These bounds exclude the initial HVG construction.

4.6. Results of the GPDs

The GPD results presented in this section correspond specifically to the EGG-1 and EGG-2 worked examples and illustrate the construction and interpretation of the grouped distributions at motif sizes $m=3,4,5$, and 8. For EGG-1 at motif size $m=8$, a total of $K_8=21$ overlapping induced subgraphs are obtained. Over the fixed theoretical support $S_8=\{g_0,\dots,g_6\}$, their occurrence counts are $n^8=(4,9,2,4,1,1,0)$, yielding the empirical GPD $P^{(8)}=(0.1905,0.4286,0.0952,0.1905,0.0476,0.0476,0.0000)$. Thus, g_1 has the highest observed relative frequency (0.4286), followed by g_0 and g_3 (0.1905 each), whereas g_2, g_4, and g_5 occur less frequently and the admissible class g_6 is not observed in this segment (Table 5).

Table 5. Empirical GPDs obtained for EGG-1 and EGG-2 across motif sizes $m=3,4,5$, and 8. Each GPD is defined over the fixed theoretical support $g_0,\dots,g_{m-2}$, with $N_m=m-1$, and represents the relative occurrence probabilities of the structural connectivity classes defined by the number of additional visibility edges. Admissible classes not observed in a given signal are retained with probability zero, whereas a dash indicates a class outside the theoretical support for that motif size.

Motif Size (m)	Signal	g_0	g_1	g_2	g_3	g_4	g_5	g_6
3	EGG-1	0.8462	0.1538	—	—	—	—	—
	EGG-2	0.7692	0.2308	—	—	—	—	—
4	EGG-1	0.6800	0.3200	0.0000	—	—	—	—
	EGG-2	0.5600	0.3200	0.1200	—	—	—	—
5	EGG-1	0.5417	0.3750	0.0833	0.0000	—	—	—
	EGG-2	0.3750	0.3330	0.1670	0.1250	—	—	—
8	EGG-1	0.1905	0.4286	0.0952	0.1905	0.0476	0.0476	0.0000
	EGG-2	0.0000	0.1429	0.1905	0.2381	0.4286	0.0000	0.0000

The organization of the observed subgraph occurrences into structural classes according to their number of additional visibility edges, as represented by the GPDs, follows the grouped-HVG methodology previously proposed in [2]. The grouped ITQs GSEHVG, GHVGSCSh, GTEHVG, and GHVGSC are likewise adopted from that methodological

foundation rather than introduced here. In the present study, these established constructs are implemented within the EGG-specific multiscale workflow described above and evaluated for the EGG-1 and EGG-2 worked examples across the selected motif sizes $m = 3, 4, 5,$ and 8 (Table 6).

For the EGG-1 and EGG-2 worked examples, both raw and normalized Shannon- and Tsallis-based entropies were computed for each analyzed motif size m. The normalized entropies were evaluated over the fixed theoretical support of $N_m = m - 1$ structural classes, including admissible but unobserved classes with zero probability. The corresponding statistical-complexity measures were calculated from the normalized entropy and the standard Hellinger distance between each empirical GPD and the uniform distribution defined over the same theoretical support. The Shannon- and Tsallis-based formulations provide complementary information-theoretic descriptions of the empirical GPDs, with the Tsallis formulation evaluated at $q = 0.25$. The resulting values of GSEHVG, GTEHVG, GHVGSCSh, and GHVGSC are summarized in Table 6 and subsequently used for the multiscale descriptive analysis and complexity–entropy representations

Table 6. Grouped entropy and statistical-complexity measures obtained for EGG-1 and EGG-2 across motif sizes $m = 3, 4, 5,$ and 8. GSEHVG and GTEHVG denote the Shannon- and Tsallis-based grouped entropy measures, respectively, whereas GHVGSCSh and GHVGSC denote their corresponding Hellinger-based statistical-complexity measures. Normalized entropy values were computed over the fixed theoretical support of $N_m = m - 1$ structural classes for each motif size, including admissible but unobserved classes with zero probability. Statistical complexity was computed as the product of the corresponding normalized entropy and the standard Hellinger distance used as the disequilibrium term. Tsallis-based quantities were calculated using $q = 0.25$.

Motif Size (m)	Signal	GSEHVG	GSEHVG Normalized	GHVGSCSh	GTEHVG	GTEHVG Normalized	GHVGSC
3	EGG-1	0.42932	0.61938	0.16644	0.78051	0.85859	0.23071
3	EGG-2	0.54020	0.77935	0.15615	0.83948	0.92346	0.18502
4	EGG-1	0.62687	0.57060	0.25346	0.88028	0.51599	0.22920
4	EGG-2	0.94375	0.85904	0.17469	1.60766	0.94235	0.19163
5	EGG-1	0.90698	0.65425	0.27872	1.57029	0.64412	0.27440
5	EGG-2	1.29257	0.93239	0.14516	2.36790	0.97129	0.15121
8	EGG-1	1.50873	0.77533	0.29119	3.49356	0.79315	0.29788
8	EGG-2	1.29865	0.66738	0.34051	2.37742	0.53975	0.27539

Worked Numerical Example and Traceability

As a worked example enabling direct verification of the complete analytical sequence, EGG-2 at $m = 8$ can be traced across Tables 3–6. For this case, the 28-point signal generates $K_8 = 28 - 8 + 1 = 21$ overlapping induced subgraphs. As shown in Table 3, their classification according to the number of additional visibility edges yields 3, 4, 5, and 9 occurrences of classes g_1, g_2, g_3 and g_4, respectively, whereas the admissible classes g_0, g_5 and g_6 are not observed. The resulting count vector over the fixed theoretical support $S_8 = \{g_0, \ldots, g_6\}$ is therefore (0,3,4,5,9,0,0), reproducing the corresponding entry in Table 4. Normalizing these structural-class counts by the total number of available subgraphs, $K_8 = 21$, gives the empirical GPD reported in Table 5,

$$P_g^{(8)} = \left(0, \frac{3}{21}, \frac{4}{21}, \frac{5}{21}, \frac{9}{21}, 0, 0\right) = (0, 0.142857, 0.190476, 0.238095, 0.428571, 0, 0)$$

The probability space is defined by the complete theoretical support for $m = 8$, which contains $N_8 = m - 1 = 7$ structural classes, including admissible classes with zero empirical probability. Accordingly, the reference uniform distribution is

$$P_e = \left(\frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}\right).$$

Application of Equations (2)–(7) gives $GSEHVG = 1.29865$, $H_{Sh} = 0.66738$, $GTEHVG_{q=0.25} = 2.37742$, and $H_q = 0.53975$. Using the standard Hellinger distance in Equation (8) gives $D_{He} = 0.510218$. Equation (9) consequently yields

$$C_{Sh,He} = 0.66738(0.510218) \approx 0.34051,$$

and

$$C_{q,He} = 0.53975(0.510218) \approx 0.27539.$$

These values reproduce the corresponding EGG-2, $m = 8$ results reported in Table 6, providing a directly verifiable numerical trace from the induced-subgraph classification in Table 3 through the structural-class counts and empirical GPDs in Tables 4 and 5 to the final ITQs.

4.7. Phase 5: Multiscale Information-Theoretic Quantification of Grouped HVG Motifs

For the EGG-1 and EGG-2 worked examples, the HVG is traversed using a one-sample sliding window at motif sizes $m = 3,4,5,$ and 8 to generate the corresponding induced subgraphs. Each subgraph is assigned to a structural connectivity class g_k according to its number of additional visibility edges. For each motif size, the occurrence frequencies are represented over the fixed theoretical support $S_m = \{g_0, \ldots, g_{m-2}\}$, with $N_m = m - 1$ structural classes; admissible classes that are not observed in a given signal are retained with zero probability. Normalization of the class-occurrence counts by the number of available subgraphs $K_m = n - m + 1$ yields the empirical GPD for that signal segment and motif size. The resulting GPD provides the probability distribution from which the grouped ITQs, including the Shannon- and Tsallis-based entropy and statistical-complexity measures considered in this study, are computed.

Within this framework, entropy quantifies the informational content associated with the distribution of connectivity classes, whereas statistical complexity combines normalized entropy with a disequilibrium term that measures departure from the uniform distribution defined over the same theoretical support. In the present analysis, this disequilibrium is quantified using the standard Hellinger distance, without additional dimension-dependent normalization. Their joint use therefore provides complementary information about the organization of the empirical GPDs derived from the HVG [15].

Once the empirical GPD has been constructed for each worked example and motif size m, the corresponding ITQs are computed from that distribution. The overlapping sliding windows are used exclusively to estimate the structural-class frequencies underlying each GPD and are not treated as independent observational units. For the present analysis, the metric values obtained from EGG-1 and EGG-2 at the same motif size are subsequently averaged, as reported in Table 7. These averages summarize the two representative segments and are not used for statistical inference. In the expanded methodological evaluation, each complete EGG recording is retained as an independent recording-level unit; future applications involving comparisons between experimental or clinical conditions should likewise preserve the individual recording or subject as the observational unit for appropriate statistical inference.

Table 7. Arithmetic means of the ITQs obtained from the two representative EGG worked examples (EGG-1 and EGG-2) at each motif size m. GSEHVG and GTEHVG denote the Shannon- and Tsallis-based grouped entropy measures, respectively, whereas GHVGSCSh and GHVGSC denote their corresponding Hellinger-based statistical-complexity measures. Normalized entropy values were computed over the fixed theoretical support of $N_m = m - 1$ structural classes for each motif size, including admissible but unobserved classes with zero probability. Statistical complexity was computed as the product of the corresponding normalized entropy and the standard Hellinger distance

used as the disequilibrium term. The Tsallis-based GTEHVG and GHVGSC measures were calculated using $q = 0.25$. These averages are presented for descriptive purposes only and are not used for statistical inference.

Motif Size (m)	GSEHVG	GSEHVG Normalized	GHVGSCSh	GTEHVG Normalized (q = 0.25)	GHVGSC (q = 0.25)
3	0.484760	0.699365	0.161295	0.891025	0.207865
4	0.785310	0.714820	0.214075	0.729170	0.210415
5	1.099775	0.793320	0.211940	0.807705	0.212805
8	1.403690	0.721355	0.315850	0.666450	0.286635

Consider the two representative EGG worked examples, EGG-1 and EGG-2. At motif size $m = 3$, their empirical GPDs yield GSEHVG values of 0.42932 and 0.54020, respectively (Table 6). The arithmetic mean of the two GSEHVG values is 0.48476. Similarly, the normalized GSEHVG values of 0.61938 and 0.77935 yield a mean of 0.69937, while the corresponding GHVGSCSh values of 0.16644 and 0.15615 yield a mean of 0.16130 (Table 7). These values illustrate how the segment-specific ITQs can be descriptively summarized at a common motif size. The same procedure is applied independently at each motif size considered for the worked examples using its corresponding fixed theoretical support; the averages obtained from EGG-1 and EGG-2 are used only for descriptive purposes and not for statistical inference.

When the framework is extended to comparative studies, the data can be organized according to experimental condition and motif size m. For each signal or subject, the overlapping induced-subgraph occurrences are assigned to their corresponding structural classes over the fixed theoretical support for that motif size, defining a signal-specific GPD from which the ITQs are computed. The resulting metric values can then be summarized within experimental groups while retaining the individual signal or subject as the observational unit for statistical analysis. This organization enables the behavior of the ITQs to be examined across motif sizes without treating the overlapping subgraphs used to construct each GPD as independent observations.

As the motif size m increases, each induced subgraph contains a larger number of consecutive HVG vertices and therefore represents connectivity relationships within progressively larger structural configurations. At the same time, the theoretical grouped support expands according to $N_m = m - 1$, while the number of available overlapping subgraphs decreases according to $K_m = n - m + 1$ for a fixed signal length n. Accordingly, examining the empirical GPDs and their associated ITQs across different values of m provides a multiscale structural description of the signal in the HVG domain. Here, "scale" refers specifically to the motif size m, defined by the number of consecutive vertices included in each induced subgraph, and should not be interpreted as a physiological timescale or as direct evidence of distinct underlying dynamical mechanisms.

4.8. Results of the Multiscale Information-Theoretic Analysis

The multiscale information-theoretic results presented in this section correspond specifically to the EGG-1 and EGG-2 worked examples at motif sizes $m = 3, 4, 5$, and 8 and are intended to illustrate the behavior of the proposed quantifiers across the selected motif sizes. The normalized Shannon- and Tsallis-based entropy values exhibit motif-size-dependent variations rather than a common monotonic trend across the analyzed scales. The mean normalized GSEHVG values are 0.699365, 0.714820, 0.793320, and 0.721355 for $m =$ 3,4,5, and 8, respectively, whereas the corresponding normalized GTEHVG values at $q =$ 0.25 are 0.891025, 0.729170, 0.807705, and 0.666450. The associated statistical-complexity measures also vary with motif size: GHVGSCSh ranges from 0.161295 at $m = 3$ to 0.315850 at $m = 8$, while GHVGSC ranges from 0.207865 to 0.286635 over the same analyzed motif sizes. Within the scope of the two EGG-1 and EGG-2 worked examples

considered here, these variations describe how the grouped connectivity distributions and their information-theoretic quantifiers change with motif size under the fixed theoretical-support formulation. They should be interpreted as a descriptive multiscale structural characterization rather than as evidence of comparative sensitivity, physiological discrimination, or distinct underlying physiological mechanisms.

Beyond the worked examples presented in this section, the expanded methodological evaluation applies the grouped-HVG framework to 19 complete EGG recordings while preserving each recording as an independent observational unit. This evaluation examines finite-sample behavior across increasing signal lengths, compares the grouped representation with the corresponding ungrouped HVG representation, and assesses recording-level associations with conventional EGG spectral descriptors. These analyses extend the methodological assessment of the proposed framework but are not intended to evaluate its ability to differentiate physiological or pathological conditions, which would require appropriately designed cohort-based comparisons and subject-level statistical inference.

5.Expanded Methodological Evaluation Using Complete EGG Recordings

The expanded methodological evaluation included 19 complete EGG recordings, with one recording per participant. Recording durations ranged from 34.64 to 154.35 min, with a mean duration of 129.35 ± 24.42 min. These complete recordings were analyzed independently across motif sizes $m = 3,\ldots,50$, yielding one grouped probability distribution and the corresponding normalized Tsallis entropy and Hellinger-based statistical complexity for each recording and motif size. This recording-level design was used for all subsequent finite-sample, grouped-versus-ungrouped, and conventional EGG descriptor analyses.

5.1. Finite-Sample Effects

Finite-sample effects were assessed by comparing the grouped-HVG quantifiers obtained at progressively increasing signal lengths with those obtained at the common operational reference length $N_{ref} = 16{,}000$ samples. Across all recordings and motif sizes, the mean absolute deviation in normalized Tsallis entropy decreased from 0.2360 at N = 100 to 0.1476, 0.0991, 0.0744, 0.0422, 0.0239, and 0.0135 at N = 250, 500, 1000, 2500, 5000, and 10,000, respectively. The corresponding mean absolute deviation in Hellinger-based statistical complexity decreased from 0.0515 at N = 100 to 0.0331, 0.0255, 0.0204, 0.0144, 0.0094, and 0.0049 over the same sequence of signal lengths. By definition, both deviations were zero at N_{ref}.

Finite-sample deviations also depended on motif size. At shorter signal lengths, deviations were generally smaller for the lowest motif sizes and increased as larger motifs were considered. For example, at N = 100, the mean absolute deviations in GTEHVG were 0.0363, 0.1211, 0.2272, 0.2868, 0.3369, and 0.3460 for $m = 3, 10, 20, 30, 40$, and 50, respectively. The corresponding deviations in Hellinger-based statistical complexity were 0.0117, 0.0272, 0.0445, 0.0537, 0.0733, and 0.0942. At N = 10,000, the entropy deviations at the same motif sizes were reduced to 0.0041, 0.0098, 0.0137, 0.0185, 0.0189, and 0.0184, respectively. These results show that finite-sample deviations decreased as the available signal length increased and that the effect of limited signal length was more pronounced at larger motif sizes, particularly for the shortest signals. To further characterize between-recording variability, the standard deviation of the absolute deviations relative to the operational reference was calculated across the 19 recordings. At N = 100, the mean absolute deviation in GTEHVG was 0.0363 ± 0.0275 (mean ± SD) for $m = 3$ and 0.3460 ± 0.0758 for

$m = 50$. The corresponding values for GHVGSC were 0.0117 ± 0.0096 and 0.0942 ± 0.0432, respectively. At N = 10,000, the GTEHVG deviations decreased to 0.0041 ± 0.0037 for $m = 3$ and 0.0184 ± 0.0214 for $m = 50$. These results illustrate between-recording variability and the greater finite-sample sensitivity of larger motifs, particularly for short recordings, without implying convergence to asymptotic values.

Figure 3 summarizes the finite-sample deviations across the complete range of motif sizes and evaluated signal lengths. The heatmaps show the mean absolute deviations relative to $N_{ref} = 16000$ samples for normalized Tsallis entropy (Figure 3A) and Hellinger-based statistical complexity (Figure 3B), with the grouped probability distributions defined over the fixed theoretical support $S_m = \{g_0, \ldots, g_{m-2}\}$. The reference row at N = 16000 is zero by definition and represents the common operational benchmark used for comparison rather than an asymptotic or theoretically sufficient signal length.

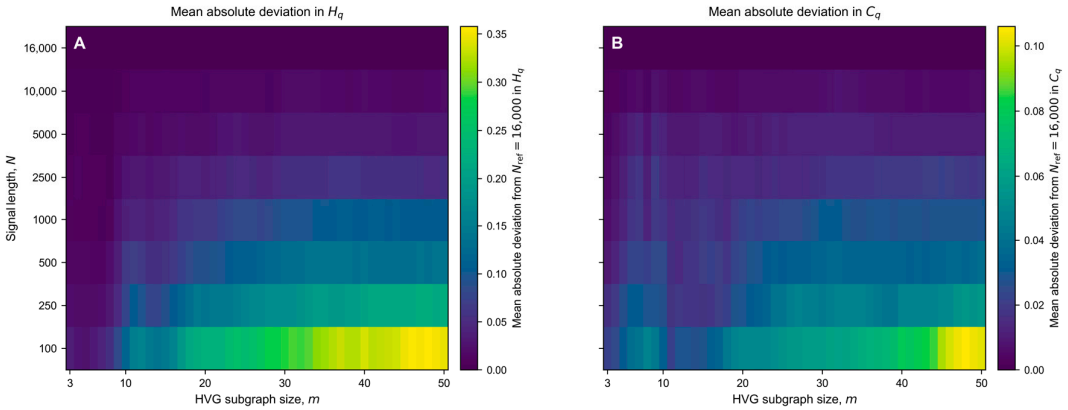


Figure 3. Finite-sample behavior of the grouped-HVG information-theoretic quantifiers across signal length and subgraph size. (**A**) Mean absolute deviation of normalized Tsallis entropy H_q and (**B**) mean absolute deviation of Hellinger-based statistical complexity C_q, calculated across the 19 complete EGG recordings relative to the corresponding values obtained at the common operational reference length $N_{ref} = 16{,}000$ samples. Grouped probability distributions were defined over the fixed theoretical support $S_m = \{g_0, \ldots, g_{m-2}\}$. Rows correspond to the evaluated signal lengths $N = 100, 250, 500, 1000, 2500, 5000, 10{,}000,$ and $16{,}000$, and columns correspond to HVG subgraph sizes $m = 3, \ldots, 50$. The zero deviation at $N = 16{,}000$ follows by definition from its use as the operational reference and should not be interpreted as evidence of asymptotic convergence.

5.2. Comparison with the Ungrouped HVG Representation

The effect of connectivity-based grouping on the representational state space was evaluated by comparing the grouped and ungrouped HVG representations obtained from the same 19 complete EGG recordings across motif sizes $m = 3, \ldots, 50$. In the grouped representation, sequential HVG subgraphs were assigned to connectivity classes according to their number of additional visibility edges, whereas in the ungrouped representation each distinct observed sequential HVG topology was retained as an individual state. This comparison was based on the number of states actually observed in each recording and motif size and therefore does not require enumeration of the complete theoretical support of the ungrouped HVG representation.

The number of observed states increased with motif size in both representations, but the increase was substantially greater for the ungrouped representation. At $m = 3$, both representations contained 2 observed states on average. At $m = 10$, the grouped representation contained 9 observed connectivity classes, whereas the ungrouped representation contained an average of 2843.84 distinct topologies across recordings. This difference increased further at larger motif sizes: at $m = 20$, the corresponding mean numbers of

observed states were 19.00 and 39,742.16, respectively, and at $m = 50$ they were 46.84 and 61,551.47. Thus, connectivity-based grouping markedly reduced the number of distinct observed states required to represent the sequential HVG subgraphs, particularly at larger motif sizes.

The proportion of available sequential subgraphs exhibiting distinct ungrouped topologies also increased with motif size. The mean unique-topology ratio increased from 0.000036 at $m = 3$ to 0.0486 at $m = 10$, 0.6446 at $m = 20$, 0.9288 at $m = 30$, 0.9806 at $m = 40$, and 0.9927 at $m = 50$. At $m = 50$, 18 of the 19 recordings exhibited a unique-topology ratio of at least 0.95, and 6 recordings reached a ratio of 1.0, indicating that almost every sequential subgraph extracted at this motif size had a distinct ungrouped topology in most recordings. This behavior reflects the rapid proliferation of observed topological states in the ungrouped representation within the finite recordings analyzed here and should not be interpreted as saturation of the theoretical ungrouped-HVG state space.

Figure 4 summarizes these differences across the complete range of motif sizes. Figure 4A shows the number of observed states in the grouped and ungrouped representations, whereas Figure 4B shows the proportion of available sequential subgraphs exhibiting unique ungrouped topologies across the 19 recordings. Together, these results quantify the representational effect of connectivity-based grouping: as motif size increases, the ungrouped representation generates a rapidly expanding set of distinct observed topologies, while the grouped representation maps these subgraphs onto a substantially smaller set of connectivity classes. This comparison characterizes state-space reduction and does not, by itself, establish that either representation is more informative or more suitable for EGG analysis.

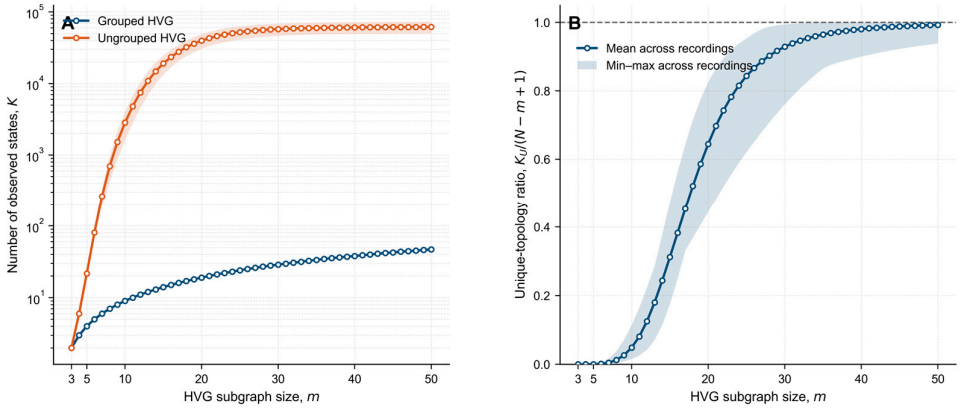


Figure 4. Comparison of grouped and ungrouped HVG representations across subgraph size. (**A**) Number of observed states as a function of HVG subgraph size m for the grouped and ungrouped representations. Lines indicate the mean across the 19 complete EGG recordings, and shaded regions indicate mean ± standard deviation. The vertical axis is shown on a logarithmic scale. (**B**) Proportion of available sequential subgraphs exhibiting distinct ungrouped topologies, calculated as $K_U/(N - m + 1)$, where K_U is the number of distinct ungrouped topologies observed and $N - m + 1$ is the number of sequential subgraphs available for a recording of length N. The line represents the mean across recordings and the shaded region represents the minimum-to-maximum range. The dashed horizontal line at 1 indicates the case in which every available sequential subgraph has a distinct observed topology.

5.3. Comparison with Conventional EGG Spectral Descriptors

Across the 19 control recordings, 2386 valid spectral windows were analyzed, of which 32 (1.34%) were unclassified because their dominant frequencies fell within the gaps between the predefined frequency bands. The number of unclassified windows

ranged from 0 to 20 per recording (0–15.50%). These windows were retained in the denominator when calculating bradygastria, normogastria, and tachygastria percentages, rather than being reassigned to any frequency band. Recording-level counts and percentages are provided in Supplementary Table S1.

The relationship between the grouped-HVG information-theoretic quantifiers and conventional EGG spectral descriptors was evaluated across the 19 complete recordings. For each motif size $m = 3, \ldots, 50$, Spearman correlations were calculated between GTEHVG and GHVGSC, obtained from GPDs defined over the fixed theoretical support, and four conventional EGG descriptors: mean dominant frequency and the percentages of bradygastria, normogastria, and tachygastria. Across the resulting 384 correlations, nine associations between GHVGSC and the percentage of normogastria, corresponding to motif sizes $m = 42 - 50$, reached statistical significance after BH-FDR correction within each family of 48 motif sizes.

For these nine GHVGSC–normogastria associations, Spearman correlation coefficients ranged from ($\rho = -0.5818$) at $m = 42$ to ($\rho = -0.6222$) at $m = 48$, with within-family BH-FDR-adjusted p-values ranging from 0.0393 to 0.0478. Individually estimated 95% percentile bootstrap confidence intervals ranged from $[-0.7990, -0.1585]$ at $m = 42$ to $[-0.8211, -0.2440]$ at $m = 48$, based on 10,000 recording-level resamples. These intervals were not adjusted for multiple comparisons. Importantly, when BH-FDR correction was applied jointly across all 384 correlations as a sensitivity analysis, none remained statistically significant (global adjusted $p = 0.2152$ for the nine GHVGSC–normogastria associations). No associations involving GTEHVG or the remaining conventional EGG descriptors reached significance under the original family-wise correction.

Figure 5 summarizes the correlation patterns across the complete range of motif sizes. The negative GHVGSC–normogastria correlations observed at larger motif sizes suggest an exploratory recording-level association between grouped-HVG statistical complexity and the proportion of spectral windows classified as normogastric within this dataset. However, these associations did not survive the global multiple-testing correction, and their interpretation is further limited by the sample of 19 recordings and the dependence among neighboring motif sizes. Consequently, the findings should be regarded as exploratory and do not establish a causal or physiological mechanism or demonstrate diagnostic or predictive performance. Correlation analysis alone does not fully establish the relationship between graph-based complexity and conventional EGG spectral descriptors. Likewise, the absence of globally significant associations should not be interpreted as evidence of independence between the corresponding measures.

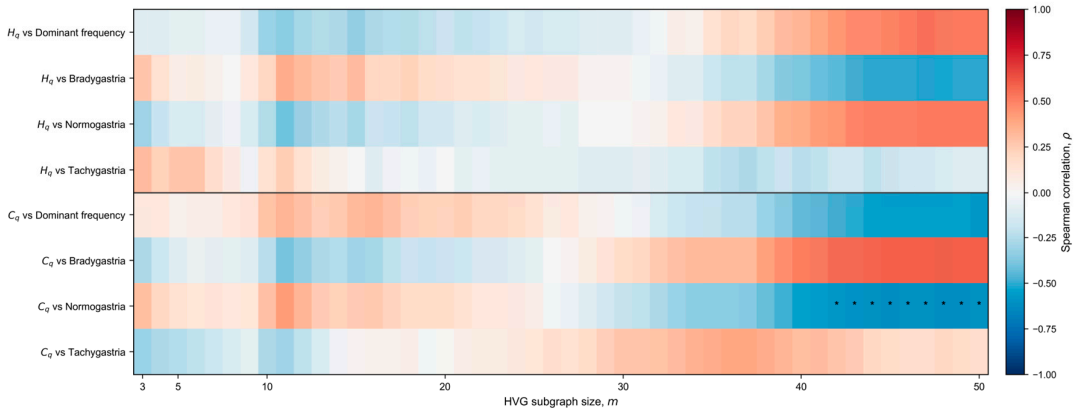


Figure 5. Associations between grouped-HVG information-theoretic quantifiers and conventional EGG spectral descriptors across subgraph size. Spearman correlation coefficients ρ between normalized Tsallis entropy H_q and Hellinger-based statistical complexity C_q, calculated from grouped

probability distributions defined over the fixed theoretical support, and four conventional EGG descriptors: mean dominant frequency and the percentages of bradygastria, normogastria, and tachygastria. Correlations were calculated across the 19 complete EGG recordings independently for each HVG subgraph size $m = 3, \ldots, 50$. Positive and negative correlations are represented by red and blue tones, respectively. Statistical significance was assessed using the Benjamini–Hochberg false discovery rate correction across the 48 motif sizes within each HVG-quantifier–EGG-descriptor pair. Asterisks indicate FDR-significant associations ($p_FDR < 0.05$), which were observed for C_q versus normogastria at $m = 42 - 50$.

6. Discussion

In this study, grouped information-theoretic metrics derived from the HVG representation were employed, including GSEHVG, GTEHVG, GHVGSCSh, and GHVGSC. GSEHVG and GTEHVG quantify the information content of the GPDs using the Shannon and Tsallis formulations, respectively, whereas GHVGSCSh and GHVGSC combine the corresponding normalized entropy with Hellinger-based disequilibrium. These quantifiers provide complementary information-theoretic descriptions of the probability organization encoded by the GPDs, each defined over the fixed theoretical support associated with the corresponding motif size.

In the analyzed EGG segments, the GPDs exhibit heterogeneous probability profiles whose composition changes with motif size, reflecting scale-dependent differences in the relative occurrence of connectivity groups. Within this framework, Shannon and Tsallis entropies provide complementary information-theoretic characterizations of these distributions. In particular, the entropic index q in the Tsallis formulation controls the relative weighting of probability states: for $q < 1$, lower-probability states receive greater relative weight, whereas the limit $q \to 1$ recovers the Shannon entropy [27].

Therefore, the use of GTEHVG provides a parameter-dependent characterization of the grouped motif distributions without requiring the observed EGG dynamics themselves to be interpreted as nonextensive.

The standard Hellinger distance [28], without additional dimension-dependent normalization, was used in the present implementation to quantify the disequilibrium between each empirical GPD $P^{(s,m)}$ and the uniform reference distribution P_e defined over the same fixed theoretical support. This Hellinger-based disequilibrium, together with the normalized Shannon or Tsallis entropy, was used to construct the corresponding statistical complexity measures. In terms of the generalized formulation introduced in Algorithm 1, Hellinger therefore represents the specific measure $\mathrm{d} \in \mathfrak{D}$ selected for the present implementation.

In this study, $q = 0.25$ was retained as the entropic index [29] for the Tsallis-based formulation, assigning relatively greater weight to less frequent probability states than values of q closer to unity. To examine the dependence of the proposed quantifiers on this parameter, an additional sensitivity analysis was performed for $q = 0.10, 0.25, 0.50,$ and 0.75. For each value of q, normalized GTEHVG and its associated GHVGSC were recalculated for all analyzed motif sizes $m = 3, 4, 5,$ and 8 using the corresponding fixed theoretical support. As summarized in Table 8, the dependence on q varies with motif size rather than following a common monotonic pattern across all scales. For $m = 3$, both normalized GTEHVG and GHVGSC decrease as q increases, whereas for $m = 8$ both quantities increase over the evaluated range. At $m = 4$ and $m = 5$, the variations are smaller and are not uniformly monotonic for both quantifiers. Thus, the sensitivity analysis confirms that the Tsallis-based quantities depend on the selected entropic index, while also showing that the direction and magnitude of this dependence are motif-size specific within the evaluated range. Accordingly, $q = 0.25$ is retained as a reference value

within the low-parameter range explored in previous entropy-based EGG studies [30,31], rather than as an optimized or uniquely preferred value.

Table 8. Sensitivity analysis of normalized GTEHVG and the corresponding Hellinger-based statistical complexity (GHVGSC) for different values of the Tsallis entropic index q. GTEHVG denotes the normalized Tsallis-based grouped entropy, whereas GHVGSC denotes the corresponding statistical complexity computed as the product of normalized GTEHVG and the standard Hellinger distance used as the disequilibrium term. For each motif size, the calculations were performed over the fixed theoretical support of $N_m = m - 1$ structural classes, retaining admissible but unobserved classes with zero probability. Values correspond to the arithmetic means obtained from EGG-1 and EGG-2 at each motif size m.

Motif Size (m)	Metric	q = 0.10	q = 0.25	q = 0.50	q = 0.75
3	Normalized GTEHVG	0.952839	0.891028	0.808207	0.745796
3	GHVGSC	0.222981	0.207869	0.187682	0.172527
4	Normalized GTEHVG	0.740397	0.729170	0.717664	0.713429
4	GHVGSC	0.211523	0.210413	0.210143	0.211479
5	Normalized GTEHVG	0.821413	0.807700	0.794276	0.790257
5	GHVGSC	0.216481	0.212808	0.209759	0.209732
8	Normalized GTEHVG	0.665041	0.666446	0.677372	0.696541
8	GHVGSC	0.284481	0.286633	0.293548	0.303635

For a given motif size m, let $P_g^{(m)} = \left(p_1^{(m)}, \ldots, p_N^{(m)}\right)$ denote the grouped probability distribution used in the information-theoretic calculation. The distribution is defined over the fixed theoretical support $S_m = \{g_0, \ldots, g_{m-2}\}$, containing $N_m = m - 1$ structural classes. Admissible classes that are not observed in a given signal are retained in $P_g^{(m)}$ with zero probability. The associated uniform reference distribution is $P_e = \left(p_1^e, \ldots, p_{N_m}^e\right)$, with $P_i^e = \frac{1}{N_m}$. For both entropy formulations, the normalized entropy H is obtained by dividing the corresponding entropy by its maximum value over this same fixed theoretical support. Accordingly, $H \in [0,1]$, with $H = 1$ for the uniform distribution. The Shannon grouped horizontal visibility entropy and its normalization are defined in Equations (2)–(4), whereas the corresponding Tsallis quantities are given in Equations (5)–(7).

$$\mathrm{GSEHVG}\left(P_g^{(m)}\right) = -\sum_{i=1}^{N_m} p_i^{(m)} * \ln\left(p_i^{(m)}\right) \tag{2}$$

$$\mathrm{GSEHVG}_{max} = \ln(N_m) \tag{3}$$

$$H_{Sh} = \mathrm{GSEHVG}_{norm} = \frac{\mathrm{GSEHVG}\left(P_g^{(m)}\right)}{ln(N_m)} \tag{4}$$

$$\mathrm{GTEHVG}_q\left[P_g^{(m)}\right] = \frac{1 - \sum_{i=1}^{N_m}(p_i^m)^q}{q - 1} \tag{5}$$

$$\mathrm{GTEHVG}_{q,max} = \frac{1 - {N_m}^{1-q}}{q - 1} \tag{6}$$

$$H_q = \mathrm{GTEHVG}_{q,norm} = \frac{\mathrm{GTEHVG}_q\left[P_g^{(m)}\right]}{\mathrm{GTEHVG}_{q,max}} \tag{7}$$

For each motif size m, $N_m = m - 1$ denotes the number of structural classes in the fixed theoretical support S_m, independently of how many of these classes are observed in a particular signal. The corresponding uniform reference distribution is therefore defined

over the same support as $P_e = \left(\frac{1}{N_m}, \ldots, \frac{1}{N_m}\right)$, with $p_i^e = 1/N_m$. The Hellinger distance between the empirical GPD and this uniform reference distribution is then given by

$$D_{He}\left[P_g^{(m)}, P_e\right] = \sqrt{\frac{1}{2}\sum_{i=1}^{N_m}\left(\sqrt{p_i^m} - \sqrt{p_i^e}\right)^2} \quad (8)$$

$$C_{e,He} = H_e D_{He}\left(P_g^{(m)}, P_e\right) \quad (9)$$

In the empirical results reported in the present study, the standard Hellinger distance D_{He} was used directly as the disequilibrium term. Accordingly, the Shannon- and Tsallis-based statistical-complexity measures were computed as $C_{Sh,He} = H_{Sh} D_{He}$ and $C_{q,He} = H_q D_{He}$.

For extensions involving comparisons across probability spaces of different dimensionality, a dimension-normalized Hellinger disequilibrium can alternatively be defined as

$$Q_{He}\left(P_g^{(m)}, P_e\right) = \frac{D_{He}\left[P_g^{(m)}, P_e\right]}{D_{He}^{max}(N_m)} = \frac{D_{He}\left[P_g^{(m)}, P_e\right]}{\sqrt{1 - 1/\sqrt{N_m}}} \quad (10)$$

This alternative normalization yields a disequilibrium measure bounded between 0 and 1, with $Q_{He} = 0$ at the uniform distribution and $Q_{He} = 1$ at the maximally concentrated distribution relative to the uniform reference. Although this dimension-dependent normalization was not used in computing the statistical-complexity values reported in the present study, it provides a suitable extension for future comparisons involving probability distributions defined over different numbers of states.

Taken together, the proposed framework provides a structured information-theoretic characterization of the signal in the visibility domain by combining the probability organization encoded by the GPDs with normalized entropy and Hellinger-based disequilibrium. In the present study, two short EGG segments are used as worked examples to establish and describe the methodological workflow, whereas the expanded evaluation uses 19 complete EGG recordings to assess finite-sample behavior, the representational effect of connectivity-based grouping, and relationships with conventional spectral descriptors.

The expanded methodological evaluation provides a direct comparison between the grouped and ungrouped HVG representations using the same 19 complete EGG recordings and motif sizes $m = 3, \ldots, 50$. The ungrouped representation exhibited a rapid increase in the number of observed topological states as motif size increased, with the proportion of available subgraphs exhibiting distinct topologies approaching unity at larger m. In contrast, connectivity-based grouping retained a substantially smaller set of structural classes. This comparison quantifies the representational effect of the grouping strategy but should not be interpreted as evidence of superior analytical or physiological performance of the grouped-HVG approach. No comparative validation against alternative nonlinear measures, such as sample entropy or permutation entropy, was performed; therefore, the observed reduction in structural classes does not establish superior analytical performance. Furthermore, because motifs with different edge arrangements, including non-isomorphic configurations, may belong to the same connectivity class, the grouped representation does not retain their individual topological identities and cannot fully characterize motif-level structural diversity.

The comparison with conventional EGG spectral descriptors provides an additional reference for examining how the grouped-HVG quantifiers relate to established frequency-domain characteristics of the same recordings. Negative associations between

GHVGSC and the percentage of normogastria were observed at larger motif sizes ($m = 42 - 50$) and reached significance under BH-FDR correction within families of 48 tests. However, none of the 384 correlations remained significant after global BH-FDR correction, indicating that these findings should be interpreted as exploratory. Although the observed patterns suggest a possible relationship between grouped-HVG complexity and conventional spectral characteristics, they do not establish that the two representations provide complementary or independent information. Furthermore, these recording-level associations were obtained from 19 control recordings and do not describe temporal changes within individual recordings or establish a physiological mechanism, diagnostic performance, or predictive value. Further studies involving independent datasets and larger cohorts are needed to determine whether grouped-HVG quantifiers provide information beyond conventional EGG spectral analysis.

Finally, to provide a reproducible structural reference context for the EGG analysis, six synthetic time series representing different dynamical regimes were generated, each with a final length of 100,000 samples. The Logistic map was generated as $x_{(n+1)} = r\, x_{n(1-x_n)}$, with r = 4.0 and $x_0 = 0.4$, whereas the Hénon map was generated as $x_{(n+1)} = 1 - a\, x_n^2 + y_n$ and $y_{(n+1)} = b\, x_n$, with $a = 1.4, b = 0.3$, $x_0 = 0.1$, and $y_0 = 0.3$; only the x component was retained. The Schuster map was generated as $x_{(n+1)} = (x_n + x_n^z)$, with z = 2.5 and $x_0 = 0.7$. For these three deterministic maps, the first 1000 iterations were discarded as transients before retaining the 100,000-point series. Pink noise was generated from zero-mean, unit-variance Gaussian white noise using the fixed IIR filter coefficients $b = [0.049922035, -0.095993537, 0.050612699, -0.004408786]$ and $a = [1, -2.494956002, 2.017265875, -0.522189400]$. The first 1000 filtered samples were discarded and the retained series was standardized to zero mean and unit variance; a fixed random seed of 12,345 was used.

Fractional Gaussian noise ($fGn, H = 0.9$) and fractional Brownian motion ($fBm, H = 0.5$) were generated using the Davies–Harte method with total interval length 1 and fixed random seeds of 23,456 and 34,567, respectively. For fBm, the initial sample was excluded so that the final realization contained 100,000 observations. These parameter settings and generation procedures were fixed before the grouped-HVG analysis to ensure exact computational reproducibility [32–35].

Each synthetic series was then processed using the same grouped-HVG framework over motif sizes $m = 3, \ldots, 50$. At each motif size, the GPD was defined over the fixed theoretical support $S_m = \{g_0, \ldots, g_{m-2}\}$, including zero-probability states, and the normalized Tsallis entropy H_q at $q = 0.25$ and the corresponding statistical complexity $C_{q,He} = H_q D_{He}$, based on the standard Hellinger distance, were calculated. For visualization in Figure 3, each synthetic reference was represented by a single descriptive multiscale point obtained by averaging its 48 entropy and complexity coordinates over $m = 3, \ldots, 50$, i.e., $\bar{H}_q = \left(1/48\right) \sum_{m=3}^{50} H_q(m)$ and $\bar{C}_{q,He} = \left(1/48\right) \sum_{m=3}^{50} C_{q,He}(m)$. These synthetic points provide only a qualitative reference context and are not theoretical bounds of the complexity–entropy plane; likewise, proximity between an EGG point and a synthetic reference should not be interpreted as evidence that the EGG signal belongs to, or is generated by, the corresponding dynamical regime.

Figure 3 shows the complexity–entropy plane constructed from the synthetic reference time series and the EGG signal segments. The yellow triangles represent the arithmetic means of the normalized GSEHVG and GHVGSCSh values obtained from EGG-1 and EGG-2 at motif sizes m = 3,4,5, and 8. Across these motif sizes, mean normalized GSEHVG ranges from 0.69937 to 0.79332, reaching its highest value at $m = 5$, whereas GHVGSCSh ranges from 0.16130 to 0.31585 and reaches its highest value at $m = 8$. These variations describe the motif-size-specific positions of the EGG segments in the Shannon-based complexity–entropy representation.

Because the grouped state space and the set of admissible structural configurations depend on m, these cross-m variations should be interpreted descriptively within the corresponding motif-size-specific representations rather than as direct evidence of an intrinsic increase in signal irregularity. These positions provide a descriptive representation of changes in the grouped connectivity probability distributions as m varies, where m denotes the structural scale defined by the number of consecutive vertices in each induced subgraph.

The purple circles represent the arithmetic means of the normalized GTEHVG and GHVGSC values obtained from EGG-1 and EGG-2 at the same motif sizes using $q = 0.25$. Mean normalized GTEHVG ranges from 0.66645 to 0.89103 across the analyzed motif sizes, whereas GHVGSC ranges from 0.20787 to 0.28664. These motif-size-dependent positions differ from those obtained with the Shannon-based formulation, reflecting the different entropy formulations applied to the same grouped probability distributions. However, these descriptive differences should not be interpreted as evidence of greater sensitivity, discriminatory ability, or superior performance of either entropy formulation. The relative positions of the Shannon- and Tsallis-based EGG representations with respect to the synthetic reference series provide a common graphical context for examining their structural organization across motif sizes within the grouped-HVG framework.

The 28-point EGG segments analyzed above are retained as worked examples for the detailed illustration of the proposed procedure, whereas finite-sample behavior was evaluated separately using the 19 complete EGG recordings. For a segment of length n, the number of overlapping induced subgraphs available at motif size m is $n - m + 1$; consequently, the number of subgraph occurrences contributing to the estimation of the empirical GPD decreases as m increases. In the present examples, this corresponds to 26, 25, 24, and 21 overlapping subgraphs for $m = 3, 4, 5$, and 8, respectively. These subgraphs are used to estimate the connectivity-group frequencies underlying each signal-specific GPD and should not be interpreted as statistically independent observational units.

The synthetic reference series shown in Figure 6 are substantially longer than the two 28-sample EGG segments used as worked examples and therefore provide considerably larger numbers of subgraphs for estimating their corresponding GPDs; accordingly, their locations in the complexity–entropy plane should be interpreted as descriptive references for those worked examples rather than as statistically matched comparisons. Applications aimed at statistical comparisons between experimental conditions should evaluate signal length in relation to the motif-size range of interest and retain the individual signal or subject as the observational unit, because the present finite-sample analysis demonstrates progressively smaller deviations with increasing signal length but does not establish a universally sufficient recording length. The expanded evaluation performed using the complete EGG recordings directly examined the influence of signal length and motif size on the finite-sample behavior of the derived entropy and complexity estimates, showing progressively smaller deviations with increasing signal length while also revealing greater finite-sample sensitivity at larger motif sizes.

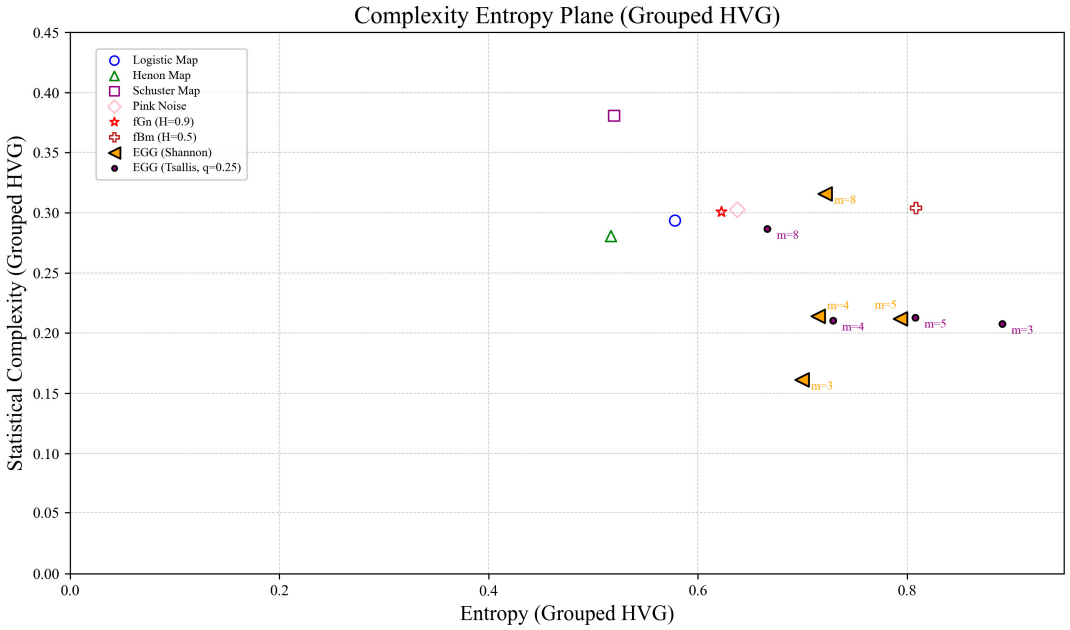


Figure 6. Complexity–entropy plane for the grouped-HVG analysis of the EGG segments and synthetic reference series. EGG triangles represent the mean normalized Shannon entropy (GSEHVG) and the corresponding Hellinger-based statistical complexity (GHVGSCSh), whereas EGG circles represent the mean normalized Tsallis entropy (GTEHVG, $q = 0.25$) and the corresponding complexity (GHVGSC) for motif sizes $m = 3, 4, 5$, and 8. For each synthetic reference series (Logistic map, Hénon map, Schuster map, pink noise, fGn with H = 0.9, and fBm with H = 0.5), the displayed coordinate is the descriptive multiscale mean of the normalized Tsallis entropy and its standard-Hellinger-based statistical complexity across the 48 motif sizes $m = 3, \ldots, 50$. All motif-size-specific GPDs were defined over the fixed theoretical support $N_m = m - 1$, including zero-probability states. Synthetic reference points are included only for qualitative contextualization and do not represent theoretical bounds or classification targets.

7. Conclusions

The present study focuses on the methodological development and evaluation of the proposed approach, combining two short EGG segments used as worked examples with an expanded analysis of 19 complete EGG recordings. Its central contribution is a grouped-HVG framework in which time series are mapped into HVGs, consecutive induced subgraphs are organized according to structural connectivity, and signal- and motif-size-specific GPDs are constructed and retained for subsequent information-theoretic characterization. This procedure preserves the temporal ordering inherent in the HVG representation while enabling the structural organization of the signal to be examined across multiple motif sizes.

The proposed methodology, based on grouped motifs, GPDs, and ITQs, provides a multiscale framework for characterizing the structural organization of EGG signals in the visibility domain. The motif-size-specific GPDs, defined over the corresponding fixed theoretical support, capture changes in the relative occurrence of connectivity groups as m varies, whereas the Shannon- and Tsallis-based formulations provide alternative information-theoretic characterizations of these distributions. Their representation in the complexity–entropy plane provides a common descriptive context for examining the EGG segments alongside synthetic reference series exhibiting different dynamical characteristics. Together with the expanded evaluation using complete EGG recordings, these results provide a methodological basis for subsequent studies designed to assess the reproducibility

and comparative behavior of the HVG–motif–GPD procedure across independent experimental or clinical conditions.

Among the relevant properties of the HVG are its affine invariance and its ability to encode structural relationships among observations at different positions in a time series. Through the proposed procedure—including HVG construction, extraction of consecutive overlapping induced subgraphs, grouping according to structural connectivity, and construction of the corresponding GPDs—the resulting probability organization can be quantified using GSEHVG and GTEHVG entropies and their normalized forms. The normalized entropy is subsequently combined with the standard Hellinger distance used as the disequilibrium term to obtain the corresponding statistical-complexity measures, GHVGSCSh and GHVGSC. These quantifiers provide numerical descriptions of the same motif-size-specific GPDs under the respective Shannon- and Tsallis-based formulations.

The resulting quantifiers describe variations in the information-theoretic organization of the GPDs across motif sizes, while their representation in the complexity–entropy plane provides a common descriptive space for contextualization relative to the synthetic reference series. In this framework, the multiscale character refers specifically to changes in the structural scale defined by motif size m, rather than to distinct physiological timescales. Accordingly, the present results demonstrate the application of the grouped-HVG framework to both representative worked examples and complete EGG recordings, providing motif-size-dependent quantitative descriptions and an expanded methodological evaluation of finite-sample behavior and state-space representation, without establishing physiological discrimination or diagnostic performance.

The expanded finite-sample analysis confirmed that the grouped-HVG quantifiers are sensitive to the amount of signal available for estimating the GPDs, particularly as motif size increases. Deviations in GTEHVG and GHVGSC relative to the common operational reference decreased progressively with increasing signal length, while larger motif sizes generally exhibited greater finite-sample sensitivity for shorter signals. These findings emphasize that signal length and motif size should be considered jointly when interpreting the resulting quantifiers. The overlapping subgraphs contribute collectively to the construction of each recording- and motif-size-specific GPD and are not treated as statistically independent observational units.

The comparison with conventional EGG spectral descriptors revealed negative associations between GHVGSC and the percentage of normogastria at larger motif sizes ($m = 42 - 50$), which reached significance under within-family BH-FDR correction but did not survive global correction across all 384 correlations. These exploratory findings suggest a possible recording-level relationship between grouped-HVG complexity and a conventional spectral descriptor, without establishing physiological equivalence, independence, or diagnostic value. Moreover, because the expanded methodological evaluation was restricted to EGG recordings from control participants, the results do not provide direct evidence of the framework's ability to distinguish between physiological or pathological conditions.

The validity of the present methodological evaluation should be considered within its specific scope. Construct validity remains constrained by the fact that the proposed quantifiers characterize grouped-HVG representations and therefore should not be interpreted as direct measures of physiological or pathological properties. External validity is likewise limited because the expanded evaluation involved recordings from a single participant category and does not establish generalizability to broader populations, independent cohorts, or different clinical conditions. Consequently, the present findings cannot be generalized to patients with gastrointestinal disorders without further evaluation and independent validation in appropriately characterized clinical populations.

The synthetic reference series used for contextualization in the complexity–entropy plane were substantially longer than the two 28-sample EGG segments used as worked examples; consequently, their locations should be regarded as descriptive references for those examples rather than as statistically matched comparators. Further studies should extend the present finite-sample evaluation to independent datasets and examine the reproducibility of the estimated GPDs and derived quantifiers across different recording characteristics, populations, and experimental or clinical conditions.

On the basis of the methodological developments presented in this study, several directions for future research can be identified. The explicit retention of the signal- and motif-size-specific GPDs enables subsequent studies to evaluate alternative entropy formulations and disequilibrium or distance measures within the same grouped-HVG representation, without modifying the preceding graph and motif-grouping stages. Further work may also investigate alternative visibility-graph architectures and grouping strategies, as well as broader sets of synthetic reference series for descriptive contextualization.

The present comparison with conventional EGG spectral descriptors provides an initial basis for exploring possible relationships between structural and frequency-domain descriptions of the same recordings. Future studies should determine whether the exploratory GHVGSC–normogastria associations are reproducible in larger independent datasets and whether grouped-HVG quantifiers provide information complementary to conventional spectral descriptors across different physiological or clinical conditions. Comparisons may also be extended to additional established time-series measures, including sample entropy and permutation entropy. Such studies should employ appropriately designed subject-level analyses and independent validation to distinguish complementary information from redundancy between analytical representations.

Finally, integration with complementary time-series analysis and machine-learning approaches could be explored in future studies to determine whether grouped-HVG representations provide informative features beyond those captured by conventional signal descriptors. Such applications would require independent validation and appropriately designed predictive analyses before any conclusions regarding classification or diagnostic performance could be established.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/bioengineering13111302/s1: Supplementary Code S1, validate_hvg_motif_tables.py: Python script for the independent validation of HVG motif configuration counts in Tables 1 and 2 and verification of the illustrative EGG results in Tables 3 and 4; Supplementary Table S1, Supplementary_Table_S1.csv: Recording-level counts and percentages of unclassified spectral windows across the 19 control EGG recordings, using anonymized recording identifiers.

Author Contributions: Conceptualization: L.G.G.A., and M.C.P. Data curation: L.G.G.A. Formal analysis: L.G.G.A. Investigation: M.C.P. and L.G.G.A. Methodology: L.G.G.A. Project administration: M.C.P. Software: L.G.G.A. Supervision: M.C.P. Writing—original draft: L.G.G.A. Writing—review and editing: L.G.G.A. and M.C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Chilean National Agency for Research and Development (ANID), through ANID FONDECYT, Grant No. 1241202.

Institutional Review Board Statement: The study protocol involving the EGG recordings was reviewed and approved by the Scientific Ethics Committee of the Clinical Hospital of the University of Chile (Approval Act No. 184; approval granted on 25 June 2008). The study was conducted in accordance with the applicable national and international ethical principles governing biomedical research involving human participants

Informed Consent Statement: Not applicable, as the manuscript does not contain any identifiable personal data. Informed consent was obtained in writing from all participants before the EGG records were acquired.

Data Availability Statement: The datasets analyzed in this study consist of previously acquired EGG recordings from human participants, including the representative signal segments used as worked examples and the complete recordings used for the expanded methodological evaluation. Owing to ethical and privacy restrictions related to human physiological data, the raw recordings are not publicly available. Access requests may be directed to the corresponding author and will be considered subject to applicable ethical and institutional requirements and approvals. The software implementation developed for this study forms part of an ongoing collaborative research project involving multiple institutions and is therefore not publicly available at this stage. Nevertheless, to facilitate independent implementation of the proposed methodology, the manuscript provides the complete analytical workflow, detailed pseudocode, representative signal segments, mathematical formulations, parameter settings, motif search strategy, grouped probability distribution (GPD) construction, normalization procedures, and entropy and complexity calculations.

Acknowledgments: The authors thank the Clinical Hospital of the University of Chile for facilitating access to the EGG recordings used in this study. We especially thank Ana María Madrid Silva for her valuable support in facilitating access to these recordings.

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

References

1. Gómez Acosta, L.G.; Chacón Pacheco, M. Entropy and Statistical Complexity in Bioelectrical Signals: A Literature Review. *Signals* **2025**, *6*, 7. https://doi.org/10.3390/signals6010007.
2. Choudhary, G.I.; Aziz, W.; Khan, I.R.; Rahardja, S.; Franti, P. Analysing the dynamics of interbeat interval time series using grouped horizontal visibility graph. *IEEE Access* **2019**, *7*, 9926–9934. https://doi.org/10.1109/ACCESS.2018.2890542.
3. Chang, F.Y. Electrogastrography: Basic knowledge, recording, processing and its clinical applications. *J. Gastroenterol. Hepatol.* **2005**, *20*, 502–516. https://doi.org/10.1111/j.1440-1746.2004.03751.x.
4. Schamberg, G.; Varghese, C.; Calder, S.; Waite, S.; Erickson, J.; O'Grady, G.; Gharibans, A.A. Revised spectral metrics for body surface measurements of gastric electrophysiology. *Neurogastroenterol. Motil.* **2023**, *35*, e14491. https://doi.org/10.1111/nmo.14491.
5. Ravetti, M.G.; Carpi, L.C.; Gonçalves, B.A.; Frery, A.C.; Rosso, O.A. Distinguishing noise from chaos: Objective versus subjective criteria using horizontal visibility graph. *PLoS ONE* **2014**, *9*, e108004. https://doi.org/10.1371/journal.pone.0108004.
6. Lv, D.; Dang, W.; Min, R.; Wu, M.; Guo, W.; Li, M.; Wang, R.; Gao, Z. Multilayer visibility graph-based ordinal network for revealing gas–liquid nonlinear dynamic flow behaviors. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–10. https://doi.org/10.1109/TIM.2024.3413135.
7. Liu, H.R.; Li, M.X.; Zhou, W.X. Visibility graph analysis of the grains and oilseeds indices. *Phys. A* **2024**, *650*, 130004. https://doi.org/10.1016/j.physa.2024.130004.
8. Ma, P.; Liang, W.; Zhang, H.; Wang, C.; Li, X. Multiscale permutation entropy based on natural visibility graph and its application to rolling bearing fault diagnosis. *Struct. Health Monit.* **2025**, *24*, 313–326. https://doi.org/10.1177/14759217241229999.
9. Mohammadpoory, Z.; Nasrolahzadeh, M.; Haddadnia, J. Epileptic seizure detection in EEGs signals based on the weighted visibility graph entropy. *Seizure* **2017**, *50*, 202–208. https://doi.org/10.1016/j.seizure.2017.07.001.
10. Wen, T.; Chen, H.; Cheong, K.H. Visibility graph for time series prediction and image classification: A review. *Nonlinear Dyn.* **2022**, *110*, 2979–2999. https://doi.org/10.1007/s11071-022-08002-4.
11. Muñoz-Diosdado, A.; Aguilar-Molina, A.M.; Solis-Montufar, E.E.; Zamora-Justo, J.A. Analysis of aftershocks from California and synthetic series by using Visibility Graph Algorithm. *Entropy* **2025**, *27*, 178. https://doi.org/10.3390/e27020178.
12. Mohammadpoory, Z.; Nasrolahzadeh, M.; Amiri, S.A. Patient-independent epileptic seizure detection using weighted visibility graph features and wavelet decomposition. *Multimed. Tools Appl.* **2025**, *84*, 3197–3221. https://doi.org/10.1007/s11042-025-20594-8.

13. Mohammadpoory, Z.; Nasrolahzadeh, M.; Amiri, S.A.; Haddadnia, J. A non-invasive approach for early Alzheimer's detection through spontaneous speech analysis using deep visibility graphs. *Cogn. Comput.* **2025**, *17*, 42. https://doi.org/10.1007/s12559-024-10398-7.
14. Gao, M.; Ge, R. Mapping time series into signed networks via horizontal visibility graph. *Phys. A* **2024**, *633*, 129404. https://doi.org/10.1016/j.physa.2023.129404.
15. Luque, B.; Ballesteros, F.J.; Robledo, A.; Lacasa, L. Entropy and renormalization in chaotic visibility graphs. In *Mathematical Foundations and Applications of Graph Entropy*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2016; pp. 1–39. https://doi.org/10.1002/9783527693245.ch1.
16. Yin, Y.; Wang, W.; Li, Q.; Ren, Z.; Shang, P. Jensen–Shannon divergence based on horizontal visibility graph for complex time series. *Fluct. Noise Lett.* **2021**, *20*, 2150013. https://doi.org/10.1142/S0219477521500139.
17. Nasrolahzadeh, M.; Mohammadpoory, Z.; Haddadnia, J. Small-world networks propensity in spontaneous speech signals of Alzheimer's disease: Visibility graph analysis. *Sci. Rep.* **2025**, *15*, 4860. https://doi.org/10.1038/s41598-025-88947-9.
18. Muñoz-Diosdado, A.; Solís-Montufar, É.E.; Zamora-Justo, J.A. Visibility graph analysis of heartbeat time series: Comparison of young vs. Old, healthy vs. Diseased, rest vs. Exercise, and sedentary vs. Active. *Entropy* **2023**, *25*, 677. https://doi.org/10.3390/e25040677.
19. Choudhary, G.I.; Aziz, W.; Fränti, P. Detection of time irreversibility in interbeat interval time series by visible and nonvisible motifs from horizontal visibility graph. *Biomed. Signal Process. Control* **2020**, *62*, 102052. https://doi.org/10.1016/j.bspc.2020.102052.
20. Azizi, H.; Sulaimany, S. A Review of Visibility Graph Analysis. *IEEE Access* **2024**, *12*, 93517–93530. https://doi.org/10.1109/ACCESS.2024.3401485.
21. Cai, Q.; An, J.; Gao, Z. A multiplex visibility graph motif-based convolutional neural network for characterizing sleep stages using EEG signals. *Brain Sci. Adv.* **2020**, *6*, 355–363. https://doi.org/10.26599/BSA.2020.9050016.
22. Zhan, T.; Xiao, F. A novel weighted approach for time series forecasting based on visibility graph. *Pattern Recognit.* **2024**, *155*, 110720. https://doi.org/10.1016/j.patcog.2024.110720.
23. Zeng, S.; Chen, Y.; Li, X.; Zhu, J.; Shen, Y.; Shiratori, N. Visibility graph entropy based radiometric feature for physical layer identification. *Ad Hoc Netw.* **2022**, *127*, 102780. https://doi.org/10.1016/j.adhoc.2022.102780.
24. Conejero, J.A.; Velichko, A.; Garibo-i Orts, Ò.; Izotov, Y.; Pham, V.T. Exploring the entropy-based classification of time series using visibility graphs from chaotic maps. *Mathematics* **2024**, *12*, 938. https://doi.org/10.3390/math12070938.
25. Madl, T. Network analysis of meditative states in highly skilled meditators using EEG and horizontal visibility graphs. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; IEEE: Piscataway, NJ, USA, 2024; Volume 2024, pp. 1–3. https://doi.org/10.1109/EMBC53108.2024.10782024.
26. Angelidis, A.K.; Goulas, K.; Bratsas, C.; Makris, G.C.; Hanias, M.P.; Stavrinides, S.G.; Antoniou, I.E. Distinction of chaos from randomness is not possible from the degree distribution of the visibility and phase space reconstruction graphs. *Entropy* **2024**, *26*, 341. https://doi.org/10.3390/e26040341.
27. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. https://doi.org/10.1007/BF01016429.
28. Andersen, E.; Chiarandini, M.; Hassani, M.; Janicke, S.; Tampakis, P.; Zimek, A. Evaluation of probability distribution distance metrics in traffic flow outlier detection. In *Proceedings of the 2022 23rd IEEE International Conference on Mobile Data Management (MDM)*; IEEE: Piscataway, NJ, USA, 2022. https://doi.org/10.1109/MDM55031.2022.00030.
29. Gómez Acosta, L.G.; Chacón Pacheco, M. Entropy as a tool for quantifying biomedical signals: Gaps and opportunities in clinical methodological approaches-A scoping review. *Biomed. Eng. Comput. Biol.* **2026**, *17*, 11795972261438666. https://doi.org/10.1177/11795972261438666.
30. Alagumariappan, P.; Krishnamurthy, K.; Kandiah, S.; Ponnuswamy, M.J. Effect of electrode contact area on the information content of the recorded electrogastrograms: An analysis based on Rényi entropy and Teager-Kaiser Energy. *Pol. J. Med. Phys. Eng.* **2017**, *23*, 37–42. https://doi.org/10.1515/pjmpe-2017-0007.
31. Alagumariappan, P.; Rajagopal, A.; Krishnamurthy, K. Complexity analysis on normal and abnormal electrogastrograms using Tsallis entropy. In *Proceedings of the 3rd International Electronic and Flipped Conference on Entropy and Its Applications*; Multidisciplinary Digital Publishing Institute: Basel, Switzerland, 2016; Volume 10. Available online: https://scispace.com/pdf/complexity-analysis-on-normal-and-abnormal-234lro4tuu.pdf (accessed on 15 March 2026).
32. Rosso, O.A.; Larrondo, H.A.; Martin, M.T.; Plastino, A.; Fuentes, M.A. Distinguishing noise from chaos. *Phys. Rev. Lett.* **2007**, *99*, 154102. https://doi.org/10.1103/PhysRevLett.99.154102.

33. Xu, M.; Shang, P.; Zhang, S. Complexity analysis of the time series using inverse dispersion entropy. *Nonlinear Dyn.* **2021**, *105*, 499–514. https://doi.org/https://doi.org/10.1007/s11071-021-06528-7.
34. Zhang, Z.; Wu, J.; Chen, Y.; Wang, J.; Xu, J. Distinguish between stochastic and chaotic signals by a local structure-based entropy. *Entropy* **2022**, *24*, 1752. https://doi.org/10.3390/e24121752.
35. Chagas, E.T.C.; Queiroz-Oliveira, M.; Rosso, O.A.; Ramos, H.S.; Freitas, C.G.S.; Frery, A.C. White noise test from ordinal patterns in the entropy–complexity plane. *Int. Stat. Rev.* **2022**, *90*, 374–396. https://doi.org/10.1111/insr.12487.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.